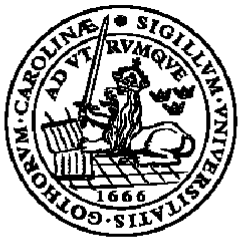
Power Electronics

Devices, Converters, Control and Applications

2019

Mats Alaküla, Per Karlsson and Hans Bängtsson



LUND UNIVERSITY

Table of Contents

1 Converter Basics

In this chapter the basics of power electronics in terms of applications and operation is introduced. First, the reasons for utilizing power electronics are highlighted. Then, the most commonly converter topologies are presented. Here, basic single-phase diode rectifiers are investigated. The investigation continues to also cover three phase thyristor bridges, which forms the classical HVDC link. From thyristor-based converters, the attention is instead focused on transistor-based power electronic converters and the concept of switching is introduced. The classical step-down (Buck), step-up (boost) and Buck-boost converters are introduced. The widely used flyback- and forward-converter principles, derived from the Buck converter, are discussed. From the Buck-boost converter AC-converters are derived. A typical application of three-phase transistor converters is the recently introduced HVDC light concept, based on IGBT technology, which is discussed last in this chapter.

1.1 Why power electronics?

Electric energy is converted to and from other energy forms in many application areas. Motor drives is one, acoustics another, radio a third etc…. In most modern forms of such energy conversion a control process is involved to enhance the result/performance or to facilitate the conversion at all. The hardware used to perform the control is called power electronics, and is constituted of silicon-based components like diodes, transistors, thyristors etc. in various forms. The development of these hardware components is fast; new types appear on the market frequently, pushing up the speed and power with which power electronics can operate. Power electronics is today used almost everywhere, from the drive system in the semi analogue wristwatch or portable DVD-player to large scale filters and power converters in the power system, where the power rating can be 100's of Megawatts.

The motivation to use the power electronic converters varies between applications:

Rotating Machinery. Even the simplest to control type of electrical machine, the DC machine, needs a variable voltage to control the torque or the speed. In more difficult to control types of machines, like synchronous, induction or reluctance machines, a variable voltage and frequency is necessary to accomplish control of quantities like current, torque, flux, speed or even position. Power electronic conversion is a viable solution to all these problems, with very high efficiency and time delays in demanding application that are almost negligible.

Power System. An increased use of nonlinear loads and small scale energy production as well as reduced margins in the transmission capacity of the power system contributes to an increased need of power electronic energy converters, active filters, power flow controllers, DC links etc. Another application area is systems for ride-through capability for sensitive load, i.e. systems that protects loads from short duration irregularities in the supply voltage. 100 ms voltage sag is sometimes enough to stop an expensive production system that can take hours or even days to restart with very high costs as consequence. Power System applications of power electronics are today an emerging market where very high power, very high efficiency and very high reliability is combined.

Power supply. The gigantic market of signal electronics, with mobile phones, PDA's, all kinds of computers and gadgets consumes a large amount of power supplies that are mounted on the PCB together with the signal electronics. Some of the key words in this field are the same as in power systems; high efficiency and very high reliability and a trend is visible where the PCB producers to a growing extent build their own power supplies.

Automotive. A vehicle like a private car contains a large and increasing amount of electronic and electromechanical energy conversion, using power electronics as well of course. Loads like servo steering, servo throttle, seat, mirror and window adjustments, ventilation and even brakes are today already, or soon going to be controlled by electrical motors, a development referred to as "drive by wire". There can be up to 200 electrical motors in one vehicle, but usually only one ICE. In hybrid vehicles there are also electrical machines involved in the traction work with rated power from 10..100 kW. Thus, modern vehicles are already today, and even more so tomorrow, and inferno of power electronic controlled loads.

1.2 The single-phase diode rectifier

The single-phase diode rectifier is used in a wide variety of electronic apparatus. Almost all power electronic converters operate on a rectified AC supply, for example the switch mode power supplies (SMPS) used in computers. In this type of low power applications a voltage source DC link is normally utilized, which means that the DC link is formed by capacitors. In diode and thyristor high power conversion applications, a current source DC link, i.e. inductive, is used instead. Both types are investigated here.

Capacitive DC link

Figure 1.1shows a single-phase diode rectifier with a capacitive DC link. This type of line interface is often used in consumer electronics.

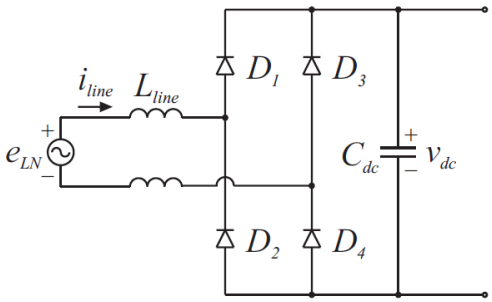


Figure 1.1 A single-phase diode rectifier with a capacitive DC link.

The DC link capacitor stores energy needed for the smoothing action, which means that the voltage across it remains essentially constant. The average DC voltage and the voltage variation is load dependent. For light loads the average DC voltage is close to the peak level of the line-to-neutral voltage and the voltage ripple is very small. For high loads the average DC voltage is equal to the absolute mean of the line-to-neutral voltage and the voltage ripple is considerably higher. For a single-phase diode rectifier

$$V_{dc} = \frac{1}{T/2}\int_{T/2}\hat{e}_{LN}\cos(\omega t)dt = 2\frac{\hat{e}_{LN}}{2\pi}\int_{-\pi/2}^{\pi/2}\cos(\omega t)d(\omega t) = \frac{2}{\pi}\hat{e}_{LN} = \frac{2\sqrt{2}}{\pi}E_{LN} \tag{1.1}$$

Figure 1.2 shows the relation between the line-to-neutral voltage and the DC side voltage.

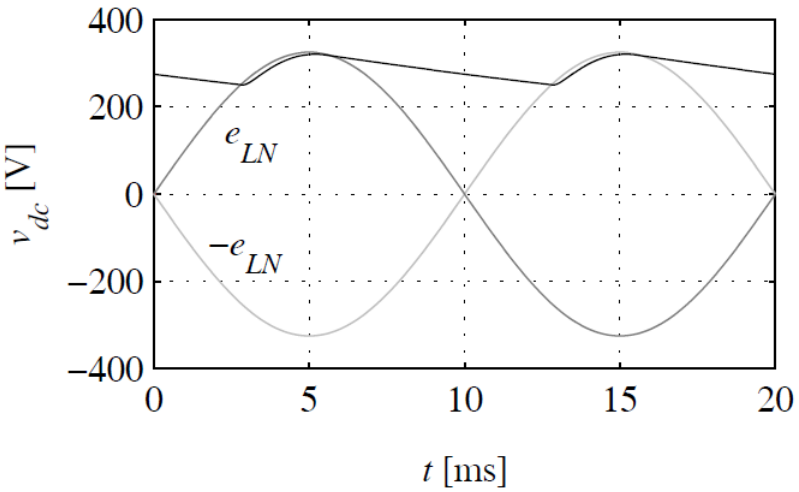


Figure 1.2 Line-to-neutral voltage and DC side voltage for a single-phase diode rectifier with a capacitive DC link.

Note that when $|e_{LN}(t)|$ is higher than v_{dc} two of the rectifier diodes are forward biased and consequently, current can flow. For the positive half-period of e_{LN} the diodes D_1 and D_4 conducts and for the negative D_2 and D_3 conducts. The line current will not be sinusoidal due to the fact that there are time intervals where no diode is forward biased. The line current is shown in Figure 1.3. Figure 1.3 also shows the harmonic spectrum of the line current. The frequencies at which the harmonics appear are predictable. However, the amplitude of each harmonic

component is not possible to calculate in an analytical way, due to the fact that both the width and the amplitude of the current pulses vary with the actual loading of the rectifier. Since the single-phase diode rectifier is common in consumer electronics the harmonic current could cause severe voltage distortion via the line impedance. This distortion could lead to malfunction and even failure of sensitive electronic equipment. Therefore, EMC and EMI regulations and standards have been discussed. In the EU it is not allowed to market TV-sets, computers and computer screens with simple single-phase diode rectifiers equipped only with capacitive DC links. Therefore, most manufacturers equip their single-phase diode rectifier with a power factor corrector (PFC), which consists of a step-up transistor converter connected between the diode rectifier and the DC link capacitor. The transistor is switched at several tens of kHz and is controlled in such a way that the line current contains no low order harmonics and has a power factor of one. This means that the current out from the rectifier has the same shape as the voltage on the DC side of the rectifier.

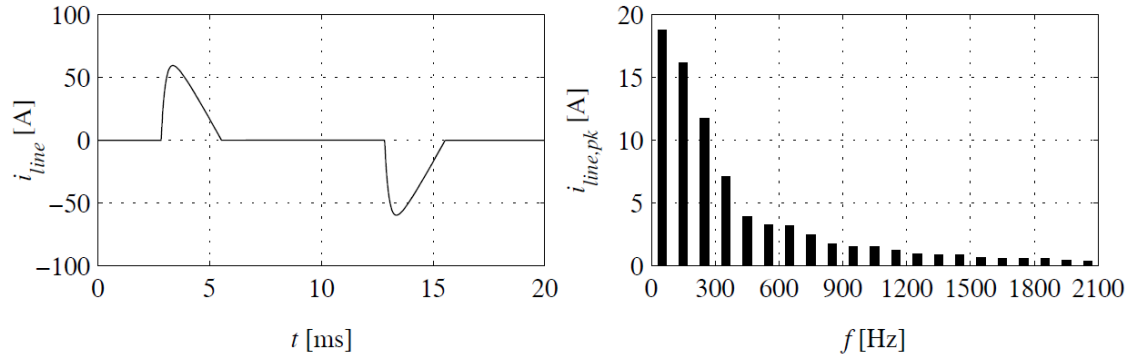


Figure 1.3 *Line current (left) and its frequency spectrum (right), for a single-phase diode rectifier with a capacitive DC link.*

Inductive DC link

Figure 1.4 shows a single-phase diode rectifier with an inductive DC link. Inductive DC links are not common for low power or single-phase diode rectifiers. Instead, inductive DC links has been used for three phase high power applications. Still, it is interesting to investigate the differences compared to the capacitive DC link counterpart. This investigation also forms the basis for the investigation of three phase thyristor converters used for example in HVDC power transmission.

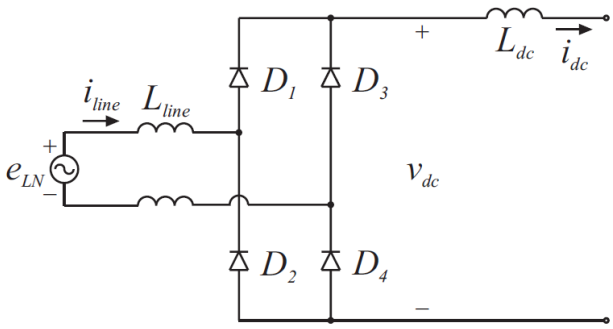


Figure 1.4 *A single-phase diode rectifier with an inductive DC link.*

In the case of an inductive DC link the average voltage at the rectifier output also equals the absolute mean of the line-to-neutral voltage, i.e. it is given by (1.1). However, since the DC link appears as a current source, the DC current is of more importance. If the time constant of the inductor and the load (or its linearized equivalent) is considerably longer than the period time of the grid frequency, the DC current will be essentially constant. In this case the conduction interval of the diodes will be 180° each, and therefore the line current is in principle square-shaped. Figure 1.5 shows a typical line current. Note that the current edges in Figure 1.5 have finite slope, which makes the line current not entirely square. These finite slopes are due to the presence of line inductance on the AC side, slowing down the commutations or transitions between the diodes conducting. Together with the line current, Figure 1.5 also shows the corresponding harmonic spectrum. Again, the frequencies at which the harmonics appear is well defined. For the inductive DC link operated with 180° conduction interval, the amplitudes of the different harmonics are also well defined, due to the fact that the line current is essentially square. Therefore, the amplitude of the harmonics falls off approximately as $1/k$, where k is the harmonic order of the component. The finite slope of the edges of the line current reduce the amplitudes of the higher order harmonics slightly compared to what is expected from a true square shape.

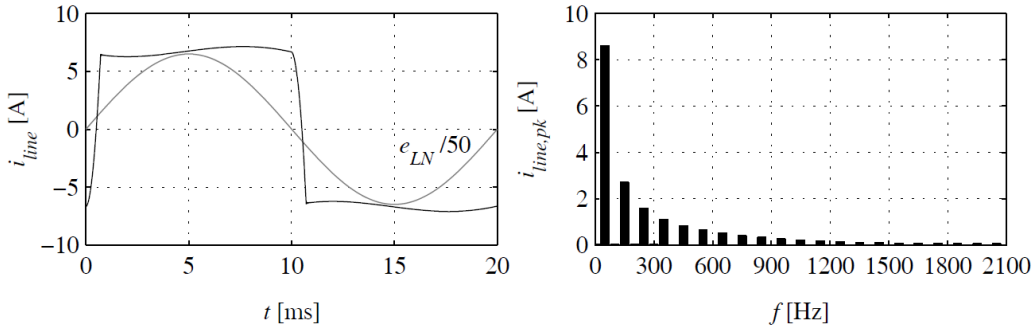


Figure 1.5 *Line current (left) and its frequency spectrum (right), for a single-phase diode rectifier with an inductive DC link.*

The three-phase diode rectifier with capacitive DC link

The three-phase diode rectifier with a capacitive DC link shown in Figure 1.6, is mainly used in the power range from 1 to 100 kW together with voltage source transistor converters. For high power applications the three-phase diode rectifier with capacitive DC link is not desirable due to the large harmonic content consumed, which will be further discussed later on.

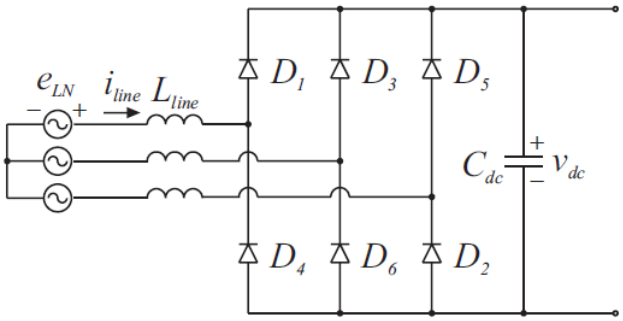


Figure 1.6 Three-phase diode rectifier with capacitive DC link.

The DC side voltage is related to the line-to-line voltage according to Figure 1.7.

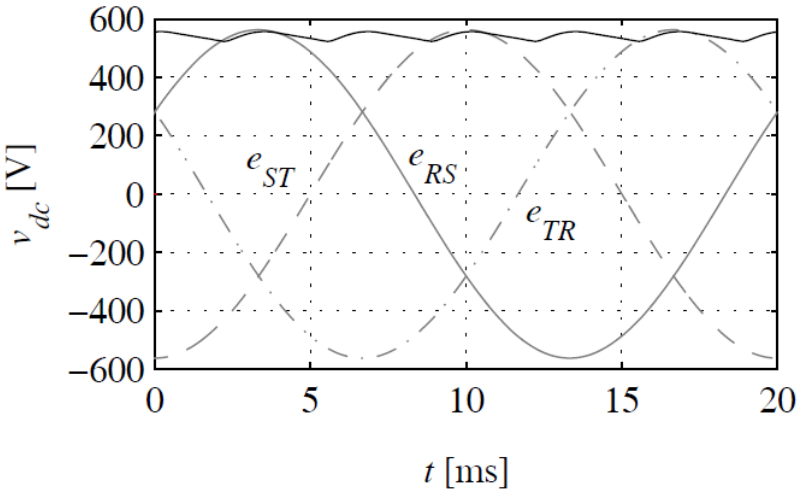


Figure 1.7 The DC side voltage of a three-phase diode rectifier with capacitive DC link.

The operation of the capacitive DC link is similar to the single-phase case but for the three-phase diode rectifier the maximum conduction interval is 60°. Consequently, the DC link voltage ripple has a fundamental frequency equal to six times the line frequency. Therefore, three phase diode and thyristor rectifiers are sometimes referred to as six pulse rectifiers.

The average DC side voltage is calculated from

$$V_{dc} = \frac{1}{T/6}\int_{T/6} \hat{e}_{LL}\cos(\omega t)\,dt = 6\frac{\hat{e}_{LL}}{2\pi}\int_{-\pi/6}^{\pi/6}\cos(\omega t)\,d(\omega t) = \frac{3}{\pi}\hat{e}_{LL} = \frac{3\sqrt{2}}{\pi}E_{LL} \qquad (1.2)$$

The typical line current resulting from this type of rectifier is shown in Figure 1.8 Note that the current pulses seen in Figure 1.8 are a consequence of the

capacitor charging current. It is obvious that the harmonic content of the line current is high, which is also shown in Figure 1.8.

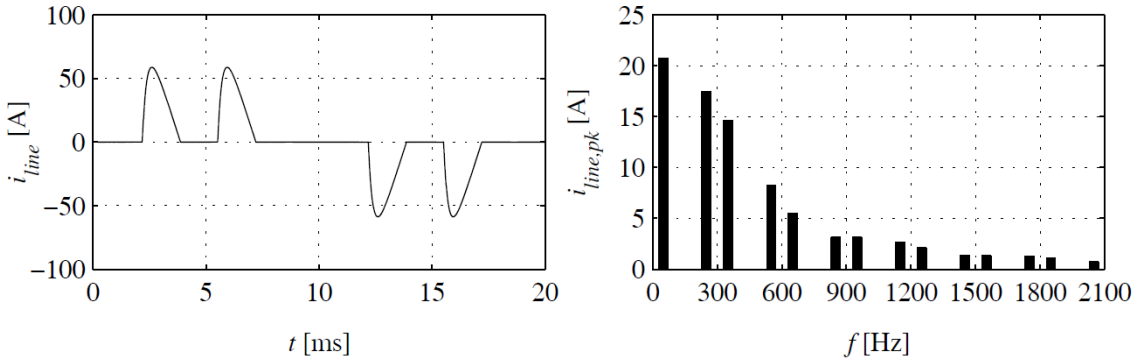


Figure 1.8 *R-phase current (left) and the harmonic spectrum (right) of a three-phase diode rectifier with capacitive DC link.*

Since the diode rectifier is a non-linear element, the amplitude of each harmonic is complicated to compute analytically. Therefore, simulation is the best way to find the harmonic content for a specific design. This can also be used to calculate the harmonic content of the capacitor current needed for proper thermal design. Thermal design of the DC link capacitors is often made based on the fundamental and only a few harmonics. This results in a poor lifetime of the DC link capacitors. Consequently, it is often stated that electrolytic DC link capacitors have a high failure rate and that they are the most common reason of converter failure. However, if the thermal design is made in accordance with the manufacturer recommendations, the estimated lifetime is in most cases longer than 10 years operation.

1.3 The three-phase thyristor rectifier with inductive DC link

In the case of inductive DC links, the diode rectifier is a special case of the thyristor rectifier, since control angle equal to zero corresponds to diode rectifier operation. Therefore, only the thyristor rectifier shown in Figure 1.9 is investigated here.

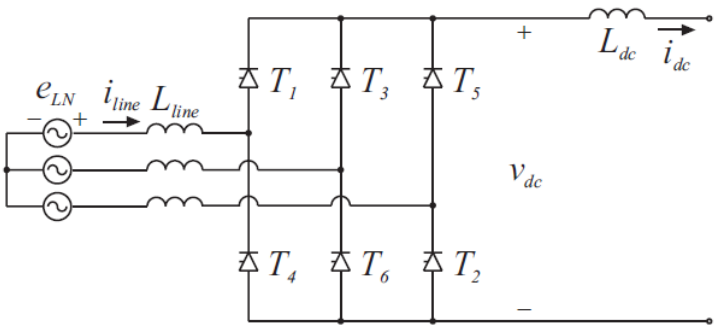


Figure 1.9 *Three-phase thyristor rectifier with an inductive DC link.*

The average DC link voltage for different levels of firing angle α is calculated from

$$V_{dc} = \frac{1}{T/6}\int_{T/6}\hat{e}_{LL}\cos(\omega t)\,dt = 6\frac{\hat{e}_{LL}}{2\pi}\int_{-\pi/6+\alpha}^{\pi/6+\alpha}\cos(\omega t)\,d(\omega t) = \frac{3}{\pi}\hat{e}_{LL}\cos(\alpha) = \\ = \frac{3\sqrt{2}}{\pi}E_{LL}\cos(\alpha) = V_{dc0}\cos(\alpha) \tag{1.3}$$

where V_{dc0} is the average DC link voltage in the case of a three phase diode rectifier. Note α is defined as the angular displacement from the angle where the corresponding diode rectifier would have experienced a commutation, to the angle where it takes place. For a single-phase diode rectifier this angle coincides with the zero crossing of the phase voltage. For a three-phase rectifier this angle is where two line-to-line voltages v_{LL} intersect.

The centre of the line current pulse is displaced, by an angle equal to the control angle, relative to the line voltage. This means that the line current exhibits a phase-lag given by $\varphi=\alpha$. Therefore

$$\begin{cases} P = \sqrt{3}E_{LL}I_{line,1}\cos(\alpha) \\ Q = \sqrt{3}E_{LL}I_{line,1}\sin(\alpha) \end{cases} \tag{1.4}$$

Assuming a constant DC side current, I_{dc}, the average DC power is given by

$$P = V_{dc}I_{dc} = \frac{3\sqrt{2}}{\pi}E_{LL}I_{dc}\cos(\alpha) \tag{1.5}$$

If the losses are neglected this means that

$$P = \sqrt{3}E_{LL}I_{line,1}\cos(\alpha) = \frac{3\sqrt{2}}{\pi}E_{LL}I_{dc}\cos(\alpha) \;\Leftrightarrow I_{line,1} = \frac{\sqrt{6}}{\pi}I_{dc} \tag{1.6}$$

If the line current is assumed to be square (60° conduction interval) the harmonics are given by

$$I_{line,k} = \frac{2\sqrt{2}}{k\pi}I_{dc}\sin\left(\frac{k\pi}{3}\right) \tag{1.7}$$

This is an important feature of three-phase diode and thyristor rectifiers with inductive DC links. Typical line current spectrums are shown in Figure 1.12 and Figure 1.13 where an HVDC transmission system is investigated.

1.4 HVDC

In this section two three-phase thyristor converters are connected together with an inductive DC link between them. This is the principal structure of an HVDC transmission link. However, in real installations more complicated converter structures are utilized. Also, passive filters for suppression of harmonics and shunt reactive compensators are used. This is further discussed at the end of this section. Figure 1.10 shows a principal schematic of an HVDC transmission system.

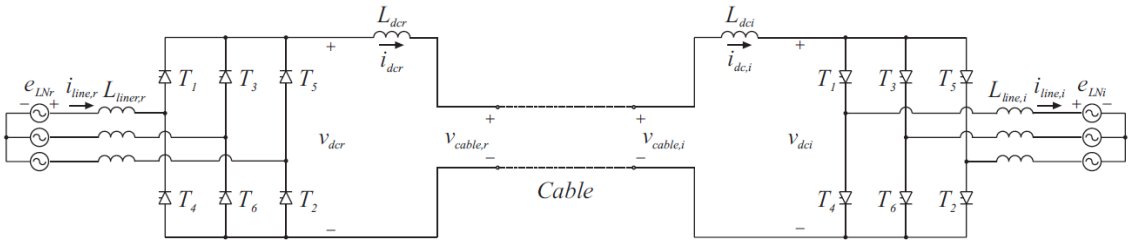


Figure 1.10 Principal schematic of a HVDC power transmission system.

The leftmost converter in Figure 1.10 operates as rectifier and the rightmost as inverter. The rectifier operates at control angles between 0° and 90° and therefore its thyristors are connected in such a way that the power is positive flowing into the DC cable. The inverter operates at control angles between 90° and 180° and therefore its thyristors are connected in such a way that the power is positive flowing out of the DC cable.

Figure 1.11 shows the line-to-line voltages, DC side voltage and control angle for the rectifier and inverter, respectively.

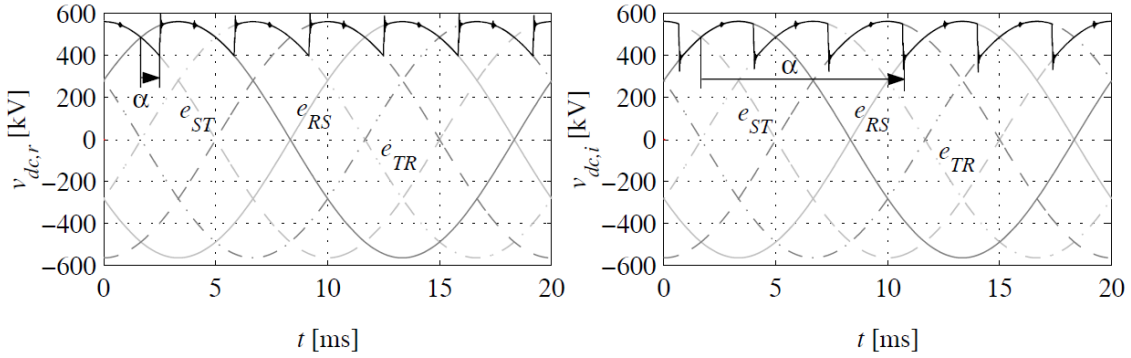


Figure 1.11 Line-to-line voltages, DC side voltage and control angles for the rectifier (left) and inverter (right) of an HVDC transmission system.

In Figure 1.12 and Figure 1.13 the *R*-phase line current and its spectrum are shown for the rectifier and inverter, respectively.

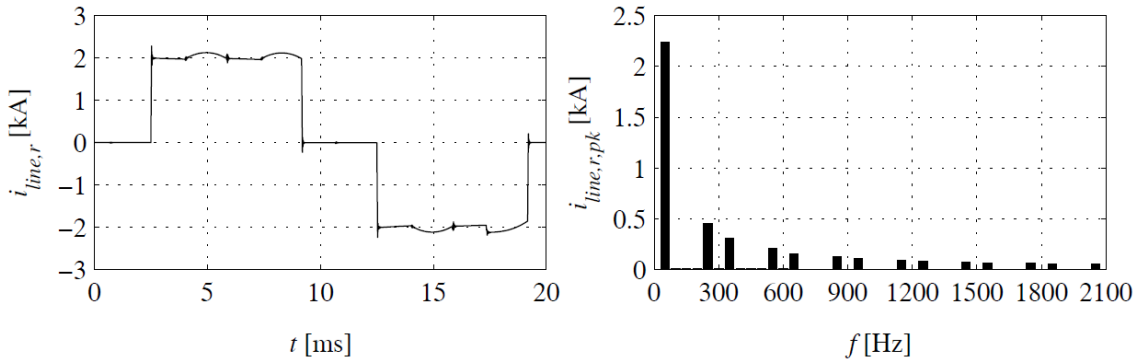


Figure 1.12 R-phase line current (left) and its spectrum (right) for the rectifier of an HVDC transmission system.

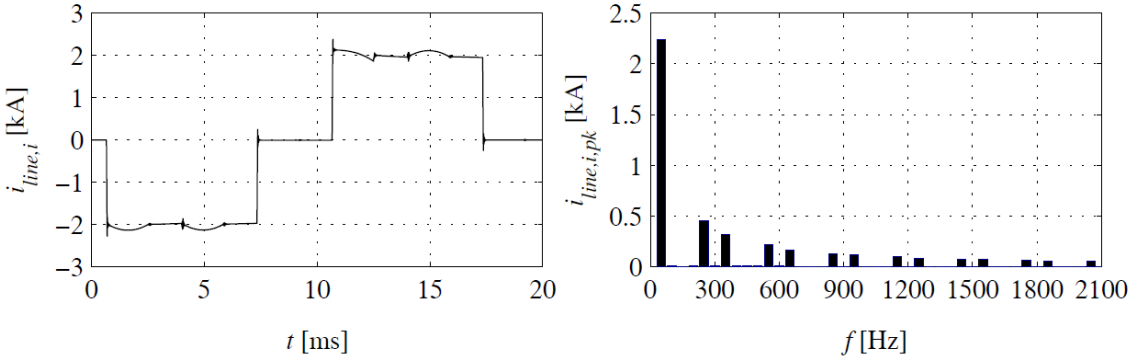


Figure 1.13 R-phase line current (left) and its spectrum (right) for the inverter of an HVDC transmission system.

Note that the commutation overlap affects the spectrum of a diode or thyristor converter, since the edges are slightly smoother due to the finite current derivatives. However the reduction in harmonic content is negligible at least for low order harmonics. Therefore it is still a good approximation to assume that the amplitude of the harmonics fall of as $1/k$ where k is the harmonic order. Figure 1.14: shows a typical commutation between thyristor 1 and 3 of the rectifier.

Note that during the commutation the DC side voltage of the rectifier equals the average of the line-to-line voltage levels of $-e_{TR}$ and e_{ST}. This corresponds to a reduction of the average DC side voltage, which will be calculated in the following.

Assume that during the commutation the DC side current is equal to I_{dc} (constant). Since the thyristors T_{1r} and T_{3r} operates in the on-state simultaneously, for this short period, it is found that

$$i_{T1r}+i_{T13}=0 \;\Rightarrow\; \frac{di_{T1r}}{dt}+\frac{di_{T3r}}{dt}=0 \;\Leftrightarrow\; \frac{di_R}{dt}+\frac{di_S}{dt}=0 \qquad (1.8)$$

This also means that Kirchoff´s voltage law, for the R and S phases, is written

$$e_R-L_{line}\cdot\frac{di_R}{dt}+L_{line}\cdot\frac{di_S}{dt}-e_S=0 \;\Leftrightarrow\; e_R-e_S=e_{RS}=2L_{line}\cdot\frac{di_R}{dt} \qquad (1.9)$$

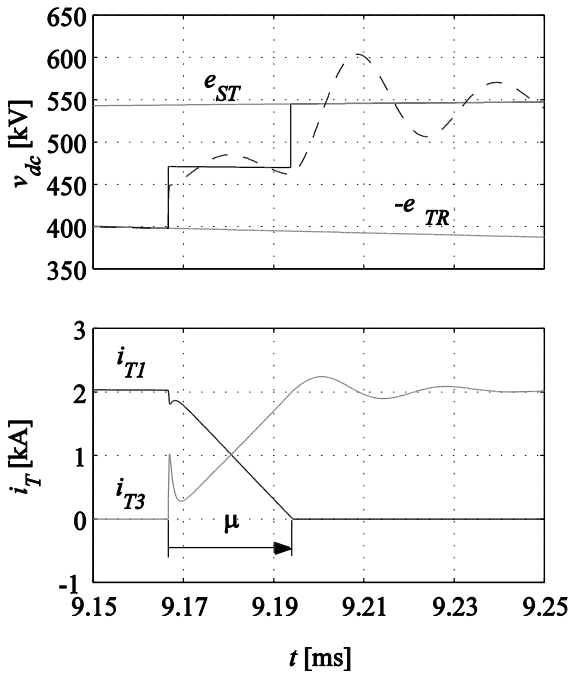


Figure 1.14 Commutation voltage (top) and current (bottom) for the rectifier.

Consequently, the potential of the positive supply rail, during the commutation, equals

$$v_+ = e_R - L_{line} \cdot \frac{di_R}{dt} = L_{line} \cdot \frac{di_S}{dt} - e_S = \frac{e_R + e_S}{2} \tag{1.10}$$

and the rectifier DC side voltage is therefore

$$v_{dc} = v_+ - e_T = \frac{e_R + e_S - 2e_T}{2} = \frac{e_{ST} - e_{TR}}{2} \tag{1.11}$$

whereas for an ideal commutation, the rectifier DC side voltage would equal e_{ST} directly after the triggering of thyristor T_{3r}. Consequently, the commutation represents a loss of voltage-time area and therefore also a loss of average voltage. The average voltage is

$$\begin{aligned} \Delta V_{dc} &= \frac{1}{T/6} \int_{5\pi/6+\alpha/\omega}^{5\pi/6+(\alpha+\mu)/\omega} \left((e_S - e_T) - (v_+ - e_T) \right) dt = \frac{6}{T} \int_{5\pi/6+\alpha/\omega}^{5\pi/6+(\alpha+\mu)/\omega} (e_S - v_+) dt = \\ &= \frac{6}{2\pi} \int_{\alpha}^{\alpha+\mu} \left(\frac{\hat{e}_{LL}}{2} \sin(\omega\tau) \right) d(\omega\tau) = \\ &= \frac{1}{2} \cdot \frac{3\sqrt{2}}{\pi} E_{LL} \left(\cos\alpha - \cos(\alpha+\mu) \right) = \frac{V_{dc0}}{2} \left(\cos\alpha - \cos(\alpha+\mu) \right) \end{aligned} \tag{1.12}$$

Note that the point where the corresponding diode of a diode rectifier, i.e. D_{3r}, starts to conduct is at $\omega t_0 = 5\pi/6 \pm k2\pi$ where k is an integer, which is the reason why this point show up in the integration limits.

The current derivative for the *R*-phase current during the commutation is expressed as

$$\frac{di_R}{dt}=\frac{e_R-e_S}{2L_{line}}=\frac{e_{RS}}{2L_{line}}=\frac{\hat{e}_{LL}}{2L_{line}}\sin\left(\omega t+\frac{\pi}{6}\right) \tag{1.13}$$

The actual current is found by integration. Note that the initial *R*-phase current equals I_{dc}.

$$\begin{aligned} i_R &= I_{dc}+\int_{5\pi/6+\alpha}^{5\pi/6+\alpha+\mu}\frac{\hat{e}_{LL}}{2\omega L_{line}}\sin\left(\omega t+\frac{\pi}{6}\right)d(\omega t)= \\ &= I_{dc}-\frac{\pi}{3\omega L_{line}}\cdot\frac{V_{dc0}}{2}\left(\cos\alpha-\cos(\alpha+\mu)\right)=I_{dc}-\frac{\pi}{3\omega L_{line}}\Delta V_{dc}=0 \end{aligned} \tag{1.14}$$

This gives

$$\begin{aligned} V_{dcr} &= V_{dc0}\cdot\cos(\alpha)-\Delta V_{dc}=V_{dc0}\cdot\cos(\alpha)-\frac{3}{\pi}\omega L_{line}I_{dc}= \\ &= V_{dc0}\cdot\cos(\alpha)-\frac{3}{\pi}X_{line}I_{dc}=V_{dc0}\cdot\cos(\alpha)-R_{dcr}I_{dc} \end{aligned} \tag{1.15}$$

which means that the line inductance acts as an equivalent resistance seen from the DC side. This is true also for the inverter side converter of the HVDC transmission line. By specifying

$$\alpha+\mu=\pi-\gamma \;\Leftrightarrow\; \cos(\alpha+\mu)=-\cos(\gamma) \tag{1.16}$$

for inverter operation it is found that

$$I_{dc}=\frac{\pi}{3\omega L_{line}}\cdot\frac{V_{dc0}}{2}\left(\cos\alpha-\cos(\alpha+\mu)\right)=\frac{\pi}{3\omega L_{line}}\cdot\frac{V_{dc0}}{2}\left(\cos\alpha+\cos(\gamma)\right) \tag{1.17}$$

which gives the inverter side voltage according to

$$V_{dci}=-V_{dc0}\cdot\cos(\gamma)+\frac{3}{\pi}\omega L_{linei}I_{dc}=-V_{dc0}\cdot\cos(\gamma)+R_{dci}I_{dc} \tag{1.18}$$

Practical HVDC installations

Several modifications are made for practical HVDC installations, Figure 1.15, compared to the principal discussed previously. The main reasons for these modifications are related to the shape of the line currents, shown in Figure 1.12 and Figure 1.13. First, the centre of the current pulse is delayed compared to the centre of the sinusoidal pulses of the line voltages. This corresponds to a phase lag with angle $\varphi=\alpha$. This means that the reactive power is consumed. To

circumvent this, switched shunt capacitor banks are installed on the AC line sides of the HVDC transmission.

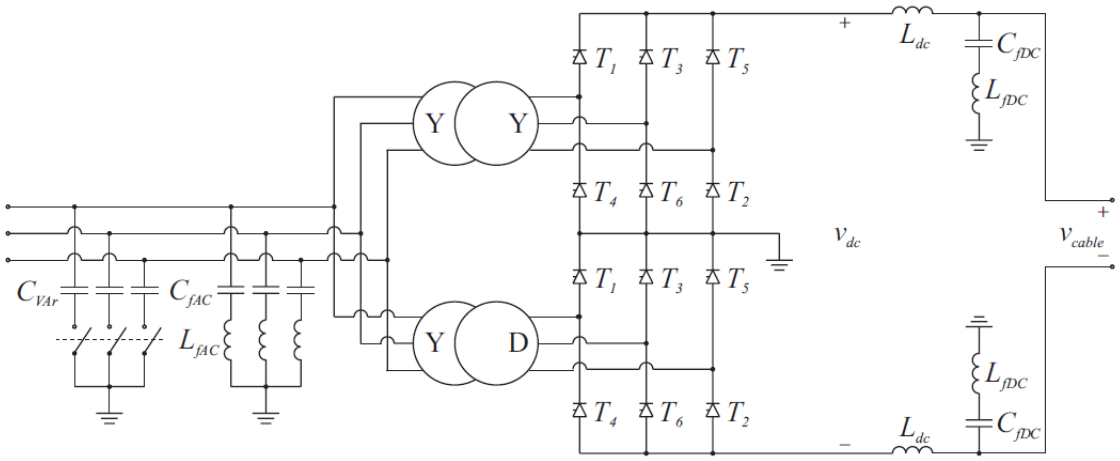


Figure 1.15 Realistic schematic of an HVDC power station.

A switched capacitor bank is formed by several capacitors connected to the line through circuit breakers in such a way that the reactive power delivered can be varied in steps. This is needed since the reactive power consumed by a thyristor rectifier is

$$Q = \sqrt{3} E_{LL} I_{line,1} \sin(\alpha) \tag{1.19}$$

which means that the reactive power consumed is dependent both on the RMS value of the line current fundamental and the control angle. The second issue regarding the AC line currents relates to its harmonic content. Two methods are used to reduce the harmonics. First, twelve pulse thyristor converter structures are used instead of the basic six-pulse thyristor converter discussed here. A twelve-pulse converter has two six pulse bridges with the DC sides connected in series. The six pulse bridges are connected to the same AC grid via a three winding three-phase transformer. One of the thyristor converters is connected to a delta (Δ) winding and the other to a wye (Y) winding. This means that there will be 30° phase-shift between the AC voltage connected to the two converters. This implies that there will be a 30° phase-shift between the currents fed to the converters also. Since the transformer connected to the converters has a common primary side for both secondary windings the line current on the primary side will exhibit a lowest harmonic around the 12^{th} harmonic, i.e. the lowest order harmonics that will show up in the line currents are the 11^{th} and 13^{th}. The next pair of harmonic currents will show up at the 23^{rd} and 25^{th} harmonics. Passive shunt filters tuned to these specific harmonics are also included to further reduce the harmonic content of the AC line currents.

1.5 The Buck (Step Down) converter

The Buck converter is the most simple transistor based power electronic converter. The purpose of the Buck converter (Figure 1.16) is to convert the DC voltage applied between the input terminals to an average output voltage that is lower than or equal to the input voltage. However the instantaneous voltage at the terminal before the output filter (inductor choke) is ideally either equal to the input voltage or zero since the converter operates in switch mode which means that the transistor behaves as a short circuit (on-state) or open circuit (off-state). The purpose of the inductor at the output is to absorb the voltage pulses so that the output voltage after the inductor becomes smooth.

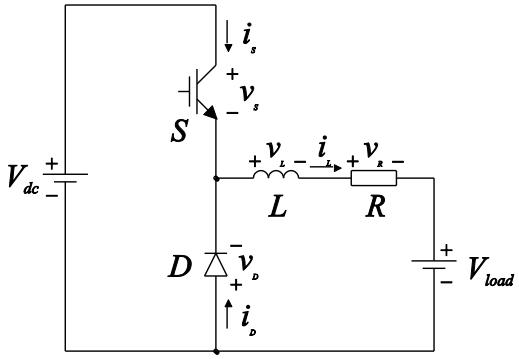


Figure 1.16 Buck converter.

Switching and commutation

If a positive non-zero current flows through the inductor, energy is stored in the inductors according to

$$W_L = \frac{1}{2} L i_L^2 \tag{1.20}$$

For the inductor current to decrease to zero immediately when the transistor is turned off would result in an infinite (or at least very high) negative voltage since for an inductor the voltage equation is

$$v_L = L \frac{di_L}{dt} \tag{1.21}$$

This means that the inductor sets up a voltage across its terminal with such a sign (negative with the references in Figure 1.16) and magnitude that the instantaneous current continues to flow. This would in turn result in an infinite (or at least very high) voltage across the transistor (here denoted "*s*" like in "**s**witch") since

$$v_S = v_{dc} - v_L - v_{load} \tag{1.22}$$

where v_{dc} and v_{load} act as voltage sources and are constant and v_R is disregarded since the resistance of the inductor usually is very low. This would cause a transistor permanent failure. To overcome this problem the freewheeling diode is added. When the transistor turns off the inductor current commutates from the transistor to the freewheeling diode. At first sight it seems strange that the freewheeling diode becomes forward biased at any time instant. However, it is the inductor current that forces the freewheeling diode since the diode voltage is

$$v_D = -v_L - v_{load} \tag{1.23}$$

This means that if the inductor current is forced towards zero with a high derivative, e.g. at transistor turn-off, the inductor voltage becomes very high (negative) which means that the diode becomes forward biased and the inductor currents starts to flow through the freewheeling diode.

Basic operation

As mentioned earlier the instantaneous output voltage before the output inductor is equal to EITHER the DC link voltage V_{dc} OR zero depending on whether the transistor is in the on-state or off-state. For carrier wave modulation (Chapter **Error! Reference source not found.**) the switching frequency is constant and the average output voltage is controlled by varying the ratio between the on-state times in relation to the switching frequency period. This way of controlling the average voltage is commonly referred to as pulse width modulation (PWM) and the ratio is referred to as duty-cycle D. In the following the switching frequency is denoted f_{sw} and the corresponding period is $T_{sw} = 1/f_{sw}$. Figure 1.17 shows ideal waveforms of the Buck converter.

The transistor operates in the on-state for a duration given by DT_{sw} for each switching period. When the transistor conducts the following equation holds

$$V_{dc} = v_S + v_L + v_{load} \tag{1.24}$$

Typically v_S is only a few Volts and $V_{dc} >> v_S$. Therefore can in most cases be neglected, i.e. $v_S = 0$ may be assumed. Also the time constant of the load voltage is usually much longer than the switching period so the load voltage may be treated as constant, i.e. $v_{load} = V_{load}$. Note that the current ripple reaches its maximum at $\Delta t = DT_{sw}$. This means that

$$V_{dc} - V_{load} = v_L = L\frac{di_L}{dt} \;\Rightarrow\; \Delta i_L = \frac{V_{dc} - V_{load}}{L} \cdot DT_{sw} \tag{1.25}$$

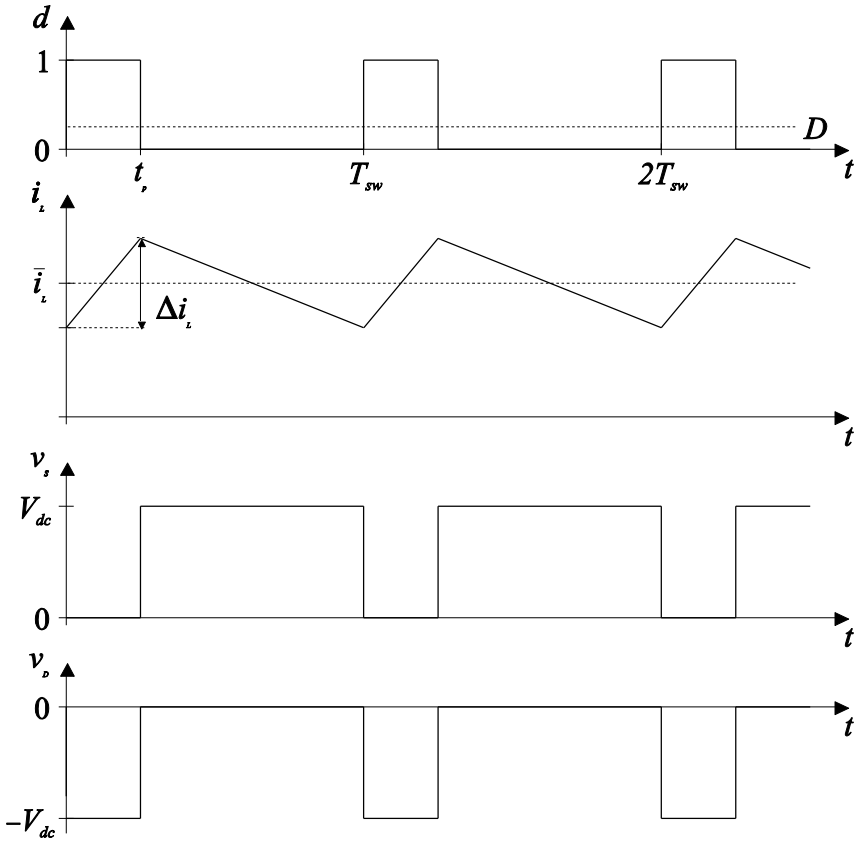


Figure 1.17 Ideal waveforms of the Buck converter.

The transistor operates in the off-state for a time duration equal to (1-*D*)T_{sw} for each switching period. This means that the freewheeling diode carries the inductor current. Consequently, the following equation holds

$$0 = v_D + v_L + v_{load} \tag{1.26}$$

Similarly to the transistor forward voltage drop, the diode forward voltage drop can be ignored (v_D = 0) and the load voltage assumed to be constant which means that

$$-V_{load} = v_L = L\frac{di_L}{dt} \;\Rightarrow\; \Delta i_L = -\frac{V_{load}}{L}\cdot(1-D)T_{sw} \tag{1.27}$$

In steady state, the current ripple must be equal, but with opposite sign, for the transistor on- and off-states. This implies

$$\frac{V_{dc}-V_{load}}{L}\cdot DT_{sw} = -\frac{V_{load}}{L}\cdot(1-D)T_{sw} \;\Leftrightarrow\; D = \frac{V_{load}}{V_{dc}} \tag{1.28}$$

This means that the current ripple is given by

$$\Delta i_L = \frac{V_{dc}-V_{load}}{L}\cdot DT_{sw} = \frac{V_{dc}-V_{load}}{L}\cdot\frac{V_{load}}{V_{dc}}T_{sw} \tag{1.29}$$

The maximum current ripple therefore appears for $V_{load} = V_{dc}/2$. The maximum current ripple is

$$\Delta i_{L,max} = \frac{V_{dc} T_{sw}}{4L} \tag{1.30}$$

Remark

Usually the resistance of the inductive filter and wires is included in the analysis. Typically, in power electronic applications, the time constant of the output circuit, mainly determined by the inductor, is much longer than the switching period, i.e. $\tau_L = L/R_L >> T_{sw}$. This means that the approximate current derivative with respect to time can still be expressed as

$$\frac{di_L}{dt} \approx \frac{\Delta i_L}{\Delta t} \tag{1.31}$$

even if the current increases and decreases according to an exponential function (solution to a first order differential equation) in this case. This can be understood from the MacLaurin expansion of the exponential function, i.e.

$$e^{-t/\tau_L} = 1 + \left(-\frac{t}{\tau_L}\right) + \frac{1}{2!}\left(-\frac{t}{\tau_L}\right)^2 + \frac{1}{3!}\left(-\frac{t}{\tau_L}\right)^3 + \ldots \approx 1 + \left(-\frac{t}{\tau_L}\right) \tag{1.32}$$

for $t << \tau_L$. If this is applied to the solution to the differential equation in the case that the transistor is on and assuming that the initial current is zero the following is obtained

$$i_L(t) = \frac{V_{dc} - V_{load}}{R}\left(1 - e^{-t/\tau_L}\right) \approx \frac{V_{dc} - V_{load}}{R}\left(1 - 1 - \left(-\frac{t}{\tau_L}\right)\right) = \frac{V_{dc} - V_{load}}{L} t \tag{1.33}$$

At $t = DT_{sw}$ the same current ripple as earlier derived results.

The average resistive voltage drop is often included in the duty-cycle, i.e.

$$D = \frac{V_{load} + R i_L}{V_{dc}} \tag{1.34}$$

This last expression shows the difficulties of open-loop converter current control. Typically the DC link voltage is rather high say 500 V and the load voltage may be for example 250 V. The resistance is always low for example 0.1 Ω. In this case a duty-cycle of 0.501 will result in an average inductor current of 5 A and a duty-cycle of 0.502 will result in an average inductor current of 10 A. Therefore, closed-loop current control is usually applied.

1.6 The Boost (Step Up) converter

In the previous section, the Buck converter was introduced as the simplest converter for supplying a load from a higher voltage than the load requires; that is called a step-down converter since the output voltage is lower than or equal to the input voltage.

The Boost converter (Figure 1.18) investigated in this section is a step-up converter which means that the output voltage is always higher than or equal to the input voltage. Typically, the boost converter is used in battery powered applications where high voltage is needed for example for a camera flash.

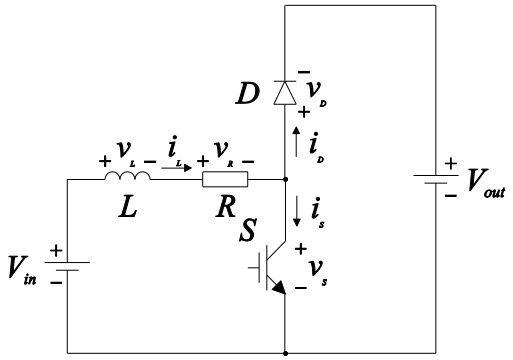


Figure 1.18 Boost converter.

Basic operation

When the transistor *S* is closed, current flows through the inductor on the input side of the converter. The current increases according to the following expression (again assuming zero resistive voltage drop v_R).

$$L\frac{di_L}{dt} = v_L = V_{in} - v_S \approx V_{in} \tag{1.35}$$

This means that the inductor current is increasing and energy is stored in the inductor when the transistor is on. The proportion of the switching period the transistor is on is denoted duty-cycle (*D*) as previous. This means that the current increases during the transistor on-state according to

$$\Delta i_L = \frac{V_{in}}{L} \cdot DT_{sw} \tag{1.36}$$

When the transistor is turned off the inductor current commutates from the transistor to the freewheeling diode for the same reasons as for the Buck converter, i.e. the inductor current forward biases the freewheeling diode. This circuit with the transistor operating in the off-state is described by the following differential equation

$$L\frac{di_L}{dt} = v_L = V_{in} - v_D - V_{out} \approx V_{in} - V_{out} \tag{1.37}$$

which gives the current ripple

$$\Delta i_L = \frac{V_{in} - V_{out}}{L} \cdot (1-D)T_{sw} \tag{1.38}$$

Note that in steady-state operation $V_{out} \geq V_{in}$, meaning that the current derivative is negative in this interval. Therefore it is more convenient rewrite the last expression into

$$\Delta i_L = \frac{V_{in} - V_{out}}{L} \cdot (1-D)T_{sw} = -\frac{V_{out} - V_{in}}{L} \cdot (1-D)T_{sw} \tag{1.39}$$

In steady-state operation the increasing and decreasing current ripples are equal but with opposite sign, i.e.

$$\frac{V_{in}}{L} \cdot DT_{sw} = \frac{V_{out} - V_{in}}{L} \cdot (1-D)T_{sw} \quad \Leftrightarrow \quad D = \frac{V_{out} - V_{in}}{V_{out}} = 1 - \frac{V_{in}}{V_{out}} \tag{1.40}$$

Therefore if a certain duty-cycle is applied the resulting load voltage is

$$V_{out} = \frac{1}{1-D} V_{in} \tag{1.41}$$

Note that the waveforms shown in Figure 1.17 are valid also for the boost converter if V_{dc} in Figure 1.17 is replaced with V_{out}. As a matter of fact the circuits are also equal if V_{dc} and V_{load} of the Buck converter in Figure 1.16 are replaced with V_{out} and V_{in} of the boost converter in Figure 1.18, except for due to the location of the transistor and freewheeling diode, power is transferred from the high voltage side to the low voltage side in the Buck converter and vice versa for the boost converter. This is utilized in the Buck-boost converter.

1.7 The Buck-Boost converter

The Buck-boost converter (Figure 1.19) is a combination of the Buck and boost converters. Despite the name Buck-boost converter it does not provide the possibility to output voltage that can be both higher and lower than the input voltage. Instead, the Buck-boost converter provides the possibility of bidirectional current flow. The name Buck-boost converter comes from that the converter acts as a Buck converter in one load current direction (positive in Figure 1.19) and a boost converter in the other direction (negative load current with reference to Figure 1.19). However, there are power electronic converters that exhibit the possibility of output voltage that can be both higher and lower

than the input voltage, for example the Cûk converter [14] but those are outside the scope of this course.

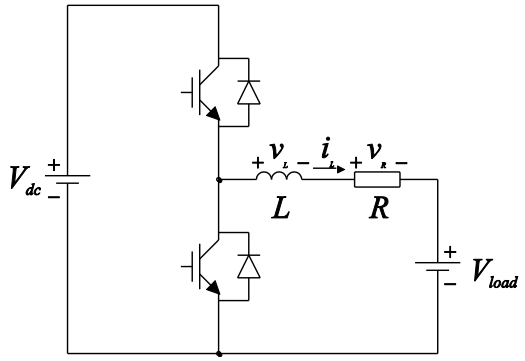


Figure 1.19 Buck-boost converter.

The Buck-boost converter is often used as a module of other converter types where it is termed a transistor half-bridge. The most common example of this is the three-phase transistor converter that is formed by three transistor half-bridges with a common DC link. This type of converter is typically used in electrical machine drives (see Chapters 10, 11and 12) but also in modern power systems applications for example VSC-based HVDC, discussed at the end of this chapter.

1.8 DC-DC Switch-Mode Power Supplies

Switch mode power supplies are common in consumer electronics. For example, the widely spread use of computers have opened an enormous market for switch mode power supplies, since all computers are equipped with one. Several types of communication interfaces also require galvanic isolation, resulting in that both senders and receivers need to be equipped with DC-DC converters providing a galvanically isolated power supply. On the whole, galvanically isolated power supply is required in many applications, both for reasons of personal safety but also for immunity against electrical noise. Figure 1.20 shows the principal schematic of a switch-mode power supply including input rectifier. The input rectifier is usually a single phase diode rectifier which can be equipped with a power factor corrector to avoid low order harmonics being drawn from the power grid. In this chapter only the DC-DC converter of switch mode power supplies are addressed since diode rectifiers are treated in Chapter 4.2). Two widely used switch mode power supplies are investigated, the flyback and the forward converters. Both these converters are equipped with a transformer to provide galvanic isolation and therefore the transformer is treated first.

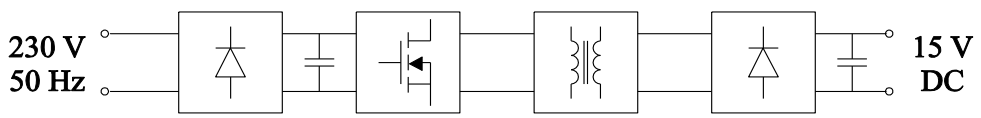


Figure 1.20 Principal schematic of a switch-mode power supply.

The switch-transformer

The transformer that provides galvanic isolation is an important component in both the flyback and forward converters. Both the investigated switch-mode power supplies are primary switched which means that the switching element, i.e. the transistor, is connected to the primary of the transformer so that the transformer transfers the switched voltage. The voltage that is switched is usually rectified mains voltage. For a switch-mode power supply where switching takes place on the secondary side of the transformer the rectifier is instead connected to the secondary. This means that the transformer (needed to provide galvanic isolation) operates at mains frequency, i.e. 50 Hz (in Europe). However, the main reason to use a primary switching is that the transformer can be made geometrically smaller if operating at a high frequency, a fact discussed in more detail in Chapter 5.1**Error! Reference source not found.**. Therefore, switching frequencies from 30 kHz to 10 MHz are often utilized in switch-mode power supplies. Figure 1.21 shows a principal sketch of a transformer with one primary and one secondary winding.

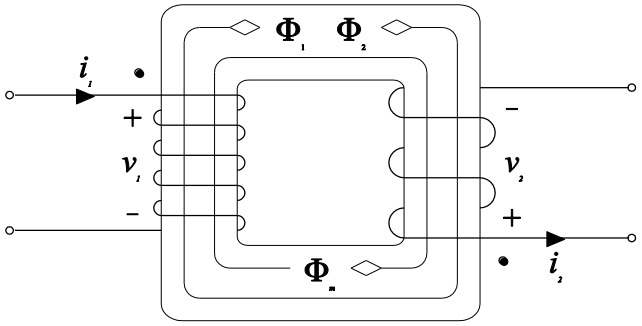


Figure 1.21 Principal transformer with co-ordinated reference directions marked.

Note the sign convention for voltage and current on both sides of the transformer and that the magnetic flux direction is coordinated with the current direction of each winding (Figure 1.21). In Figure 1.21, the core flux Φ_m and the flux through the primary winding Φ_1 since it is assumed that the core is magnetized from the primary. This is however not necessary. The magnetic flux of the windings, Φ_1 and Φ_2, are of equal magnitude but has opposite directions and therefore do not magnetize the core.

For an ideal transformer the magnetizing flux equals zero, i.e. $\Phi_m = 0$, which implies that no energy is stored in the core. Also for an ideal transformer, the leakage inductance and the winding resistance are neglected, which implies that the induced voltage *e* equals the terminal voltage *v*. According to Lenz law, the induced voltage is proportional to the flux linkage Ψ derivative with respect to time, which gives

$$\begin{cases} v_1 \approx e_1 = \dfrac{d\Psi_1}{dt} = \dfrac{d}{dt}\left(N_1 \cdot \Phi_1\right) = N_1 \cdot \dfrac{d\Phi_1}{dt} \\ v_2 \approx e_2 = \dfrac{d\Psi_2}{dt} = \dfrac{d}{dt}\left(N_2 \cdot \Phi_2\right) = N_2 \cdot \dfrac{d\Phi_2}{dt} \end{cases} \tag{1.42}$$

As stated earlier, the magnetic flux of the primary and secondary are equal for an ideal transformer, i.e.

$$\Phi_1 = \Phi_2 \;\Rightarrow\; \frac{v_1}{N_1} = \frac{v_2}{N_2} \;\Rightarrow\; \frac{v_1}{v_2} = \frac{N_1}{N_2} \tag{1.43}$$

Power balance also holds for an ideal transformer since no energy is stored in the core. Thus

$$v_1 \cdot i_1 = v_2 \cdot i_2 \Rightarrow \frac{i_1}{i_2} = \frac{N_2}{N_1} \tag{1.44}$$

A resistance (or impedance in the general case) on the secondary side can be expressed in the quantities of the primary side, and vice versa, according to the relationship

$$R_2 = \frac{v_2}{i_2} = \frac{v_1}{i_1} \cdot \left(\frac{N_2}{N_1}\right)^2 \;\Rightarrow\; R_2' = \frac{v_1}{i_1} = \left(\frac{N_1}{N_2}\right)^2 \cdot \frac{v_2}{i_2} = \left(\frac{N_1}{N_2}\right)^2 \cdot R_2 \tag{1.45}$$

In the expression above <’> is used to denote that the quantity or parameter is transferred to the primary. For a quantity or parameter on the primary that is transferred to the secondary <’’> is used.

In the investigations of the flyback and forward converters, the transformer is considered as being formed by an ideal transformer and a magnetizing inductance L_m. The leakage inductance and winding resistance that usually are included in the equivalent circuit (Figure 1.22) of a transformer are neglected in the analysis.

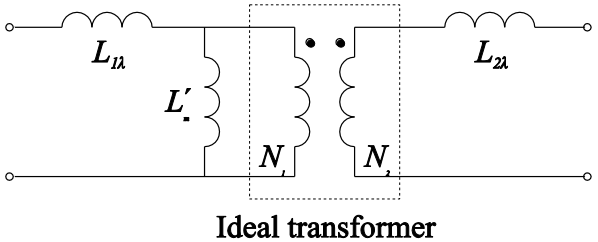


Figure 1.22 *The equivalent circuit of a transformer. Note that the winding resistances have been neglected. In the full equivalent circuit they show up in series with the leakage inductances, L1λ and L2λ. The core loss equivalent resistance has also been omitted. This is located in parallel to the magnetizing inductance.*

Note that especially the leakage inductance is important to consider in design of power electronic equipment, since the high current derivatives with respect to

time induce high voltage and as such expose the power electronic devices, especially the semiconductors, for excessive stress.

The resistive components of the equivalent circuit of the transformer, i.e. the winding resistance and the core loss equivalent resistance can usually be neglected in analysis of power electronic equipment. However, in design of inductive elements (Chapter **Error! Reference source not found.**) they are of course very important. Note that in Figure 1.22 the magnetizing inductance is placed on the primary side of the transformer, even though it cannot be said to belong to any side of the transformer. However, it is convenient to place the magnetizing inductance on the supply side, in this case the primary.

The flyback converter

The principal schematic of a flyback converter is shown in Figure 1.23. In a typical power supply application, the flyback converter is supplied from a diode rectifier connected to the AC mains. Also, protective circuits, snubbers, are included to protect the power semiconductors against over-voltage. The devices that are important for understanding the flyback principle are thus; the transformer, the switch *S*, the flyback diode *D* and the capacitive output filter *C*. Sometimes a more complex *CLC* output filter is utilized to obtain increased damping at high frequencies but to understand the operation a purely capacitive *C* filter is sufficient. The switch *S* is usually a MOSFET due to the high switching frequency (Chapter 4.6**Error! Reference source not found.**) required minimizing the volume of the transformer. Due to the high switching frequency, the transformer core material is usually ferrite (Chapter5.2).

Figure 1.24 shows the simplified flyback converter when the switch *S* is closed. When *S* is closed the voltage across the secondary of the transformer will obtain a polarity so that *D* is reverse biased and consequently operates in the off-state. Therefore, the secondary current equals zero when *S* operates in the on-state. Since the secondary current is zero the flux through the secondary winding is also zero, i.e. $\Phi_2 = 0$. Since the transformer is ideal this means that the flux through the primary winding is also zero, i.e. $\Phi_1 = 0$. However the magnetizing flux is not zero and therefore the magnetizing current is neither zero. Seen from the terminals of the transformer the currents $i_1 = i_m$ and $i_2 = 0$ are flowing, which means that all the current flowing into the transformer (at the primary) magnetizes the core.

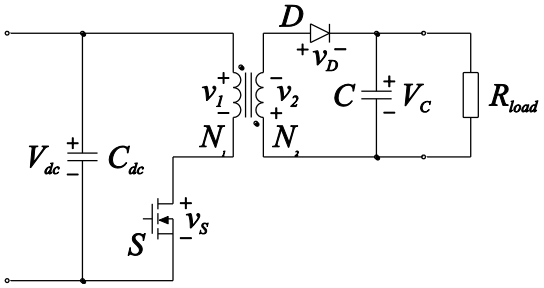


Figure 1.23 *Principal schematic of a flyback converter. Only the devices needed to understand the operation are included.*

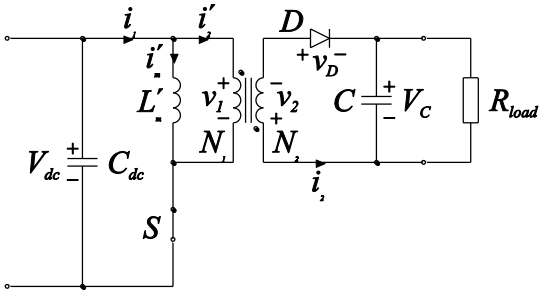


Figure 1.24 *The flyback converter when the switch S is closed. Since the transformer is magnetized from the primary in this case, the magnetizing inductance is placed on the primary.*

Since the transistor is operating in switch-mode it is assumed that the forward voltage drop $V_{S(on)}$ is very low compared to the DC link voltage V_{dc}. For analysing the operation of power electronic converters it is common practice to neglect the on-state forward voltage drop of power electronic devices if $V_{S(on)} << V_{dc}$, which typically holds. Therefore, it is assumed that $V_{S(on)} = 0$ in this case, which implies that the voltage applied across the transformer primary winding equals V_{dc}. Provided that the electrical time constant of the electrical circuit formed on the primary side of the transformer is much longer than the switching frequency fundamental time, i.e. $T_{sw} << \tau_{el}$, the following expression is derived from Kirchoff's voltage law

$$V_{dc} = L'_m \cdot \frac{di'_m}{dt} \approx L'_m \cdot \frac{\Delta i'_m}{\Delta t} \quad \Rightarrow \quad i'_m(t) = i'_m(t_0) + \frac{V_{dc}}{L'_m} \cdot (t - t_0) \quad t > t_0 \qquad (1.46)$$

The requirement on the electrical time constant implies that the resistance of the circuit should be low since $\tau_{el} = L/R$ for an *RL* circuit. In power electronic applications this is almost always true since the losses also should be kept low.

When the switch *S* is turned off the circuit on the transformer primary becomes open (Figure 1.25). Since magnetic energy is stored in the transformer core (in the magnetizing inductance) the magnetic flux transfers to the secondary winding in such a way that its direction and magnitude are preserved. The secondary current (of the transformer terminals) this core flux corresponds to has such a direction that the flyback diode *D* becomes forward biased and enters the on-state. The voltage across the secondary winding (and thereby the primary voltage) is determined by the sum of the voltage across the output filter

capacitor *C* and the forward voltage drop of the freewheeling diode *D*. In this case the voltage drop of the semiconductor may not be neglected since the difference between V_C and $V_{D(on)}$ can be small. In the analysis here it is assumed that $V_{D(on)} << V_C$ for simplicity so that the diode forward voltage drop can be neglected. The magnetizing current of the transformer decreases as specified by the output filter capacitor voltage according to

$$-V_C = L_m'' \cdot \frac{di_m''}{dt} \approx L_m'' \cdot \frac{\Delta L_m''}{\Delta t} \quad \Rightarrow \quad i_m''(t) = i_m''(t_1) - \frac{V_C}{L_m''} \cdot (t - t_1) \quad t > t_1 \tag{1.47}$$

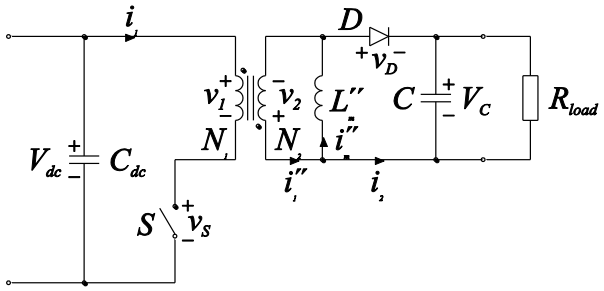


Figure 1.25 *The flyback converter when the switch S is open. Since the transformer is demagnetized from the secondary, the magnetizing inductance is placed on the secondary.*

The initial magnetization current of the secondary is given by transformer winding ratio and the final magnetizing current on the primary, i.e. the current when *S* is turned off

$$L_m' = \left(\frac{N_1}{N_2}\right)^2 \cdot L_m'', i_{m,max}' = \frac{N_2}{N_1} \cdot i_{m,max}'' \tag{1.48}$$

Note that the stored magnetic energy is not altered when the switch is turned off since

$$W_{m,max} = \frac{1}{2} \cdot L_m' \cdot i_{m,max}'^2 = \frac{1}{2} \cdot L_m'' \cdot i_{m,max}''^2 \tag{1.49}$$

Ideal semiconductor voltage stress

When the switch *S* is conducting, the flyback diode *D* must block

$$V_{D,max} = \frac{N_2}{N_1} \cdot V_{dc} + V_C \tag{1.50}$$

When the switch S is blocking it must be able to withstand the sum of the voltage across the output filter capacitor *C* and the forward voltage drop of the freewheeling diode *D*, transferred to the primary. Thus

$$V_{S,max} = V_{DC} + \frac{N_1}{N_2} \cdot \left(V_{D(ON)} + V_C\right) \tag{1.51}$$

This means that *S* has to be able to block a voltage exceeding the DC link voltage. Note that the voltage stress calculated for S and D here should be treated as the absolute minimum level since leakage and stray inductance of the circuit will add significant contributions to voltage stress due to the high current derivatives with respect to time.

Ideal waveforms

The waveforms for ideal operation of a flyback converter are shown in Figure 1.26. Note that current is flowing only on one side of the transformer in each interval. The maximum and minimum levels of the magnetizing current on the primary and secondary shown in Figure 1.26 can be derived from the expressions presented earlier

$$i'_{m,max} = i'_{m,min} + \frac{V_{dc}}{L'_m} \cdot t_p \quad \text{and} \quad i''_{m,min} = i''_{m,max} - \frac{V_C}{L''_m} \cdot \left(T_{sw} - t_p\right) \tag{1.52}$$

Since the energy is preserved during the switching operations, the duty cycle is derived from the magnetizing currents above

$$\frac{t_p}{T_{sw}} = \frac{V_C}{V_{dc} + \frac{N_1}{N_2} \cdot V_C} \tag{1.53}$$

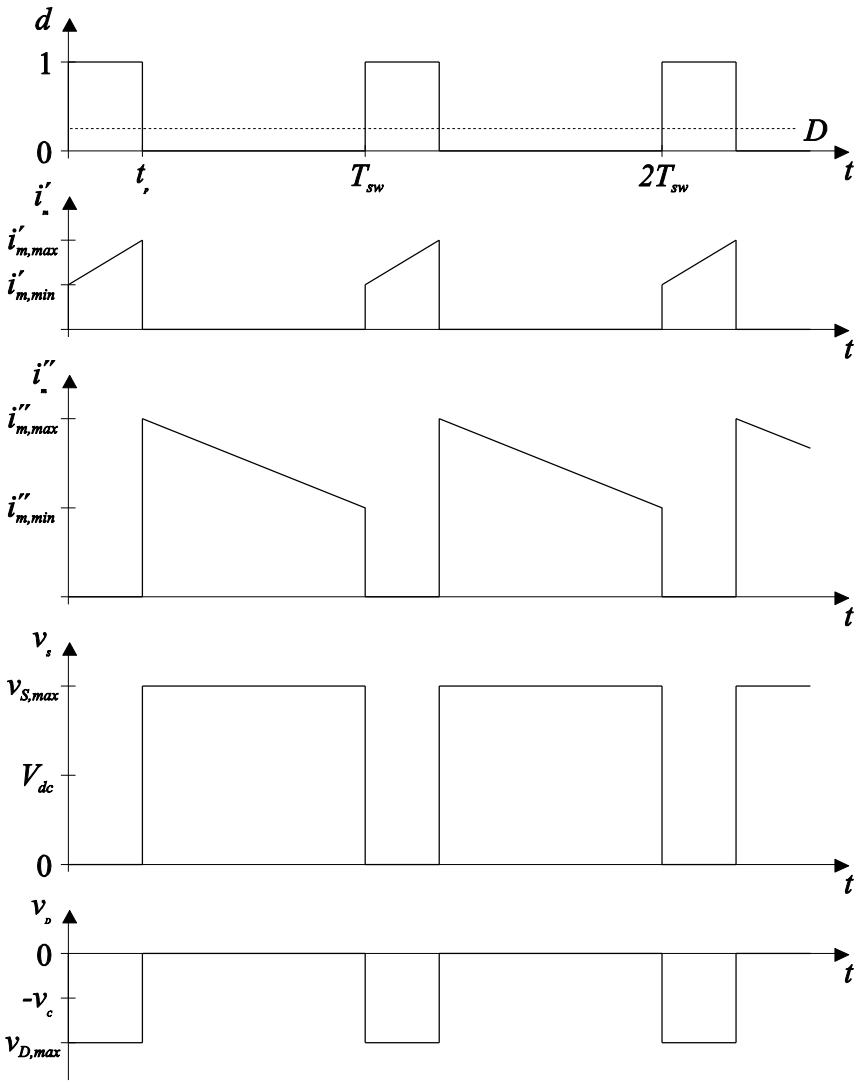


Figure 1.26 *Waveforms for ideal operation of a flyback converter. Note that the switch S must be able to block a voltage that is considerably higher than the DC link voltage also in the ideal case.*

The forward converter

The forward converter has a schematic similar to the one of the flyback converter (Figure 1.27). Several different layouts exist for the flyback converter but only one is presented here, e.g. with a zener diode connected across the transformer primary. Also since the output filter is inductive, a freewheeling diode (D_2) is needed on the secondary. If the output filter is capacitive, this freewheeling diode is not needed. Note the polarity of the diode D_1, which will be forward biased when the switch *S* operates in the on-state. To analyse the forward converter, a transformer equivalent composed of an ideal transformer and magnetization inductance is applied in the same way as in the analysis of the flyback converter. For the forward converter the magnetizing inductance is considered as being located on the primary for the entire switching period. An equivalent circuit for the analysis is shown in Figure 1.28.

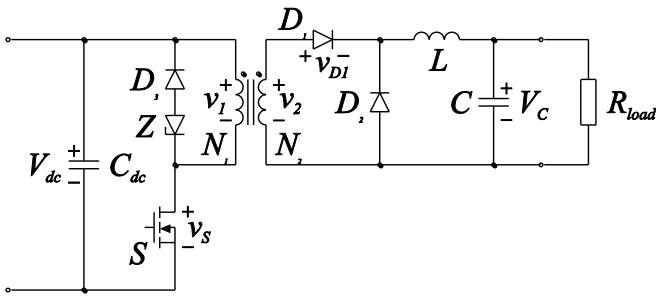


Figure 1.27 *Principal schematic of a forward converter. Only the devices needed to understand the operation are included. Note that in this schematic a zener diode is connected across the primary, which is rarely used in practice.*

When the switch *S* turns on (traverses from the blocking- to the on-state), the secondary voltage will increase so that D_1 becomes forward biased and starts to conduct. The freewheeling diode, D_2, will be reverse biased and therefore operate in the blocking state. If D_2 was conducting immediately prior to the turn-on of *S*, the current commutates from D_2 to D_1. Anyway, turning on S results in that current starts to flow through the secondary winding (of the ideal transformer), which results in a corresponding current flowing in the primary winding (of the ideal transformer). The total primary current of the (real) transformer will also have a component corresponding to the magnetizing current. As in the analysis of the flyback converter, the on-state voltage drop of the switch can in most cases be neglected, i.e. $V_{S(on)} = 0$, which means that the magnetizing current is given by

$$V_{dc} = L'_m \cdot \frac{di'_m}{dt} \approx L'_m \cdot \frac{\Delta i'_m}{\Delta t} \quad \Rightarrow \quad i'_m(t) = i'_m(t_0) + \frac{V_{dc}}{L'_m} \cdot (t - t_0) \quad t > t_0 \qquad (1.54)$$

The total primary current (of the real transformer, i.e. including magnetizing inductance) is

$$i_1 = i_m + i'_2 = i_m + \frac{N_2}{N_1} \cdot i_2 = i_m + \frac{N_2}{N_1} \cdot i_{load} \qquad (1.55)$$

This implies that the primary current will continue to increase as long as the switch S is operating in the on-state. When the switch *S* is turned off, the diode D_1 becomes reverse biased and start to block. The current of the inductive output filter forward biases the freewheeling diode D_2, which means that the load current commutates from D_1 to D_2. This means that an alternative path also must be provided for the magnetizing current. This is provided by the zener diode *Z*. The diode D_3 is included to assure that the transformer primary is not bypassed by the zener when *S* is in the on-state. If it is assumed that the forward voltage drop of D_3 is negligible compared to the zener voltage V_Z, the magnetizing current falls according to

$$-V_Z = L'_m \cdot \frac{di'_m}{dt} \approx L'_m \cdot \frac{\Delta i'_m}{\Delta t} \quad \Rightarrow \quad i'_m(t) = i'_m(t_1) - \frac{V_Z}{L'_m} \cdot (t - t_1) \quad t > t_1 \qquad (1.56)$$

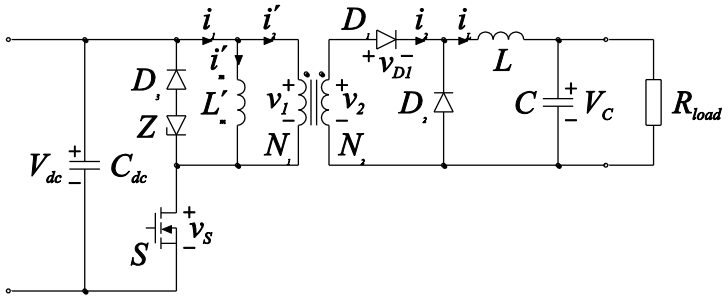


Figure 1.28 The forward converter with the simplified transformer equivalent included.

Ideal semiconductor voltage stress

The switch S must also for ideal conditions, i.e. neglecting leakage and stray inductance, be able to withstand a blocking voltage equal to

$$V_{S,max} = V_{dc} + V_Z \tag{1.57}$$

The diode D_1 must be able to block a voltage

$$V_{D1,max} = \frac{N_2}{N_1} \cdot V_Z \tag{1.58}$$

Note that in the last expression leakage and stray inductance is neglected. Also, the forward voltage drop of the freewheeling diode D_2 is neglected.

Switching waveforms

The switching waveforms of a forward converter are shown in Figure 1.29 below. Assuming that the transformer is completely demagnetized in each switching period, the magnetizing current is

$$i'_m(0) = 0 \;\Rightarrow\; i'_{m,max} = \frac{V_{dc}}{L'_m} \cdot t_p \tag{1.59}$$

$$i'_m(T_{sw}) = 0 \;\Rightarrow\; i'_m(T_{sw}) = i'_m(t_p) - \frac{V_Z}{L'_m} \cdot (T_{sw} - t_p) \tag{1.60}$$

The maximum duty cycle is thus

$$\frac{t_p}{T_{sw}} \leq \frac{V_Z}{V_{dc} + V_Z} \tag{1.61}$$

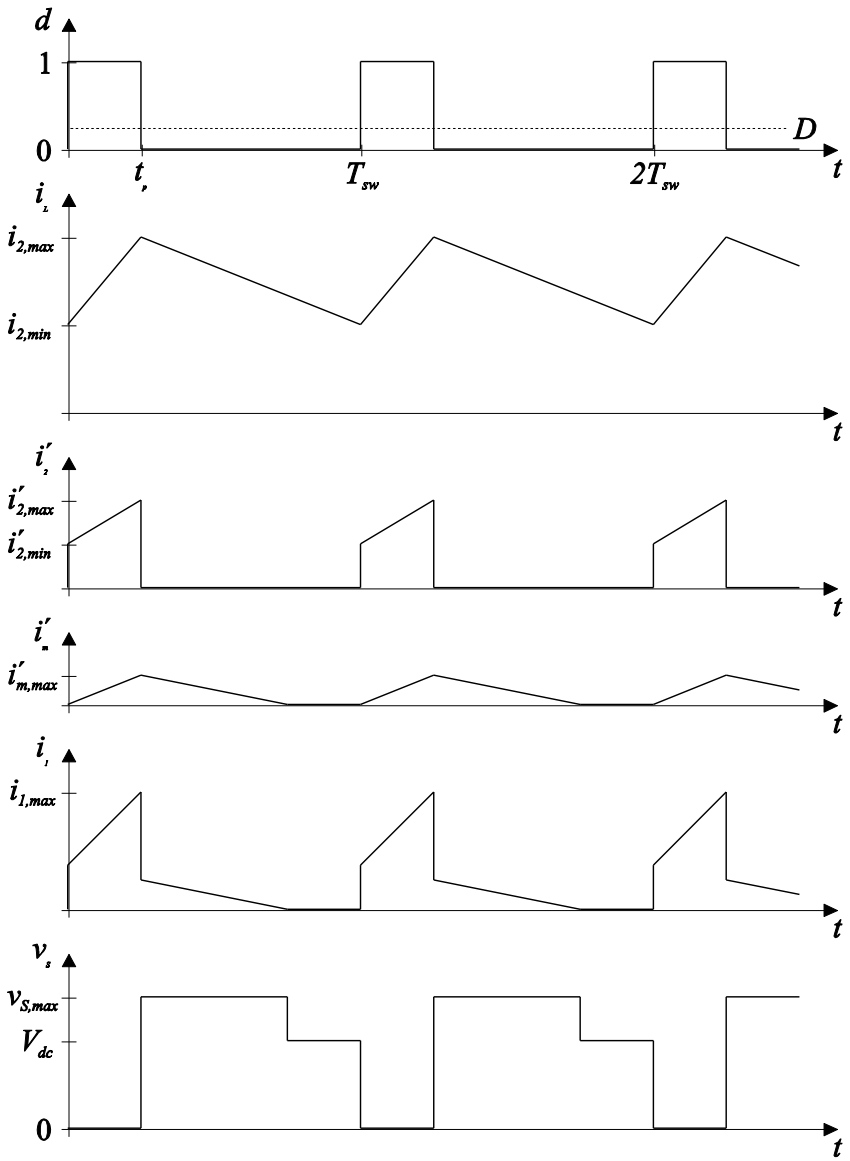


Figure 1.29 Waveforms for ideal operation of a forward converter. Note that the switch S must be able to block a voltage that is considerably higher than the DC link voltage also in the ideal case.

One problem with the forward converter analysed above is that it results in substantial power dissipation in the zener diode, since it operates under high voltage and current simultaneously. In most practical forward converters, a third winding is wound upon the transformer core. This additional winding together with diode replaces Z and D_3.

1.9 AC inverters

Voltage source AC inverters are utilized to transfer DC power into AC power (single- or three-phase). A single-phase AC inverter can be formed by two transistor half-bridges with the load connected between the output terminals of the half-bridges. This arrangement facilitates the possibility to provide the load with bidirectional voltage and current. A three-phase AC inverter is formed by having three transistor half-bridges with one of the phase connectors of the load

connected to each half-bridge output terminal. To show the operation of three-phase AC inverters an application example is investigated: (Voltage source converter (VSC) based HVDC.

HVDC light (trademark of ABB) or HVDC plus (trademark of Siemens) are basically an extension of the IGBT voltage source converter (VSC) operating back-to-back with higher power and voltage levels and with an intermediate cable between the converters. Even though this might seem as small step from a development point of view it has resulted in a very high number of new patents. The main reason for this is that the development of VSC based HVDC has really forced the limits of power electronic converter technology. For example, IGBT valves withstanding more than 100 kV (through series connection) have been developed, light triggered gate or drive circuits have been developed, new cable technologies have emerged and so on. Of course this introduction cannot provide all the details but instead show the basic operation. A principal scheme of a VSC based HVDC transmission system is shown in Figure 1.30.

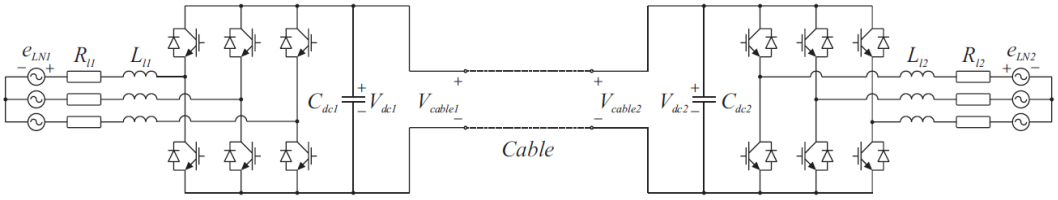


Figure 1.30 Basic VSC based HVDC power transmission interconnection.

From Figure 1.30 it is clear that voltage source converters (VSCs) are used. Three-phase voltage source converters inherently provide the possibility of bi-directional power flow. Both the sending and receiving end converters are current controlled on the AC side, since these sides are inductive. The current controllers operate in the *dq*-frame in the same way as described in Section 3.3 and Section 3.7. The sending end converter is also responsible of controlling the DC bus voltage. To increase the dynamic properties of the DC bus voltage control the receiving end converter AC power is fed forward to the DC bus voltage controller. The DC bus voltage controller for the converter operated as rectifier is shown in Figure 1.31.

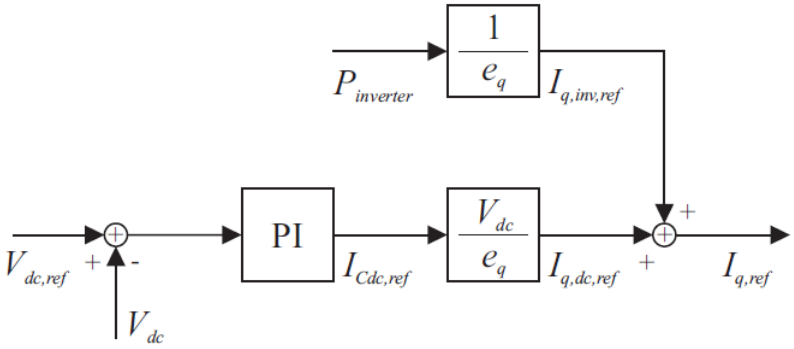


Figure 1.31 DC bus voltage controller for VSC based HVDC.

Figure 1.32 shows a more realistic converter station for VSC based HVDC where additional AC side filters are included. This is needed due to the fact that both sides of each converter are grounded and still third order injection modulation (See Section 2.7, Symmetrical modulation) is used, resulting in a zero sequence voltage.

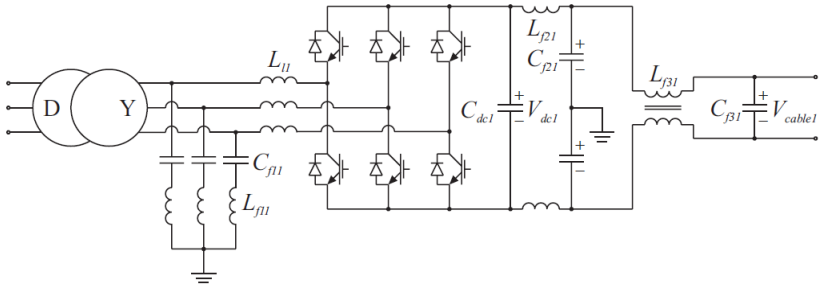


Figure 1.32 Realistic VSC based HVDC converter station.

To attenuate the zero sequence currents, shunt filters tuned to the specific third order harmonic are inserted between the converter AC side interface and the transformer. Figure 1.32 also shows high frequency filters (100 kHz and above) inserted at the DC side of the converter. This is needed to avoid disturbances transmitted from the cable, which could interfere with broadcasting and other radio frequency equipment. Following the high frequency DC side filter, a common mode choke is inserted.

A VSC based HVDC transmission system consisting of two converters according to Figure 1.32 and a T-model of an HVDC cable is investigated by means of simulations. The converters have a rated power of 330 MVA each and the nominal DC cable voltage is 300 kV (±150 kV). The inductive line filter equals 0.14 pu and the equivalent time constant of the each converter including DC link capacitor is 2 ms. This is the known data for the HVDC light system, with the highest rated power, offered by ABB Power Systems. For the simulations, the rated line-to-line voltage is set to 150 kV. The third order harmonic filter is tuned to switching frequency (1950 Hz) in such a way that the fundamental current component equals 0.01 pu, which gives the appropriate values of both L_{f11} and C_{f11} The high frequency filter on the DC side is based on a characteristic frequency of 100 kHz and $C_{f21} = C_{f11}/10$. If the capacitance of C_{f21} is higher, it will affect the characteristic frequency of the third order filter. The mutual inductance of the common mode choke is set to the same value as L_{f21} and the magnetic coupling factor is set to 0.99. The capacitance of C_{f31} is set to $C_{f31} = 2 \cdot C_{f21}$. These component values most likely defer from the ones intended for the real installation. Also the DC bus voltage controller and the converter current controllers are not the ones used in the real installation. Probably the current controller used here has a much higher gain, for example.

Figure 1.33 and Figure 1.34 shows the converter power and DC bus voltage, respectively, for a time simulation, starting at no load condition. The inverter receives an output power command equal to 165 MW (half rated power) at 20.0 ms. The current controller becomes aware of the increased power reference at time 20.25 ms, and calculates the desired AC side voltage reference which is applied one sample later, at time 20.5 ms. The current immediately starts to increase, but this is not noticed by the current controller before the next sample instant at 20.75 ms. At this time the current controller of the inverter also sends a power reference to the converter acting as a rectifier which also starts to increase its power demand. Due to the time needed to establish a rectifier input power corresponding to the inverter output power, the DC bus voltage is somewhat reduced, see Figure 1.34.

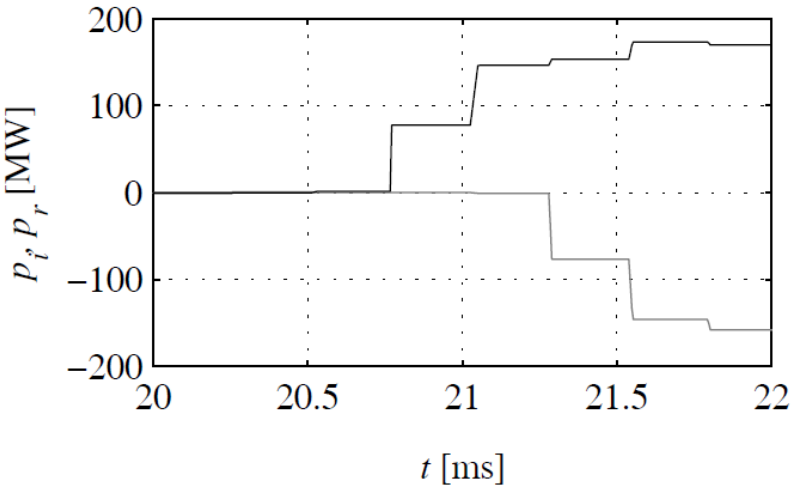


Figure 1.33 Inverter AC side power (black) and rectifier AC side power (grey).

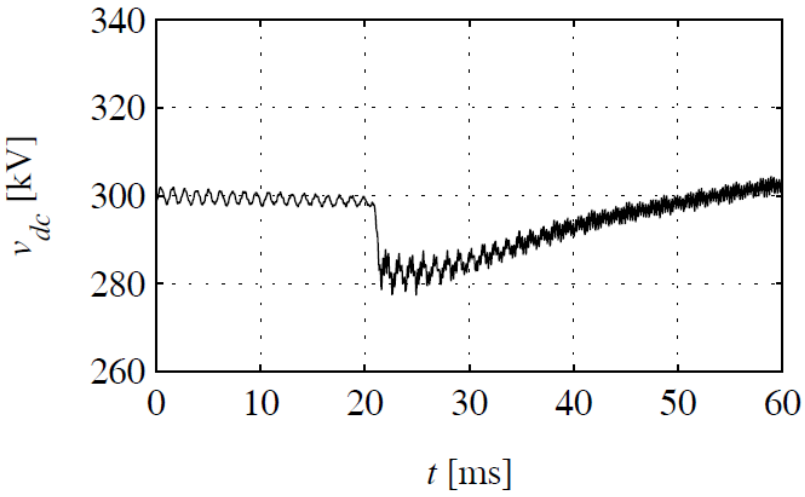


Figure 1.34 Rectifier DC side voltage.

Figure 1.35 shows the line and converter currents for both the rectifier and inverter. It is seen that the currents contain a rather high ripple current, resulting from the switching. Especially, the converter currents exhibit a large ripple as the power increase. These ripple currents are mainly zero sequence currents, and

therefore do not show up in the line currents, which have high impedance grounding. The high gain of the current controllers makes this problem worse. Figure 1.36 shows the steady state line and converter currents for the inverter, together with their harmonic spectrum.

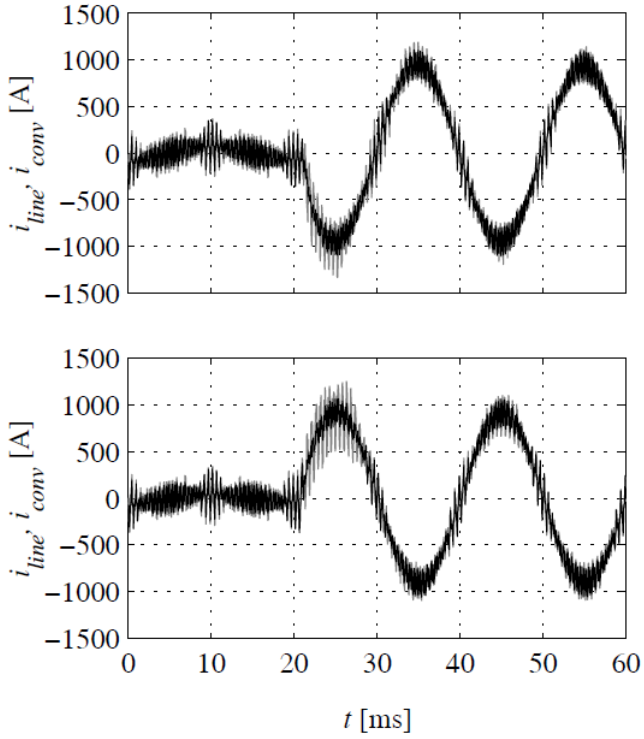


Figure 1.35 *Line current (black) and converter current (grey) for the rectifier (top) and line current (black) and converter current (grey) for the inverter (bottom) during load power increase.*

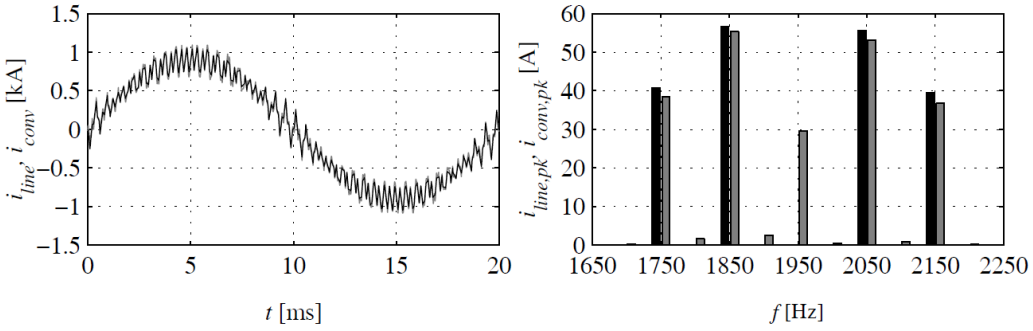


Figure 1.36 *Line current (black) and converter current (grey) for the inverter (left) and line current spectrum (black) and converter current spectrum (grey) for the inverter at frequencies close to switching frequency (right).*

2 Modulation of Power Electronic Converters

The power electronic energy converter operates as a power amplifier in an electrical control system. The energy conversion of the process can be influenced by control of the voltage fed to the load. The converter is capable to convert one voltage type to another, i.e. from dc to ac. The conversion can also be from one dc level to another or from one ac voltage to another.

If the converted power is low, a continuous control of the output voltage can be used, just like in an audio amplifier to a loudspeaker. The transistors operate then in their linear mode and significant parts of the converted power become losses. Such a control method needs large components and heat sinks already at small converted powers.

At higher power ratings the output power is controlled by modulation and the transistors are used as switches. The switch valves (power transistors, GTO-thyristors etc.) are either short circuit or open circuit. Due to the low voltage drop when conducting, typical 1-3 volts, and the short switching periods, typical <1 µs, the losses relative to the converted power is considerably reduced. The power density of the converter, i.e. the converted power divided with weight or volume, is in this way radically increased.

There are several drawbacks with the modulation of the output voltage. The large harmonic content results in a certain current ripple in the load current. A common consequence of this is radio interference and acoustic noise and vibrations. Torque pulsations and increased losses are the results in electrical machines. Therefore it is important to develop methods to decrease these consequences.

In a complete control system usually the converter and it's control are the lesser part. Therefore it is important to achieve a good co-operation with other parts of the system that the power converter acts like a module with the help of suitable interfaces. At commissioning, every module can be tested independently. One such interface is to design the converter and its control to act like a voltage or current generator.

This chapter describes the modulation of several different power converters to obtain a desired output voltage. The modulation methods described are of carrier wave type, whereas other types of modulation that involve direct control of a secondary quantity like current, torque or magnetic flux linkage are describe in

chapter 4. The goal with the modulation is to obtain linearity between the reference of the output voltage and the average of the output voltage, as measured over one carrier wave period.

2.1 The converter topology

Every converter is built according to the functionality of the power semiconductors. A power semiconductor (= valve) should not be used to connect two strong voltages, i.e. two capacitors with different charge, since the current transient through the valve would be very high. The valve should neither be used to connect two inductive circuits with different current, since the voltage transient would be extremely high. Every valve must be used to connect one voltage stiff circuit with one current stiff. Power converters are usually built with one capacitive side and one inductive. In Figure 2.1**Error! Reference source not found.** this is shown for four different types of power converters. They have been drawn with the voltage stiff side (capacitive) to the left and the current stiff side (inductive) to the right.

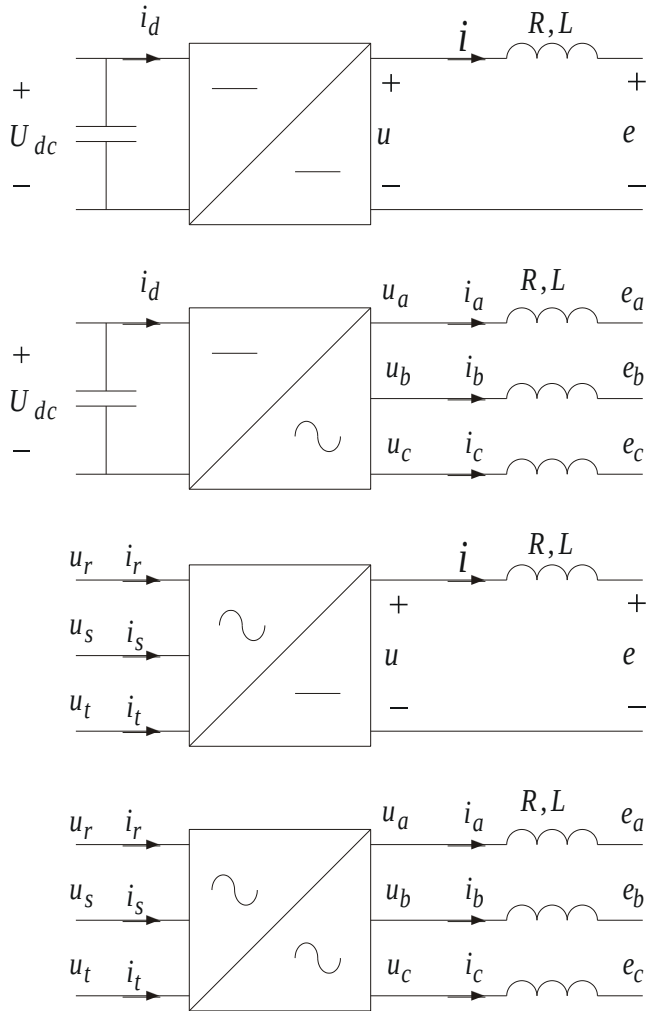


Figure 2.1 Different converter types.

The voltage on the inductive side of the converters can be modulated, since the inductance gives high impedance to the voltage transients and balances the voltage difference between the converter output voltage and the load voltage. The voltage ripple is due to the output voltage ripple whereas *e* varies with a much lower frequency then the modulation frequency. The variation of *e* is assumed to be negligible during one modulation period.

Correspondingly, the current can be modulated on the voltage stiff side (capacitive) since that side has low impedance to the current pulses. Self-commutated converters have distinct inductive and capacitive sides whereas grid or load commutated converters usually have a small commutation inductance on the voltage stiff side as well. This inductance is usually neglected in the first approximation.

Since the voltage modulation means that the voltage pulsates, the power flow will also pulsate. The converters discussed here are assumed to be ideal, and they do not contain any energy storing components such as capacitors and inductances. The momentary power will therefore be equal on the primary and secondary side. The inductances on the current stiff side and the capacitances on the voltage stiff side are used to even out the power pulsations between the source and the load.

Different types of converters have in common that the voltage can be modulated on the inductive side. Voltage modulation means, that the momentary value of the output voltage alternates between well-defined levels, i.e. two dc voltage levels or voltage curves defined by the grid voltage. The momentary value of the voltage is thus not controllable in every time instant. On the other hand, since the output voltage is modulated periodically, the average value and the voltage-time-area can be controlled for every modulation period.

The output voltage from grid- or load-commutated converters has a saw tooth like shape with a constant commutation frequency being a multiple of the line frequency. For the most common three phase connection, the modulation frequency is 300 Hz. Self-commutated converters can select the modulation frequency more freely and it doesn't have to be constant. For transistor converters it is possible to have modulation frequencies in the 10-100 kHz range, whilst a large GTO converter limits the modulation frequency to about 1000 Hz.

To arrange a unified approach, valid for both line commutated and self-commutated converters, this description will be limited to modulation with constant frequency. The modulated voltage has the period time *T* and the control is aimed at the voltage time area *y* in every modulation period. The relation

between the voltage time area and the average voltage during the switch interval is:

$$\bar{u} = \frac{y}{T} \tag{2.1}$$

The voltage time area is controlled after a reference *y**, and the goal is that the real area and the reference should be equal.

$$y = y^* \tag{2.2}$$

In the next section of this chapter, a description of the control of the voltage time area is given.

2.2 Modulation Control of voltage time area

The most important principles for modulation can be illustrated with the circuit in Figure 2.2. A two-position switch controls the potential in *c*. When the switch switches periodically, the output voltage is modulated periodically.

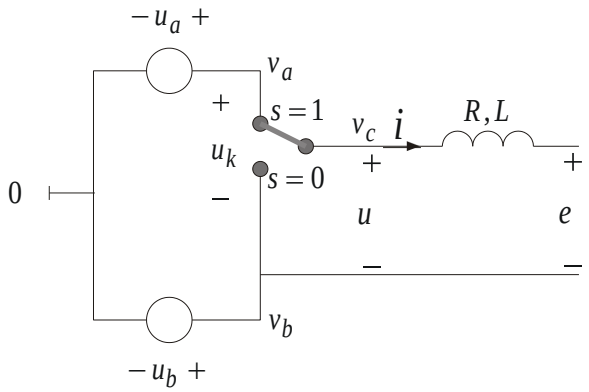


Figure 2.2 Circuit with switch to give pulse modulated output voltage. [1]

The potential v_c can alter between the potentials v_a and v_b and the relation with the switch position *s* is

$$v_c = \begin{cases} v_a \text{ when } s = 1 \\ v_b \text{ when } s = 0 \end{cases} \tag{2.3}$$

The output voltage *u*, connected to a load with a resistance (R), inductance (L) and the counter emf (*e*) in series, is (pulse-) modulated as a function of the switch position

[1] Note, this circuit doesn't describe the case with discontinuous current.

$$u = s \cdot (v_a - v_b) = s \cdot u_k = \begin{cases} u_k \text{ when } s = 1 \\ 0 \text{ when } s = 0 \end{cases} \tag{2.4}$$

The voltage u_k is the voltage step achieved when the switch is moved from one position to another. In this context the step is assumed to be equal to the difference between *a* and *b*.

Figure 2.3 gives an example of modulation of output voltage. The potential v_c alternates between the levels v_a and v_b whilst u alternates between 0 and u_k. A modulated voltage like the one in Figure 2.3 can be controlled with the ratio of the pulse width in relation to the pulse gaps. In addition, the supply voltage influences the surfaces.

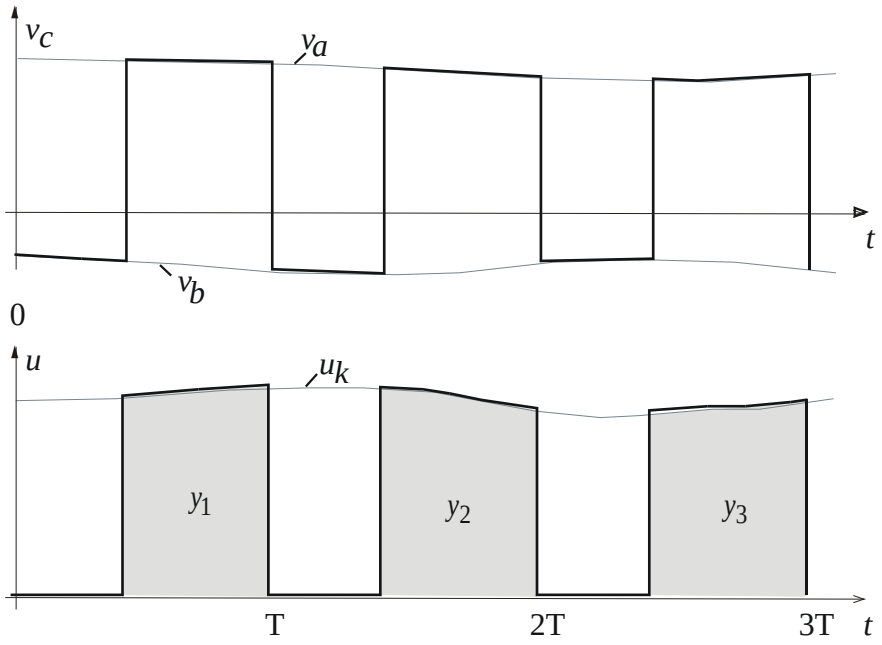


Figure 2.3 The modulation of the potentials va and vb and the modulation of the output voltage between 0 and uk.

When the modulation is limited within a period *T*, and a slowly varying voltage u_k, there are three fundamental ways to influence the surface

1. Variation of the position of the positive flank
2. Variation of the position of the negative flank
3. Variation of both the negative and the positive flank.

These three cases are shown in Figure 2.4. The maximum Y_0 that can be achieved during the period *T* is according to Figure 2.4 a.

$$Y_0 = \int_0^T u_k(t) \cdot dt \tag{2.5}$$

To control the area the position τ_+ for the positive flank and τ_- for the negative flank is varied. The relation between a small variation $d\tau_+$ or $d\tau_-$ and a small change in the output area *dy*, can be written

$$u_k(\tau_+) = -\frac{dy(\tau_+)}{d\tau_+}$$
$$u_k(\tau_-) = \frac{dy(\tau_-)}{d\tau_-} \qquad (2.6)$$

With control of the positive flank, see Figure 2.4 b, the output surface is,

$$y(\tau_+) = \int_{\tau_+}^{T} u_k \cdot dt = Y_0 - \int_{0}^{\tau_+} u_k \cdot dt \qquad (2.7)$$

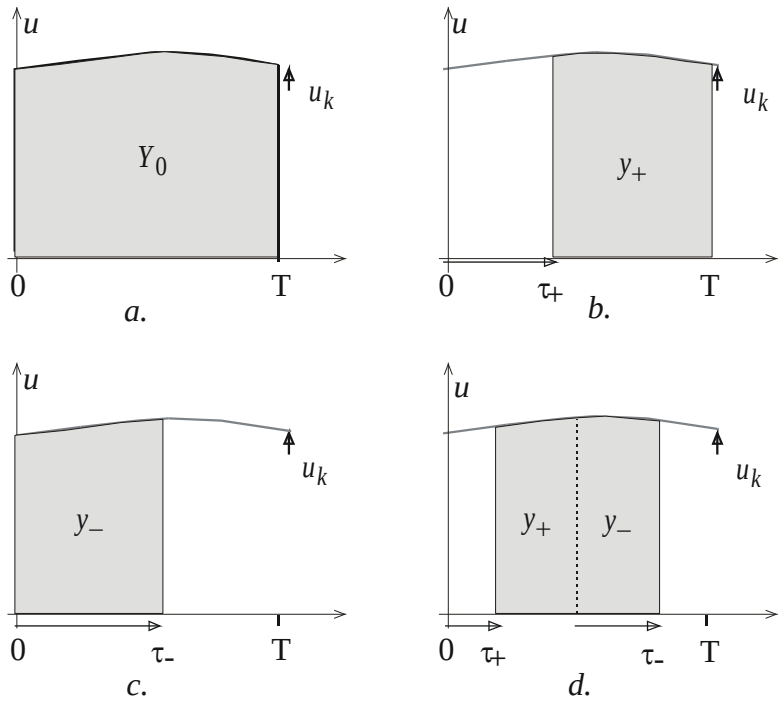


Figure 2.4 Control of voltage time area. a) Maximum area, b) variation with positive flank, c) variation of negative flank and d) variation with both flanks.

This equation describes how the voltage time area is reduced more the longer the positive flank is delayed. If Y_0 is known during the period, the surface $y(\tau_+)$ is determined already at τ_+ i.e. before the voltage pulse really occurs. This is an important property that later on will be used for control of the surface.

Control with only the negative flank, see Figure 2.4 c, gives the surface

$$y(\tau_+) = \int_{0}^{\tau_-} u_k \cdot dt \qquad (2.8)$$

The surface grows in this case with τ_- and when the negative flank comes, the output surface is ended. In the last case, see Figure 2.4 d, both parts of the surface are controlled separately

$$y(\tau_+, \tau_-) = y_+ + y_- \tag{2.9}$$

Every part of the surface can be written as

$$y_+ = \int_{\tau_+}^{T/2} u_k \cdot dt = Y_0 - \int_0^{\tau_+} u_k \cdot dt \tag{2.10}$$

and

$$y_- = \int_{T/2}^{T/2+\tau_-} u_k \cdot dt \tag{2.11}$$

where Y_0 in this case is

$$Y_0 = \int_0^{T/2} u_k \cdot dt \tag{2.12}$$

The surfaces under a given voltage u_k in a modulation period T can thus be described with the variables τ_+ and τ_- according to the equations (2.5) to (2.12). When controlling the output surface, the logics are the opposite. From a given reference for the expected voltage time area (y^*) the switch times τ_+ and τ_- have to be determined. Using the expressions for the voltage time areas above can do this translation. Note that τ_+ and τ_- have to be determined "on-line" which is possible, since both variables implies waiting for the right time for the flank, i.e. for the switch to change position. The method for doing this calculation is developed in the following chapters.

2.3 Carrier wave modulation

A reference value y^* for the desired average voltage over one switching period is calculated by an external control system (current controller, see chapter 4). The pulse width must be controlled to make the voltage time area equal to the reference value. One method is to use a modulating wave y_m that is compared with the reference y^*.

To accomplish the desired result, the carrier wave has to estimate the surface as a function of the position of the flank. If the time function of the carrier wave within the modulation period is

$$y(\tau) = y_m(\tau) \tag{2.13}$$

where $y(\tau)$ is determined by the equations (2.5) to (2.12), it is also a signal that for every time instant τ has a value that, if the switching takes place in that time instant, corresponds to the resulting surface. By detecting when the carrier wave and the reference value are equal, and by at that time instant flip the switch, the problem is solved. There is however difficulties related to the generation of a carrier wave that estimates the real surface. Y_0 has to be known in the beginning of the interval. For the three cases shown in Figure 2.4, the corresponding carrier waves are given in Figure 2.5.

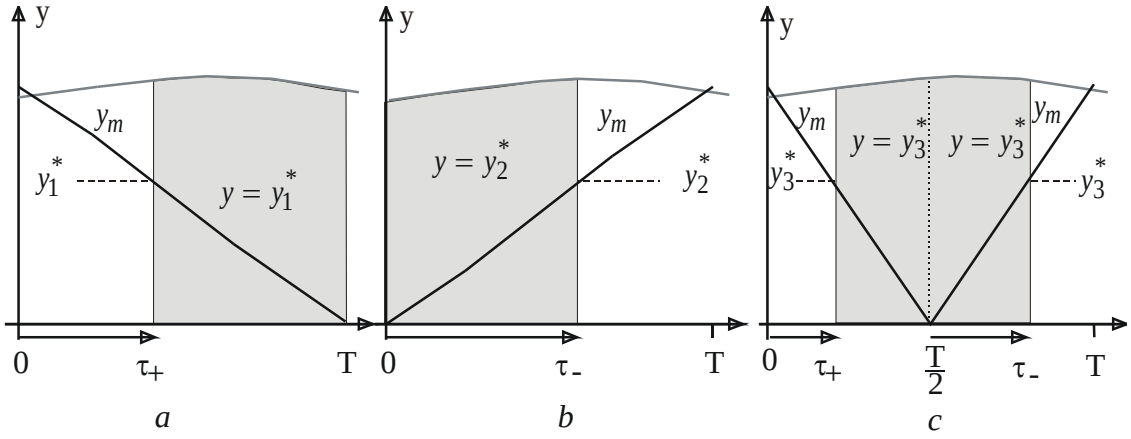


Figure 2.5 Carrier wave for three different types of modulation.

When the carrier wave and the reference are equal, the switch is switched. This is repeated every modulation period. Note that the amplitude scale in Figure 2.5 c is different from the ones in figure Figure 2.5 a, b, since the maximum voltage time area (Y_0) is calculated based on half the interval length as compared to the two cases where only the position of one flank is controlled.

The methodology with carrier wave modulation, which is focused on in this book, is fundamental for the control of several types of converters. It is however not the only method possible. With µ-computers, other methods based on prediction and other calculations are possible. It will however be proven in chapter 4 that when used properly carrier wave modulation in most cases is the best method for modulation of power electronic converters.

There are also methods where voltage modulation and current control cannot be separated. This type of control is often called tolerance band control of flux, current or torque. When torque and flux are controlled that way, the method is referred to as “DTC”-Direct Torque Control. Such methods are described in detail in Chapter 3.9.

Voltage time area vs. average voltage in the modulation

This far the modulation has been described with voltage time areas, both regarding the estimate of the output voltage time area as a function of the switching time instant, i.e. the modulating wave y_m, and the references for the output voltage time areas y^*.

Voltage time area as an entity is very straightforward since it emphasizes that the shape of the output voltage is not relevant, only the average value. Of course this can also be expressed as average voltages rather that voltage time areas. Experience of teaching this subject for a decade has shown that students generally prefer the average voltage, probably since voltage is easier to comprehend, (can be measured with any hand held universal instrument) than voltage time area.

Thus, in the following sections and chapters, voltage time area is replaced with average voltages. No particular notation is used; the assumption of average values is implicitly applied to modulating waves and reference values according to

$$\begin{aligned} u_m &= \frac{y_m}{T_s} \\ u^* &= \frac{y^*}{T_s} \end{aligned} \tag{2.14}$$

The implicit assumption of average values is only valid for modulating waves and reference values according to(2.14), whereas output voltages and potentials are instantaneous unless otherwise stated.

The modulator

The method with a carrier wave u_m that together with the reference value u^* determines the position of the switch, is implemented in a unit called *the modulator*. Based on the previous description, a method for modulation will be developed.

The modulator consists of a unit to compare the carrier wave with the reference value. This can either be done in an analogue modulator with a triangular wave generator and comparators, or in digital modulator using a timer circuit or an up/down counter with a digital comparator.

The carrier wave has the period time T and the amplitude depends on Y_0, which is proportional to u_k. The carrier wave amplitude will vary as a function of u_k, something that also the amplitude of the voltage pulse does. For a given (e.g.

constant) reference value for the average output voltage, small (and slow) variations in the dc link voltage will not affect the real average output voltage.

PWM-controlled dc converters

In Figure 2.6 four different dc converters are shown. In the right column of the figure, the corresponding equivalent circuit is shown. The modulation method use is called Pulse Width Modulation (PWM). The dc link voltage varies slowly in the interval [0, U_{dmax}].

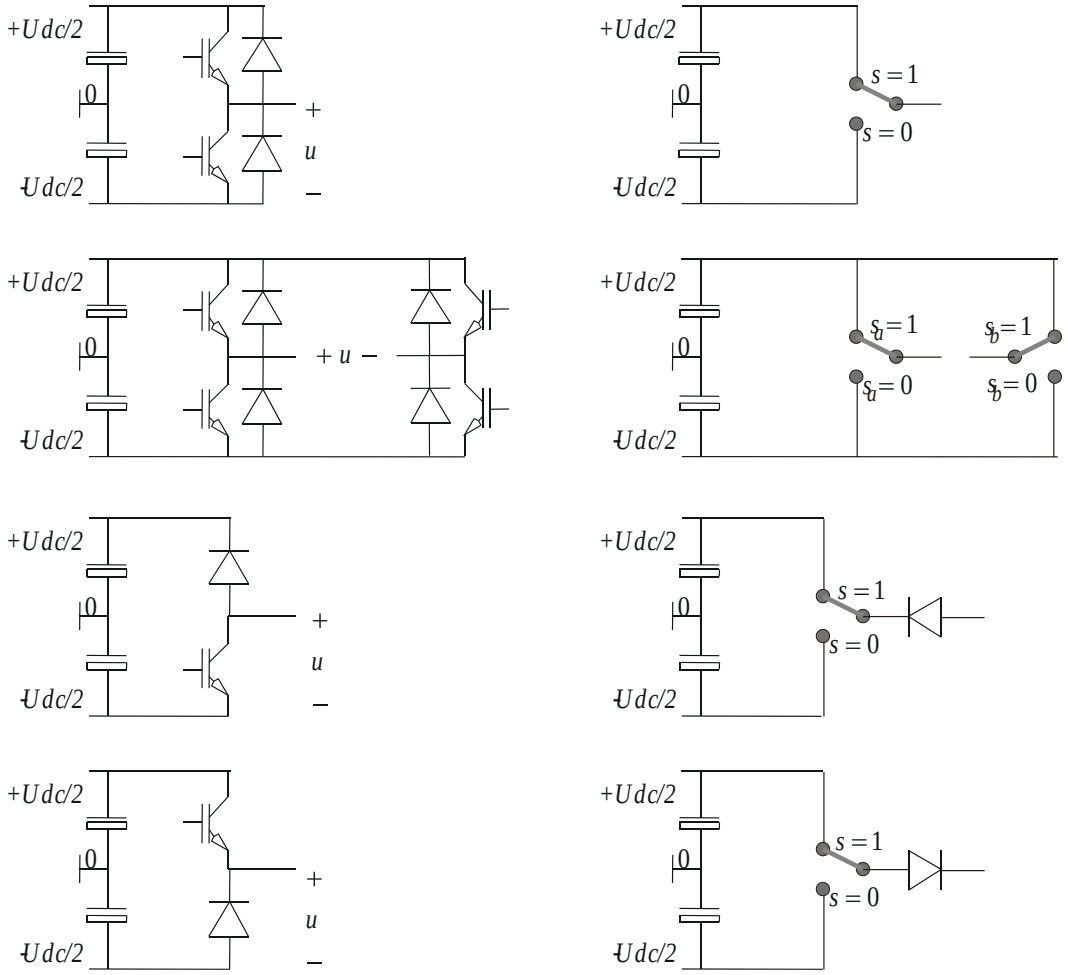


Figure 2.6 DC converters. From top: 2-quadrant, 4-quadrant, 1-quadrant step up and 1-quadrant step down

2.4 Two quadrant DC converters and one quadrant DC converters with continuous current

The current can flow in both directions in a 2-quadrant DC converter (Figure 2.7) whereas the current has only one possible direction in a 1-quadrant DC converter. As long as the current is continuous in a 1-quadrant DC converter, it works in the same way as a 2-quadrant converter with respect to the modulation.

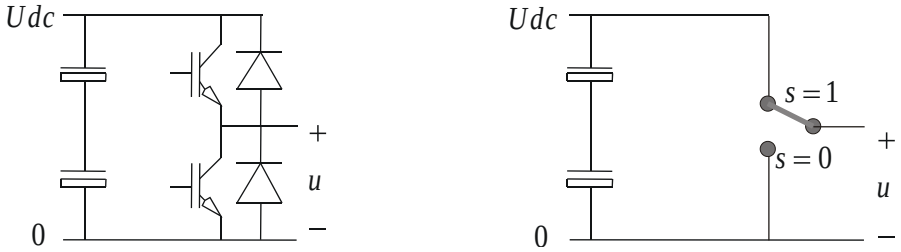


Figure 2.7 *2Q converter. The switch equivalent is also relevant for the 1Q converter if the current is continuous.*

In the first case, with control of only the positive voltage flank, the carrier wave becomes a saw tooth wave according to

$$u_m(\tau_+) = U_{dc} \cdot \left(1 - \frac{\tau_+}{T}\right) \tag{2.15}$$

The principle for the modulation is displayed in Figure 2.8. A signal generator generates a carrier wave with the amplitude U_{dc} and the desired frequency.

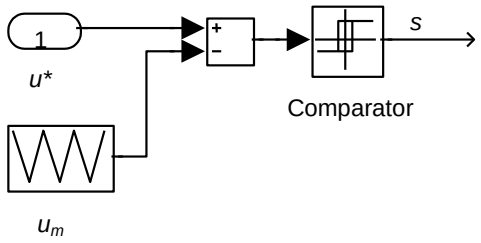


Figure 2.8 *Generic structure of a modulator. Note the use of Simulink® symbols.*

One example of pulse width modulation with this type of carrier wave is shown in Figure 2.9 below. The value of the reference signal in the cross over instant will determine the whole voltage time area achieved in that certain interval.

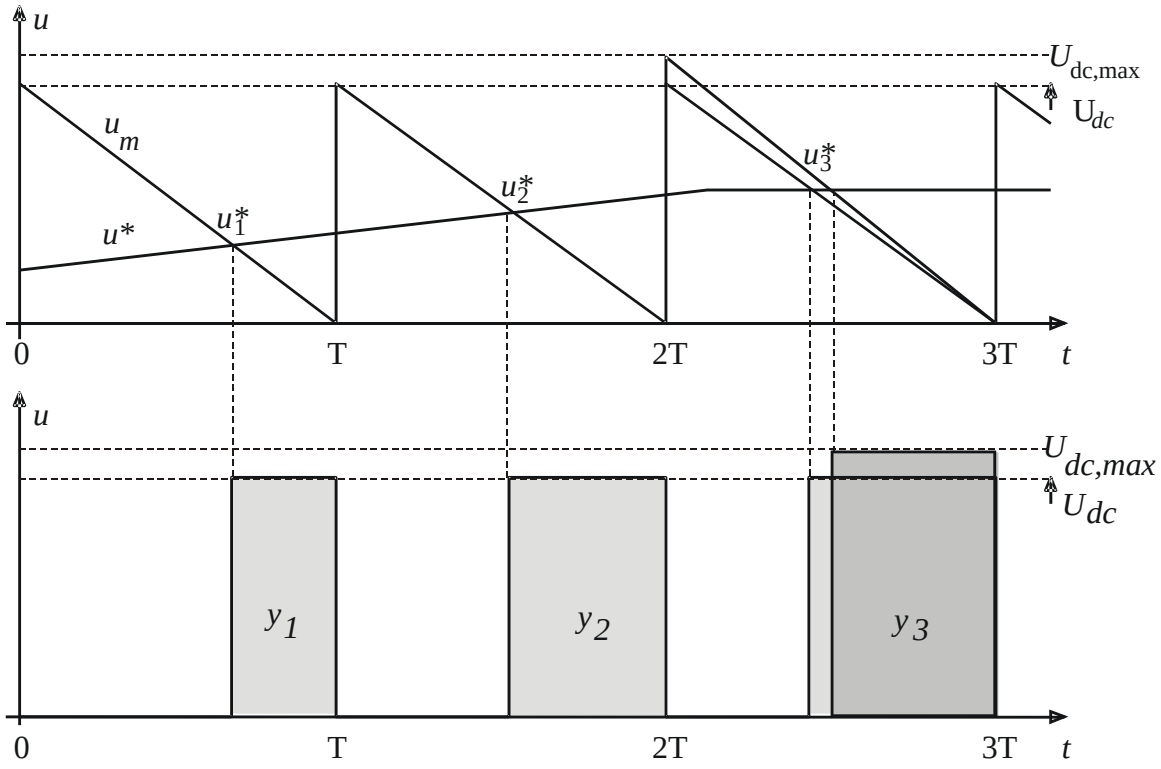


Figure 2.9 Example of control of the position of the positive flank

The maximum value of the dc link voltage and corresponding values of the carrier wave are also shown in the figure. Since the carrier wave amplitude is proportional to the dc link voltage, the resulting voltage time area is independent of the dc link voltage. An increase in the dc link voltage, e.g., results in a higher but shorter voltage pulse with maintained area.

If instead the position of the negative flank is controlled, the saw tooth wave will change shape slightly according to

$$u_m(\tau_-) = U_{dc} \cdot \frac{\tau_-}{T} \tag{2.16}$$

i.e. a saw tooth with positive slope. The whole modulation in all other aspects, are equal to the case where the positive slope is controlled.

The third possibility is to control both the positive and the negative flanks. In this case the carrier wave becomes triangular. According to the equation below it is built up of two parts:

$$u_m(\tau_+,\tau_-) = \begin{cases} u_{dc}\left(1-\dfrac{\tau_+}{T/2}\right) \\ \qquad\qquad 0<\tau_+,\tau_-<T/2 \\ u_{dc}\dfrac{\tau_-}{T/2} \end{cases} \tag{2.17}$$

With this method, the voltage pulse can be influenced two times every switching period according to the example in Figure 2.10.

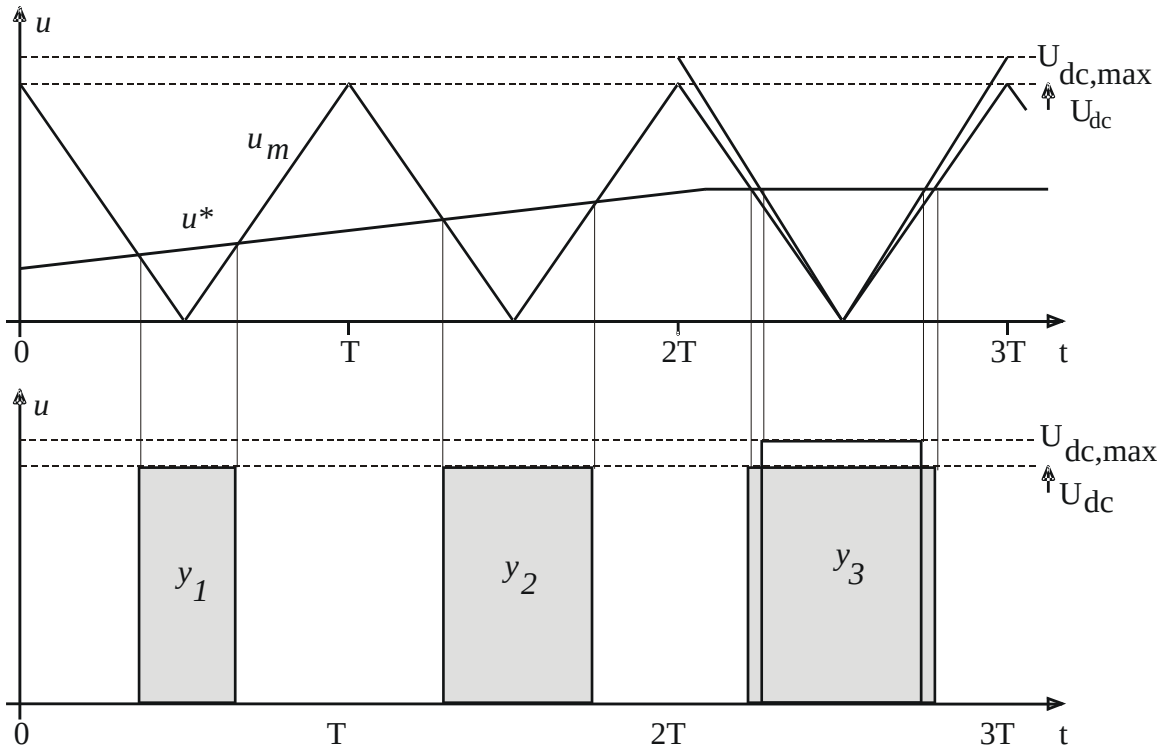


Figure 2.10 Example of control by variation of the position of both flanks.

The differences between these three methods might seem small, but they will turn out to be very important when it comes to current control. Furthermore the cases with discontinuous current with the 1-quadrant connections must be treated separately. It will be shown that the same carrier wave with small changes can be used even for 4-quadrant converters and 3 phase converters.

2.5 Four quadrant (bridge connected) converters

In the bridge connected DC converter (Figure 2.11), each branch (*a* and *b*) is controlled independently. The mid-point of the DC link voltage is used as reference for the calculation of voltage time area.

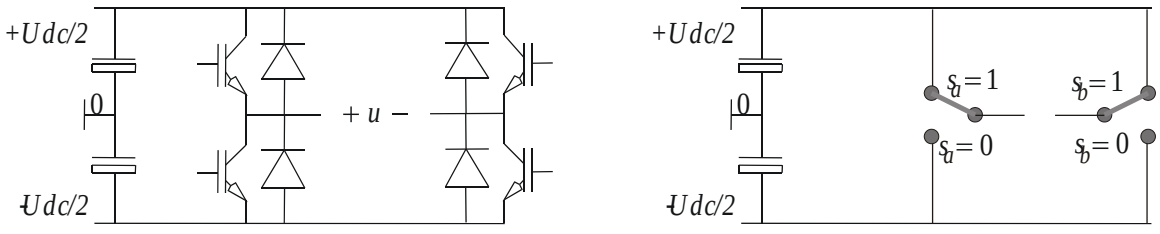


Figure 2.11 4Q converter.

The phase potentials v_a and v_b are modulated between $\pm U_{dc}/2$, see Figure 2.12.

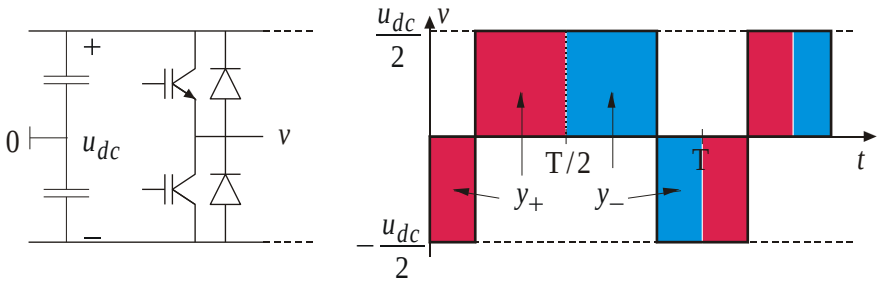


Figure 2.12 Modulation of a phase potential.

In this case the modulation is designed to control the position of both flanks. The maximum voltage time area for one half switching period is given by (2.18). Note that the output average voltage in this case can be both negative and positive with reference to the mid-point in the DC link.

$$Y_0 = \frac{u_{dc}}{2} \cdot \frac{T}{2} \tag{2.18}$$

The voltage time area y can be controlled between Y_0 and $-Y_0$. According to Figure 2.12, the modulating wave can be written as

$$u_m(\tau_+, \tau_-) = \begin{cases} \frac{u_{dc}}{2}\left(1 - \frac{\tau_+}{T/4}\right) \\ \qquad\qquad 0 < \tau_+, \tau_- < T/2 \\ \frac{u_{dc}}{2}\left(\frac{\tau_-}{T/4} - 1\right) \end{cases} \tag{2.19}$$

This is a triangular wave varying between ± $U_{dc}/2$. When the dc link voltage varies, the amplitude of the triangular carrier wave varies in proportion. The differences between the phase potentials create the output voltage of the converter. It can be expressed as:

$$u = v_a - v_b \tag{2.20}$$

The idea is to control the voltage time area of the output voltage (u) by controlling the voltage time area of the phase potentials (v_a, v_b). The pulse modulation can be done in several different ways, and it is therefore important to select the best possible. That method is shown in Figure 2.13.

Base d on the reference value u^* for the average of the output voltage (u), two new reference values are calculated, one for each phase. To achieve an optimal control the calculation of the two final reference values should be done in the following way:

$$u_a^* = \frac{u^*}{2}$$
$$u_b^* = -\frac{u^*}{2} \qquad (2.21)$$

where u_a* controls the phase potential v_a and u_b* controls v_b.

By means of this way of calculating the phase voltage time area references, and by the fact that both phases are modulated by the same carrier wave, the output voltage will have a pulse frequency of twice the modulation frequency. The relatively high pulse frequency minimizes the current ripple and increases the number of control possibilities.

Two units build up the modulator for a 4-quadrant DC converter. Each unit controls its own phase potential, but the carrier wave generator is shared. Figure 2.13 shows the structure of the modulator.

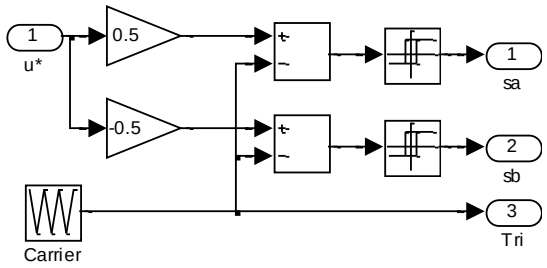


Figure 2.13 Modulator for a 4-quadrant DC converter

Figure 2.14 shows a four-quadrant DC-DC converter where the transistors are denoted T_1-T_4, and the freewheeling diodes D_1-D_4.

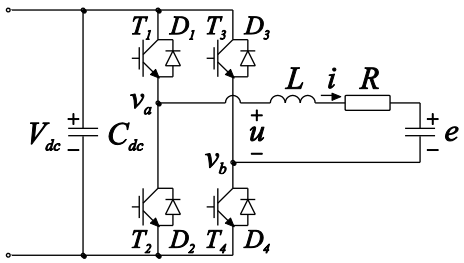


Figure 2.14 Four-quadrant DC-DC converter utilized in the study of modulation.

In the case of DC-DC conversion, the current controller inherently provides only a single voltage reference, the load voltage reference. However, for the converter in Figure 2.14, two transistor half bridges should be modulated. To solve this, the original load voltage reference is split in two phase potential references, one for each half bridge. Here, these phase potential references are denoted $v_{a,\,ref}$ and $v_{b,\,ref}$, which must be selected so that $v_{a,\,ref} - v_{b,\,ref} = v_{ref}$. These two references can be selected in any way as long as the difference between them equals the load voltage reference (in engineering units). Often, the phase potential references are chosen so that the risk of entering over-modulation is

minimized. Over-modulation occurs when one or several of the phase potential references exceeds the peak of the carrier wave. A suitable choice is therefore to choose the phase potential references in such a way that two (both for the four-quadrant DC-DC converter) phase potential references reach over-modulation simultaneously for a certain load voltage reference. This means that the phase potential references should be selected according to

$$v_{a,ref} = \frac{v_{ref}}{2} \quad \text{and} \quad v_{b,ref} = -\frac{v_{ref}}{2} \tag{2.22}$$

This selection of modulator references where they are equal but with opposite sign, i.e. symmetrical with respect to the time axis, is referred to as symmetrical modulation. Over-modulation occurs when

$$v_{a,ref} = \hat{v}_{tri} \quad \Leftrightarrow \quad v_{b,ref} = -\hat{v}_{tri} \tag{2.23}$$

Since the maximum load voltage that can be obtained is equal to V_{dc}, this corresponds to the boundary between linear modulation and over-modulation, i.e.

$$\hat{v}_{tri} = v_{a,ref} = \frac{V_{dc}}{2} \tag{2.24}$$

The triangular carrier should thus have amplitude of $V_{dc}/2$. Usually, the current controller operates in engineering units (for example full scale of transducers often correspond to 10 V), which means that the triangular carrier must be scaled in the same way. Figure 2.15 shows typical modulator waveforms and converter output voltage and current, in the case of an ideal four-quadrant DC-DC converter.

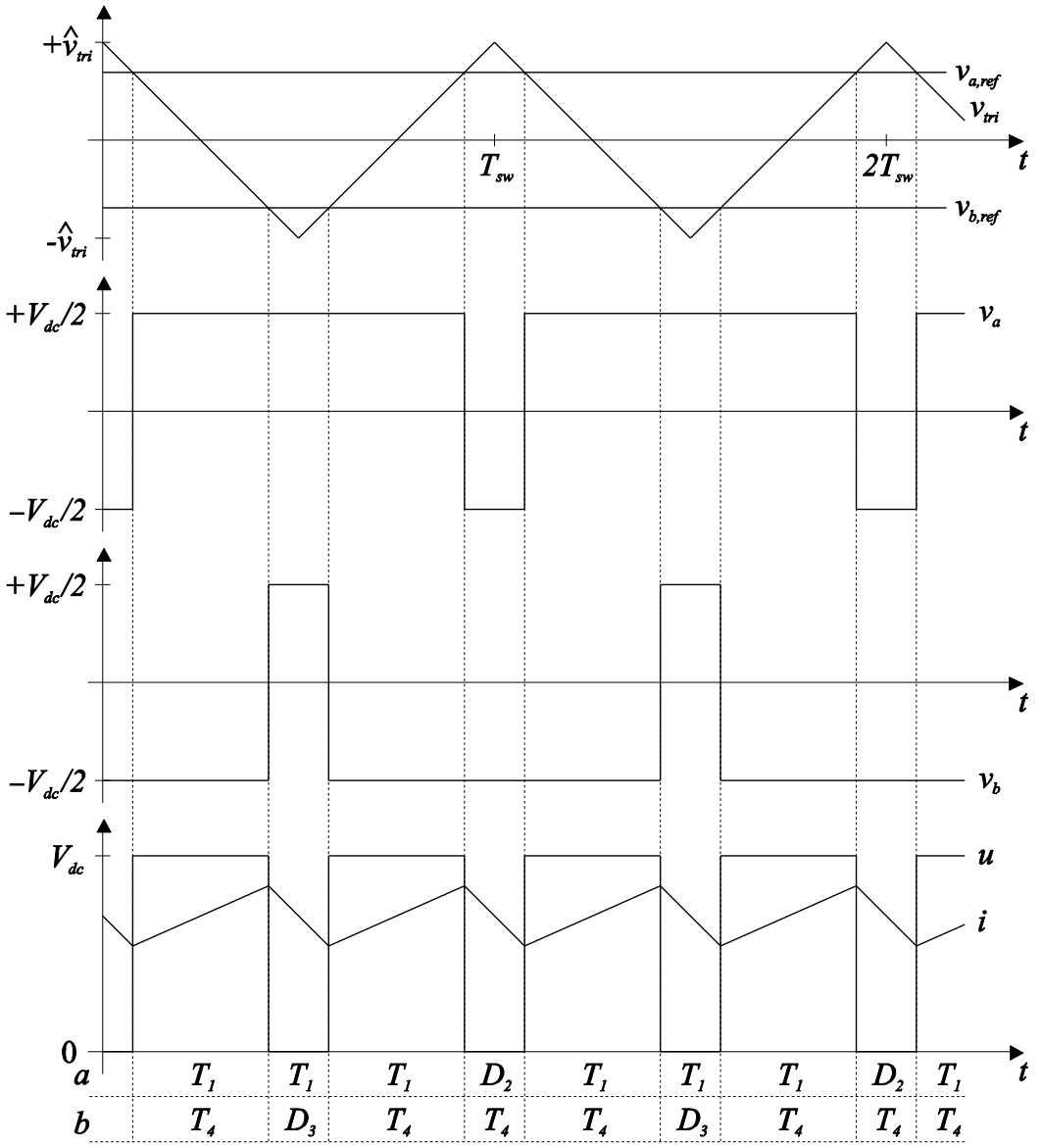


Figure 2.15 Modulation Waveforms and converter output voltage and current for an ideal four-quadrant DC-DC converter.

The output potentials of the half bridges are denoted v_a and v_b. The voltage across the load is denoted u, and is given by

$u = v_a - v_b$. The load current is denoted i. During one period of the triangular carrier, each half bridge output exhibits one positive and one negative step. The switching frequency is therefore equal to the frequency of the triangular carrier. If the period time of the triangular carrier is considerably shorter than the time constant of the load, the load current will exhibit a triangular shape as in Figure 2.15. In Figure 2.15 it is also shown which among the eight power semiconductors that are conducting in each sub-interval of the switching period. Note that the direction of the current play as important role in determining this.

As earlier stated, Figure 2.15 shows the modulator waveforms and the output voltage and current of the converter in the case of an ideal four-quadrant DC-DC

converter. In reality, power semiconductor devices cannot traverse from off-state to on-state and vice versa instantly. If both transistors of a half bridge operate in the on-state simultaneously there will be a extremely high power dissipation inside both since the DC link voltage will be divided between the two conducting devices but also since the current will be very high due to the low on-state resistance of the devices. To prevent this from occurring, during switching transients, transistor turn-on is delayed but not transistor turn-off. The duration of this time delay is termed blanking time or interlock time and the circuit performing the delay (which is a part of the transistor driver) is called blanking time or interlock circuit. Figure 2.16 shows the resulting modulation waveforms together with the output voltage and current of the four-quadrant DC-DC converter. Note that the average output voltage and current becomes lower than expected due to blanking time. Therefore, converter current controllers should be equipped with an integral part.

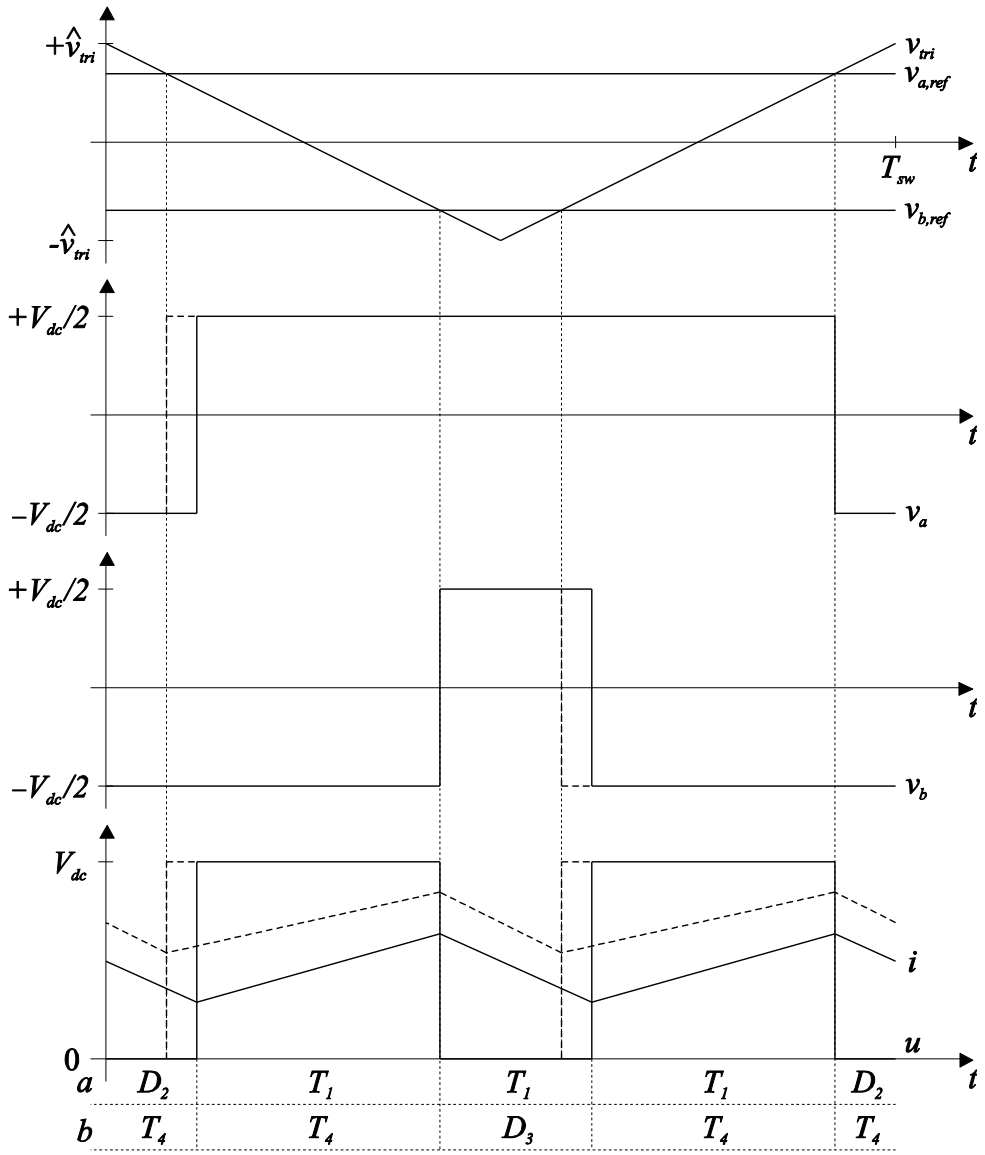


Figure 2.16 Modulation waveforms and converter output voltage and current for a four-quadrant DC-DC converter including blanking time effects. Note that the effect is strongly exaggerated.

This converter is called a DC converter, but as stated in Chapter 2.5 it can also be used for generation of a single-phase AC voltage, since the output voltage can have any polarity (Figure 2.17).

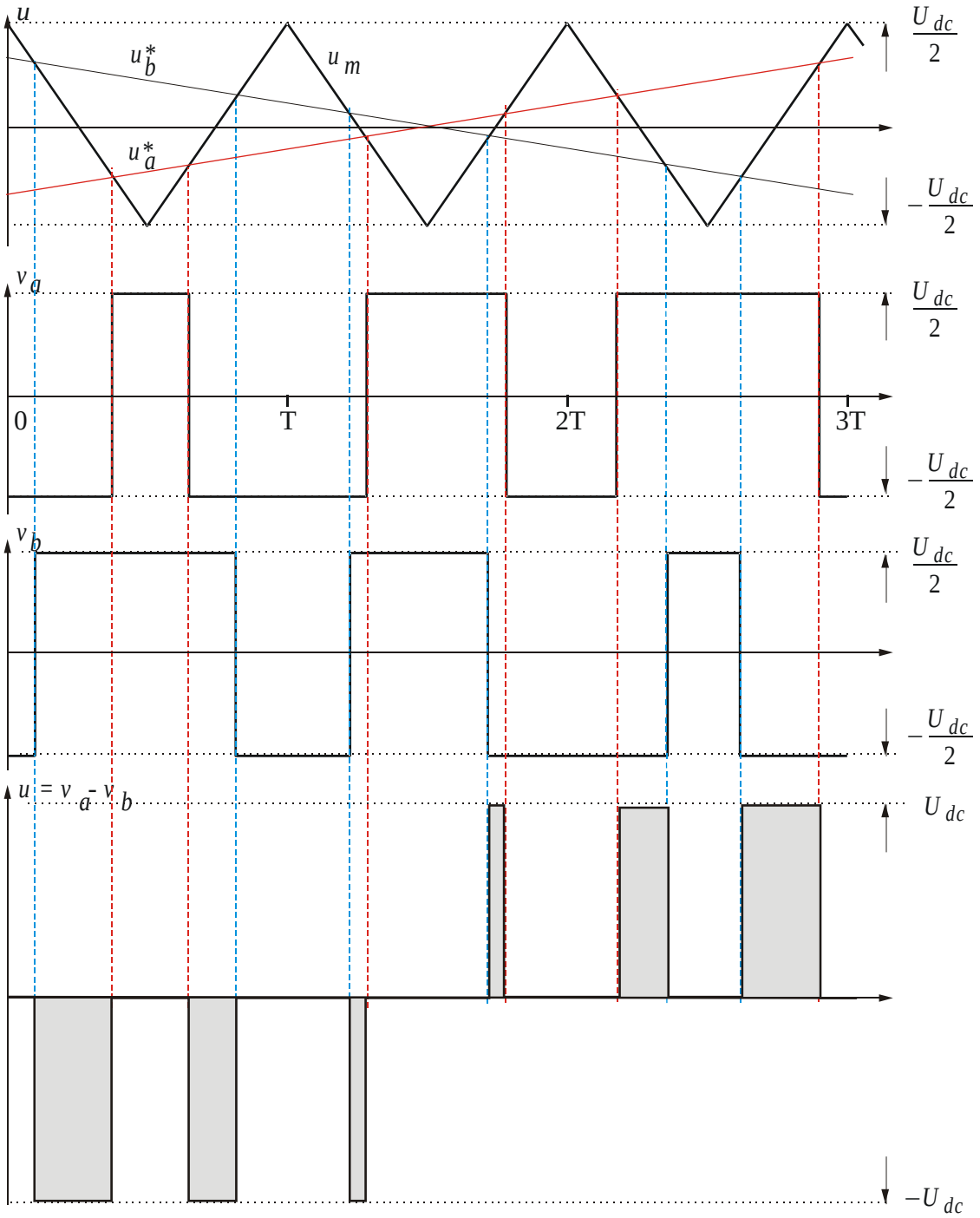


Figure 2.17 Example of modulation with 4 quadrant DC converter.

2.6 Modulation of three-phase converters

The three-phase converter is basically an extension of the 4-quadrant DC converter with one phase added to facilitate the generation of three individual

phase potentials, used to control three-phase loads. The output voltage created can be supplied to a specific machine or constitute a three phase system, e.g. for uninterruptible power supplies (UPS). It is also possible to connect it to the utility grid (Figure 2.18). The three-phase converter consists of three transistor half-bridges in principle of operation corresponding to three switches s_a, s_b and s_c that are independently controlled.

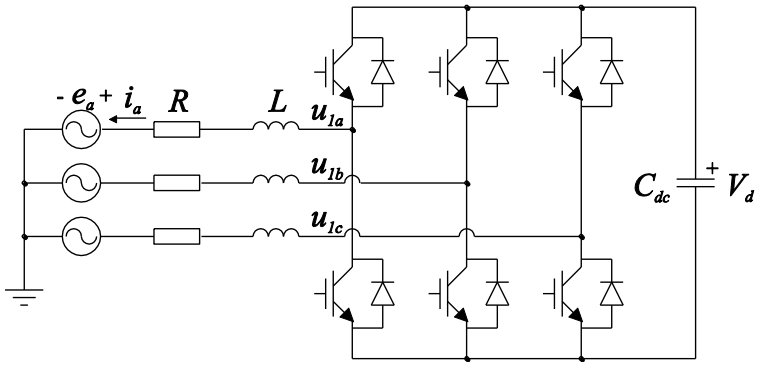


Figure 2.18 The investigated converter and grid connection.

Modulation schemes

The only difference between the investigated carrier wave modulation schemes for three-phase converters is the reference waveforms. To be more correct, the only difference is whether a zero-sequence signal is added or not, and if added, what the zero-sequence looks like [7]. A zero-sequence component in the references does not show up in the output voltage since the converter is not connected to the neutral. Therefore, the converter cannot produce any voltage affecting the neutral. A zero-sequence voltage per definition results in a zero-sequence current and since no neutral is connected there is no path for the zero-sequence (or sometimes called neutral) current to flow through. The converter investigated and its connection to the power grid, are shown in Figure 2.18.

This means that all the reference waveforms considered can be treated as being originally sinusoidal but later altered with a zero-sequence component. The reference waveforms are altered to affect the harmonic current spectrum or to increase the utilisation of the converter by increasing the maximum allowable modulation index for which linear modulation is obtained. The modulation index m_a is defined as

$$m_a = \frac{u_{i,ref}}{V_{dc}/2} \quad \text{where} \quad i = a,b,c \tag{2.25}$$

where u is the converter output phase potential and V_{dc} the DC bus voltage. The reason why m_a is calculated using $V_{dc}/2$ is that in most cases the triangular carrier has a DC level equal to zero and amplitude proportional to $V_{dc}/2$. Here, the amplitude is equal to $V_{dc}/2$. This is because, otherwise, the reference signals

should be scaled by the same factor as the carrier peak-to-peak level in order to get an accurate modulation. Note that for full-bridge converters, the carrier DC level is not critical since the output voltages are differential, i.e. line-to-line voltages.

When m_a higher than 1 occurs, the switching of the particular phase is inhibited, since no crossings between the carrier and the reference waveforms occur. This situation is termed over-modulation. Over-modulation is equal to actuator saturation and should be avoided at least for stationary operation. The problem is not only that the controller command signals are not obeyed by the power electronic converter, but also that the output current harmonic content increases strongly [7][14].

The most important issue for all the investigated modulation methods is that the line-to-line output voltage should be equal for the different reference signals. As mentioned earlier, the voltage references are modified by a zero-sequence signal, in such a way that the maximum m_a is altered. The different voltage references are shown in Figure 2.19 for a time equal to one period of the fundamental.

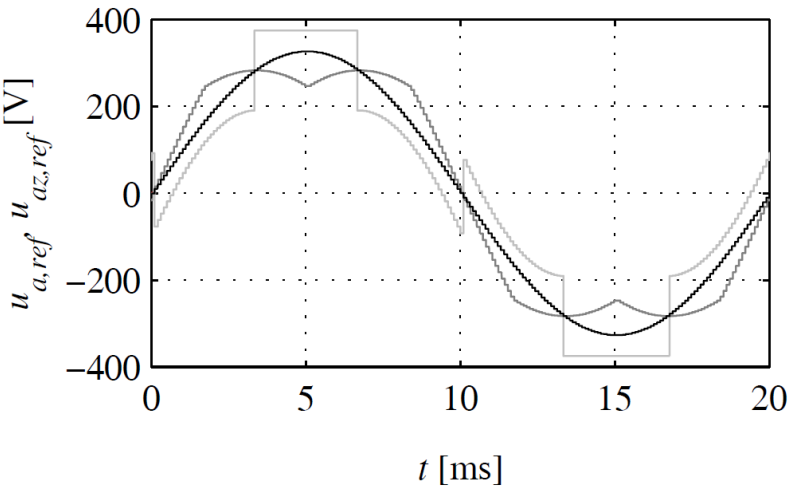


Figure 2.19 Carrier wave modulation: sinusoidal (black), symmetrical (grey) and bus-clamped (light-grey) voltage references.

The line-to-line output voltage resulting from these references equals 400 V, 50 Hz in all the cases. For sinusoidal modulation, the references are kept as they are, i.e. as sinusoidal phase potentials. For symmetrical references, the references are altered in such a way that the most positive and most negative reference is made equal but with opposite sign. For bus-clamped modulation the reference with the highest instantaneous value (positive or negative) is clamped to $\pm V_{dc}/2$, dependent on which is the closest.

For sinusoidal references it is clear that over-modulation, i.e. m_a=1, occurs when the reference peak equals half the DC link voltage. Therefore,

$$u_{i,ref,max} = \frac{V_{dc}}{2} \quad \text{where} \quad i = a,b,c \tag{2.26}$$

for sinusoidal modulation. For symmetrical modulation the zero-sequence signal is calculated from

$$u_z = \frac{1}{2}\left(\max\left(u_{a,ref}, u_{b,ref}, u_{c,ref}\right) + \min\left(u_{a,ref}, u_{b,ref}, u_{c,ref}\right)\right) \tag{2.27}$$

For 1/6th of the fundamental period where $u_{a,ref}$ is the most positive and $u_{b,ref}$ is the most negative reference in stationary conditions, the resulting reference for the half-bridge connected to the *a*-phase is written

$$u_{az,ref} = u_{a,ref} - u_z = \frac{\sqrt{3}}{2} \cdot \hat{u} \cdot \cos\left(\omega_1 t - \frac{\pi}{6}\right) \tag{2.28}$$

As in the previous case, over-modulation occurs when the reference equals half the DC link voltage, which gives

$$u_{iz,ref,max} = \frac{1}{\sqrt{3}} \cdot V_{dc} \quad \text{where} \quad i = a,b,c \tag{2.29}$$

For bus-clamped modulation, the transition between two bus-clamped references is investigated. Assume that the *b*-phase has been clamped to $-V_{dc}/2$ and that the *a*-phase is going to be clamped to $+V_{dc}/2$. When this occurs

$$u_{a,ref} = \frac{\sqrt{3}}{2} \cdot \hat{u} \quad \text{and} \quad u_{b,ref} = -\frac{\sqrt{3}}{2} \cdot \hat{u} \tag{2.30}$$

This means that the zero-sequence voltage to be set equals

$$u_z = \frac{\sqrt{3}}{2} \cdot \hat{u} - \frac{V_{dc}}{2} < 0 \tag{2.31}$$

This gives

$$u_{az,ref} = \frac{\sqrt{3}}{2} \cdot \hat{u} - \left(\frac{\sqrt{3}}{2} \cdot \hat{u} - \frac{V_{dc}}{2}\right) = \frac{V_{dc}}{2} \tag{2.32}$$

and

$$u_{bz,ref} = -\frac{\sqrt{3}}{2} \cdot \hat{u} - \left(\frac{\sqrt{3}}{2} \cdot \hat{u} - \frac{V_{dc}}{2}\right) = \frac{V_{dc}}{2} - \sqrt{3} \cdot \hat{u} \tag{2.33}$$

The limit for over-modulation is given by

$$u_{bz,ref} = \frac{V_{dc}}{2} - \sqrt{3} \cdot \hat{u} \geq -\frac{V_{dc}}{2} \tag{2.34}$$

which gives

$$u_{iz,ref,max} = \frac{1}{\sqrt{3}} \cdot V_{dc} \quad \text{where} \quad i = a,b,c \tag{2.35}$$

This means that modulation with symmetrical or bus-clamped references provides a 15 % higher margin to over-modulation compared to modulation with sinusoidal references. The harmonic content of the output voltage depends on the waveform of the reference voltage. Both the differential mode (DM) harmonics, i.e. the harmonic content in

$$u_{ij} = u_{iz} - u_{jz} \quad \text{where} \quad i, j = a,b,c \tag{2.36}$$

and the common mode (CM) harmonic content, i.e. the harmonic content in u_z, are of great importance. For example, the DM harmonics result in losses in filter inductors etc. The CM harmonics can also result in parasitic current through capacitive coupling between conductors and ground. In turn, this results in malfunction of earth protectors and the occurrence of bearing currents in motors. The DM and CM harmonic content for the different modulation references in Figure 2.19 are shown in Figure 2.20 and Figure 2.21, respectively.

For a purely inductive filter, the damping is proportional to the frequency, since the impedance is given by

$$Z_L = j\omega_1 L \tag{2.37}$$

The resulting harmonic content in the output currents is shown in Figure 2.22, in the case of L=L_{line}=1.7 mH. This filter inductance is suitable for a 100 kW converter with 750 V DC bus voltage connected to a 400 V three-phase AC grid [2]. Note that the harmonic current content is calculated from the DM spectrum of the output voltage.

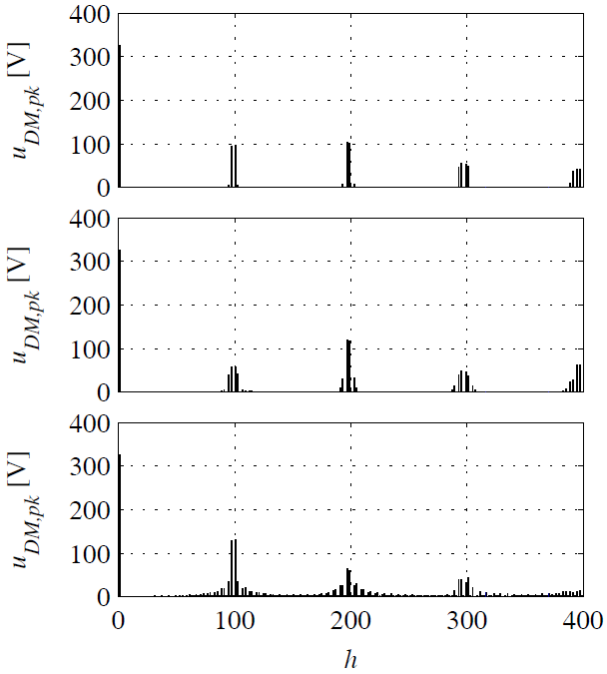


Figure 2.20 *Differential mode output voltage spectrum for sinusoidal (top), symmetrical (middle) and bus-clamped modulation (bottom). The fundamental frequency is 50 Hz and the spectrums are shown as a function of harmonic order h, where 1 denotes the fundamental.*

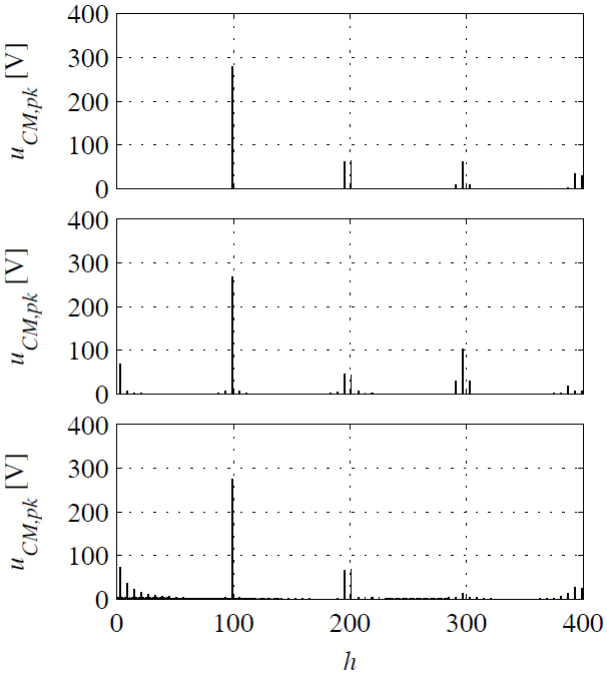


Figure 2.21 *Common mode output voltage spectrum for sinusoidal (top), symmetrical (middle) and bus-clamped modulation (bottom). The fundamental frequency is 50 Hz and the spectrums are shown as a function of harmonic order h, where 1 denotes the fundamental.*

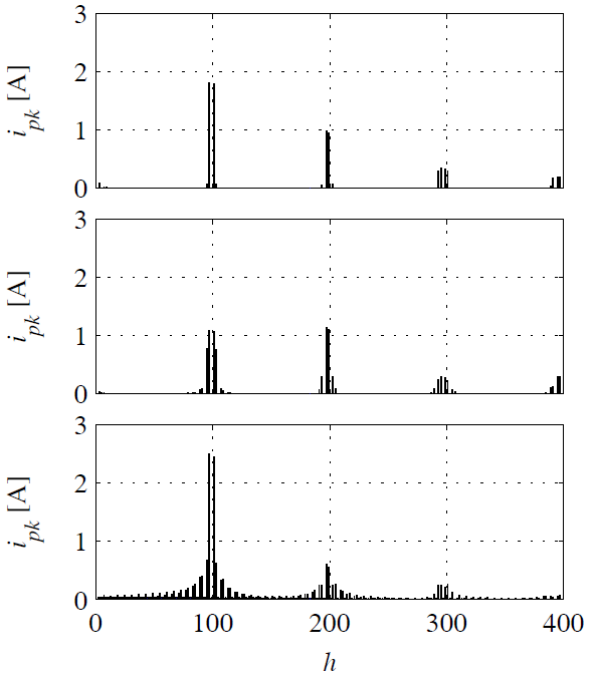


Figure 2.22: Output current spectrum for sinusoidal (top), symmetrical (middle) and bus-clamped modulation (bottom). The fundamental frequency is 50 Hz and the spectrums are shown as a function of harmonic order h, where 1 denotes the fundamental.

2.7 Vector representation

In the three phases self-commutated converter with modulation of the phase potential in two levels ($\pm U_{dc}/2$) the switches can be combined in eight patterns. Each combination (s_a, s_b, s_c) is defined as a state. In Figure 2.23 are all combinations and corresponding states are shown. To every combination a set of phase voltages is related.

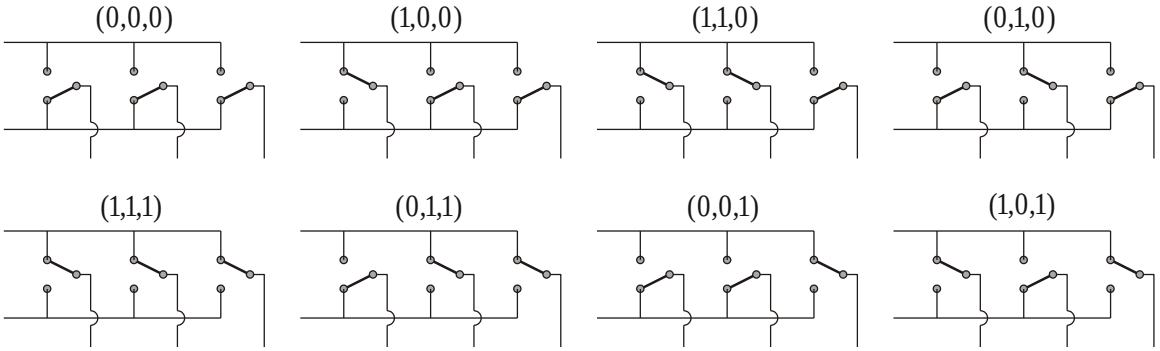
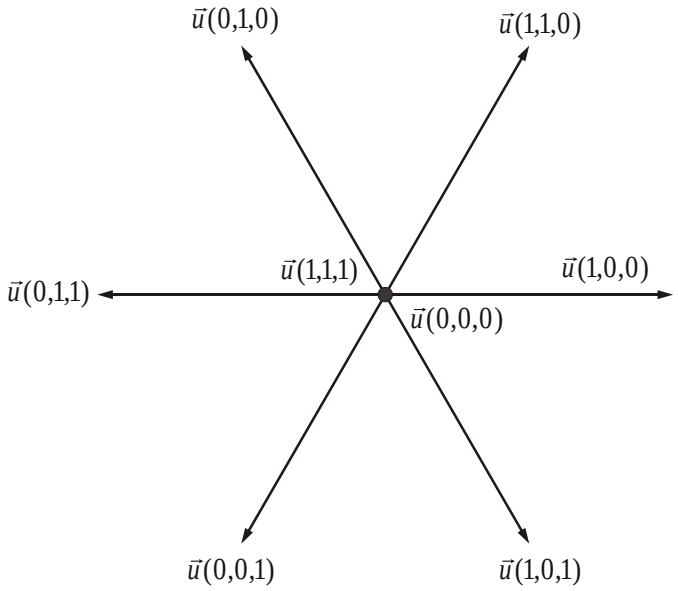


Figure 2.23 Eight combinations….

The combinations of phase voltages in Figure 2.23 can be expressed as complex space vectors. A space vector $\vec{x}^{\alpha\beta}$ is defined from instantaneous three-phase values x_a, x_b and x_c according to

$$\vec{x}^{\alpha\beta} = x_a \cdot e^{j0} + x_b \cdot e^{j\frac{2\pi}{3}} + x_c \cdot e^{j\frac{4\pi}{3}} = x_\alpha + jx_\beta \qquad (2.38)$$

For the converter output voltage the expression above results in eight possible output voltage vectors. Two of the vectors are zero vectors and the other six create a star, see Figure 2.24.



$$\vec{u}(1,0,0) = \sqrt{\frac{2}{3}} U_{dc} = -\vec{u}(0,1,1)$$

$$\vec{u}(0,1,0) = \sqrt{\frac{2}{3}} U_{dc} \cdot e^{j\frac{2\cdot\pi}{3}} = -\vec{u}(1,0,1)$$

$$\vec{u}(0,0,1) = \sqrt{\frac{2}{3}} U_{dc} \cdot e^{j\frac{4\cdot\pi}{3}} = -\vec{u}(1,1,0)$$

$$\vec{u}(0,0,0) = 0 = \vec{u}(1,1,1)$$

Figure 2.24 Voltage vectors from the 3-phase 2-level converter.

For general symmetrical three-phase quantities, i.e. if

$$\begin{cases} x_a = \hat{x} \cdot \cos(\omega_1 t) \\ x_b = \hat{x} \cdot \cos\left(\omega_1 t - \frac{2\pi}{3}\right) \\ x_c = \hat{x} \cdot \cos\left(\omega_1 t - \frac{4\pi}{3}\right) \end{cases} \tag{2.39}$$

the power-invariant transformation is given by

$$\begin{cases} x_\alpha = \sqrt{\frac{3}{2}} \cdot x_a \\ x_\beta = \frac{1}{\sqrt{2}}(x_b - x_c) \end{cases} \tag{2.40}$$

It is important to understand that the last expression cannot be used for switched quantities since these are not symmetrical. In the stationary $\alpha\beta$-frame, the same calculation methods as used for regular three-phase calculations can be used. Kirchhoff's voltage law in space coordinates applied to a three-phase converter with inductive filter connected to a three-phase grid (Figure 2.18) gives

$$\vec{u}^{\alpha\beta} - L\frac{d}{dt}\vec{i}^{\alpha\beta} - R\vec{i}^{\alpha\beta} - \vec{e}^{\alpha\beta} \approx \vec{u}^{\alpha\beta} - L\frac{d}{dt}\vec{i}^{\alpha\beta} - \vec{e}^{\alpha\beta} = 0 \tag{2.41}$$

This means that the current derivatives are expressed as

$$\begin{cases} \dfrac{di_\alpha}{dt} = \dfrac{1}{L}\left(u_\alpha - e_\alpha\right) \\ \dfrac{di_\beta}{dt} = \dfrac{1}{L}\left(u_\beta - e_\beta\right) \end{cases} \tag{2.42}$$

Assuming that the switching frequency period time T_{sw} is much longer than the time constant of the inductive filter, i.e. $T_{sw} = 1/f_{sw} << \tau_{el} = L/R$, implies that the current derivatives can be approximated according to

$$\begin{cases} \dfrac{\Delta i_\alpha}{\Delta t} = \dfrac{1}{L}\left(u_\alpha - e_\alpha\right) \\ \dfrac{\Delta i_\beta}{\Delta t} = \dfrac{1}{L}\left(u_\beta - e_\beta\right) \end{cases} \Leftrightarrow \begin{cases} \Delta i_\alpha = \dfrac{1}{L}\left(u_\alpha - e_\alpha\right)\Delta t \\ \Delta i_\beta = \dfrac{1}{L}\left(u_\beta - e_\beta\right)\Delta t \end{cases} \tag{2.43}$$

Since triangular carrier modulation is utilised, the duration of each interval can be calculated from their carrier crossings. For one half period (negative slope) of the carrier, the following expression is valid

$$u_{i,ref} = \frac{V_{dc}}{2} - \frac{V_{dc}}{T_s} \cdot t_i \quad \Leftrightarrow \quad t_i = \left(\frac{1}{2} - \frac{u_{i,ref}}{V_{dc}}\right) \cdot T_s \tag{2.44}$$

Note Δt is the time duration for which a certain voltage vector is applied, in other words the distance between two consecutive carrier crossings for the different reference voltages. Actually, for the investigated carrier wave modulation strategies, the only difference is the time duration for which each different voltage vector is applied.

The sampling time T_s is selected as half the switching frequency time period, i.e. $T_s = T_{sw}/2$. Furthermore, the sampling instants are synchronised with the carrier in such a way that the sampling instants coincide with the carrier peaks. This is done in order to perform the samplings when the currents pass their average to avoid sampling distortion.

The grid is considered as being symmetrical, i.e.

$$\begin{cases} e_a = \hat{e} \cdot \cos\left(\omega_1 t\right) \\ e_b = \hat{e} \cdot \cos\left(\omega_1 t - \dfrac{2\pi}{3}\right) \\ e_c = \hat{e} \cdot \cos\left(\omega_1 t - \dfrac{4\pi}{3}\right) \end{cases} \tag{2.45}$$

The instantaneous grid voltages for $\omega_1 t = \pi/4$ are given by

$$\begin{cases} e_a = \dfrac{1}{\sqrt{2}}\hat{e} \\ e_b = \dfrac{\sqrt{3}-1}{2\sqrt{2}}\hat{e} \\ e_c = -\dfrac{\sqrt{3}+1}{2\sqrt{2}}\hat{e} \end{cases} \tag{2.46}$$

This gives

$$\begin{cases} e_\alpha = \sqrt{\dfrac{3}{2}} \cdot e_a = \dfrac{\sqrt{3}}{2} \cdot \hat{e} \\ e_\beta = \dfrac{1}{\sqrt{2}}(e_b - e_c) = \dfrac{\sqrt{3}}{2} \cdot \hat{e} \end{cases} \tag{2.47}$$

To calculate the converter output voltage (which is not a symmetrical three-phase quantity) equation (2.38) has to be used. The switches are considered as being ideal, i.e. the output voltage assumes the two discrete levels $\pm V_{dc}/2$. The voltage vectors applied for $\omega_1 t = \pi/4$ are shown in Figure 2.25.

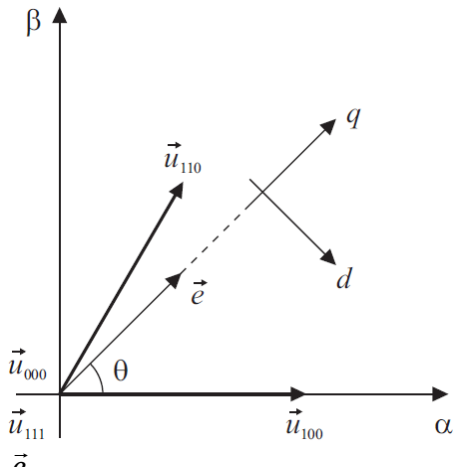


Figure 2.25 The grid voltage vector $\vec{e}$ and the converter output voltage vectors applied at $\omega 1t = \pi/4$.

The voltage vectors are calculated in Table 2.1.

Table 2.1 Voltage vectors used at $\omega_1 t = \pi/4$

Switch state	uα	uβ
000	0	0
100	$\sqrt{2/3} \cdot V_{dc}$	0

110	$1/\sqrt{6}\cdot V_{dc}$	$1/\sqrt{2}\cdot V_{dc}$
111	0	0

In the following sections the switching behaviour is investigated for the different modulation strategies. As in the previous section, the investigation is performed under the conditions listed in Table 2.2.

Table 2.2 Converter specification.

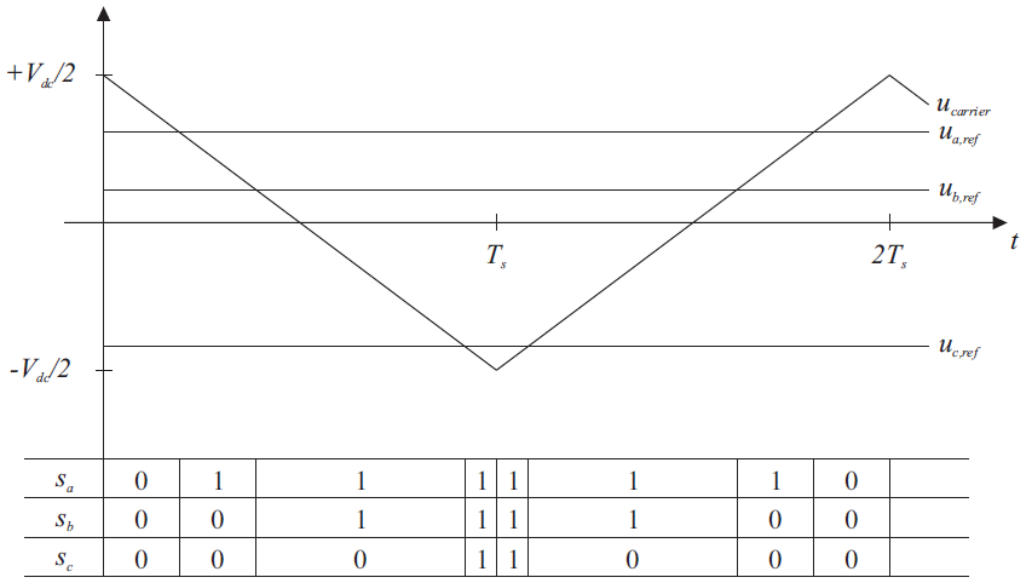
DC link voltage	Vdc	750 V
Grid peak voltage	ê	325 V
Grid frequency	f1	50 Hz
Switching frequency	fsw	5000 Hz
Line filter inductance	L	1.7 mH

Sinusoidal modulation

In the case of sinusoidal references in no-load condition, the stationary voltage reference for each phase is given by the instantaneous grid voltage. At $\omega_1 t = \pi/4$ this means that

$$\begin{cases} u_{a,ref} = \dfrac{1}{\sqrt{2}}\hat{e} = 229.8\ \text{V} \\ u_{b,ref} = \dfrac{\sqrt{3}-1}{2\sqrt{2}}\hat{e} = 84.1\ \text{V} \\ u_{c,ref} = -\dfrac{\sqrt{3}+1}{2\sqrt{2}}\hat{e} = -313.9\ \text{V} \end{cases} \tag{2.48}$$

The carrier and the converter output voltage references in the case of sinusoidal modulation are shown in Figure 2.26. Figure 2.26 also shows the switching states applied. Note that only the vectors depicted in Figure 2.25 are used, as expected.



s_a	0	1	1	1	1	1	1	0	
s_b	0	0	1	1	1	1	0	0	
s_c	0	0	0	1	1	0	0	0	

Figure 2.26 *The modulation carrier and the voltage references at $\omega_1 t = \pi/4$ in the case of modulation with sinusoidal references. The lower part shows the applied switch states.*

The corresponding carrier wave crossings occur at (calculated for the negative slope)

$$\begin{cases} t_a = \left(\frac{1}{2} - \frac{u_{a,ref}}{V_{dc}}\right)\cdot T_s = 0.194\cdot T_s = 19.4\,\mu s \\ t_b = \left(\frac{1}{2} - \frac{u_{b,ref}}{V_{dc}}\right)\cdot T_s = 0.388\cdot T_s = 38.8\,\mu s \\ t_c = \left(\frac{1}{2} - \frac{u_{c,ref}}{V_{dc}}\right)\cdot T_s = 0.919\cdot T_s = 91.9\,\mu s \end{cases} \tag{2.49}$$

The corresponding duration of each voltage vector (during the negative slope of the carrier) are, thus, equal to

$$\begin{cases} \Delta t_{000} = t_a - 0 = \left(\frac{1}{2} - \frac{u_{a,ref}}{V_{dc}}\right)\cdot T_s = 19.4\,\mu s \\ \Delta t_{100} = t_b - t_a = \left(\frac{u_{a,ref}}{V_{dc}} - \frac{u_{b,ref}}{V_{dc}}\right)\cdot T_s = 19.4\,\mu s \\ \Delta t_{110} = t_c - t_b = \left(\frac{u_{b,ref}}{V_{dc}} - \frac{u_{c,ref}}{V_{dc}}\right)\cdot T_s = 53.1\,\mu s \\ \Delta t_{111} = T_s - t_c = \left(\frac{1}{2} + \frac{u_{c,ref}}{V_{dc}}\right)\cdot T_s = 8.1\,\mu s \end{cases} \tag{2.50}$$

Applying this to equation (2.43) gives current vector variation according to Table 2.3.

Table 2.3 Current vector increase.

Switch state	Δiα [A]	Δiβ [A]
000	-3.21	-3.21
100	3.78	-3.21
110	0.77	7.77
111	-1.34	-1.34

Note that the sum of the variations equals zero. This is approximately true if the pulse number m_f, i.e. the ratio between switching frequency and fundamental of the output voltage, is high. Also, if m_f is high, the grid voltage vector does not change much during the following carrier half period since the sampled grid voltages are almost unaltered. This yields that the same converter output voltage vectors appear in the opposite order for approximately the same duration as for the previous sampling interval. The space vector current trajectory for one switching interval is shown in Figure 2.27.

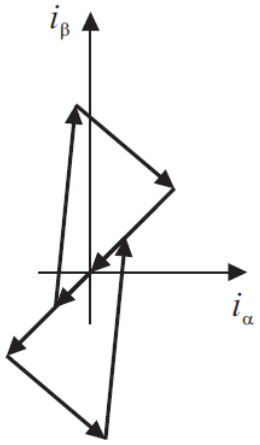


Figure 2.27 Current vectors for modulation with sinusoidal references at $\omega_1 t = \pi/4$.

Symmetrical modulation

For symmetrical references, in no-load operation, the stationary voltage reference for each phase is not given by the instantaneous grid voltage. Instead, the most positive reference and the most negative reference are altered to have the same magnitude. This is done in the following manner and gives for $\omega_1 t = \pi/4$

$$\begin{aligned} u_z &= \frac{1}{2}\left(\max\left(u_{a,ref}, u_{b,ref}, u_{c,ref}\right) + \min\left(u_{a,ref}, u_{b,ref}, u_{c,ref}\right)\right) = \\ &= \frac{1}{2}\left(u_{a,ref} + u_{c,ref}\right) = \frac{1}{2}\left(229.8 - 313.9\right)\text{V} = -42.05\text{ V} \end{aligned} \tag{2.51}$$

This gives the voltage references

$$\begin{cases} u_{az,ref} = u_{a,ref} - u_z = 271.85\text{V} \\ u_{bz,ref} = u_{b,ref} - u_z = 126.15\text{V} \\ u_{cz,ref} = u_{c,ref} - u_z = -271.85\text{V} \end{cases} \tag{2.52}$$

The corresponding switching times (calculated for the negative slope of the carrier) are

$$\begin{cases} t_a = \left(\frac{1}{2} - \frac{u_{az,ref}}{V_{dc}} \right) \cdot T_s = 0.138 \cdot T_s = 13.8\,\mu\text{s} \\ t_b = \left(\frac{1}{2} - \frac{u_{bz,ref}}{V_{dc}} \right) \cdot T_s = 0.332 \cdot T_s = 33.2\,\mu\text{s} \\ t_c = \left(\frac{1}{2} - \frac{u_{cz,ref}}{V_{dc}} \right) \cdot T_s = 0.863 \cdot T_s = 86.3\,\mu\text{s} \end{cases} \tag{2.53}$$

The duration of each voltage vector (during the negative slope of the carrier) is thus

$$\begin{cases} \Delta t_{000} = t_a - 0 = \left(\frac{1}{2} - \frac{u_{az,ref}}{V_{dc}} \right) \cdot T_s = 13.8\,\mu\text{s} \\ \Delta t_{100} = t_b - t_a = \left(\frac{u_{az,ref}}{V_{dc}} - \frac{u_{bz,ref}}{V_{dc}} \right) \cdot T_s = 19.4\,\mu\text{s} \\ \Delta t_{110} = t_c - t_b = \left(\frac{u_{bz,ref}}{V_{dc}} - \frac{u_{cz,ref}}{V_{dc}} \right) \cdot T_s = 53.1\,\mu\text{s} \\ \Delta t_{111} = T_s - t_c = \left(\frac{1}{2} + \frac{u_{cz,ref}}{V_{dc}} \right) \cdot T_s = 13.8\,\mu\text{s} \end{cases} \tag{2.54}$$

Applying this to equation (2.43) gives current variation according to Table 2.4.

Table 2.4 Current vector increase.

Switch state	$\Delta i\alpha$ [A]	$\Delta i\beta$ [A]
000	-2.28	-2.28
100	3.78	-3.21
110	0.77	7.77
111	-2.28	-2.28

Also in this case the sum of the variations equals zero. The space vector current trajectory for one switching interval, i.e. both the positive and the negative slope of the carrier, is shown in Figure 2.28.

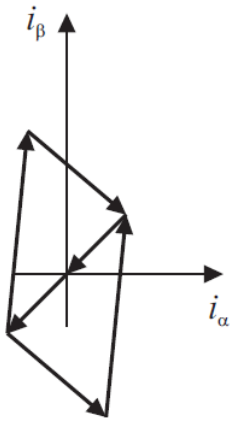


Figure 2.28 Current vectors for modulation with symmetrical references at $\omega_1 t = \pi/4$.

Bus-clamping modulation

For bus-clamped references, in no-load condition, the stationary voltage reference for each phase is not given by the instantaneous grid voltage. Instead, the reference with the most positive or negative instantaneous voltage is clamped to a level equal plus or minus half the DC link voltage, i.e. $\pm V_{dc}/2$. As for symmetrical modulation, this is equal to calculating a neutral point voltage differing from zero. This is done in the following manner

$$u_z = \begin{cases} \max\left(u_{a,ref}, u_{b,ref}, u_{c,ref}\right) - \dfrac{V_{dc}}{2} & \text{if} \quad \max\left(u_{i,ref}\right) > -\min\left(u_{i,ref}\right) \\ \min\left(u_{a,ref}, u_{b,ref}, u_{c,ref}\right) + \dfrac{V_{dc}}{2} & \text{if} \quad \max\left(u_{i,ref}\right) < -\min\left(u_{i,ref}\right) \end{cases} \tag{2.55}$$

For $\omega_1 t = \pi/4$ this yields

$$u_z = \min\left(u_{a,ref}, u_{b,ref}, u_{c,ref}\right) + \frac{V_{dc}}{2} = u_{c,ref} + \frac{V_{dc}}{2} = 61.1\,\text{V} \tag{2.56}$$

The resulting voltage references are calculated according to

$$\begin{cases} u_{az,ref} = u_{a,ref} - u_z = 168.7\,\text{V} \\ u_{bz,ref} = u_{b,ref} - u_z = 23.0\,\text{V} \\ u_{cz,ref} = u_{c,ref} - u_z = -375.0\,\text{V} \end{cases} \tag{2.57}$$

The corresponding switching times (calculated for the negative slope of the carrier):

$$\begin{cases} t_a = \left(\frac{1}{2} - \frac{u_{az,ref}}{V_{dc}}\right) \cdot T_s = 0.275 \cdot T_s = 27.5\,\mu s \\ t_b = \left(\frac{1}{2} - \frac{u_{bz,ref}}{V_{dc}}\right) \cdot T_s = 0.469 \cdot T_s = 46.9\,\mu s \\ t_c = \left(\frac{1}{2} - \frac{u_{cz,ref}}{V_{dc}}\right) \cdot T_s = 1.000 \cdot T_s = 100.0\,\mu s \end{cases} \tag{2.58}$$

The corresponding duration of each voltage vector (during the negative slope of the carrier) is equal to

$$\begin{cases} \Delta t_{000} = t_a - 0 = \left(\frac{1}{2} - \frac{u_{az,ref}}{V_{dc}}\right) \cdot T_s = 27.5\,\mu s \\ \Delta t_{100} = t_b - t_a = \left(\frac{u_{az,ref}}{V_{dc}} - \frac{u_{bz,ref}}{V_{dc}}\right) \cdot T_s = 19.4\,\mu s \\ \Delta t_{110} = t_c - t_b = \left(\frac{u_{bz,ref}}{V_{dc}} - \frac{u_{cz,ref}}{V_{dc}}\right) \cdot T_s = 53.1\,\mu s \\ \Delta t_{111} = T_s - t_c = \left(\frac{1}{2} + \frac{u_{cz,ref}}{V_{dc}}\right) \cdot T_s = 0.0\,\mu s \end{cases} \tag{2.59}$$

Applying this to equation (2.43) gives current increase according to Table 2.5.

Table 2.5 Current vector increase.

Switch state	Δi_α [A]	Δi_β [A]
000	-4.55	-4.55
100	3.78	-3.21
110	0.77	7.77
111	0.0	0.0

As before, the sum of the variations equals zero. The space vector current trajectory for one switching interval, i.e. both the positive and the negative slope of the carrier, is shown in Figure 2.29.

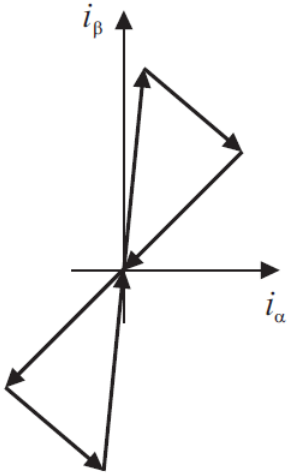


Figure 2.29 Current vectors for modulation with bus-clamped references at $\omega_1 t = \pi/4$.

2.8 Multi level converters

2-level converter.

In an ideal motor drive with a three phase AC-motor, the motor is supplied with a sinusoidal shaped voltage with the requested voltage and frequency (and phase when applicable). However, a converter with a sinusoidal shaped voltage as output is an analogue amplifier. In an analogue amplifier the voltage over the switches is high and the current through the switch is high, and a high power loss is generated. The power loss can be compared to the load power.

The traditional solution to solve this problem is to use a two level converter, see Figure 2.30, which has three different states,

Conducting. In this state the voltage over the switches is low, some volts, and the current through the switch is the load current, which is high. However, as the power loss in the switch equals the switch voltage times the current the power loss is low.

Not conducting. In this state the voltage over a switch is high, but its current is almost zero, and as the power loss equals the voltage times the current the power loss is close to zero.

Switching. A short time when the switch is changing from conducting to not conducting or, the opposite, changing from not conducting to conducting. In this state both the voltage over the switch and the current through the switch are high. As the power loss equals the voltage times the current the power loss is high. However, as the duration of the switching time is low, the total dissipated switch energy is low.

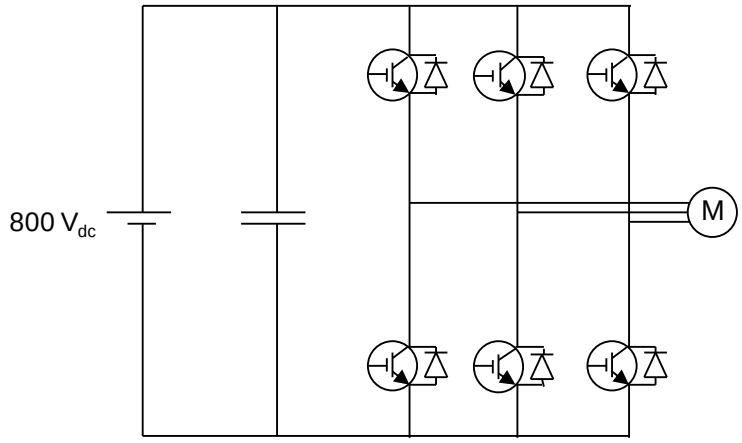


Figure 2.30 The converter with its dc-supply voltage and the single AC-phase motor

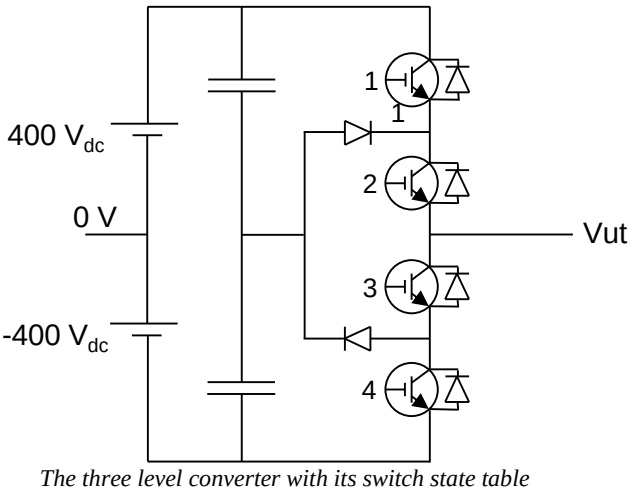
Often the load, e.g. a motor, connected to a two level converter is inductive. If the duration of the voltage pulses are adopted in a certain way, the converter phase current can be rather sinusoidal, and thus the motor torque ripple is low. The technique with voltage pulse duration is called pulse width modulation (PWM) and is well covered for 2-level converters in previous sections of this chapter. The drawback is the harmonics in the converter output voltage. These harmonics generate unwanted power loss.

To avoid this we will study some ways to generate a more sinusoidal shaped voltage

The dc supply voltage to all examples in in this section is 800 V_{dc}, in order för the different results to be compared.

3-level converter.

The first example is the 3-level converter, see Figure 2.31.



Vout(V)	400	0	-400
switch 1	on	off	off
switch 2	on	on	off
switch 3	off	on	on
switch 4	off	off	on

Figure 2.31 The three level converter with its switch state table

When the output voltage equals 400 V switch 1 and switch 2 are conducting, while switch 3 and switch 4 are not conducting. The upper diode between switch

1 and switch 2 is reversed biased so no current is flowing through this diode. The lower diode anode is connected between switch 3 and switch 4, which both are not conducting, so no current is flowing through this diode. When the output voltage equals -400 V switch 1 and switch 2 are not conducting while switch 3 and switch 4 are conducting. The lower diode between switch 3 and switch 4 is reversed biased so no current is flowing through this diode. The upper diode cathode is connected between switch 1 and switch 2 which both are not conducting, and no current is flowing through this diode. When the output voltage equals 0 V switch 2 and switch 3 are conducting, and switch 1 and switch 4 are not conducting. Both the diodes which are connected to the midpoint between the two series connected dclink capacitors can be conducting, the upper diode when the phase current is positive and the lower when the phase current is negative.

This converter is called a neutral point clamped (NPC) three level converter.

In Figure 2.32 the circuit of a three phase three level NPC converter is displayed.

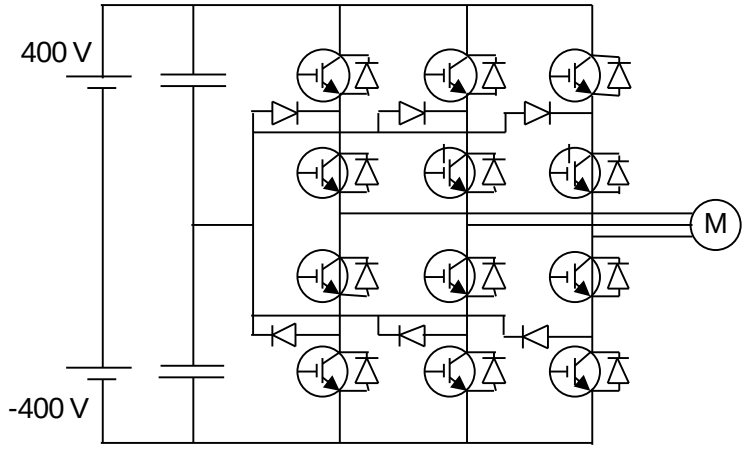


Figure 2.32 A three phase three level NPC converter

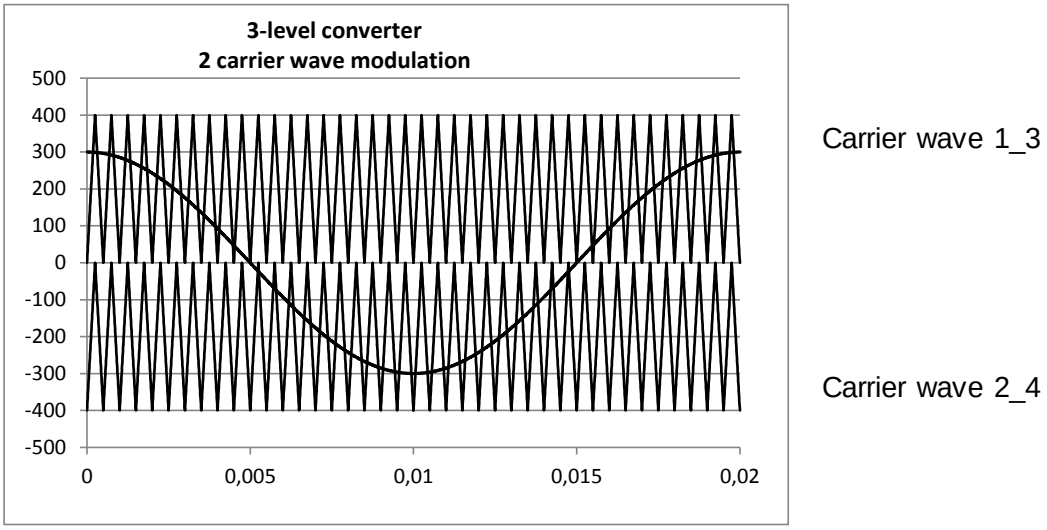


Figure 2.33 The three level converter carrier wave modulation

In Figure 2.33 the carrier wave modulation of a three level NPC converter is displayed. The upper carrier wave 1_3 modulates switch 1 and switch 3, when switch 1 is conducting switch 3 is not conducting, and vice versa. The lower carrier wave 2_4 modulates switch 2 and switch 4 when switch 2 is conducting switch 4 is not conducting, and vice versa. Note that the way carrier wave modulation is derived accounts for the "commutation voltage" and thus one carrier wave each can be applied to the two "switches" connecting an output phase of the 3-level converter EITHER to the upper or the middle DC link level OR to the middle or lower level see section 2.2.

In the diagrams in Figure 2.34 the resulting phase voltage and phase-to- phase voltage are displayed.

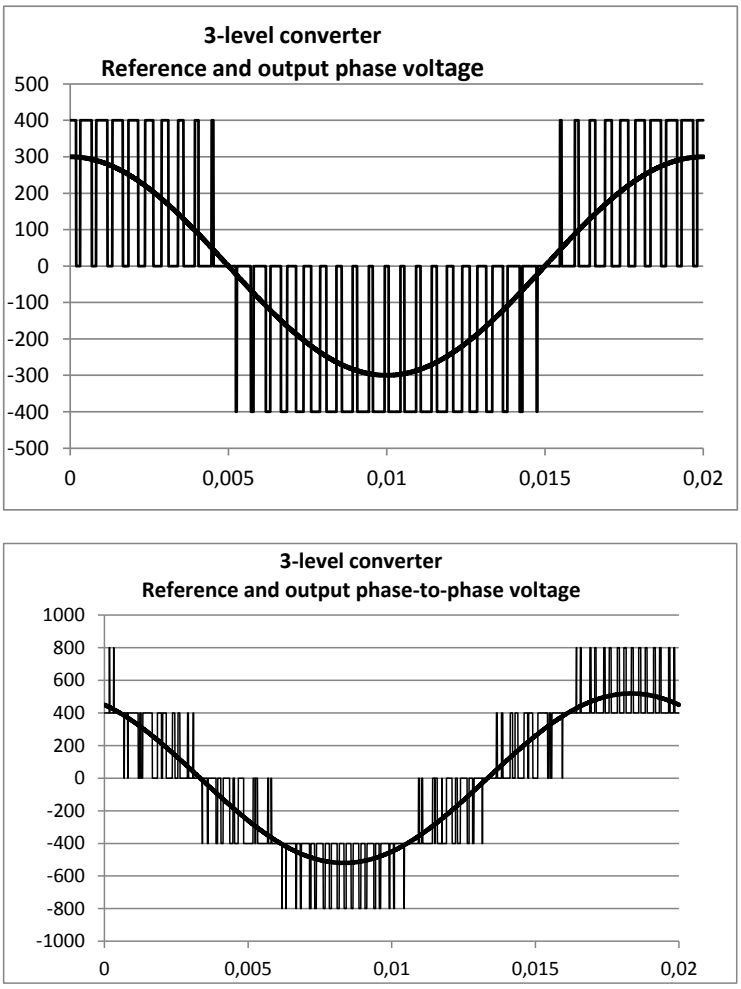
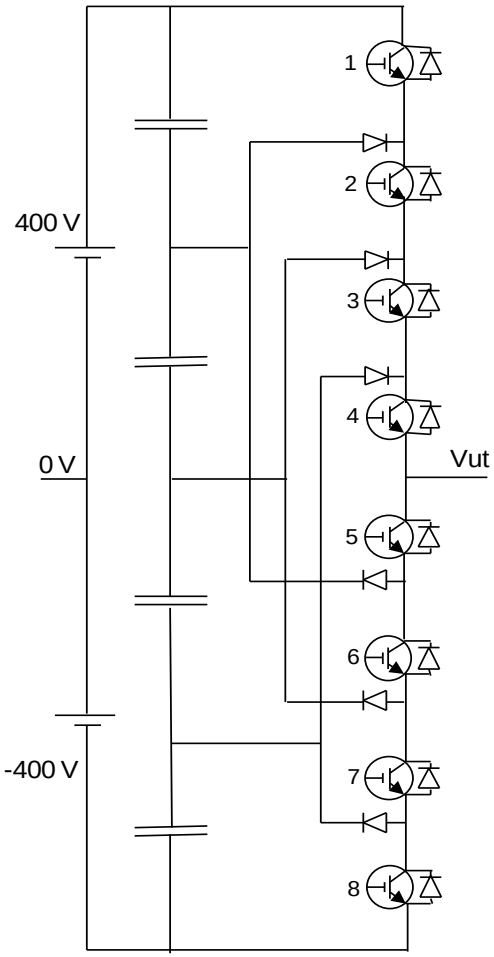


Figure 2.34 The three level converter output phase voltage and phase –to- phase voltage.

5-level converter

In a similar way as the 3-level NPC converter a 5-level NPC converter is designed, see Figure 2.35. In the 5-level converter the DC-link has four series connected capacitors, and between two capacitors a neutral point is found, and thus in the 5-level converter there will be 3 neutral points. In our case with 800

volt supply the three neutral points potential are 200 V, 0 V and -200 V. The 5 possible output voltages will be 400 V, 200 V, 0 V, -200 V and -400 V.



Vout(V)	400	200	0	-200	-400
switch 1	on	off	off	off	off
switch 2	on	on	off	off	off
switch 3	on	on	on	off	off
switch 4	on	on	on	on	off
switch 5	off	on	on	on	on
switch 6	off	off	on	on	on
switch 7	off	off	off	on	on
switch 8	off	off	off	off	on

Figure 2.35 The five level NPC converter with its switch state table

With the output voltage 400 V switches 1 to 4 are conducting, and switch 5 to 8 are not conducting, and with the output voltage -400 V switches 1 to 4 are not conducting, and switch 5 to 8 are conducting. With the output 200 V, a neutral point voltage, switch 2, switch 3, switch 4 and switch 5 are conducting and the other 4 switches are not conducting. As a general rule, when a neutral point potential is the output voltage, all 4 switches between the two diodes connected to that neutral point are conducting, and the other four are not conducting.

In Figure 2.36 a three phase five level NPC converter is displayed.

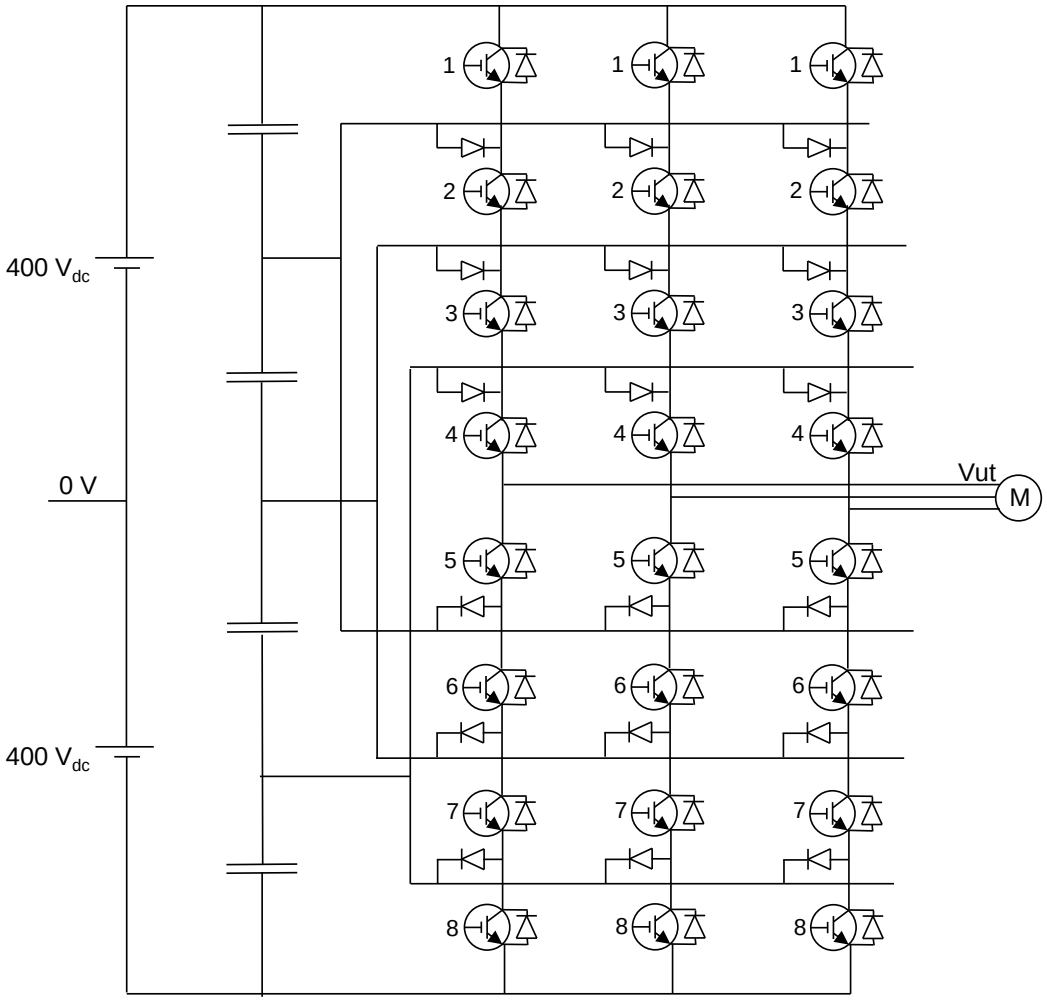


Figure 2.36 The three phase five level converter.

In Figure 2.37 the carrier wave modulation of a five level NPC converter is displayed.

- carrier wave 1_5. When switch 1 is conducting switch 5 is not conducting, and vice versa,
- carrier wave 2_6. When switch 2 is conducting switch 6 is not conducting, and vice versa,
- carrier wave 3_7. When switch 3 is conducting switch 7 is not conducting, and vice versa,
- carrier wave 4_8. When switch 4 is conducting switch 8 is not conducting, and vice versa,

See the switch state table in Figure 2.36. In Figure 2.37 the output phase voltage and the output phase-to-phase voltage in 5-level NPS converter are displayed.

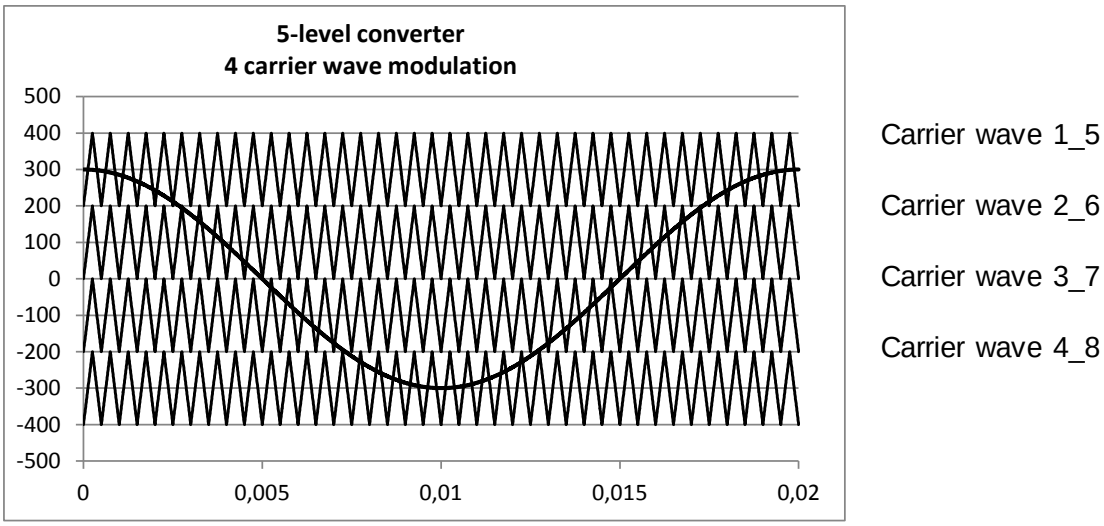


Figure 2.37 The five level converter with its four carrier waves.

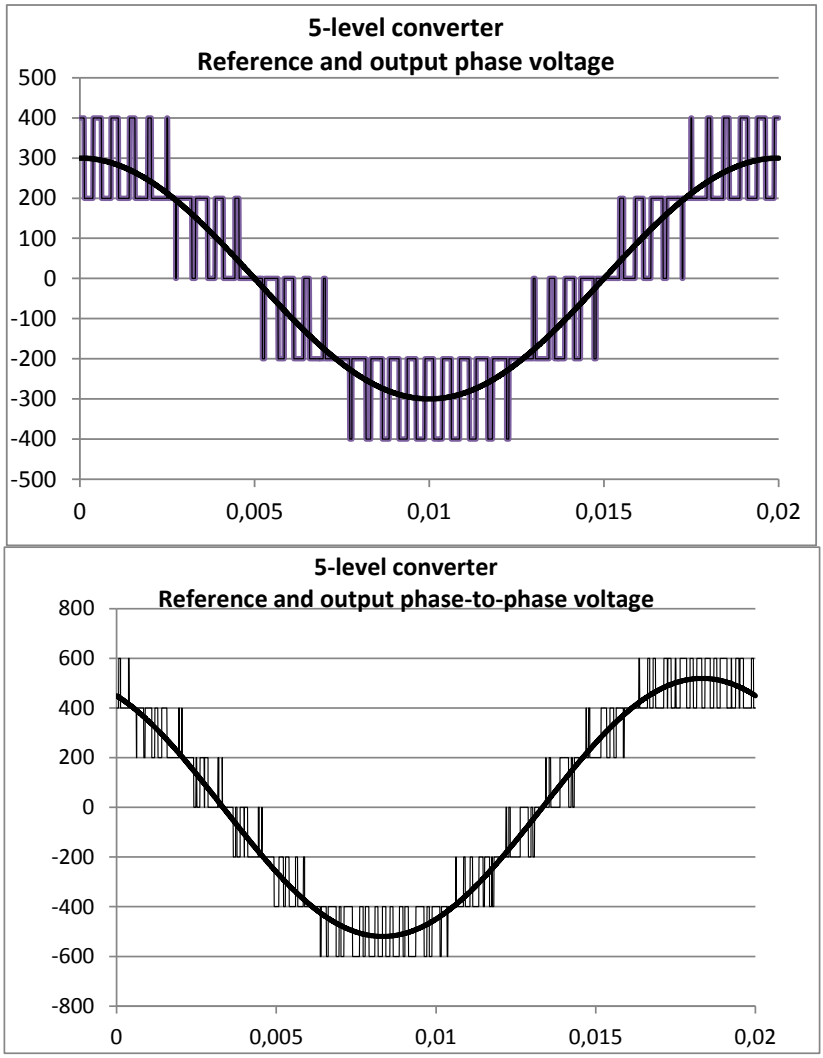


Figure 2.38 The five level converter output phase voltage and phase –to- phase voltage.

Submodules

When studying the 3 level converter and the 5 level converter, the circuit becomes more and more complex when the number of levels increase. If the

number of level shall be increased further, the circuits become more and more complex.

Here another, more general multilevel technique is described, and it is based on so called sub modules, see Figure 2.39. The sub module looks like an ordinary two level converter, with the DC-link voltage regarded constant. It is kept constant either by means of a separate supply to the capacitor via a transformer, or the modulation of the converter is done in a way so the average current through each capacitor is zero, after initially have being charged. If the upper switch is conducting and the lower is not conducting, the output voltage U_{sub} equals U_{dc}., and vice versa if the upper switch is not conducting and the lower is conducting, the output voltage U_{sub} equals zero.

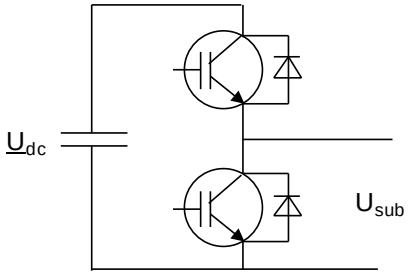


Figure 2.39 The submodule.

The benefit with this technique is that an arbitrary number of submodules can be stacked on top of each other to form the desired output voltage with the desired number of levels, without increasing the circuit complexity. With few submodules each step will be high and the output voltage shape will not be very sinusoidal, but if the number of submodules increases, and thus each step voltage will be lower the result will be a better sinusoidal shaped output voltage, see Figure 2.40. It is rather easy to increase the number of steps. Compared to 3- or 5-level converters the submodule circuit is preferred when the number of requested levels increase.

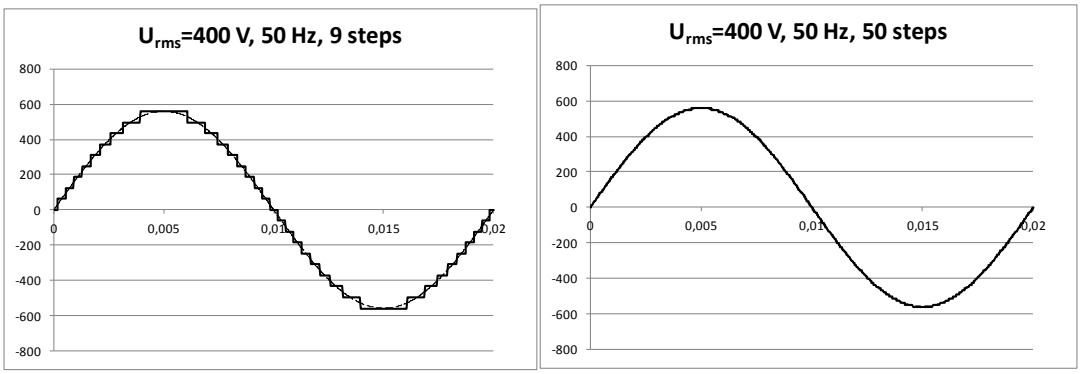
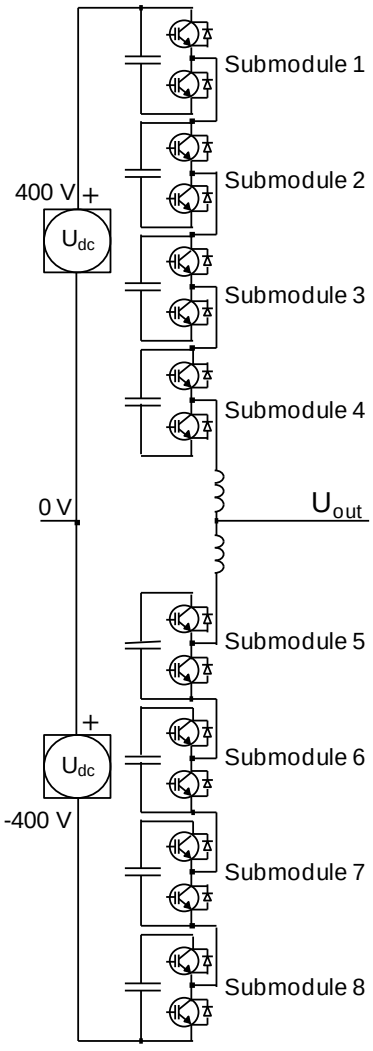


Figure 2.40 Two converter based on submodules, the left with 9 steps and the right with 50 steps.

See figure xii, with a 9-level phase leg. The displayed circuit has 8 sub modules and two series connected inductors at the output.

Example1s, see Figure 2.41with the circuit and the state table.

1. U_{out} = 400 V.
2. All submodule1 to sub module 4 Usub = 0 V, and all sub module 5 to sub module 8 Usub=200V.
3. Uout = 300 V.
4. All submodule1 to sub module 4 remain Usub = 0 V, but one of the sub modules 5 to sub modules 8 must change to 0 V, it does not matter which, and the total voltage from the sub modules 5 to sub modules 8 equals 600 V. As the dc supply equals 800 V 200 V falls over the two series connected inductors. These two inductors form a voltage divider. The voltage over each inductor equals 100 V. The output voltage Uout equals 400 V-100 V=300 V
5. Uout = 200 V. One of the submodule1 to sub module 4 must change to 200 V, it doesn't matter which, and one of the sub module 5 to sub module 8 must change to 0 V.



Vout(V)	400	300	200	200	100	0	-100	-200	-300	-400
submod 1	0	0	200	0	200	200	200	200	200	200
submod 2	0	0	0	0	0	200	200	200	200	200
submod 3	0	0	0	200	0	0	0	200	200	200
submod 4	0	0	0	0	0	0	0	0	0	200
submod 5	200	0	0	0	0	0	0	0	0	0
submod 6	200	200	200	200	200	0	0	0	0	0
submod 7	200	200	200	200	0	200	0	0	0	0
submod 8	200	200	200	200	200	200	200	200	0	0

Figure 2.41 A 9 level converter, based on submodule, and a table with the submodules voltage

"Star C" based on current pulses

At the description of the two level converter it is said that a fairly sinusoidal shaped load current is generated even with a non-sinusoidal shaped pulsed voltage by utilizing the inductance in the load, following equations are used

$$u = L \cdot \frac{di}{dt}$$

Which after integration is written as

$$i(t) = \frac{u}{L} \cdot t$$

With a certain voltage *u* the current varies linearly with the time. By changing polarity of the voltage and by varying the duration of the voltage we can form dual polarity voltage pulses and the result can be a current with sinusoidal shape.

We can instead use following equation

$$i = C \cdot \frac{du}{dt}$$

This after integration is written as

$$u(t) = \frac{i(t)}{C} \cdot t$$

The current *i* can either be a constant current or a time varying current, and when multiplied with the duration *t* it forms the charge, with which the capacitor C is charged, and the voltage of the capacitor is changed. By changing the polarity and the duration of the current, we can form a sinusoidal shaped voltage. It does not have to be one current pulse with certain duration; it can instead be many current pulses, each with much shorter duration. Instead of varying the duration of one current pulse, the number of short pulses can be varied and form the total charge with which the capacitor is charged.

If isolation is requested, an isolation transformer can be included after the generation of the current pulses, see Figure 2.42. With a high number of current pulses i.e. high frequency, a transformer with lower weight can be used.

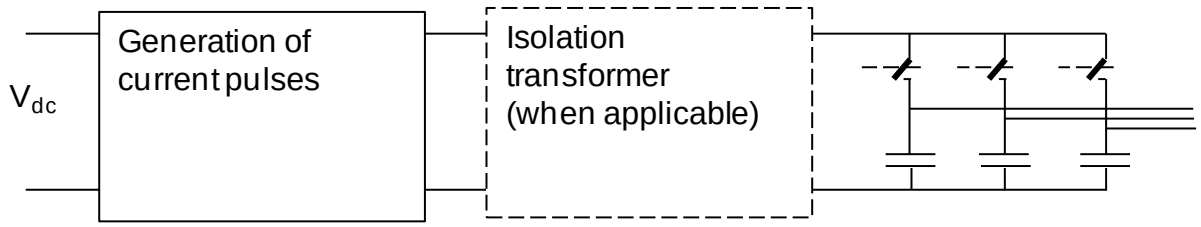


Figure 2.42 A multilevel 3-phase voltage converter

In Figure 2.43 a 3-phase converter is displayed. Each phase has its own capacitor. At a certain moment the charge, formed by the current pulse(s), is addressed to the capacitor which has the highest demand to maintain the requested phase voltage.

We can see that a large number of semiconductor switches are used in the other multilevel 3-, 5-, 9-converters

- 3-level three phase NPC converter has 12 semiconductor switches
- 5-level three phase NPC converter has 24 semiconductor switches
- 9 level three phase converter based on submodules has 48 semiconductor switches

As can be seen, the number of semiconductor switches increase rapidly with the number of levels.

In this converter, based on current pulses and output capacitors the number of semiconductor switches is low.

- A 4QC (bridge converter) for the generation of the dual polarity current pulses has 4 semiconductor switches.
- The switches, which addresses the charge to the phase where it is needed has 2 semiconductor switches per phase, in total 6 semiconductor switches.

For the complete converter the number of switches is 10 semiconductor switches, a little more than in a 2-level converter, but less than in a 3-level NPC converter.

A three phase voltage output from a real converter is displayed in the upper diagram in Figure 2.43. In the lower diagram one phase current is displayed.

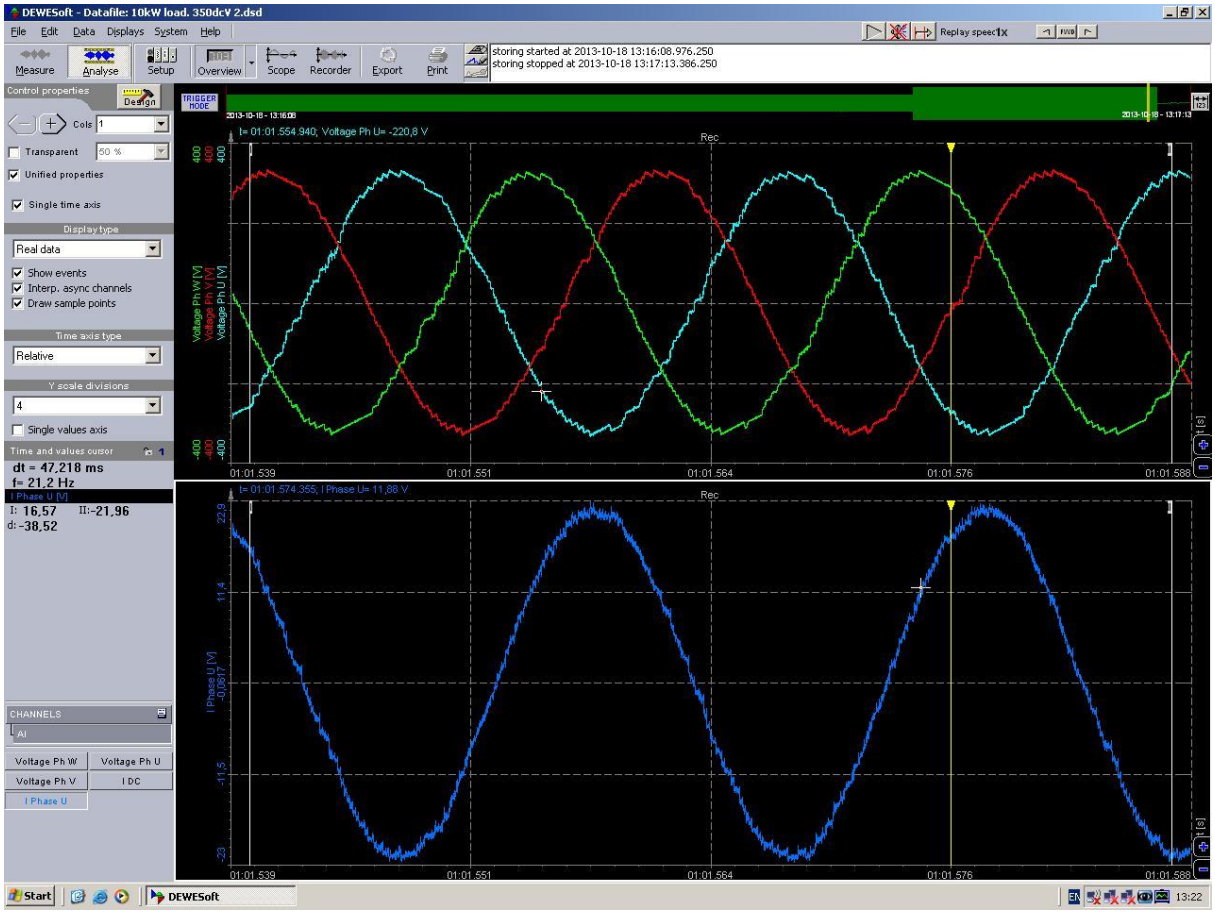


Figure 2.43 The output 3-phase output voltage(!!!) and an output phase current from a current pulses multilevel converter

(This converter type has been developed between Lund University and Bombardier Transportation, with support from the Swedish Energy Agency in Sweden. A patent application is filed)

3 Current Control

In electrical drives, the controlled quantity is ultimately the motor torque, and intermediately currents and possibly the motor flux. In power electronic applications where the power converter is connected to the grid with the purpose of injecting or controlling currents in the grid, the currents is the prime target for the control algorithms. In many other applications, such as battery charging, lighting, active anti-corrosion equipment etc. it is generally the current that is the controlled quantity.

This chapter deals with current control methods suitable for power electronic converters with idealized loads, i.e. with loads consisting of a resistance (R), inductance (L) and counter emf (*e*) in series. The methods developed for such loads are easily transferred and adapted to more complex loads such as electrical machines, but the theoretical derivation of the control algorithms are more transparent and didactic with the idealized loads.

3.1 Discrete control

In this chapter there will be three different assumptions made regarding the signal processing speed of the current control computer:

Fast computer, fix sampling interval. This means that the computer is able to provide a voltage reference for a sampling interval that is based on measurements sampled in the beginning of the sampling interval. This is relevant for analogue controllers, but generally not for digital since the processing speed is finite even with the fastest processors available today.

Slow computer, fix sampling interval. This means that the computer provides a voltage reference that is based on the measurement of currents that is made in the beginning of the previous sampling interval.

Fast computer, continuous sampling. This means that the computer samples and provides switch position selection as fast as it can. This applies to direct control methods, i.e. the switch positions are determined directly from the current errors, without going through an intermediate current control algorithm followed by a modulator.

Figure 3.1 illustrates the three different ways of using the control computer. Note that one specific computer can be used in all three different ways, depending on the sampling time and the dynamics of the application. The slow computer is usually the reality with carrier wave modulation, since the sampling

interval usually is in the range of [100…500] microseconds, and sampling, estimation of non-measured quantities and calculation of the control algorithms take several 10's of milliseconds, i.e. most of the sampling interval is passed before the voltage reference is finally calculated. As will be shown later, there are intelligent prediction methods that partially compensates for the slow calculation speed.

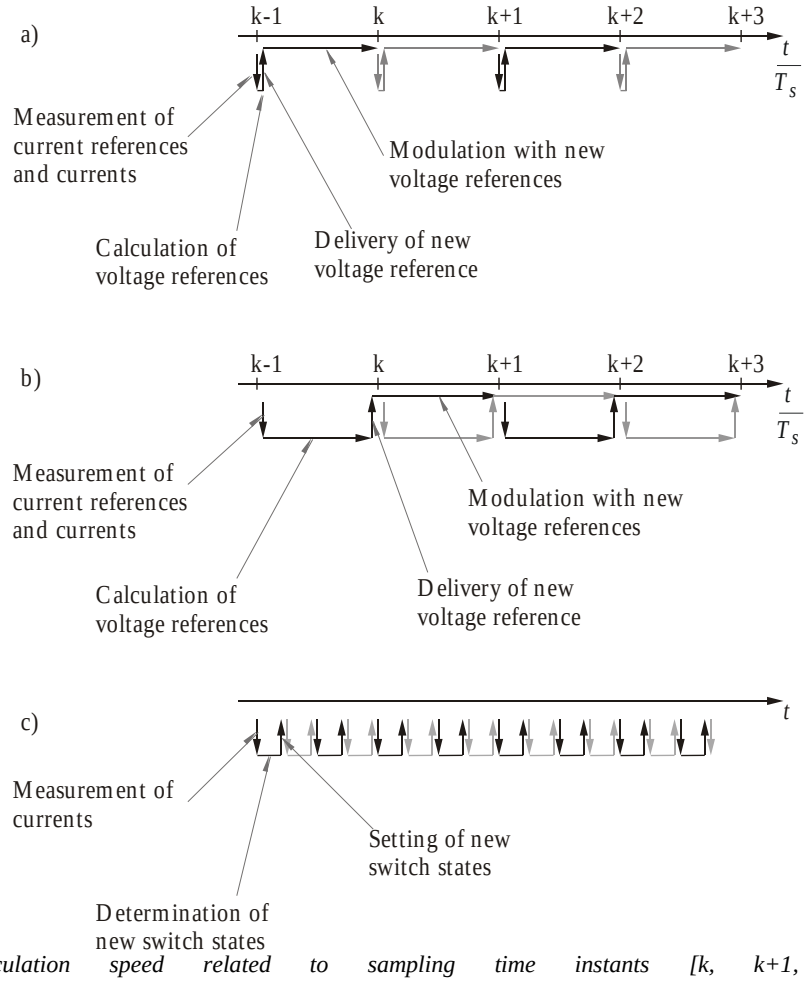


Figure 3.1 Calculation speed related to sampling time instants [k, k+1, k+2, ...] for a) a fast computer, b) a slow computer and with continuous sampling c).

When deriving current control algorithms, both the assumption of a fast computer as well as that of a slow computer will be used, and the consequence for the design of the controller parameters will be highlighted. The continuous sampling alternative is handled differently since it does not involve any current control algorithm.

3.2 Single-phase load model

The single phase load model is a resistance (R), an inductance (L) and an induced emf (e) in series according to Figure 3.2.

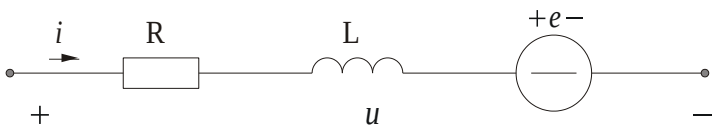


Figure 3.2 A generic 1 phase load.

The equation representing the 1-phase load is given in (3.1).

$$u = R \cdot i + L \cdot \frac{di}{dt} + e \tag{3.1}$$

Note that the parameters R and L in the 1-phase load are considered to be constant, i.e. no temperature dependence in the resistor and no saturation dependence in the inductor.

3.3 3-phase load model

The three phase load model consists of a symmetric resistive-inductive (R-L) load in series with a symmetric three phase voltage vector ($\vec{e}$).

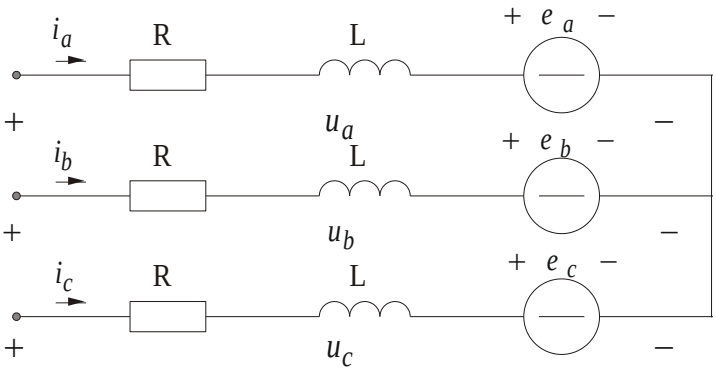


Figure 3.3 3-phase load.

The equations describing the three phase load according to Figure 3.1 can be described as three scalar equations of one vector equation according to:

$$\begin{aligned} &\sqrt{\frac{2}{3}}\left(u_a = R \cdot i_a + L \cdot \frac{di_a}{dt} + e_a\right) \\ &\sqrt{\frac{2}{3}} \cdot e^{j\frac{2\pi}{3}} \cdot \left(u_b = R \cdot i_b + L \cdot \frac{di_a}{dt} + e_b\right) \\ &+ \sqrt{\frac{2}{3}} \cdot e^{j\frac{2\pi}{3}} \cdot \left(u_c = R \cdot i_c + L \cdot \frac{di_a}{dt} + e_c\right) \\ &\vec{u} = R \cdot \vec{i} + L \cdot \frac{di}{dt} + \vec{e} \end{aligned} \tag{3.2}$$

With a symmetric load voltage $\{e_a, e_b, e_c\}$ according to equation (3.3), the load voltage vector can be derived according to equation (3.4).

$$\begin{cases} e_a = \hat{e} \cdot \cos(\omega \cdot t) \\ e_b = \hat{e} \cdot \cos\left(\omega \cdot t - \frac{2\pi}{3}\right) \\ e_c = \hat{e} \cdot \cos\left(\omega \cdot t - \frac{4\pi}{3}\right) \end{cases} \tag{3.3}$$

$$\begin{aligned} \vec{e} &= \sqrt{\frac{2}{3}} \cdot \left(e_a + e_b \cdot e^{j\frac{2\pi}{3}} + e_c \cdot e^{j\frac{4\pi}{3}} \right) = \\ &= \sqrt{\frac{2}{3}} \cdot \hat{e} \cdot \left(\cos(\omega \cdot t) + \cos\left(\omega \cdot t - \frac{2\pi}{3}\right) \cdot \left(-\frac{1}{2} + j\frac{\sqrt{3}}{2} \right) + \cos\left(\omega \cdot t - \frac{4\pi}{3}\right) \cdot \left(-\frac{1}{2} - j\frac{\sqrt{3}}{2} \right) \right) = \\ &= \sqrt{\frac{2}{3}} \cdot \hat{e} \cdot \left(\begin{array}{l} \cos(\omega \cdot t) + \left(\cos(\omega \cdot t) \cdot \cos\left(\frac{2\pi}{3}\right) + \sin(\omega \cdot t) \cdot \sin\left(\frac{2\pi}{3}\right) \right) \cdot \left(-\frac{1}{2} + j\frac{\sqrt{3}}{2} \right) \\ + \left(\cos(\omega \cdot t) \cdot \cos\left(\frac{4\pi}{3}\right) + \sin(\omega \cdot t) \cdot \sin\left(\frac{4\pi}{3}\right) \right) \cdot \left(-\frac{1}{2} - j\frac{\sqrt{3}}{2} \right) \end{array} \right) = \\ &= \sqrt{\frac{2}{3}} \cdot \hat{e} \cdot \left(\cos(\omega \cdot t) \cdot \left(1 + \frac{1}{4} + \frac{1}{4} \right) + j \cdot \sin(\omega t \cdot t) \cdot \left(\frac{3}{4} + \frac{3}{4} \right) \right) = \\ &= \sqrt{\frac{3}{2}} \cdot \hat{e} \cdot \left(\cos(\omega \cdot t) + j \cdot \sin(\omega t \cdot t) \right) = E \cdot e^{j\omega t} \end{aligned} \tag{3.4}$$

Note that in equation (3.4) E is the rms value of the phase-to-phase voltage. The load voltage is thus a vector with constant length E and constant rotating speed ω. It is useful to arrange a reference frame to the integral of this voltage vector, denoted $\vec{\psi}$ according to equation (3.5) and Figure 3.4.

$$\begin{aligned} \vec{\psi} = \int_0^t \vec{e} \cdot dt = \int_0^t E \cdot e^{j\omega \cdot t} dt = \frac{\vec{e}}{j \cdot \omega} = \\ = \frac{E}{\omega} e^{j\left(\omega \cdot t - \frac{\pi}{2}\right)} \end{aligned} \tag{3.5}$$

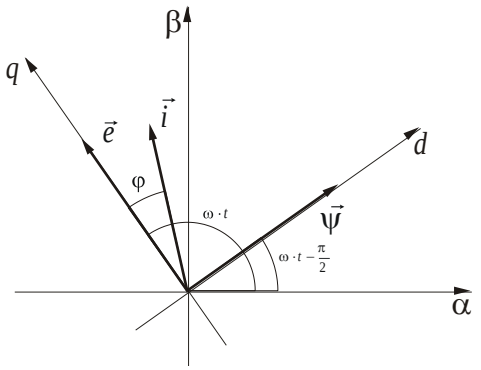


Figure 3.4 Rotating reference frame defined by the integral of the load emf vector.

The voltage equation can be expressed in the (*d*,*q*) reference frame aligned with the integral of the load voltage vector.

$$\vec{u} = R \cdot \vec{i} + L \cdot \frac{d\vec{i}}{dt} + j \cdot \omega \cdot L \cdot \vec{i} + \vec{e} \tag{3.6}$$

The active power delivered to the load, expressed in the (*d*,*q*) quantities is give by equation (3.7).

$$\begin{aligned}
\vec{u} &= R \cdot \vec{i} + L \cdot \frac{d\vec{i}}{dt} + j \cdot \omega \cdot L \cdot \vec{i} + \vec{e} \\
p(t) &= \mathrm{Re}\left\{\vec{u} \cdot \vec{i}^{\,*}\right\} = \mathrm{Re}\left\{R \cdot \vec{i} \cdot \vec{i}^{\,*} + L \cdot \frac{d\vec{i}}{dt} \cdot \vec{i}^{\,*} + j \cdot \omega \cdot L \cdot \vec{i} \cdot \vec{i}^{\,*} + \vec{e} \cdot \vec{i}^{\,*}\right\} = \\
&= \mathrm{Re}\left\{R \cdot \vec{i} \cdot \vec{i}^{\,*} + L \cdot \frac{d\vec{i}}{dt} \cdot \vec{i}^{\,*} + j \cdot \omega \cdot L \cdot \vec{i} \cdot \vec{i}^{\,*} + \vec{e} \cdot \vec{i}^{\,*}\right\} = \\
&= \underbrace{R i_d^2 + R i_q^2}_{1} + \underbrace{L \frac{di_d}{dt} i_d + L \frac{di_q}{dt} i_q}_{2} + \underbrace{e_q i_q}_{3}
\end{aligned} \tag{3.7}$$

The first term (1) in equation (3.7) is the losses in the load resistance. The second term (2) is the power delivered to the inductance upon a change of the load current. Finally the third term (3) is the power absorbed by the load emf $\vec{e}$ that by definition is oriented to the *q*-axis.

Note that the active power absorbed by the load emf $\vec{e}$ at symmetrical emf's according to equation (3.3), and with stationary symmetrical sinusoidal load currents is:

$$p(t) = E \cdot \sqrt{\frac{3}{2}} \cdot \left|\vec{i}\right| \cdot \cos(\varphi) = \sqrt{3} \cdot E \cdot I_{rms,\,phase} \cdot \cos(\varphi) \tag{3.8}$$

where $I_{rms,phase}$ is the rms value of the phase current. Note that equation (3.8) is equivalent to the conventional power expression in stationary symmetrical sinusoidal three phase systems. Note also that not equation (3.7), that is the instantaneous power without requirement of sinusoidal stationarity, nor equation (3.8), that requires sinusoidal stationarity, has any term corresponding to the conventional expression for reactive power. The expression "instantaneous reactive power" that sometimes is heard in the power engineering community is misleading; there is no such thing as instantaneous reactive power!

3.4 Sampled current control of a generic single phase load with a fast computer

It is assumed that the current is given at the sampling time instant as *i(k*) and the current reference at the same instant *i*(k)*. Furthermore, it is assumed that the current error (*i*(k)-i(k)*) will be eliminated during the next sampling interval, i.e. during the interval [*k,k+1*]. The voltage needed during this interval can thus be derived from the voltage equation (3.1) by integration over one sampling interval T_s according to equation (3.9).

$$\frac{\int_{k\cdot T_s}^{(k+1)T_s} u\cdot dt}{T_s} = \frac{R\cdot \int_{k\cdot T_s}^{(k+1)T_s} i\cdot dt + L\cdot \int_{k\cdot T_s}^{(k+1)T_s} \frac{di}{dt}\cdot dt + \int_{k\cdot T_s}^{(k+1)T_s} e\cdot dt}{T_s} = \tag{3.9}$$

$$= \bar{u}(k,k+1) = R\cdot \bar{i}(k,k+1) + L\cdot \frac{i(k+1)-i(k)}{T_s} + \bar{e}(k,k+1)$$

To convert the "average voltage"-equation (3.9) into a current control algorithm some assumptions have to be made:

$$\begin{aligned} \bar{u}(k,k+1) &= u^*(k) && (a)\\ i(k+1) &= i^*(k) && (b)\\ \bar{i}(k,k+1) &= \frac{i^*(k)+i(k)}{2} && (c)\\ \bar{e}(k,k+1) &= e(k) && (d)\\ i(k) &= \sum_{n=0}^{n=k-1}\left(i^*(n)-i(n)\right) && (e) \end{aligned} \tag{3.10}$$

Assumption (*a*) is straight forward, the reference voltage for the coming period should be the voltage needed to achieve the desired current change. Assumption (*b*) is an assumption of "dead beat" current control, i.e. the entire current error is eliminated in one sampling interval. Assumption (*c*) is an approximation, and is not always true, since the trajectory of the current will depend on the relation between the load emf *e*, the dc link voltage U_{dc} and the switching frequency. However, the assumption (*c*) is true in average, and furthermore it is related to the resistive voltage drop that normally is relatively small, thus as small error in this assumption is not expected to distort the current control performance significantly. The assumption (*d*) is usually true since the sampling frequency usually is several orders of magnitude faster than the dynamics of the load emf. One example is a DC machine where the emf corresponds to the induced voltage due to speed, which usually changes in the [10…100] millisecond range, whereas the sampling interval is in the [100-1000] microsecond range. The assumption (*e*) is a consequence of the "dead beat" assumption (*b*). The current

at instant *k* is the sum of all changes of currents until the last change depicted in the previous sampling instant. With all these assumptions a current controller can be derived from the "average voltage" equation (3.9) as:

$$\begin{aligned} u^*(k) &= R \cdot \frac{i^*(k)+i(k)}{2} + L \cdot \frac{i^*(k)-i(k)}{T_s} + e(k) = \\ &= R \cdot \frac{i^*(k)-i(k)}{2} + R \cdot i(k) + L \cdot \frac{i^*(k)-i(k)}{T_s} + e(k) = \\ &= \left(\frac{L}{T_s} + \frac{R}{2}\right)\left(i^*(k)-i(k)\right) + R \cdot \sum_{n=0}^{n=k-1}\left(i^*(n)-i(n)\right) + e(k) = \\ &= \left(\frac{L}{T_s} + \frac{R}{2}\right) \cdot \left(\underbrace{\left(i^*(k)-i(k)\right)}_{\text{Proportional}} + \underbrace{\frac{T_s}{\left(\frac{L}{R}+\frac{T_s}{2}\right)} \cdot \sum_{n=0}^{n=k-1}\left(i^*(n)-i(n)\right)}_{\text{Integral}} \right) + \underbrace{e(k)}_{\substack{\text{Feed} \\ \text{forward}}} \end{aligned} \tag{3.11}$$

Note that a "PIE" current controller now is derived, with both gain and integral time constant directly related to the circuit parameters, and with a feed forward term directly related to the emf of the load. These properties are useful since it is now possible to determine the controller sensitivity to changes of a particular parameter, e.g. saturation dependent inductances and temperature dependent resistances. A comparison to conventional PI-controller derivation is made in section 3.8.

3.5 Sampled current control of a generic single phase load with a slow computer

The case with a slow computer corresponds to a second order system, with a delay of one sample interval between the calculation of an output voltage set point and the actual use of that set point in the modulation. The most common way to get around this problem is to use a prediction of the measured current instead of the real measured current. The prediction of the current at sample instant {k} is based on the measured current at time instant {*k*-1} extrapolated with the expected change of the current due to the known applied voltage during the interval {*k*-1,*k*}, i.e. *u**(*k*-1). This can be done in at least two different ways:

1. The expected current change can be based on the applied voltage during the interval {k-1,k}, i.e. u*(k-1), the estimated resistive voltage drop and the estimated load emf, by using equation (3.1). The local current change is thus calculated, but no state is remembered for future use. The advantage of this method is the accuracy if all the input variables (voltage reference, resistive voltage drop and induced emf) are known. On the other hand, in the case that these variables are not known well enough, the errors in these may excite control errors.
2. A simplified load model can be made that does not include any induced emf, only an inductance and a resistance. In this simplified model the state of the current will be remembered, and it will be wrong, but the change of the current will be correct when the input signal, the voltage reference is changing. Since it is the state that is desired, such a simplified state model is enough.

Both these methods are variants of the same idea, often referred to as a "Smith predictor". The circuit model in the single phase load model can be derived as a discrete model since the controller is discrete, with the emf omitted:

$$u = R \cdot i + L \cdot \frac{di}{dt} \text{ (NOTE! No emf modelled)}$$

$$u^*(k) = u(k,k+1) = R \cdot i_{sp}(k) + \frac{L}{T_s} \cdot \left(i_{sp}(k+1) - i_{sp}(k) \right)$$

or, in the discrete domain;

$$\begin{aligned} u &= R \cdot i_{sp} + \frac{L}{T_s} \cdot (z-1) \cdot i_{sp} \\ \frac{i_{sp}}{u} &= \frac{1}{R + \frac{L}{T_s} \cdot (z-1)} = \frac{1}{\frac{L}{T_s} \cdot z + \left(R - L/T_s \right)} = \frac{T_s}{L} \cdot \frac{1}{z - (1 - R \cdot T_s / L)} \end{aligned} \qquad (3.12)$$

or, again in the time domain;

$$i_{sp}(k+1) = i_{sp}(k) \cdot \left(1 - R \cdot T_s / L \right) + u^*(k) \cdot \frac{T_s}{L}$$

With the transfer function of equation 3.12, a "dummy" current i_{smith} is created, that does not have the same stationary value as the real current of the circuit that we are really controlling, but has the same dynamics. In stationarity, the dummy current i_{smith} will reach a level determined by the actual voltage and the resistance only. However, when any change is made to the voltage reference there will be a change in the dummy current i_{smith} that will be identical to whatever it would have been even if the induced emf was represented in the model. Since we are only interested in the change of the current to predict the

actual circuit current based on the latest measurement and the smith-prediction, this dummy model is enough. The current estimate to be used by the controller of the "slow computer" is thus finally calculated as:

$$\hat{i}(k+1) = i(k) + \left(i_{sp}(k+1) - i_{sp}(k)\right)$$

The accuracy of the current prediction is of course dependent on the accuracy of the parameters and the estimate of the induced voltage used by the controller. A simple sensitivity analysis is presented in the next section. The current controller algorithm when using a slow computer is thus given by equation (3.13).

$$u^*(k) = \left(\frac{L}{T_s} + \frac{R}{2}\right) \cdot \left(\left(i^*(k) - \hat{i}(k)\right) + \frac{T_s}{\left(\frac{L}{R} + \frac{T_s}{2}\right)} \cdot \sum_{n=0}^{n=k-1} \left(i^*(n) - \hat{i}(n)\right) \right) + \hat{e}(k) \quad (3.13)$$

The quality of the current prediction depends on application. With a single phase load like the one used in this chapter, it is possible to measure e.g. the emf *e*, and the inductor parameters are usually rather easy to measure. When controlling e.g. a DC machine, the emf is not available for measurement and the estimate is of course more difficult to make. It remains to be determined in the specific application what methods can be used, and with what accuracy, to estimate circuit parameters and the emf.

Sampled current controller applied to a 2 quadrant converter.

The current controllers derived in the previous sections will now be applied to study some different behaviors depending on limitations in the converter. A circuit with the controller, a 2-quadrant converter and the one phase load is shown in Figure 3.5.

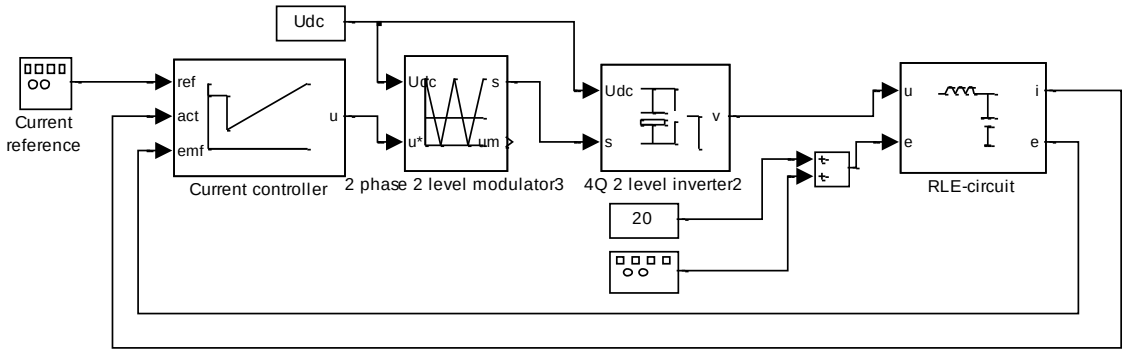


Figure 3.5 2 quadrant DC link converter.

Assume a load circuit and a 2-quadrant DC converter with parameters according to:

$$L = 10 \ mH$$
$$R = 1 \ \Omega$$
$$Ts = 0.5 \ ms$$
$$U_{dc} = 100 \ V$$

To evaluate the performance of the developed current controllers, a series of simulations are made. In all cases a series of current reference steps between +/- 5 A are used. The following situations are simulated:

3. The current controller has the right set of parameters, i.e. a perfect measurement of the load parameters and an as perfect measurement of the load emf is used. The controller is of the "fast computer" type. Three different cases are simulated, each with a unique emf e to show the dynamic performance at different emf's.
4. In the other cases the current controller is of the slow computer type, and the controller parameters are deliberately different from those of the load, to show the performance of a mistuned current controller. Six different cases are simulated; three with $R_{controller}$=[0.5 1.0 2.0]R and $L_{controller}$=L, and three with $L_{controller}$=[0.9 1.0 1.1]L and $R_{controller}$=R. The load emf is e=30 V in all cases.

Figure 3.6 shows the step response for situation 1 above. Note the difference in raise time and fall time at the three different emf's. Note how the current controller takes more than one sample interval to settle the step response. The controller is equipped with anti-windup function. Would it be possible to have the negative half period of the current in case that the emf e=0V?

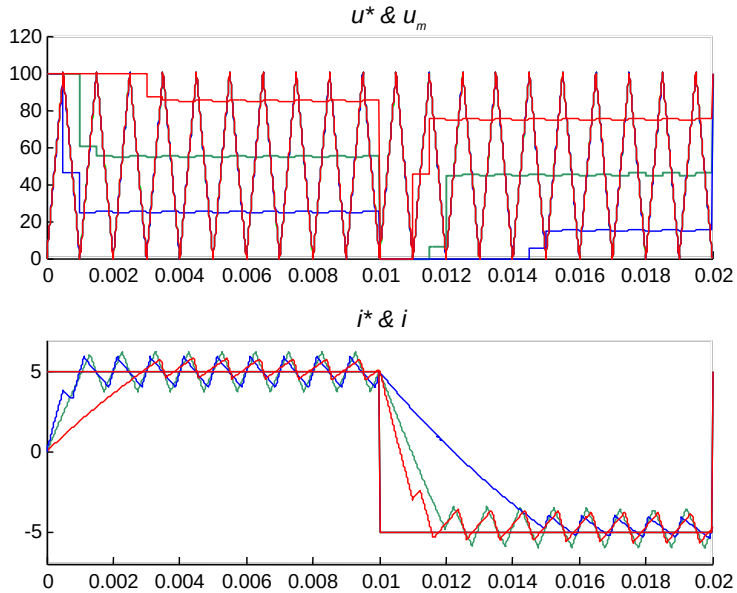


Figure 3.6 *Load current step response. Upper diagram with modulating wave and references with e=[20, 50, 80] V. Lower diagram with current reference and current.*

Figure 3.7 shows the current step response for situation II above, with the estimated load resistance tuned to 50%, 100% and 200% respectively. Note the small difference between the three cases, indicating a low sensitivity to variations in the load resistance.

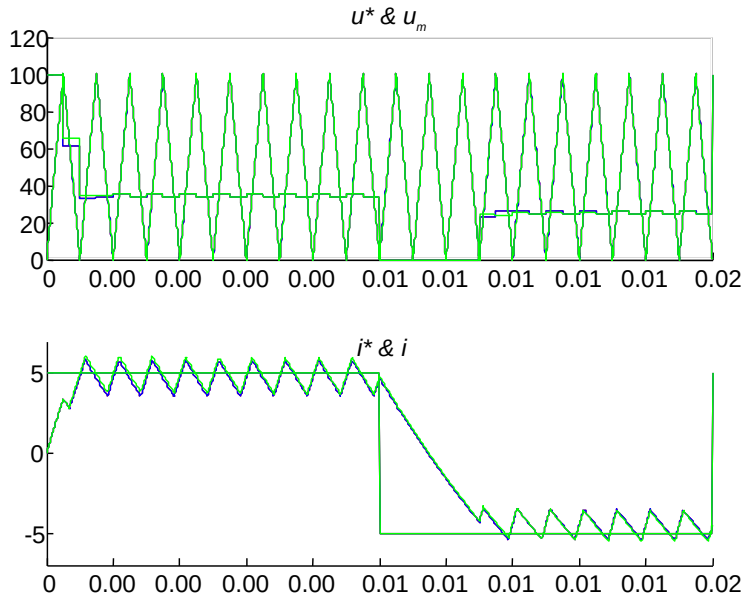


Figure 3.7 *Current step response with a "slow computer" current controller and mistuned resistance with [50%, 100%, 200%] of the load resistance used in the controller.*

Figure 3.8 shows the current step response for situation II above, with the estimated load inductance tuned to 80%, 100% and 120% respectively.

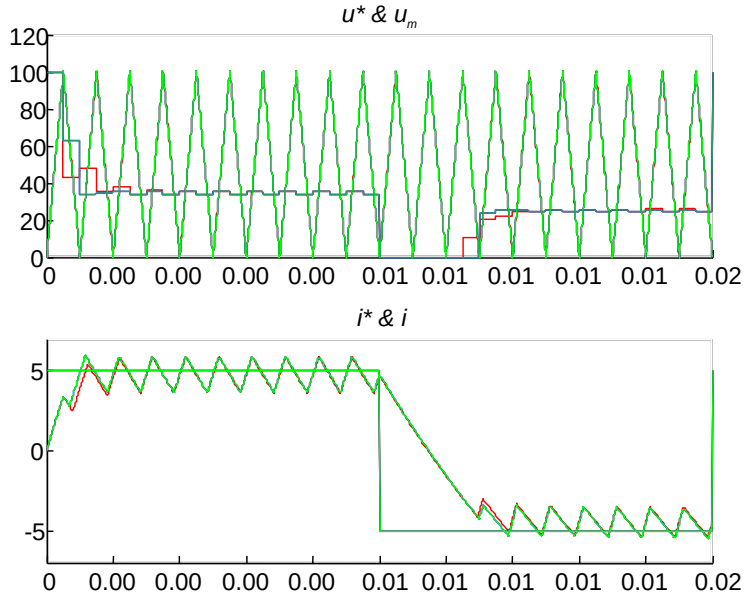


Figure 3.8 *Current step response with a "slow computer" current controller and mistuned inductance with [80%, 100%, 120%] of the load inductance used in the controller.*

Sampled current controller applied to a 4-quadrant converter.

The current controllers derived in the previous sections will now be applied to study some different behaviors depending on limitations in the converter. A

circuit with the controller, a 2-quadrant converter and the one phase load is shown in Figure 3.9.

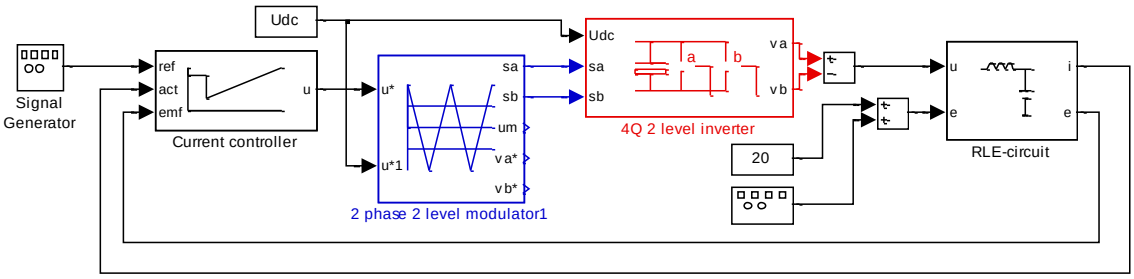


Figure 3.9 4 quadrant DC link converter.

Assume a load circuit and a 2-quadrant DC converter with parameters according to:

$$L = 10 \ mH$$
$$R = 1 \ \Omega$$
$$Ts = 0.5 \ ms$$
$$U_{dc} = 100 \ V$$

To demonstrate the difference to the 2 quadrant converter a simulation with the emf *e*=20 V is made. In the 2-quadrant the performance at the negative current step was rather slow (see Figure 3.8) but the 4-quadrant converter is able to supply a negative voltage, thus a faster current fall. Figure 3.10 shows the simulation results.

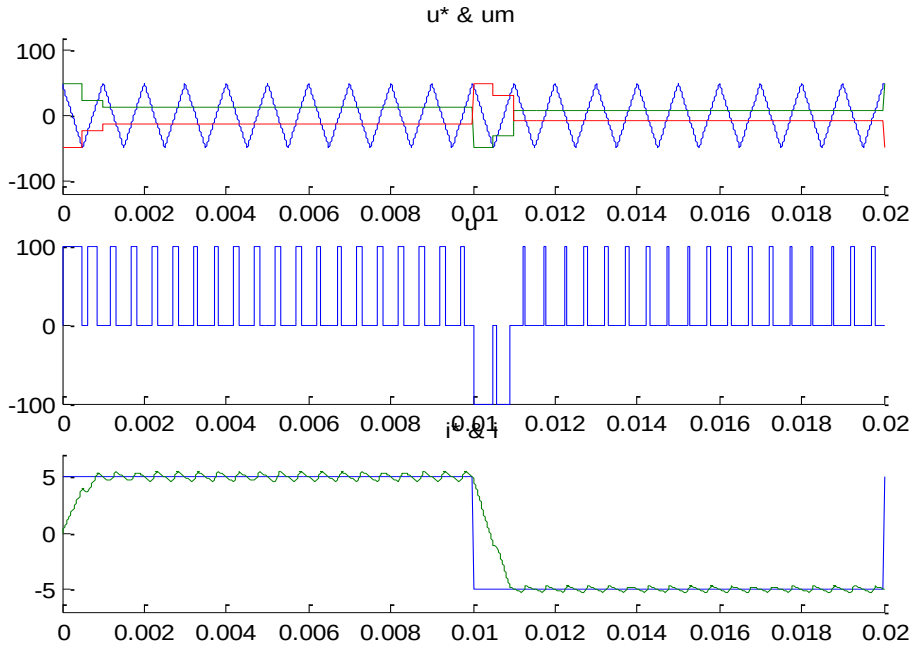


Figure 3.10 Step response with a 4-quadrant converter.

Note the fast negative current response, due to the ability of the 4-quadrant converter to supply the load with a negative load.

3.6 Direct current control of a single phase load

With a finite number of switch states it is possible to select a relevant switch state based on the instantaneous value of the current error, i.e. without using any intermediate current controller and modulator as have been used in the previous sections. This method is used in several AC motor drive types, denoted "DTC"-Direct Torque Control. It is perfectly possible to apply the same methods to DC current control with 2-quadrant or 4-quadrant converters. Some examples on how this could be done are shown in the following sections.

Direct Current Control with a 2-quadrant converter

This case is the straightest forward, since there are only two switch states to select. A switch function is applied to the current error, selecting the proper switch state according to equation (3.14). Note that the switch state is denoted [-1,1] instead of [0,1] as has been used before. This change is introduced due to simplicity in the implementation.

$$s = \begin{cases} 1 \; if \;\; i < i^* - \dfrac{\Delta i}{2} \\ -1 \; if \;\; i > i^* + \dfrac{\Delta i}{2} \\ s \; if \;\; i^* - \dfrac{\Delta i}{2} < i < i^* + \dfrac{\Delta i}{2} \end{cases} \qquad (3.14)$$

where Δi is the allowed current ripple in the load current. The switch function defined by equation (3.14) is easy implemented in a relay with hysteresis. Figure 3.11 shows the top-level block diagram of a Simulink model of a 2-quadrant DC converter with a Direct Current Controller according to equation (3.14).

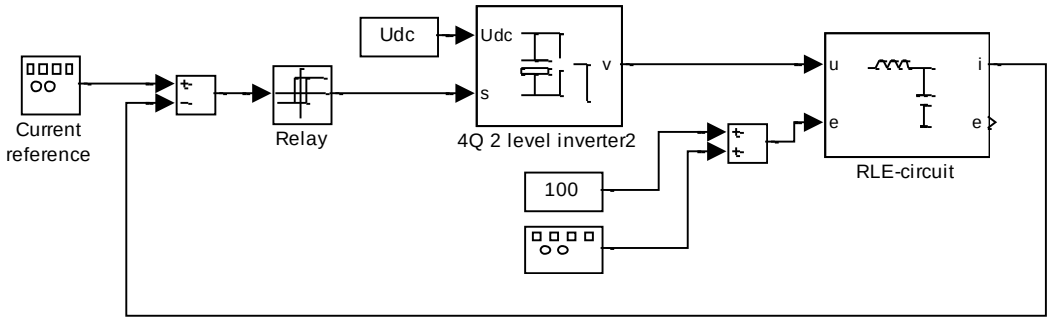


Figure 3.11 2-quadrant DC converter with a Direct Current Controller implemented as a relay.

The switching frequency is not constant with this type of controller, since the transition time between the two extreme values of the current ripple depends on the emf *e* in relation to the applied voltage *u*. When *e* approaches *u* the transition time and thus the switching frequency approaches zero. Figure 3.12 shows an example of a simulation with the system in Figure 3.11 with an emf *e* that increases monotonously from *e*=0 V to *e*=100 V and a current reference being *i**=0 A. The other parameters are:

$$L = 10 \ mH$$
$$R = 1 \ \Omega$$
$$Ts = 0.5 \ ms$$
$$U_{dc} = 100 \ V$$
$$\Delta i = 3 \ A$$

Note from Figure 3.12 that the current ripple has constant amplitude but varying frequency. The frequency variation is sometimes regarded, as a draw back since it may excite undesired resonance frequencies in nearby systems in a particular application. The audible noise from the electromagnetic forces within the reactor does have correspondingly wide frequency content. The character of the audible noise is often put forward as a benefit with this kind of modulation, but such statements are very subjective. Figure 3.13 shows a current step with the same control system.

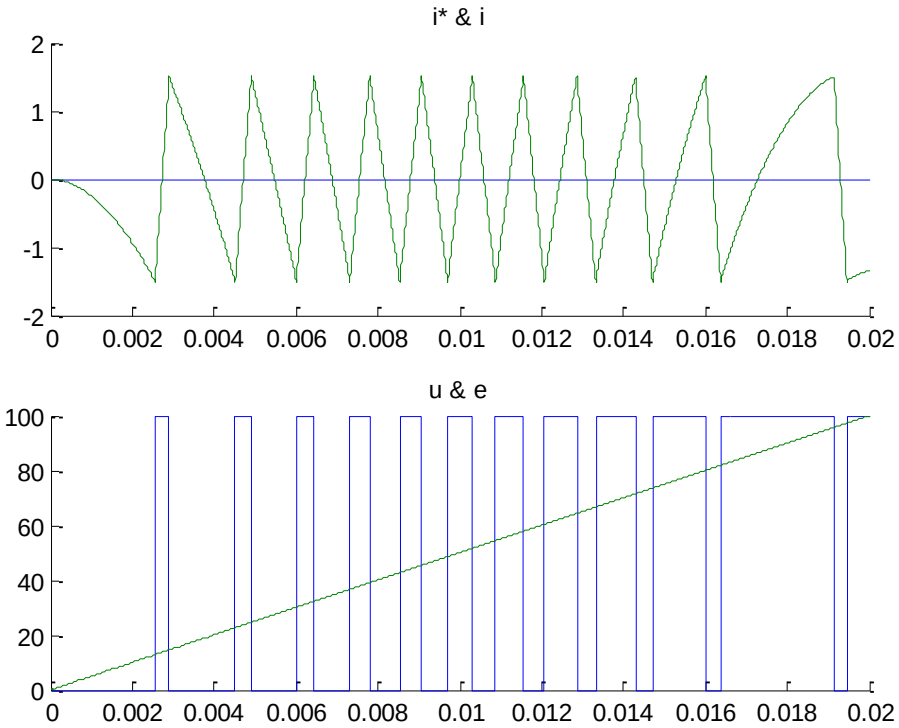


Figure 3.12 Current ripples with a varying emf e and zero current b reference.

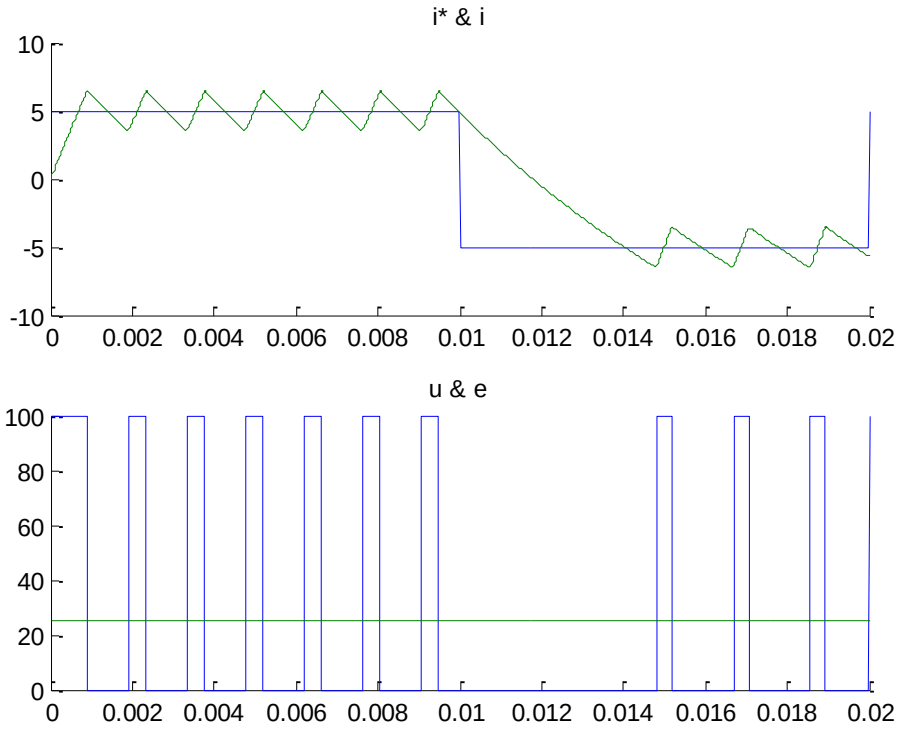


Figure 3.13 *Step response with a Direct Current Controller and a 2-quadrant DC converter.*

Note the fast step response, i.e. no time delay elapses between a change of the reference and a setting of new switch states. This is regarded as the major advantage with direct current control, but is requires that the processor that implements the control operates with high sampling frequency, since it has to follow the current all the time.

Direct Current Control with a 4-quadrant converter

This case is a bit more difficult to handle than the 2-quadrant case, since there are 4 switch states to select, but only three output voltages [U_{dc}, 0, -U_{dc}]. One solution to this problem is to apply two tolerance bands to the current ripple, one slightly wider than the other. To understand the point with this, consider what happens when controlling a current against a positive load emf *e*. To increase the current, it is necessary to use the switch state [1,-1], i.e. the left branch connected "up" to Udc/2, and the right branch connected "down" to -Udc/2. When the current thus increases it eventually hits the top of the tolerance band, and a new switch state has to be selected to force the current down again. The new switch state can be either [-1, 1] or one of the alternatives [1,1] or [-1,-1]. The two latter ones represent a zero voltage applied to the load, and would be enough since the current will decrease under the influence of the load emf *e*. The active state [-1,1] will of course also work, but has the draw back that both switches are operated at the same time. Another drawback with choosing the active state [-1,1] to reduce the current is that the decent of the current will be faster, under the influence of both the emf *e* and the applied voltage u=-U_{dc}. The

transition to the negative limit of the current ripple will be unnecessary fast and generate unnecessary high switching frequency. A switching strategy that involves the inactive states [-1,-1] and [1,1] is thus advantageous, since it will reduce the number of switching's.

Figure 3.14 shows a tolerance band that involves the passive states as well. The tolerance band is designed with two widths. At very high current errors, a suitable active state is always chosen. When the current error decreases, after the use of an active state, and eventually hits the inner tolerance band a passive state is selected.

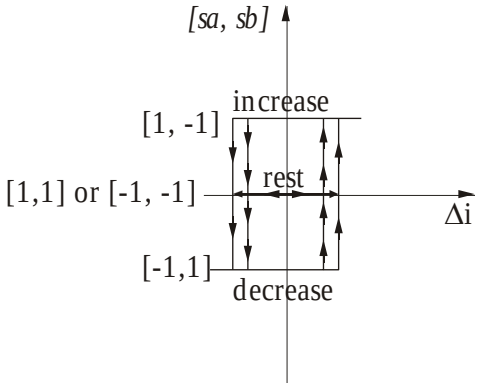


Figure 3.14 4 quadrant DCC hysteresis including the passive states.

Assume like in the previous section that the current is to be controlled against a positive emf *e*, starting with a positive step higher than the tolerance bandwidth. The sequence of states is the generated according to the following:

1. The current error is large and positive so the active state [1,-1] is selected. The current increases and the current error correspondingly decrease.
2. When the current error has become negative it eventually hits the inner tolerance band in the 2'nd quadrant in Figure 3.14. A passive state is then selected, and the current will fall under the influence of the positive emf e, and the current error will after a while become positive.
3. When the current error hits the inner tolerance band level in the 1'st quadrant of Figure 3.14, the positive active state [1,-1] will be selected again. The current increases and the current error correspondingly decrease. Proceed to step 2 above …
4. If the current error becomes negative and higher than the inner tolerance band (e.g. due to a big negative step in the current reference), the active negative state [-1,1] is selected. The current error will after a while become positive again and hit the inner tolerance band in the 4'th quadrant in Figure 3.14. A passive state is then selected. If the emf e>0, as assumed in this example, then the current will continue to decrease under the influence of e and the current error will become even more positive and finally hit the outer tolerance band and set the active state [1,-1]. Then the current will grow and the current error will decrease again.

Since there are two passive states, a rule must be made on which of the two states that should be used at a given time. It is favorable to use both passive states approximately the same number of times since the losses that inevitable will be dissipated in the semiconductors thus will be distributed evenly in all semiconductors in the bridge. A simple rule on how to do this is thus to change the passive state every time a passive state is selected.

Figure 3.15 upper shows the top level of a Simulink simulation program with a hysteresis controller implementing the double tolerance band given in Figure 3.14, and Figure 3.15 lower part shows the content of the current controller block.

Note in Figure 3.15 that the upper hysteresis has the inner (narrower) hysteresis width and the lower hysteresis block has the outer (wider) hysteresis width. Note that the block named "Toggle" toggles between state [1 1] and [-1 –1] every time a passive state is called.

In Figure 3.16 a simulation example is given with the following input data:

$$L = 10 \ mH$$
$$R = 1 \ \Omega$$
$$U_{dc} = 100 \ V$$
$$\Delta i = 4 \ A$$

Note from Figure 3.16 that the only time the negative active state [-1 1] is used is at the negative transition of the current reference. Note also that the two passive states are altered every time a passive state is used.

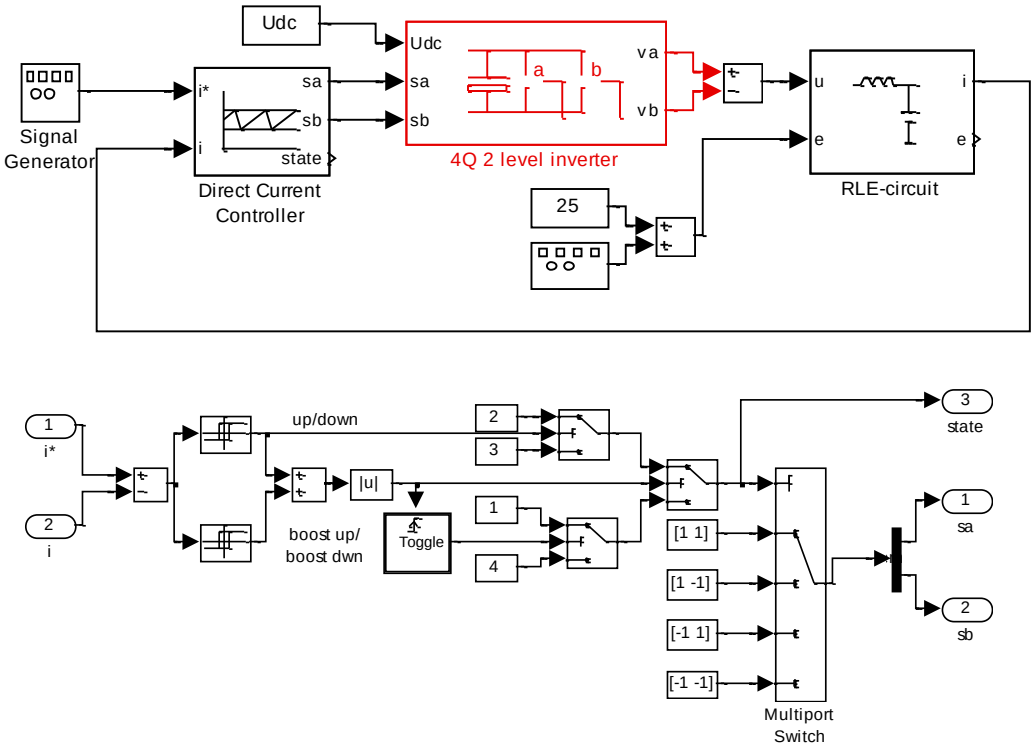


Figure 3.15 4-quadrant DC converter with a Direct Current Controller.

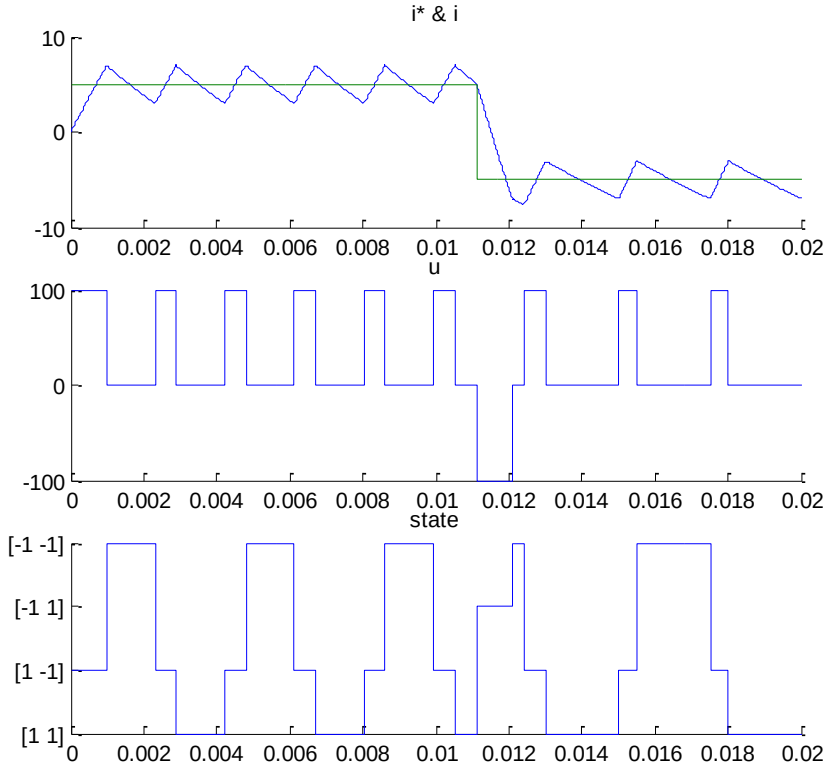


Figure 3.16 Step response with a 4 quadrant Direct Current Controller according to Figure 3.15.

3.7 Sampled current control of a three-phase converter connected to a generic three phase load

A sampled current controller for the three-phase load can be derived in the same way as with the one phase load. The main difference is that the scalar quantities used in the description of the load (3.6) and the current controller (equations (3.11) and (3.13)) is replaced with the corresponding vector quantities. The controllers derived for the 3-phase case is thus denoted vector controllers. Vector control is preferably performed in the rotating reference frame (d,q) (see Figure 3.4) instead of the stationary reference frame (α,β). The reason is that in (α,β) all signals are AC signals in stationary operation, and a PIE controller as the one derived for the scalar case (equations (3.11) and (3.13)) will give a stationary control error (eliminating the current error given at sample instant k will still leave a current error at instant (k+1) since the references are continuously changing. This could theoretically be overcome with a derivating term in the controller, but derivation of the very noisy signals that are normally the case in Power Electronic Control is not feasible. In the rotating reference frame (d,q) all quantities are DC quantities in stationary operation and no stationary control error will appear.

Assume that the current is given at the sampling time instant as $\vec{i}(k)$ and the current reference at the same instant $\vec{i}^{\,*}(k)$. Furthermore, it is assumed that the current error $\left(\vec{i}^{\,*}(k)-\vec{i}(k)\right)$ will be eliminated during the next sampling interval, i.e. during the interval [$k,k+1$]. The voltage needed during this interval can thus be derived from the voltage equation (3.6) by integration over one sampling interval T_s according to equation (3.15).

$$\frac{\int_{k\cdot T_s}^{(k+1)T_s}\vec{u}\cdot dt}{T_s}=\frac{R\cdot\int_{k\cdot T_s}^{(k+1)T_s}\vec{i}\cdot dt+L\cdot\int_{k\cdot T_s}^{(k+1)T_s}\frac{d\vec{i}}{dt}\cdot dt+j\cdot\omega\cdot L\int_{k\cdot T_s}^{(k+1)T_s}\vec{i}\cdot dt\int_{k\cdot T_s}^{(k+1)T_s}\vec{e}\cdot dt}{T_s}= \tag{3.15}$$

$$=\bar{\vec{u}}(k,k+1)=\left(R+j\cdot\omega\cdot L\right)\cdot\bar{\vec{i}}(k,k+1)+L\cdot\frac{\vec{i}(k+1)-\vec{i}(k)}{T_s}+\bar{\vec{e}}(k,k+1)$$

To convert the "average voltage"-equation (3.15) into a current control algorithm some assumptions have to be made:

$$\begin{aligned}
\vec{\bar{u}}(k,k+1) &= \vec{u}\,^*(k) && (a)\\
\vec{i}\,(k+1) &= \vec{i}\,^*(k) && (b)\\
\vec{\bar{i}}\,(k,k+1) &= \frac{\vec{i}\,^*(k)+\vec{i}\,(k)}{2} && (c)\\
\vec{\bar{e}}(k,k+1) &= \vec{e}(k) && (d)\\
\vec{i}\,(k) &= \sum_{n=0}^{n=k-1}\left(\vec{i}\,^*(n)-\vec{i}\,(n)\right) && (e)
\end{aligned} \tag{3.16}$$

Assumption (*a*) is straight forward, the reference voltage for the coming period should be the voltage needed to achieve the desired current change. Assumption (*b*) is an assumption of "dead beat" current control, i.e. the entire current error is eliminated in one sampling interval. Assumption (*c*) is an approximation, and is not always true, since the trajectory of the current will depend on the relation between the load emf $\vec{e}$, the dc link voltage U_{dc} and the sequence of switching states that is used within the interval [k,(k+1)]. However, the assumption (*c*) is true in average, and furthermore it is related to the resistive voltage drop that normally is relatively small, thus as small error in this assumption is not expected to distort the current control performance significantly. The assumption (*d*) is usually true since the sampling frequency usually is several orders of magnitude faster than the dynamics of the load emf. In the case with control of the current to the utility grid, $\vec{e}$ varies slowly or not at all when expressed in the (*d*,*q*) reference frame. The assumption (*e*) is a consequence of the "dead beat" assumption (*b*). The current at instant *k* is the sum of all changes of currents until the last change depicted in the previous sampling instant. With all these assumptions a current controller can be derived from the "average voltage" equation (3.9) as:

$$\begin{aligned}
&\vec{u}\,^*(k) = \left(R + j\cdot\omega\cdot L\right)\cdot\frac{\vec{i}\,^*(k)+\vec{i}\,(k)}{2} + L\cdot\frac{\vec{i}\,^*(k)-\vec{i}\,(k)}{T_s} + \vec{e}(k) =\\
&= R\cdot\frac{\vec{i}\,^*(k)-\vec{i}\,(k)}{2} + R\cdot\vec{i}\,(k) + L\cdot\frac{\vec{i}\,^*(k)-\vec{i}\,(k)}{T_s} + j\cdot\omega\cdot L\cdot\frac{\vec{i}\,^*(k)+\vec{i}\,(k)}{2} + \vec{e}(k) \approx\\
&\approx \left(\frac{L}{T_s}+\frac{R}{2}\right)\left(\vec{i}\,^*(k)-\vec{i}\,(k)\right) + R\cdot\sum_{n=0}^{n=k-1}\left(\vec{i}\,^*(n)-\vec{i}\,(n)\right) + j\cdot\omega\cdot L\cdot\vec{i}\,(k) + \vec{e}(k) =\\
&= \left(\frac{L}{T_s}+\frac{R}{2}\right)\cdot\left(\underbrace{\left(\vec{i}\,^*(k)-\vec{i}\,(k)\right)}_{\text{Proportional}} + \underbrace{\frac{T_s}{\left(\frac{L}{R}+\frac{T_s}{2}\right)}\cdot\sum_{n=0}^{n=k-1}\left(\vec{i}\,^*(n)-\vec{i}\,(n)\right)}_{\text{Integral}}\right) + \underbrace{j\cdot\omega\cdot L\cdot\vec{i}\,(k)+\vec{e}(k)}_{\textit{Feed forward}}
\end{aligned} \tag{3.17}$$

Note that the additional approximation introduced in equation (3.17) is the inductive voltage drop, which is relatively small compared to the proportional

part of the control algorithm in transients and to the emf in stationary operation. Note also that a "PIE" current controller has been derived, with both gain and integral time constant directly related to the circuit parameters, and with a feed forward term directly related to the inductive voltage drop and the emf of the load. These properties are useful since it is now possible to determine the controller sensitivity to changes of a particular parameter, e.g. saturation dependent inductances and temperature dependent resistances.

The controller of equation (3.17) becomes more clearly when split up in its two components according to equation (3.18).

$$\begin{aligned} u_d^*(k) &= \left(\frac{L}{T_s}+\frac{R}{2}\right)\cdot\left(\left(i_d^*(k)-i_d(k)\right)+\frac{T_s}{\left(\frac{L}{R}+\frac{T_s}{2}\right)}\cdot\sum_{n=0}^{n=k-1}\left(i_d^*(n)-i_d(n)\right)\right)-\omega\cdot L\cdot i_q(k) \\ u_q^*(k) &= \left(\frac{L}{T_s}+\frac{R}{2}\right)\cdot\left(\left(i_q^*(k)-i_q(k)\right)+\frac{T_s}{\left(\frac{L}{R}+\frac{T_s}{2}\right)}\cdot\sum_{n=0}^{n=k-1}\left(i_q^*(n)-i_q(n)\right)\right)+\omega\cdot L\cdot i_d(k)+e_q(k) \end{aligned} \tag{3.18}$$

These two components are scalar and can readily be implemented as two interrelated PIE-controllers in a vector control system for the current to a generic 3-phase load.

Sampled current control of a generic 3-phase load with a slow computer

In a previous section it was discussed how to handle the problem with a delay when controlling a 1-phase current. The identical problem occurs when controlling a three phase current with the vector controller discussed in the previous section. The solution is also the same. A simplified "dummy" system model is used, with the same dynamics as the system that we are controlling. The "dummy" model current is used to calculate the change a certain voltage reference will accomplish in the controlled current, and to use this estimated change when predicting the controlled current "one sample ahead".

$$\vec{u} = R\cdot\vec{i} + L\cdot\frac{d\vec{i}}{dt} + j\cdot\omega\cdot L\cdot\vec{i} \quad \text{(NOTE!No emf modelled)}$$

$$\vec{u}^*(k) = \vec{u}(k,k+1) = R\cdot\vec{i}_{sp}(k) + j\cdot\omega\cdot L\cdot\vec{i}_{sp}(k) + \frac{L}{T_s}\cdot\left(\vec{i}_{sp}(k+1)-\vec{i}_{sp}(k)\right)$$

or, in the discrete domain;

$$\vec{u} = R \cdot \vec{i}_{sp} + j \cdot \omega \cdot L \cdot \vec{i}_{sp} + \frac{L}{T_s} \cdot (z-1) \cdot \vec{i}_{sp}$$

$$\frac{\vec{i}_{sp}}{\vec{u}} = \frac{1}{R + j \cdot \omega \cdot L + \frac{L}{T_s} \cdot (z-1)} = \frac{T_s}{L} \cdot \frac{1}{z - (1 - R \cdot T_s - j \cdot \omega \cdot T_s)}$$

or, again in the time domain;

$$\vec{i}_{sp}(k+1) = \vec{i}_{sp}(k) \cdot (1 - R \cdot T_s / L - j \cdot \omega \cdot T_s) + \vec{u}^*(k) \cdot \frac{T_s}{L}$$

The current estimate to be used by the controller of the "slow computer" is thus finally calculated as:

$$\hat{\vec{i}}(k+1) = \vec{i}(k) + \left(\vec{i}_{sp}(k+1) - \vec{i}_{sp}(k)\right) \tag{3.19}$$

The accuracy of the current prediction is of course dependent on the accuracy of the parameters and the estimate of the induced voltage used by the controller. A simple sensitivity analysis is presented in the next section. The current controller algorithm when using a slow computer is thus given by equation (3.20).

$$\begin{aligned} u_d^*(k) &= \left(\frac{L}{T_s} + \frac{R}{2}\right) \cdot \left(\left(i_d^*(k) - \hat{i}_d(k)\right) + \frac{T_s}{\left(\frac{L}{R} + \frac{T_s}{2}\right)} \cdot \sum_{n=0}^{n=k-1} \left(i_d^*(n) - \hat{i}_d(n)\right) \right) - \omega \cdot L \cdot \hat{i}_q(k) \\ u_q^*(k) &= \left(\frac{L}{T_s} + \frac{R}{2}\right) \cdot \left(\left(i_q^*(k) - \hat{i}_q(k)\right) + \frac{T_s}{\left(\frac{L}{R} + \frac{T_s}{2}\right)} \cdot \sum_{n=0}^{n=k-1} \left(i_q^*(n) - \hat{i}_q(n)\right) \right) + \omega \cdot L \cdot \hat{i}_d(k) + e_q(k) \end{aligned} \tag{3.20}$$

The quality of the current prediction depends on application. With a 3-phase load like the one used in this section, it is possible to measure e.g. the emf $\vec{e}$, and the inductor parameters are usually rather easy to measure. When controlling e.g. a AC machine, the emf is not available for measurement and the estimate is of course more difficult to make. It remains to be determined in the specific application what methods can be used, and with what accuracy, to estimate circuit parameters and the emf.

Sampled Vector Current control applied to a 3-phase load with a 2 level converter

The vector current controller derived in the previous chapter will now be applied to a 2 level 3-phase VSI connected to a 3-phase generic load. Figure 3.17 shows a Simulink block diagram of the system under consideration.

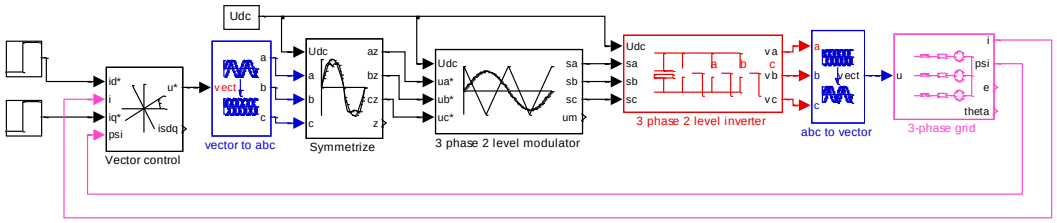


Figure 3.17 A 3-phase converter with vector control connected to a generic 3-phase load

The blocks visible in Figure 3.17 contain the following functions:

1. *Vector Control:* A vector controller according to equation (3.18) or equation (3.20), selectable.

2. *Vector to abc:* 2-phase to 3-phase transformation, power invariant.
3. *phase 2 level modulator and Inverter:* Carrier wave modulator and 3 ideal switches according to chapter 3.
4. *Abc to vector:* 3-phase to 2-phase conversion, power invariant.
5. *3-phase grid*: Ideal 3 phase load.

The contents of the current control block are shown in Figure 3.18.

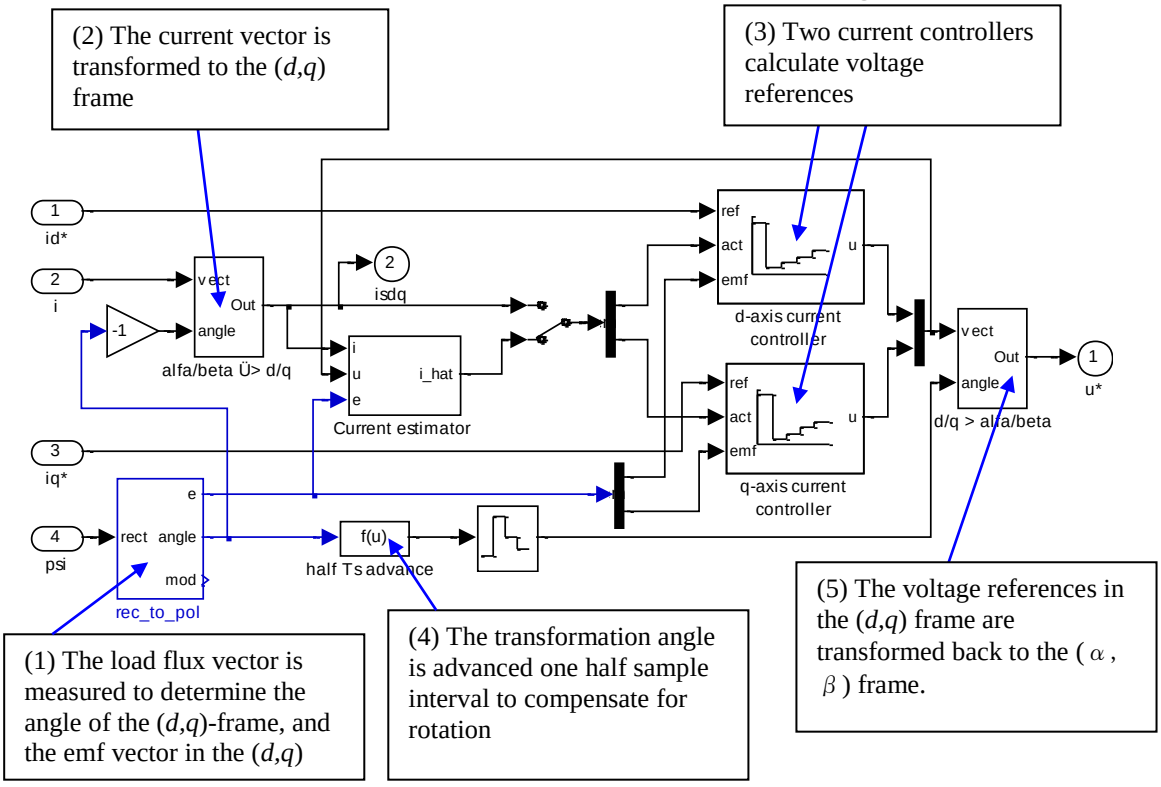


Figure 3.18 Content of the current control block.

Assume a load circuit and a 3-phase converter with parameters according to:

$$L = 10\ mH$$
$$R = 1\ \Omega$$
$$Ts = 0.5\ ms$$
$$U_{dc} = 750\ V$$

To evaluate the performance of the developed current controllers, some simulations are made. In all cases both the *d*-axis and the *q*-axis currents are making step changes between +/- 10 A. The following situations are simulated:

1. The current controller has the right set of parameters, i.e. a perfect measurement of the load parameters and an as perfect measurement of the load emf is used. The controller is of the "fast computer" type.

2. In this case the current controller is of the slow computer type, and the controller parameters are deliberately different from those of the load, to show the performance of a mistuned current controller. One case is simulated; with $R_{controller}$=1.5R and $L_{controller}$=0.5 L. This is an extreme mistuning of the controller. Normally the parameters are known with less than 10 % error, but the mistuning is chosen to show the abilities of the current estimator.

Figure 3.19 shows some simulation results from case I. Note that the phase current increases in amplitude at the q-axis current step, and that the phase is advanced as well. This is due to the increase amplitude, the modulus of the vector sum of i_d and i_q, and the fact that the current vector advances 45 degrees as the *q*-axis current increases.

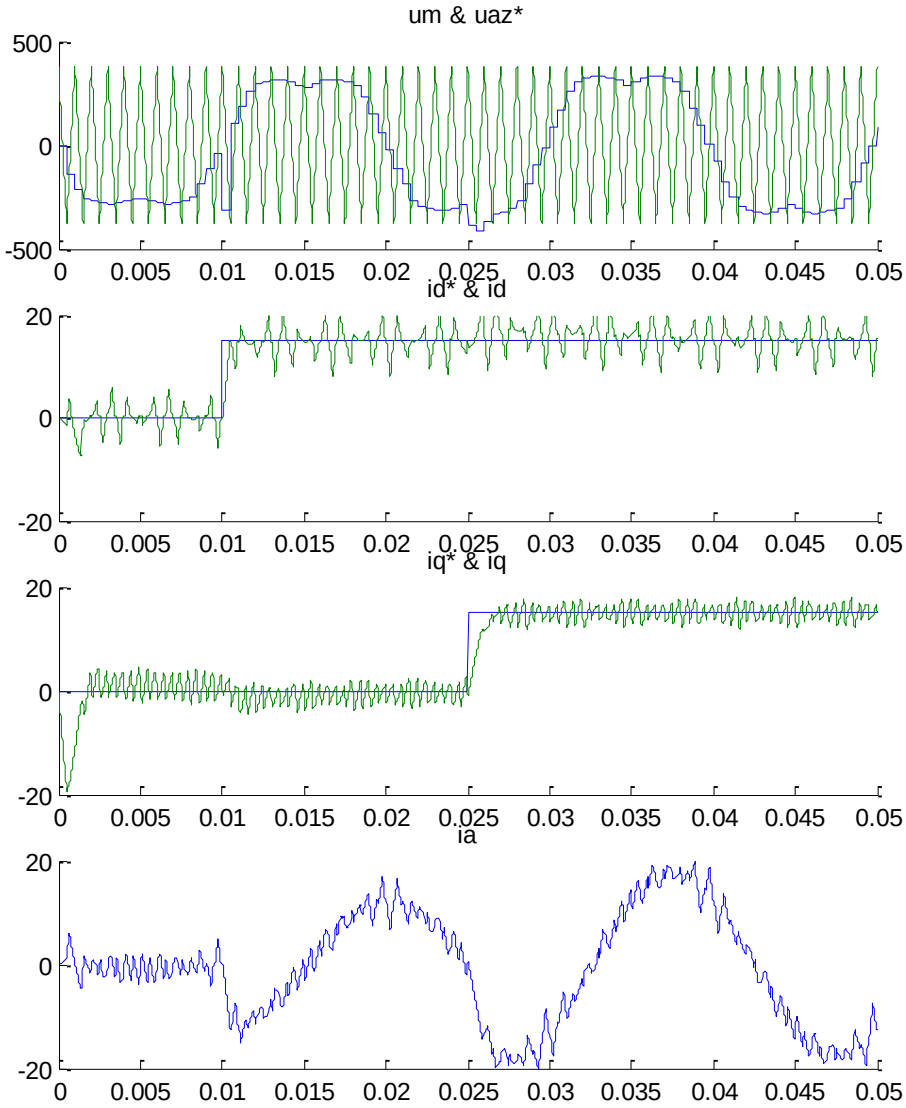


Figure 3.19 *Vector control of 3-phase converter connected to a generic 3-phase load with a fast computer.*

Figure 3.20 shows the result of case II above. Note that no significant difference in performance can be noted, in spite of the drastic mistuning of the controller parameters.

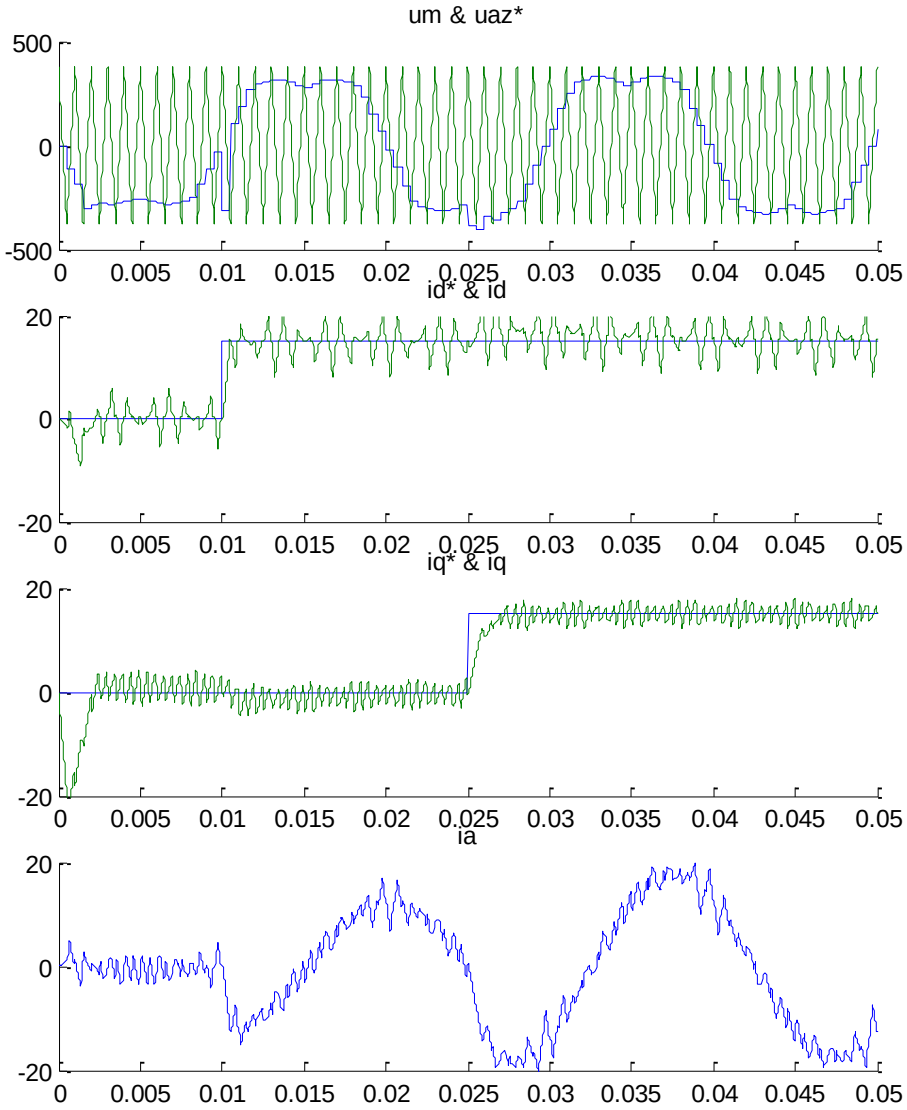


Figure 3.20 *Vector control of 3-phase converter connected to a generic 3-phase load with a fast computer using current estimation.*

The vector control system presented in this section is very similar to vector control systems used in ac motor control systems. The current control as such is not difficult to engineer. The biggest challenge is instead in the estimation of the load flux. In the generic load used in this section, the load flux is measurable, but in electrical machines it is generally more difficult to obtain a good estimate of the flux. This challenge will be addressed in later chapters on ac machine control.

3.8 Relation to traditional control theory

A model-based controller is derived for a grid connected voltage source converter (**Figure 2.18**). The controller is model-based in the sense that the controller parameters are dependent on the grid parameters, i.e. L (L_{line}) and R (R_{line}). The current controller for three-phase converters discussed here is based on the synchronously rotating (dq) reference frame. Synchronously rotating

refers to that the reference frame is rotating with the same speed as the grid voltage vector $\vec{e}^{\alpha\beta}$. The rotating reference system has two axes denoted *d* and *q*. The grid voltage vector $\vec{e}^{\alpha\beta}$ coincides with the *q*-axis in this case, which means that $e_d = 0$ (**Figure 2.25**). The transformation expressions are written

$$\vec{s}^{dq} = \vec{s}^{\alpha\beta} \cdot e^{-j\omega_1 t} \quad \text{and} \quad \vec{s}^{\alpha\beta} = \vec{s}^{dq} \cdot e^{j\omega_1 t} \tag{3.21}$$

The controller is based on a model of the grid together with the converter output filter

$$\vec{u}^{\alpha\beta} - L\frac{d}{dt}\vec{i}^{\alpha\beta} - R\vec{i}^{\alpha\beta} - \vec{e}^{\alpha\beta} = 0 \tag{3.22}$$

This is transformed to *dq*-coordinates by use of the transformation above, i.e.

$$\vec{u}^{dq} \cdot e^{j\omega_1 t} - L\frac{d}{dt}\left(\vec{i}^{dq} \cdot e^{j\omega_1 t}\right) - R\vec{i}^{dq} \cdot e^{j\omega_1 t} - \vec{e}^{dq} \cdot e^{j\omega_1 t} = 0 \tag{3.23}$$

which gives

$$\vec{u}^{dq} \cdot e^{j\omega_1 t} - L\frac{d}{dt}\vec{i}^{dq} \cdot e^{j\omega_1 t} - j\omega_1 L\vec{i}^{dq} \cdot e^{j\omega_1 t} - R\vec{i}^{dq} \cdot e^{j\omega_1 t} - \vec{e}^{dq} \cdot e^{j\omega_1 t} = 0 \tag{3.24}$$

Simplifying the last expression gives

$$\vec{u}^{dq} - L\frac{d}{dt}\vec{i}^{dq} - j\omega_1 L\vec{i}^{dq} - R\vec{i}^{dq} - \vec{e}^{dq} = 0 \tag{3.25}$$

By dividing into its components, i.e. real and imaginary parts, the resulting expressions are

$$\begin{cases} u_d = L\frac{d}{dt}i_d + Ri_d - \omega_1 L i_q + e_d \\ u_q = L\frac{d}{dt}i_q + Ri_q + \omega_1 L i_d + e_q \end{cases} \tag{3.26}$$

Controller design based on understanding

A discrete-time, dead-beat PI-controller is designed for each component based on the last two expressions. The output from the controller is two voltage references, one in *d*- and one in *q*-direction. The first step is to use backward Euler approximation [25] for the current derivatives. This gives

$$\begin{cases} \dfrac{di_d}{dt} \approx \dfrac{i_{d,k} - i_{d,k-1}}{T_s} \\ \dfrac{di_q}{dt} \approx \dfrac{i_{q,k} - i_{q,k-1}}{T_s} \end{cases} \tag{3.27}$$

Furthermore, it is assumed that the currents can be approximated with their average values, in the particular sampling interval

$$\begin{cases} i_d \approx \dfrac{i_{d,k} + i_{d,k-1}}{2} \\ i_q \approx \dfrac{i_{q,k} + i_{q,k-1}}{2} \end{cases} \tag{3.28}$$

It is also assumed that the grid voltage does not change between two samples. Consequently,

$$\begin{cases} e_{d,k} \approx e_{d,k-1} \\ e_{q,k} \approx e_{q,k-1} \end{cases} \tag{3.29}$$

This gives the following result

$$\begin{cases} u_{d,k} = L\dfrac{i_{d,k} - i_{d,k-1}}{T_s} + R\dfrac{i_{d,k} + i_{d,k-1}}{2} - \omega_1 L\dfrac{i_{q,k} + i_{q,k-1}}{2} + e_{d,k-1} \\ u_{q,k} = L\dfrac{i_{q,k} - i_{q,k-1}}{T_s} + R\dfrac{i_{q,k} + i_{q,k-1}}{2} + \omega_1 L\dfrac{i_{d,k} + i_{d,k-1}}{2} + e_{q,k-1} \end{cases} \tag{3.30}$$

Since a dead-beat controller is considered, it is assumed that

$$\begin{cases} i_{d,k} = i_{d,ref,k-1} \\ i_{q,k} = i_{q,ref,k-1} \end{cases} \tag{3.31}$$

giving

$$\begin{cases} u_{d,k} = L\dfrac{i_{d,ref,k} - i_{d,k}}{T_s} + R\dfrac{i_{d,ref,k} + i_{d,k}}{2} - \omega_1 L\dfrac{i_{q,ref,k} + i_{q,k}}{2} + e_{d,k} \\ u_{q,k} = L\dfrac{i_{q,ref,k} - i_{q,k}}{T_s} + R\dfrac{i_{q,ref,k} + i_{q,k}}{2} + \omega_1 L\dfrac{i_{d,ref,k} + i_{d,k}}{2} + e_{q,k} \end{cases} \tag{3.32}$$

This gives a P-controller according to

$$\begin{cases} u_{d,ref,k} = \left(\frac{L}{T_s} + \frac{R}{2}\right) \cdot \left(i_{d,ref,k} - i_{d,k}\right) + R \cdot i_{d,k} - \omega_1 L \frac{i_{q,ref,k} + i_{q,k}}{2} + e_{d,k} \\ u_{q,ref,k} = \left(\frac{L}{T_s} + \frac{R}{2}\right) \cdot \left(i_{q,ref,k} - i_{q,k}\right) + R \cdot i_{q,k} + \omega_1 L \frac{i_{d,ref,k} + i_{d,k}}{2} + e_{q,k} \end{cases} \tag{3.33}$$

The resistive voltage drop term can be interpreted as an integral part assuming that the current equals the sum of all the previous current errors, i.e.

$$\begin{cases} i_{d,k} = \sum_{n=0}^{k-1} \left(i_{d,ref,n} - i_{d,n}\right) \\ i_{q,n} = \sum_{n=0}^{k-1} \left(i_{q,ref,n} - i_{q,n}\right) \end{cases} \tag{3.34}$$

This gives a PI-controller

$$\begin{cases} u_{d,ref,k} = K \cdot \left(\left(i_{d,ref,k} - i_{d,k}\right) + \frac{1}{T_i} \cdot \sum_{n=0}^{k-1} \left(i_{d,ref,n} - i_{d,n}\right) \right) - K_c \cdot \frac{i_{q,ref,k} + i_{q,k}}{2} + e_{d,k} \\ u_{q,ref,k} = K \cdot \left(\left(i_{q,ref,k} - i_{q,k}\right) + \frac{1}{T_i} \cdot \sum_{n=0}^{k-1} \left(i_{q,ref,n} - i_{q,n}\right) \right) + K_c \cdot \frac{i_{d,ref,k} + i_{d,k}}{2} + e_{q,k} \end{cases} \tag{3.35}$$

where

$$\begin{cases} K = \left(\frac{L}{T_s} + \frac{R}{2}\right) \\ T_i = R \Big/ \left(\frac{L}{T_s} + \frac{R}{2}\right) = 1 \Big/ \left(\frac{L}{RT_s} + \frac{1}{2}\right) \\ K_c = \frac{\omega_1 L}{2} \end{cases} \tag{3.36}$$

RST controller design

This design of the same controller as determined previously follows the method described in Chapter 10 of [25]. The plant is the grid including line filter expressed in the rotating *dq*-frame, which is described by the first order differential equation

$$\frac{d}{dt}\vec{i}^{\,dq} = -\left(\frac{R}{L} + j\omega_1\right)\vec{i}^{\,dq} + \frac{1}{L}\vec{u}^{\,dq} - \frac{1}{L}\vec{e}^{\,dq} \tag{3.37}$$

Each component is described by

$$\begin{cases} \dfrac{di_d}{dt} = -\dfrac{R}{L} i_d + \omega_1 i_q + \dfrac{1}{L} u_d - \dfrac{1}{L} e_d \\ \dfrac{di_q}{dt} = -\dfrac{R}{L} i_q - \omega_1 i_d + \dfrac{1}{L} u_q - \dfrac{1}{L} e_q \end{cases} \tag{3.38}$$

The voltage $\vec{u}^{dq} = u_d + ju_q$ is the output of the controller and the current $\vec{i}^{dq} = i_d + ji_q$ is the controlled quantity. Therefore the term containing the grid voltage $\vec{e}^{dq} = e_d + je_q$ and the cross-coupling terms $j\omega_1 \vec{i}^{dq} = -i_q + ji_d$ can be regarded as known disturbances that are compensated by means of feed forward. This implies that the plant can be expressed as

$$\frac{d}{dt}\vec{i}^{dq} = -\frac{R}{L}\vec{i}^{dq} + \frac{1}{L}\vec{u}^{dq} + \vec{w}^{dq} \tag{3.39}$$

where the disturbance is

$$\vec{w}^{dq} = -j\omega_1 \vec{i}^{dq} - \frac{1}{L}\vec{e}^{dq} \tag{3.40}$$

The transfer function from voltage $\vec{u}^{dq}$ to current $\vec{i}^{dq}$ of the plant is thus written

$$G_{plant}(s) = \frac{B(s)}{A(s)} = \frac{1/L}{s + R/L} = (1/R)\frac{R/L}{s + R/L} \tag{3.41}$$

The next step is to discretize the plant equation. There are several methods to accomplish this for example zero-order-hold (zoh) sampling or forward Euler approximation. In case of zoh sampling the transfer function is

$$G_{plant}(q) = \frac{B(q)}{A(q)} = \frac{(1/R)\left(1 - e^{-RT_s/L}\right)}{q - e^{-RT_s/L}} \approx \frac{(1/R)(RT_s/L)}{q - (1 - RT_s/L)} = \frac{T_s/L}{q - (1 - RT_s/L)} \tag{3.42}$$

where the approximated transfer function above is the same as the one obtained with forward Euler approximation i.e.

$$\vec{i}_{k+1}^{dq} = \left(1 - \frac{RT_s}{L}\right)\vec{i}_k^{dq} + \frac{T_s}{L}\vec{u}_k^{dq} \tag{3.43}$$

Note that the converter operates in switch-mode, which means that the current can change substantially between sampling instants in the case of a change of reference current. Therefore, the resistive voltage drop is sometimes written as the average current for that particular sampling interval i.e.

$$R\vec{i}_{[k,k+1]}^{dq} \approx R\frac{\vec{i}_{k+1}^{dq} + \vec{i}_k^{dq}}{2} = \frac{R}{2}\left(\vec{i}_{k+1}^{dq} + \vec{i}_k^{dq}\right) \tag{3.44}$$

Applying this to the forward Euler approximated transfer function gives

$$\vec{i}_{k+1}^{dq} = \frac{2L - RT_s}{2L + RT_s}\vec{i}_k^{dq} + \frac{2T_s}{2L + RT_s}\vec{u}_k^{dq} \tag{3.45}$$

In this case the transfer function is written

$$G_{plant}(q) = \frac{B(q)}{A(q)} = \frac{\frac{2T_s}{2L + RT_s}}{q - \frac{2L - RT_s}{2L + RT_s}} \tag{3.46}$$

An RST controller has the form

$$\vec{u}_k^{dq,ref} = \frac{T(q)}{R(q)}\vec{i}_k^{dq,ref} - \frac{S(q)}{R(q)}\vec{i}_k^{dq} \tag{3.47}$$

If an RST controller is employed, the resulting closed loop system is

$$\vec{i}_k^{dq} = \frac{B(q)}{A(q)} \cdot \left[\frac{T(q)}{R(q)}\vec{i}_k^{dq,ref} - \frac{S(q)}{R(q)}\vec{i}_k^{dq} \right] \tag{3.48}$$

Rearranging the terms yields

$$\left[A(q)T(q) + B(q)S(q)\right]\vec{i}_k^{dq} = B(q)T(q)\vec{i}_k^{dq,ref} \tag{3.49}$$

In this case a PI controller (with feed-forward) is utilized. A PI controller is written

$$\vec{u}_k^{dq,ref} = K_i\left(\vec{i}_k^{dq,ref} - \vec{i}_k^{dq}\right) + \Sigma\Delta\vec{i}_k^{dq} \tag{3.50}$$

where $\Sigma\Delta\vec{i}_k^{dq}$ is the integral part which is updated according to

$$\Sigma\Delta\vec{i}_{k+1}^{dq} = \Sigma\Delta\vec{i}_k^{dq} + \frac{K_i}{T_i}\left(\vec{i}_k^{dq,ref} - \vec{i}_k^{dq}\right) \tag{3.51}$$

Note that integral part anti wind-up should be implemented based on back-calculation of the maximum permissible voltage vector since over-modulation may occur. The transfer function of the PI controller is thus

$$\vec{u}_k^{dq,ref} = K_i\left(1 + \frac{1/T_i}{q-1}\right)\left(\vec{i}_k^{dq,ref} - \vec{i}_k^{dq}\right) = \frac{K_i q + K_i\left(1/T_i - 1\right)}{q-1}\left(\vec{i}_k^{dq,ref} - \vec{i}_k^{dq}\right) \tag{3.52}$$

This means that

$$\begin{cases} R(q) = q - 1 \\ S(q) = T(q) = K_i q + K_i \left(1/T_i - 1\right) \end{cases} \tag{3.53}$$

According to [25] *R(q)* should be monic which it is. For causality the order of *R(q)* should be higher or equal to the order of *S(q)* and *T(q)* which is also the case here. Since *A(q)*, *R(q)* and S(q) is of the first order and B(q) is of zero order it follows that the closed loop transfer function has at most two poles (if no poles are cancelled). Since *S(q) = T(q)*, the poles introduced by *S(q)* can be cancelled if *S(q)* is a factor of *A(q)*. Assume that

$$\begin{cases} A(q) = q + a_0 \\ B(q) = b_0 \end{cases} \tag{3.54}$$

This means that (3.49) rewritten as

$$\begin{aligned} &\left[(q+a_0)(q-1)+b_0\left(K_i q+K_i\left(1/T_i-1\right)\right)\right]\vec{i}_k^{\,dq} = \\ &= b_0\left(K_i q+K_i\left(1/T_i-1\right)\right)\vec{i}_k^{\,dq,ref} \end{aligned} \tag{3.55}$$

Rearranging gives

$$\begin{aligned} &\left[q^2+\left(a_0-1+b_0K_i\right)q+\left(b_0K_i\left(1/T_i-1\right)-a_0\right)\right]\vec{i}_k^{\,dq} = \\ &= \left[b_0K_i q+b_0K_i\left(1/T_i-1\right)\right]\vec{i}_k^{\,dq,ref} \end{aligned} \tag{3.56}$$

Dead-beat control implies that the actual current should reach the reference in a time equal to the sampling interval times the order of the plant. In this case the order of the plant is one and therefore the denominator of the closed loop system should be equal to one. Furthermore, for dead-beat control the denominator should have its poles at q=0. This means that the denominator of the closed loop transfer function should be $A_m(q) = q$. Therefore, the lowest order term of the left hand expression above should be equal to zero and that *S(q) = T(q)* should be a factor of *A(q)* i.e.

$$\begin{cases} b_0K_i\left(1/T_i-1\right)-a_0 = 0 \\ a_0 = 1/T_i - 1 \end{cases} \tag{3.57}$$

which gives

$$\begin{cases} K_i = 1/b_0 \\ T_i = 1/\left(a_0+1\right) \end{cases} \tag{3.58}$$

Consequently

$$\begin{cases} a_0 = (RT_s/L) - 1 \\ b_0 = T_s/L \end{cases} \Rightarrow \begin{cases} K_i = L/T_s \\ T_i = L/RT_s \end{cases} \tag{3.59}$$

and

$$\begin{cases} a_0 = \dfrac{RT_s - 2L -}{2L + RT_s} \\ b_0 = \dfrac{2T_s}{2L + RT_s} \end{cases} \Rightarrow \begin{cases} K_i = \dfrac{L}{T_s} + \dfrac{R}{2} \\ T_i = \dfrac{L}{RT_s} + \dfrac{1}{2} \end{cases} \tag{3.60}$$

where the latter controller parameters are identical to the ones derived in the previous section. Note that in most modern applications $L/T_s >> R/2$ and therefore the two controller parameter sets give equal behaviour.

Smith-predictor

Also in the presence of plant delay the controller parameters are derived the same way. In this case the voltage reference calculated by the current controller is delayed one sample interval due to the calculation time of the DSP, which implies that the actual transfer function of the plant is not $G_p(q)$ but $q^{-1} \cdot G_p(q)$. Consequently, after selection of appropriate controller parameters a Smith-predictor is designed to predict the current based on the applied output voltage and the previous predictions, i.e. the state of the Smith-predictor. According to [25] the current estimated by the Smith-predictor is given by

$$\vec{i}_{Smith,k}^{\,dq} = \left(-1 + q^{-1}\right) G_p(q) \vec{u}_k^{\,dq,ref} \tag{3.61}$$

Shifting both sides one sample forward gives

$$\vec{i}_{Smith,k+1}^{\,dq} = \left(-q + 1\right) G_p(q) \vec{u}_k^{\,dq,ref} \tag{3.62}$$

Note that in this case the plant should be the full system since

$$\vec{i}_k = G_p(q)\left(\vec{u}_k^{\,*} - \vec{e}_k\right) = \frac{T_s/L}{q - \left(1 - T_s R/L - j\omega T_s\right)} \cdot \left(\vec{u}_k^{\,*} - \vec{e}_k\right) \tag{3.63}$$

which means that the Smith-predictor is

$$\begin{aligned}\vec{i}_{Smith,k+1}^{dq} &= (-q+1)\frac{T_s/L}{q-(1-T_sR/L-j\omega T_s)}\cdot(\vec{u}_k^* - \vec{e}_k) = \\ &= (-q+1)\frac{T_s/L}{q-(1-T_sR/L-j\omega T_s)}\cdot\vec{u}_k^* + \\ &+ \frac{T_s/L}{q-(1-T_sR/L-j\omega T_s)}\cdot(-\vec{e}_{k+1}+\vec{e}_k)\approx \\ &\approx (-q+1)\frac{T_s/L}{q-(1-T_sR/L-j\omega T_s)}\cdot\vec{u}_k^*\end{aligned} \tag{3.64}$$

Consequently, the feed forward voltage is not taken into account (if its bandwidth is much less than the sampling frequency) but the current cross-coupling terms are taken into account.

For implementation in a DSP a new state is introduced in such a way that the Smith-predictor is written

$$\begin{cases}\vec{i}_{Smith,k+1} = \left(\frac{RT_s}{L}s_k + j\omega T_s\right)\vec{s}_k - \frac{T_s}{L}\vec{u}_k^* \\ \vec{s}_{k+1} = \left(1-\frac{RT_s}{L} - j\omega T_s\right)\vec{s}_k + \frac{T_s}{L}\vec{u}_k^*\end{cases} \tag{3.65}$$

If this is divided between components the following is obtained

$$\begin{cases}i_{Smith,d,k+1} = \frac{RT_s}{L}s_{d,k} - \omega T_s s_{q,k} - \frac{T_s}{L}u_{d,k}^* \\ i_{Smith,q,k+1} = \frac{RT_s}{L}s_{q,k} + \omega T_s s_{d,k} - \frac{T_s}{L}u_{q,k}^*\end{cases} \tag{3.66}$$

$$\begin{cases}s_{d,k+1} = \left(1-\frac{RT_s}{L}\right)s_{d,k} + \omega T_s s_{q,k} + \frac{T_s}{L}u_{d,k}^* \\ s_{q,k+1} = \left(1-\frac{RT_s}{L}\right)s_{q,k} - \omega T_s s_{d,k} + \frac{T_s}{L}u_{q,k}^*\end{cases} \tag{3.67}$$

The current references above are added to the current references based on desired output power. The resulting voltage references, in *dq*–coordinates, are transformed to the fixed $\alpha\beta$-frame. Then the $\alpha\beta$-references are transformed to three-phase reference voltages. If the modulation is based on sinusoidal references the three-phase voltage references are used as they are. Otherwise, they are manipulated, for example to become symmetrical or bus-clamped, and then passed on to the modulator.

The current references above are added to the current references based on desired output power. The resulting voltage references, in *dq*–coordinates, are transformed to the fixed $\alpha\beta$-frame. Then the $\alpha\beta$-references are transformed to three-phase reference voltages. If the modulation is based on sinusoidal references the three-phase voltage references are used as they are. Otherwise, they are manipulated, for example to become symmetrical or bus-clamped, and then passed on to the modulator.

State-space model and dynamic properties

State-feedback according to Figure 3.21 is used in the implementation of the controller. The controller parameters in Figure 3.21 are set to $K_i = k_i\left((L/T_s)+(R/2)\right)$, $K_i/T_i = k_i R$, $K_{ci} = K_{ci^*} = \omega_1 L/2$ and $K_r = 0$. Note $k_i = 1$ corresponds to dead-beat response. The parameters of the Smith-predictor are set to $K_{ss} = RT_s/L$, $K_{so} = \omega_1 T_s$ and $K_{su} = T_s/L$.

The output of the controller (i.e. the converter output voltage reference where * denotes reference value)

$$\vec{u}_k^* = K_i\left(\left(\vec{i}_k^* - \vec{i}_k\right) + \sum\Delta\vec{i}_k\right) + K_r\vec{i}_k + jK_{ci}\vec{i}_k + jK_{ci^*}\vec{i}_k^* + \vec{e}_k \qquad (3.68)$$

and the integral part

$$\sum\Delta\vec{i}_{k+1} = \sum\Delta\vec{i}_k + \frac{K_i}{T_i}\left(\vec{i}_k^* - \vec{i}_k\right) \qquad (3.69)$$

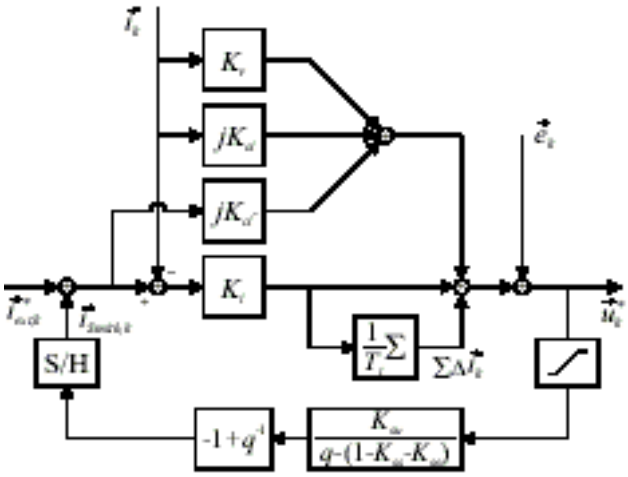


Figure 3.21 Block schematic of the converter current controller, including Smith-predictor.

where the current reference is expressed according to $\vec{i}_k^* = \vec{i}_{ext,k}^* + \vec{i}_{Smith,k}$ are rewritten

$$\begin{pmatrix} u^*_{d,k} \\ u^*_{q,k} \end{pmatrix} = L_i \begin{pmatrix} i_{d,k} \\ i_{q,k} \end{pmatrix} + L_u \begin{pmatrix} u_{d,k} \\ u_{q,k} \end{pmatrix} + L_{sp} \begin{pmatrix} s_{d,k} \\ s_{q,k} \end{pmatrix} + L_{\sum \Delta i} \begin{pmatrix} \sum \Delta i_{d,k} \\ \sum \Delta i_{q,k} \end{pmatrix} + L_e \begin{pmatrix} e_{d,k} \\ e_{q,k} \end{pmatrix} + L_{i^*,ext} \begin{pmatrix} i^*_{d,ext,k} \\ i^*_{q,ext,k} \end{pmatrix} \quad (3.70)$$

and

$$\begin{pmatrix} \sum \Delta i_{d,k+1} \\ \sum \Delta i_{d,k+1} \end{pmatrix} = I_i \begin{pmatrix} i_{d,k} \\ i_{q,k} \end{pmatrix} + I_u \begin{pmatrix} u_{d,k} \\ u_{q,k} \end{pmatrix} + I_{sp} \begin{pmatrix} s_{d,k} \\ s_{q,k} \end{pmatrix} + I_{\sum \Delta i} \begin{pmatrix} \sum \Delta i_{d,k} \\ \sum \Delta i_{q,k} \end{pmatrix} + I_{i^*,ext} \begin{pmatrix} i^*_{d,ext,k} \\ i^*_{q,ext,k} \end{pmatrix} \quad (3.71)$$

For the Smith-predictor, the state-feedback is written

$$\begin{pmatrix} s_{d,k+1} \\ s_{q,k+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1-K_{ss} & K_{so} \\ -K_{so} & 1-K_{ss} \end{pmatrix}}_{S_{sp}} \begin{pmatrix} s_{d,k} \\ s_{q,k} \end{pmatrix} + \underbrace{\begin{pmatrix} K_{su} & 0 \\ 0 & K_{su} \end{pmatrix}}_{S_u} \begin{pmatrix} u_{d,k} \\ u_{q,k} \end{pmatrix} \quad (3.72)$$

The inductive output filter has a state-space description according to

$$\frac{d}{dt} \begin{pmatrix} i_d \\ i_q \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{-R}{L} & \omega_1 \\ -\omega_1 & \frac{-R}{L} \end{pmatrix}}_{A_L} \begin{pmatrix} i_d \\ i_q \end{pmatrix} + \underbrace{\begin{pmatrix} \frac{1}{L} & 0 & \frac{-1}{L} & 0 \\ 0 & \frac{1}{L} & 0 & \frac{-1}{L} \end{pmatrix}}_{B_L = (B_{Lu} \; B_{Le})} \begin{pmatrix} u_d \\ u_q \\ e_d \\ e_q \end{pmatrix} \quad (3.73)$$

which is transferred to the discrete-time domain by zero-order-hold (zoh) sampling. The discrete-time system is represented by the matrixes Φ_L and $\Gamma_L = [\Gamma_{Lu} \; \Gamma_{Le}]$. The introduced matrixes L_i, L_u, L_{sp}, $L_{\sum \Delta i}$, L_e, $L_{i^*,ext}$, I_i, I_u, I_{sp}, $I_{\sum \Delta i}$ and $I_{i^*,ext}$ are

$$L_i = \begin{pmatrix} K_r - K_i & -K_{ci} \\ K_{ci} & K_r - K_i \end{pmatrix} \quad (3.74)$$

$$L_u = \begin{pmatrix} -K_{su}K_i & K_{su}K_{ci^*} \\ -K_{su}K_{ci^*} & -K_{su}K_i \end{pmatrix} \quad (3.75)$$

$$L_{sp} = \begin{pmatrix} K_{ss}K_i - K_{so}K_{ci^*} & -K_{ss}K_{ci^*} - K_{so}K_i \\ K_{ss}K_{ci^*} + K_{so}K_i & K_{ss}K_i - K_{so}K_{ci^*} \end{pmatrix} \quad (3.76)$$

$$L_{\sum \Delta i} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.77)$$

$$L_e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.78)$$

$$L_{i^*,ext}=\begin{pmatrix} K_i & -K_{ci^*} \\ K_{ci^*} & K_i \end{pmatrix} \tag{3.79}$$

$$I_i=\begin{pmatrix} -K_i/T_i & 0 \\ 0 & -K_i/T_i \end{pmatrix} \tag{3.80}$$

$$I_u=\begin{pmatrix} -K_{su}K_i/T_i & 0 \\ 0 & -K_{su}K_i/T_i \end{pmatrix} \tag{3.81}$$

$$I_{sp}=\begin{pmatrix} K_{ss}K_i/T_i & -K_{so}K_i/T_i \\ -K_{so}K_i/T_i & K_{ss}K_i/T_i \end{pmatrix} \tag{3.82}$$

$$I_{\sum\Delta i}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.83}$$

$$I_{i^*,ext}=\begin{pmatrix} K_i/T_i & 0 \\ 0 & K_i/T_i \end{pmatrix} \tag{3.84}$$

The state-space system for the entire converter including the line side filter and grid voltage is thus

$$\begin{pmatrix} i_{d,k+1} \\ i_{q,k+1} \\ u_{d,k+1} \\ u_{q,k+1} \\ s_{d,k+1} \\ s_{q,k+1} \\ \sum\Delta i_{d,k+1} \\ \sum\Delta i_{q,k+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \Phi_L & \Gamma_{Lu} & 0 & 0 \\ L_i & L_u & L_{sp} & L_{\sum\Delta i} \\ 0 & S_u & S_{sp} & 0 \\ I_i & I_u & I_{sp} & I_{\sum\Delta i} \end{pmatrix}}_{\Phi_c} \begin{pmatrix} i_{d,k} \\ i_{q,k} \\ u_{d,k} \\ u_{q,k} \\ s_{d,k} \\ s_{q,k} \\ \sum\Delta i_{d,k} \\ \sum\Delta i_{q,k} \end{pmatrix} + \underbrace{\begin{pmatrix} \Gamma_{Le} & 0 \\ L_e & L_{i^*,ext} \\ 0 & 0 \\ 0 & I_{i^*,ext} \end{pmatrix}}_{\Gamma_c} \begin{pmatrix} e_{d,k} \\ e_{q,k} \\ i^*_{d,ext,k} \\ i^*_{q,ext,k} \end{pmatrix} \tag{3.85}$$

If i_q is considered as the output, and $i^*_{d,ext}$ and $i^*_{q,ext}$ as inputs the output is described by

$$i_{q,k} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \end{pmatrix}}_{C_{c,iq}} \begin{pmatrix} i_{d,k} \\ i_{q,k} \\ u_{d,k} \\ u_{q,k} \\ s_{d,k} \\ s_{q,k} \\ \sum \Delta i_{d,k} \\ \sum \Delta i_{q,k} \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}}_{D_c} \begin{pmatrix} e_{d,k} \\ e_{q,k} \\ i^*_{d,ext,k} \\ i^*_{q,ext,k} \end{pmatrix} \tag{3.86}$$

The frequency response from active line current reference $i^*_{q,ext}$ to active line current i_q is shown in Figure 3.22 and the frequency response from reactive line current reference $i^*_{d,ext}$ to active line current i_q is shown in Figure 3.23. The parameters used to obtain the frequency responses in Figure 3.22 and Figure 3.23 are $L = 11.0$ mH, $R = 0.75\ \Omega$ and $T_s = 170\ \mu$s.

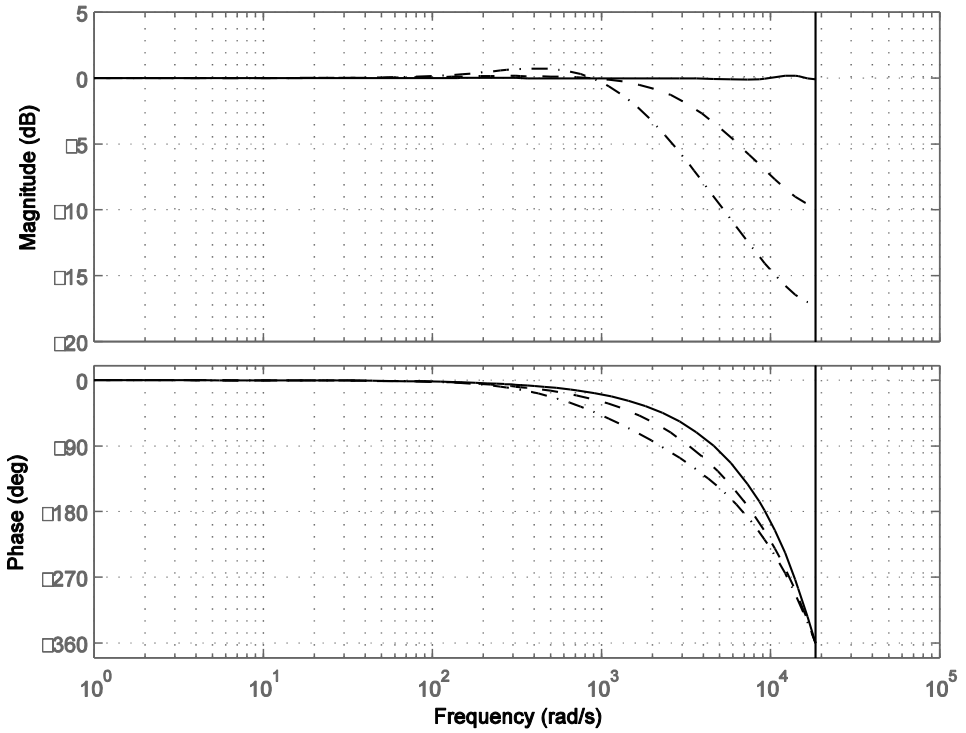


Figure 3.22 Frequency response from $i^*_{q,ext}$ to i_q for ki=1.0 (solid), ki=0.5 (dashed) and ki=0.25 (dash-dotted).

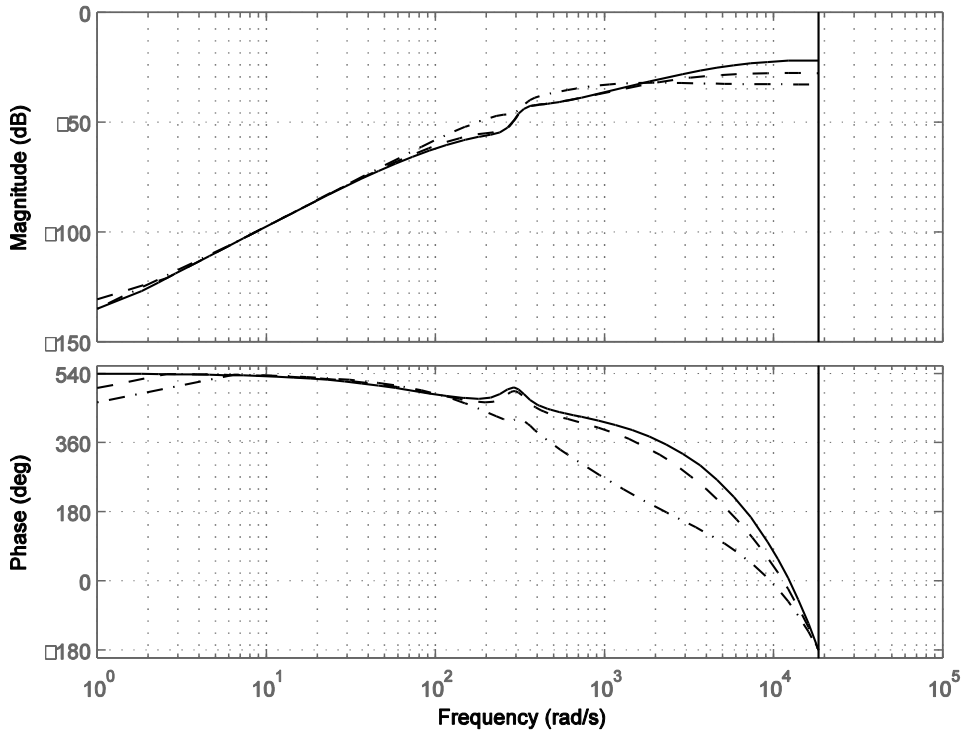


Figure 3.23 Frequency response from $i^*_{d,ext}$ *to* i_q *for ki=1.0 (solid), ki=0.5 (dashed) and ki=0.25 (dash-dotted).*

3.9 Direct Current Control (DCC) with a 3-phase converter and a generic load

The DCC methods used in previous sections together with a 1-phase load and 2- or 4-quadrant converter can be used even with the ac converter. The selection of voltage vectors as a function of the current errors must be related to the actual position of the grid flux that defines the (*d*,*q*)-axes. The derivative of the current vector in the (α,β) frame is given by equation (3.87).

$$\frac{d\vec{i}}{dt} = \frac{\vec{u} - R\cdot\vec{i} - \vec{e}}{L} \tag{3.87}$$

The choice of vector must be based on the goal to force the current vector to move in a direction that reduces the vector error of the current. Figure 3.24 shows the voltage vectors that the converter is able to create, together with the (*d*,*q*) frame. See section 3.3 and Figure 3.4 for reference to the voltage vectors and the definition of the (*d*,*q*) frame. Consider the position of the (*d*,*q*) frame as shown in Figure 3.1. All vectors affect the current according to equation (3.87). To increase the d-axis current, either voltage vector 1, 2 or 6 can be used. To decrease the *d*-axis current, either of voltage vectors 3, 4 or 5 can be used. To increase the *q*-axis current the voltage vectors 2 or 3 can be used, and to reduce the *q*-axis current the voltage vectors 5 or 6 can be used. In addition, the zero-vectors can be used, but the effect of using them is different in the d- and q-

direction. In the *d*-direction, there is no emf and a zero vector will have very little effect on the current derivative. In the *q*-direction the emf will give a significant negative derivative with a zero vector applied. Thus, the d-axis current control can be of a simpler kind that only uses active vectors in either direction to control the current, whereas the q-axis current controller benefit from using zero vectors and thus must be more advanced. The tolerance ban used in the *d*-axis is thus the same as for the DC current with the 2-quadrant converter, see Figure 3.25 left.

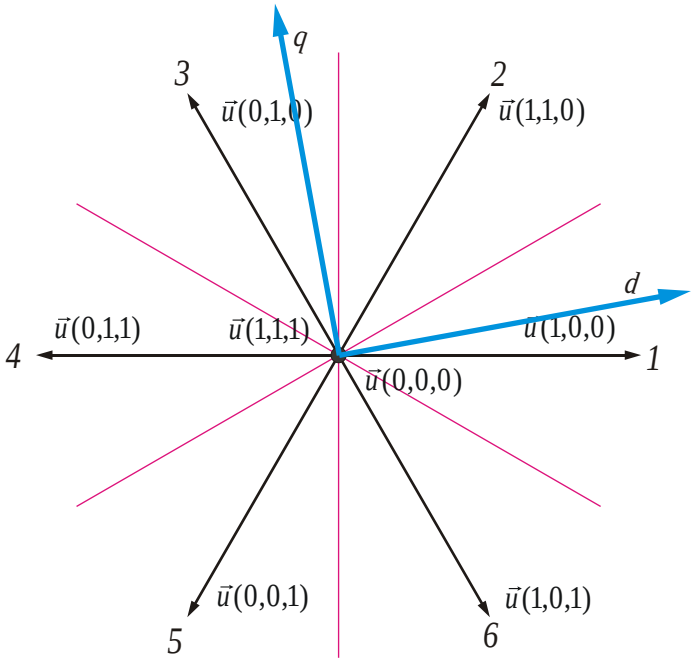


Figure 3.24 The 3-phase converter voltage vectors and the (d,q) reference frame.

The tolerance band used in the *q*-axis is the same as used for the DC current with the 4-quadrant converter, see Figure 3.25 right.

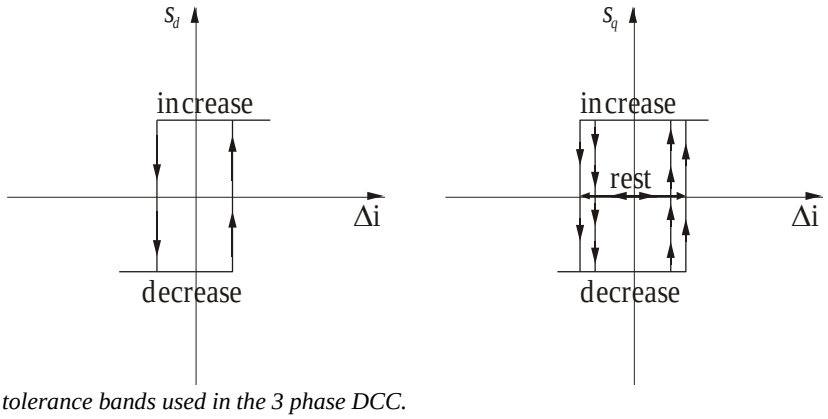


Figure 3.25 The tolerance bands used in the 3 phase DCC.

Note that the entity “rest” in Figure 3.25 implies the use of a zero vector. To proceed in defining the DCC some constraints must be introduced:

1. To control the *q*-axis current, only the active vectors closest to the *q*-axis must be involved. In Figure 3.24 this means vectors 2, 3, 5 and 6, and of

course the zero-vectors. The other two active vectors are thus excluded from the control strategy developed here.

2. To identify the possible vectors to be used it is necessary to define a set of sectors surrounding the voltage vectors with +/- π/6 radians. These sectors are indicated in Figure 3.24 and given a number from 1 to 6. Sector 1 covers the angles [-π/6 … π/6], sector 2 covers [π/6 … π/2] etc.

With these assumptions a rule can be created on which active vectors to select based on the current errors. Table 3.1gives the offset to the sector where the d-axis is located that should be added to the sector number to locate the correct active vector.

Table 3.1 Offsets to locate the correct voltage vector when the sector of the flux vector is known.

Soffset	Decrease iq	Increase iq
Decrease id	4	2
Increase id	5	1

The correct active vector can thus be calculated as

$$vector = sector + s_{offset} \tag{3.88}$$

When the *q*-axis current controller calls for a passive vector, the choice of active vectors is overruled. Figure 3.26 shows a Simulink program with a vector controller based on the above reasoning.

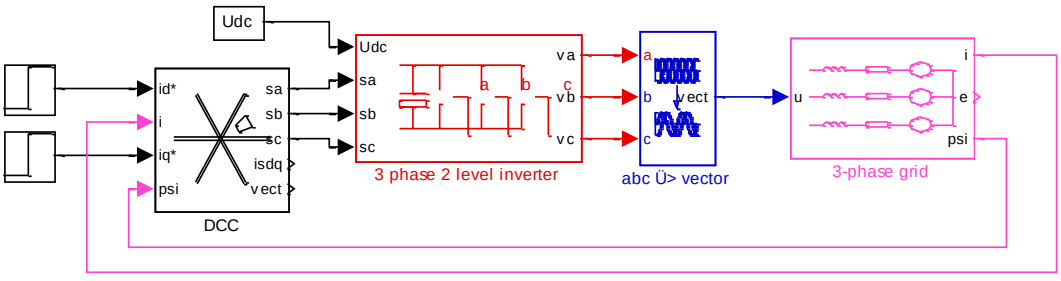


Figure 3.26: 3-phase DCC applied to a generic load.

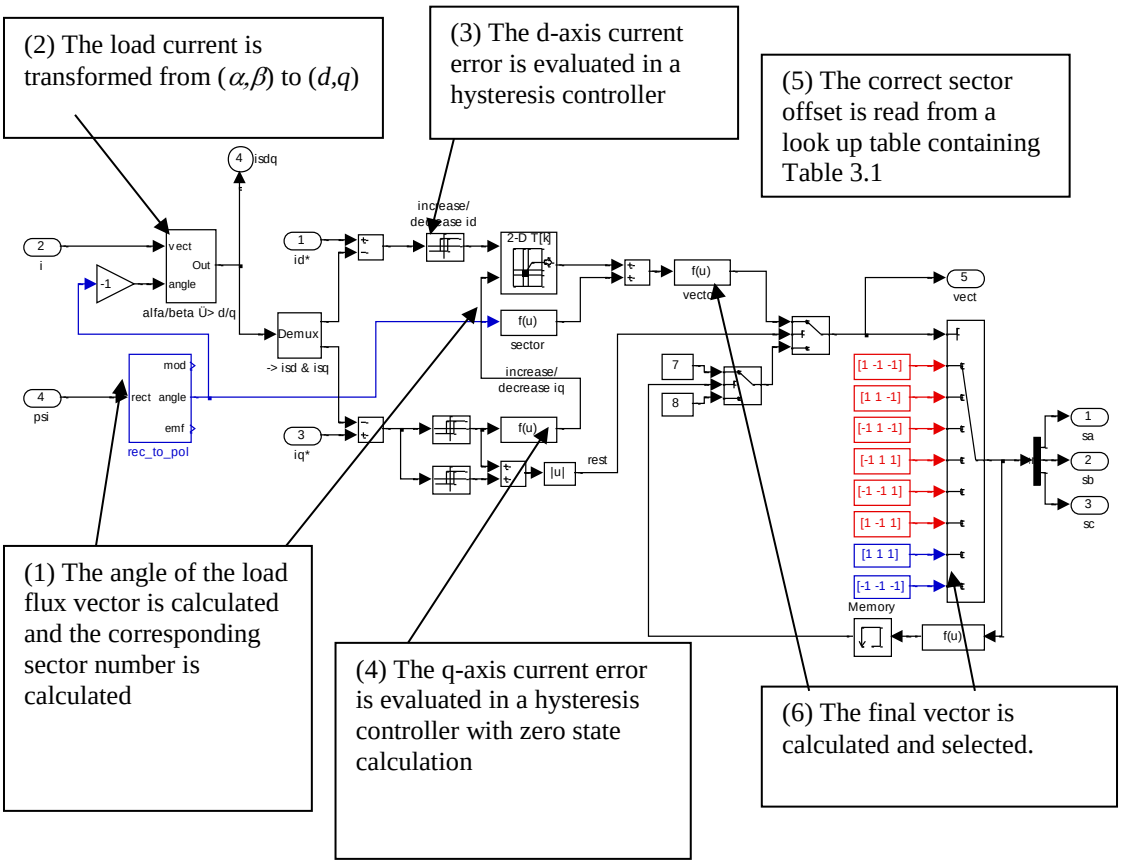


Figure 3.27 shows the content of the DCC block in Figure 3.26. With the same motivation as in the 4-quadrant DCC it is necessary to select the zero vector in a way that minimizes and distributes the losses of the semiconductor evenly. In the 3-phase case the zero vector to be used is determined by the switch state prior to the selection of a zero vector. For any given active vector, flipping one switch can set a passive vector. E.g. starting from state [1,-1,-1], the “closest”

zero vector is [-1,-1,-1] since only the switch in phase has to be flipped. Thus in

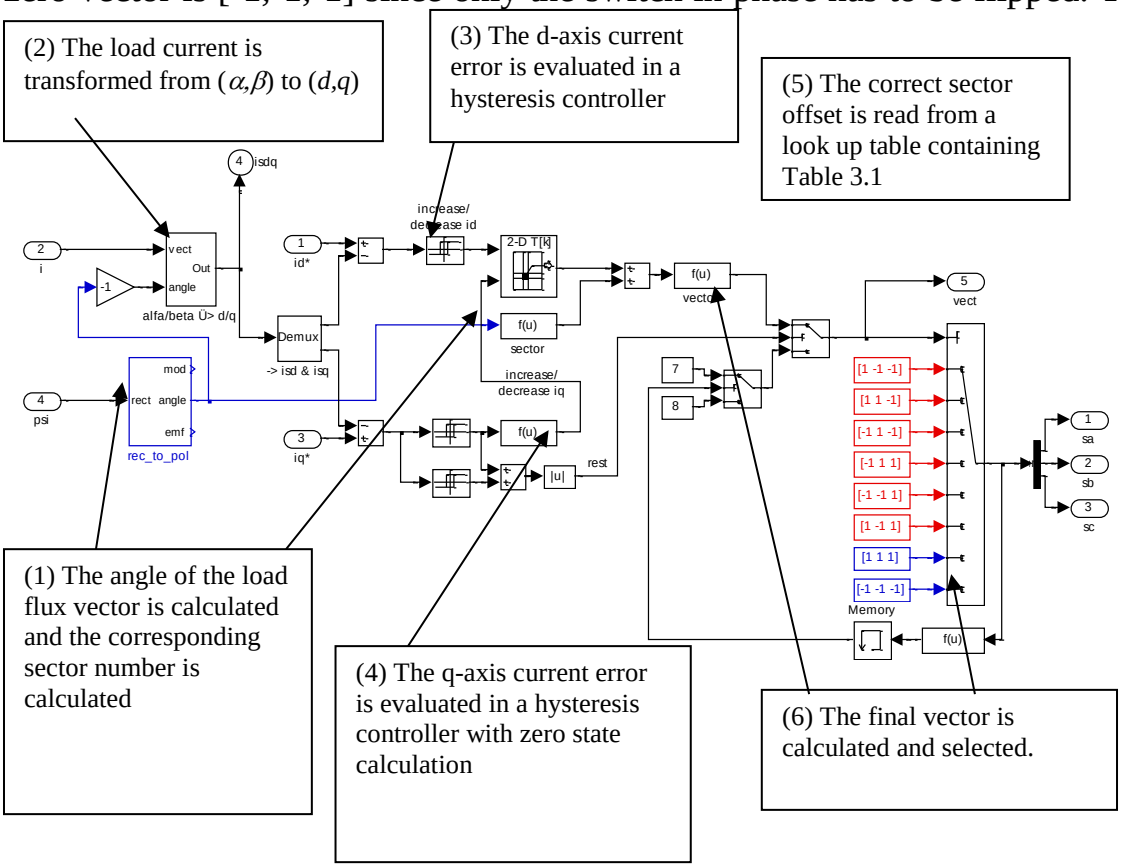


Figure 3.27 the selection of a suitable zero state is based on the last preceding active state.

Figure 3.28 shows the performance of the 3-phase DCC given the same conditions as for the simulations with the vector controller in section 3.7. Note how the two passive states (7 and 8) are both used based on the previous state. The DCC does appear to be ideally parameter insensitive, since the controller itself contains no load parameters at all. This is deceiving since the most difficult challenge usually is to estimate the quantity that defines the rotating reference frame, in this case the load flux. With the generic three phase load the emf is easily measurable and the load flux is thus easily accessed, but in most AC motor controllers the flux vector used has to be estimated which is a difficult and usually parameter sensitive task. Thus the main parameter sensitivity is not in the DCC as such, but in the determination of the rotating reference frame which is as difficult with DCC as with sampled vector control.

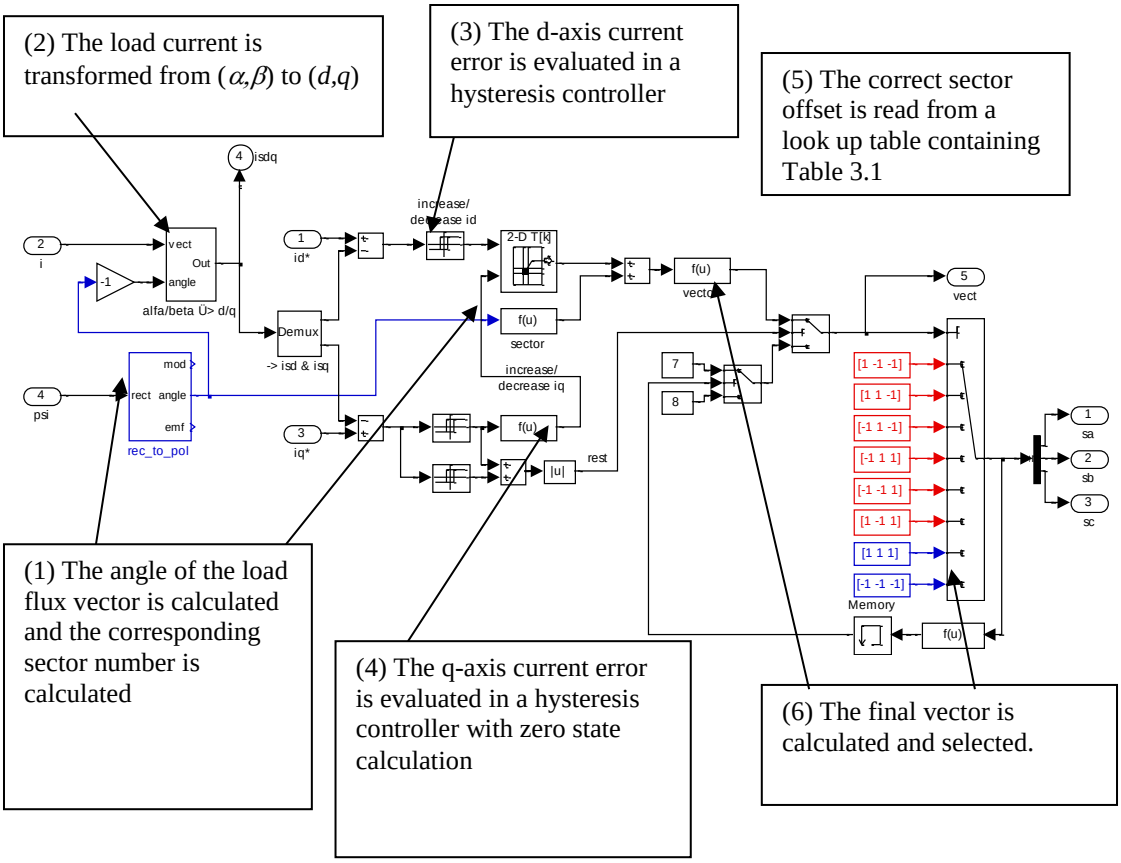


Figure 3.27 3-phase DCC and its components.

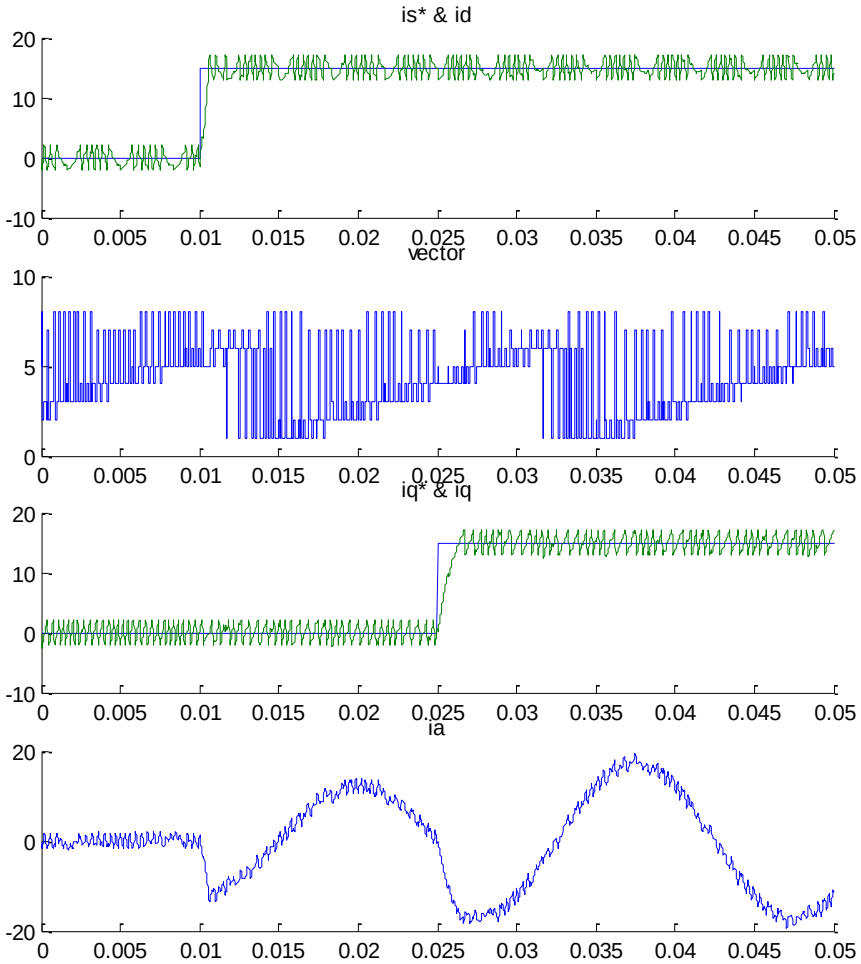


Figure 3.28 Step response of the 3-phase DCC method.

4 Power Semiconductor Devices

This chapter gives an introduction to the basic operation of power electronic semiconductor devices. For a power electronic designer at least a brief understanding of semiconductor physics and operation is important to use the devices properly. This chapter is based on the information given in [1][9][14][19][21][22].

4.1 The pn-junction

Mobility

The mobility is the constant of proportionality between carrier velocity and electric field strength. In this case carrier velocity means the carrier velocity due to the electric field excluding the thermal velocity. Still, the mobility is temperature dependent. At low temperature, the mobility is increasing with increasing temperature. At high temperature, the mobility is instead decreasing with increasing temperature. At low temperature, the carriers are affected by ionized doping atoms as the carriers pass them, and hence, the mobility is low. This phenomenon is termed impurity scattering. When the temperature increases, the effect of impurity scattering reduces, since the increased thermal velocity causes that the doping atoms are passed in a shorter time. If the temperature is further increased, the mobility reduces due to other phenomena termed lattice scattering. This phenomena is due to the fact that the atoms of the crystal oscillates around their equilibrium at high temperatures causing that the free carriers exchange energy with the atoms of the crystal, e.g. the gitter atoms.

The electron mobility is higher than the hole mobility, mainly because the electrons have lower effective mass than the holes (approximately one third). The reason why the mobility is of great importance is that the resistivity of a semiconductor is inversely proportional to the mobility. It is feasible for a semiconductor to possess high resistivity in the blocking state, since this gives a low leakage current. On the other hand, in the conduction state, the resistivity should be low since a low forward voltage gives low conduction losses.

Doping

By adding impurities to the silicon crystal, in the form of other materials with three or five valence electrons, the silicon crystal is p- or n-doped. Doping alters the thermal equilibrium density of holes and electrons, i.e. free carriers. A doping atom of a material with three valence electrons will immediately (for

temperature only slightly above 0 K) accept a free electron to create covalent bonds to four silicon atoms. The result is that the number of free electrons decreases with one while the number of free holes remains unaltered. The holes are termed majority carriers and the electrons minority carriers. The crystal is said to be p-doped and the doping material is called an acceptor.

If the atoms of the doping material instead have five valence electrons there is one additional electron after the doping atom has created bonds to four silicon atoms. Besides this free electron donated by the doping atom, there is also a fixed positive charge in the nucleus of the doping atom. This positive charge correspond hole that is not free to move in the crystal. However, space charge neutrality still holds even though the crystal is doped. In this case the electrons are majority carriers and the holes are minority carriers. The silicon crystal is n-doped and the doping material is called a donor since it donates free electrons to the crystal.

Recombination

At thermal equilibrium, the density of free charge carriers is constant, which implies that the free electron-hole pairs must disappear at the same rate as they are created. This could either be due to recombination where a covalent bond is re-established or due to trapping where the free carrier is trapped by an impurity atom or an imperfection of the crystal.

An important parameter for bipolar components is the minority carrier lifetime which is the time-constant which the excess carrier density decreases with, i.e. recombination rate. For power semiconductors there are two important parameters affecting the minority carrier lifetime. Firstly, the minority carrier lifetime increases with increasing temperature since the energy of the minority carriers increases making recombination less likely. For bipolar components this results in longer switching times since a large portion of the excess carriers are removed by recombination at turn-off. Secondly, for high excess carrier concentrations the minority carrier lifetime becomes dependent on the minority carrier concentration. When the minority carrier concentration approaches 10^{-17} cm^{-3} another recombination process commences, termed Auger recombination. For Auger recombination, the minority carrier lifetime decreases as the concentration increases. For bipolar devices, this decrease in minority carrier lifetime results in an increase of the conduction losses at high current densities (which is equivalent to high minority carrier concentrations).

Some new terms were introduced here, which may require an explanation. A recombination centre is either an imperfection in the crystal structure (could be added on purpose) or the presence of doping atoms. In both cases the recombination centre provides an energy level which the free carriers can

recombine to on the way to the next lower energy level. This implies that the recombination centre should be located at an energy level between the conduction and valence bands. The reason why a recombination centre could be favourable in power semiconductors is that Silicon has an indirect band gap structure implying that direct recombination (from the conduction to the valence band) unlikely, resulting in a very long minority carrier lifetime if the recombination centre was not present and, thus, high switching losses.

The other term, Auger recombination, is a result of interaction between free carriers at elevated doping or minority carrier injection levels, where an electron-hole pair exchanges energy and momentum with a third carrier at recombination.

Drift and diffusion

The current through a semiconductor is the sum of the hole current (in the same direction as the external current) and the electron current (in the opposite direction compared to the external current). These both current components could be further divided into two components each, drift and diffusion current. The drift current is caused by an electric field such that both free holes and free electrons are accelerated. The free holes and electrons are provided by the doping atoms and the thermally ionized electron-hole-pairs of the depletion region. Note that the hole and electron currents flow in opposite direction. The diffusion current is a consequence of the variation, or gradient, of the free carrier density, in a cross-section perpendicular to both the direction of the current flow and the doping-structure. The diffusion current flows to equalize this variation.

The pn-junction

A pn-junction is characterized by a p- and an n-region adjacent to each other on the same crystal. For example, a pn-junction could be created by diffusing acceptor atoms into an n-doped crystal. The concentration of donor atoms in the n-doped region is denoted N_d ([cm^{-3}]). The concentration of acceptor atoms in the p-doped region is denoted N_a. The pn-junction is characterized by the way the doping atom concentration varies from the p-side to the n-side (the steepness of the concentration) and how high the doping atom concentration is on each side of the junction.

In the following sections on power semiconductors, a high doping atom concentration is denoted with a + and a low with a -, for example, n^+ and n^-. A low concentration means in this case that the density of free carriers is only slightly higher than the intrinsic free carrier density, which is a function of the crystal temperature. To express doping atom densities in between high and low, the labels n and p are used, i.e. without superscripts. Note that this way of

expressing the doping atom densities on both sides of the junction is relative, even if it is not only the difference that is characterizing the junction but also the absolute levels of doping atom density.

Potential barrier

For a pn-junction that is not connected to a circuit, some majority carriers on each side will diffuse to the other side where they become minority carriers. The reason for this diffusion is, as described before, the variation in doping atom density causing a variation in the concentration of free carriers. The diffusing carriers leave an ionized doping atom behind, representing a fix charge. Together the ionized doping atoms represent a space charge region or depletion region containing fixed charge. The space charge on each side of the junction results in an electric field. To simplify the calculation of the electric field and accompanying potential barrier, the pn-junction is considered as being abrupt, which is actually the worst case regarding the magnitude of the electric field strength. The derivation of the spatial variation of the electric field given below is valid for a steep or abrupt junction.

According to Poisson's equation, the electric field, *E*, in one dimension is expressed as

$$\nabla \cdot E = \frac{\partial E}{\partial x} = \frac{\rho}{\varepsilon} \tag{4.1}$$

where ρ denotes the charge distribution in a volume element ([cm^{-3}]) and ε denotes the di-electricity constant of the volume element. For the abrupt pn-junction this means that the spatial variation of the electric field is given by

$$\frac{\partial E}{\partial x} = \begin{cases} -\dfrac{q_e \cdot N_a}{\varepsilon} & -x_p < x < 0 \\ \dfrac{q_e \cdot N_d}{\varepsilon} & 0 < x < x_n \end{cases} \tag{4.2}$$

The elementary charge (the absolute value of the charge of one electron) is denoted q_e. If the expression above is integrated, i.e. the electric field is calculated from

$$E(x) = E(x_0) + \int_{x_0}^{x} \frac{\partial E}{\partial x} dx \tag{4.3}$$

the result becomes

$$E(x)=\begin{cases} E(-x_p)-\dfrac{q_e\cdot N_a}{\varepsilon}\cdot\left(x+x_p\right) & -x_p<x<0 \\ E(0)+\dfrac{q_e\cdot N_d}{\varepsilon} & x\ 0<x<x_n \end{cases} \tag{4.4}$$

Since the depletion region should absorb the electric field, it is required that

$$E(-x_p)=0 \;\Rightarrow\; E(0)=-\frac{q_e\cdot N_a}{\varepsilon}\cdot x_p \tag{4.5}$$

Space charge neutrality (the total positive and negative charge on each side of the pn-junction should be equal but with opposite sign) yields

$$q_e\cdot N_a\cdot x_p=q_e\cdot N_d\cdot x_n \;\Rightarrow\; x_p=\frac{N_d}{N_a}\cdot x_n \;\Rightarrow\; E(0)=-\frac{q_e\cdot N_d}{\varepsilon}\cdot x_n \tag{4.6}$$

which gives an *E*-field of the depletion region according to

$$E(x)=\begin{cases} -\dfrac{q_e\cdot N_a}{\varepsilon}\cdot\left(x+x_p\right) & -x_p<x<0 \\ \dfrac{q_e\cdot N_d}{\varepsilon}\cdot\left(x-x_n\right) & 0<x<x_n \end{cases} \tag{4.7}$$

The potential barrier, or the voltage that this electric field corresponds to when no external source is applied, is calculated from

$$\nabla V=\frac{\partial V}{\partial x}=-E \tag{4.8}$$

which gives the potential barrier, V_B, according to

$$V_B=-\int_{-x_p}^{x_n}E(x)dx=-\int_{-x_p}^{0}E(x)dx-\int_{0}^{x_n}E(x)dx=\frac{q_e}{2\cdot\varepsilon}\cdot\left(N_a\cdot x_p^2+N_d\cdot x_n^2\right) \tag{4.9}$$

The expressions above are valid if the pn-junction is not connected to a voltage source. Figure 4.1 shows the extent of the depletion region and the variation of the *E*-field along the depletion region for this case. In the general case with an external voltage source applied, the width of the depletion region is affected and therefore

$$V_B-V=\frac{q_e}{2\cdot\varepsilon}\cdot\left(N_a\cdot x_p^2(V)+N_d\cdot x_n^2(V)\right) \tag{4.10}$$

where V denotes the externally applied voltage. Space charge neutrality must still hold. Hence

$$q_e \cdot N_a \cdot x_p(V) = q_e \cdot N_d \cdot x_n(V) \;\Rightarrow\; x_p(V) = \frac{N_d}{N_a} \cdot x_n(V) \tag{4.11}$$

which gives

$$V_B - V = \frac{q_e}{2 \cdot \varepsilon} \cdot \left(\frac{N_d^2}{N_a} + N_d \right) \cdot x_n^2(V) = \frac{q_e}{2 \cdot \varepsilon} \cdot \left(N_a + \frac{N_a^2}{N_d} \right) \cdot x_p^2(V) \tag{4.12}$$

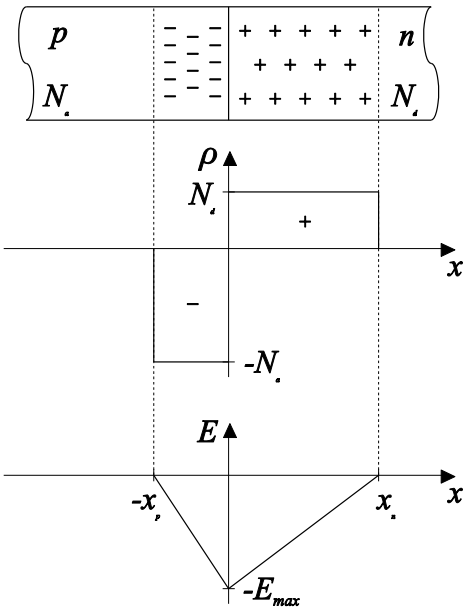


Figure 4.1 The extent of the depletion region and the variation of the E-field along the depletion region when no external voltage is applied

The width of the depletion region as a function of the applied voltage is consequently written

$$\begin{aligned} W(V) &= x_p(V) + x_n(V) = \\ &= \sqrt{\frac{2 \cdot \varepsilon \cdot V_B}{q_e} \cdot \frac{N_a \cdot N_d}{N_a + N_d}} \cdot \sqrt{1 - \left(\frac{V}{V_B} \right)} = W(0) \cdot \sqrt{1 - \left(\frac{V}{V_B} \right)} \end{aligned} \tag{4.13}$$

where *W(V)* and *W(0)* denote the width of the depletion region for voltages not equal to zero and equal to zero, respectively. An important conclusion drawn from the last expression is that the width of the depletion region increases with increasing reverse voltage (*V* negative) and decreases with increasing forward voltage (*V* positive). The width of the depletion regions in the p- and n-doped regions, respectively, is given by

$$\begin{cases} x_p(V) = \dfrac{N_d}{N_a + N_d} \cdot W(V) = \dfrac{N_d}{N_a + N_d} \cdot W(0) \cdot \sqrt{1 - \left(\dfrac{V}{V_B} \right)} \\ x_n(V) = \dfrac{N_a}{N_a + N_d} \cdot W(V) = \dfrac{N_a}{N_a + N_d} \cdot W(0) \cdot \sqrt{1 - \left(\dfrac{V}{V_B} \right)} \end{cases} \tag{4.14}$$

A very important conclusion to draw from the previous expression is that the depletion region is mainly contained in the region with the lowest doping of the pn-junction. This feature is used in all power semiconductor devices.

The maximum electric field strength of the pn-junction as a function of the externally applied voltage is written

$$|E_{max}| = \frac{q_e}{\varepsilon} \cdot \frac{N_a \cdot N_d}{N_a + N_d} \cdot \sqrt{1 - \left(\frac{V}{V_B}\right)} = \frac{2 \cdot V_B}{W(0)} \cdot \sqrt{1 - \left(\frac{V}{V_B}\right)} \qquad (4.15)$$

The maximum strength of the *E*-field is also important, since it must be below the level where avalanche breakdown occurs. Avalanche breakdown is discussed in a later section. However, note that in the expression above the electric field strength becomes low if the doping on one of the sides of the pn-junction has a low doping. Usually, in power semiconductors there is a doping layer of n-type with very low doping called the drift region. If the doping level of the adjacent p-type layer of the blocking junction is moderate to high, i.e. $N_a >> N_d$, the maximum electric field strength is given by

$$|E_{max}| = \frac{q_e}{\varepsilon} \cdot \frac{N_a \cdot N_d}{N_a + N_d} \cdot \sqrt{1 - \left(\frac{V}{V_B}\right)} \approx \frac{q_e}{\varepsilon} \cdot N_d \cdot \sqrt{1 - \left(\frac{V}{V_B}\right)} \qquad (4.16)$$

This feature is also used in all power semiconductors.

Bias

When an external voltage is applied between the p- and n-sides of a pn-junction it will essentially appear across the space charge (or depletion) region, due to its high resistance. The resistance is high as a result of the low concentration of free carriers compared to the rest of the crystal. If the applied voltage is of higher potential on the p-side compared to the n-side, the voltage counteracts the potential barrier and the pn-junction is said to be forward biased. On the contrary, if the voltage has the opposite polarity such that it has a higher potential on the n-side than on the p-side, the pn-junction is said to be reverse biased.

The width of the depletion region, i.e. space charge region, increases with reverse bias since an increase of the voltage the applied across the depletion region results in an increase of the electric field which is accomplished by a larger number of fixed charges (ionized doping atoms) included in the depletion region. Note that space charge neutrality must still hold in the depletion region, indicating that the total positive charge on the n-side of the pn-junction equals the total negative charge on the p-side. The concentration of fix charges in the depletion region of either the p- or n-side of the junction is equal to the doping

atom concentration on the same side. Therefore, the depletion region will extend mostly into the side with the lower doping. For a pn-junction not connected to a source or connected to a source with voltage $V = 0$ V, the depletion region will appear as in Figure 4.2. Note that N_a is the doping atom concentration on the p-side (acceptors) and N_d is the doping atom concentration on the n-side (donors).

Forward bias reduces the potential barrier and upsets the equilibrium between drift and diffusion current in favour of diffusion. This results in a heavy increase in the minority carrier concentrations on both sides of the depletion region. However, note that the minority carrier concentrations decrease exponentially with increasing distance away from the depletion region, as a consequence of the minority carrier lifetime.

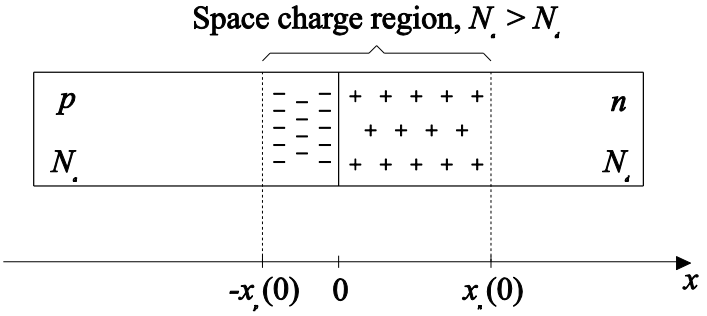


Figure 4.2 *The principal appearance of the depletion region in the case of V=0 V. Note that the depletion region extends mostly into the region of low doping of the crystal forming the pn-junction.*

The characteristic length of this decay (for partial differential equations the characteristic length corresponds to the time-constant for ordinary differential equations) is called minority carrier diffusion length. The diffusion length is associated to the minority carrier life time through the so called Einstein relation (for electrons and holes)

$$L_n = \sqrt{D_n \cdot \tau_n} \tag{4.17}$$

$$L_p = \sqrt{D_p \cdot \tau_p} \tag{4.18}$$

where L, D and τ denote diffusion length, diffusion constant and carrier lifetime for minority carriers. Index n and p denote that the minority carriers are electrons and holes, respectively. Figure 4.3 shows a sketch of the minority carrier concentration in the case of forward bias.

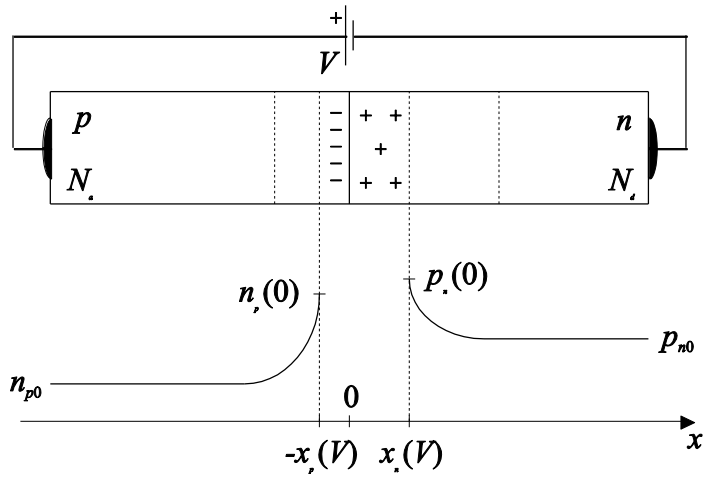


Figure 4.3 *In the case of forward bias the width of the depletion region decreases and the minority carrier concentration at the boundaries of the depletion region increases significantly.*

Note that the width of the depletion is lower for forward bias $V > 0$ V (Figure 4.3) compared to the case with an unbiased pn-junction ($V = 0$ V). The minority carrier concentration at the borders of the depletion region is exponentially depending on the level of the applied voltage and could therefore vary several orders of magnitude. This is the reason why diffusion current dominates over drift current in case of forward bias.

Reverse bias increases the potential barrier, which in turn reduces the probability of diffusion across the depletion region. The minority carrier density becomes very low (virtually zero) at the boundaries of the depletion region, see Figure 4.4.

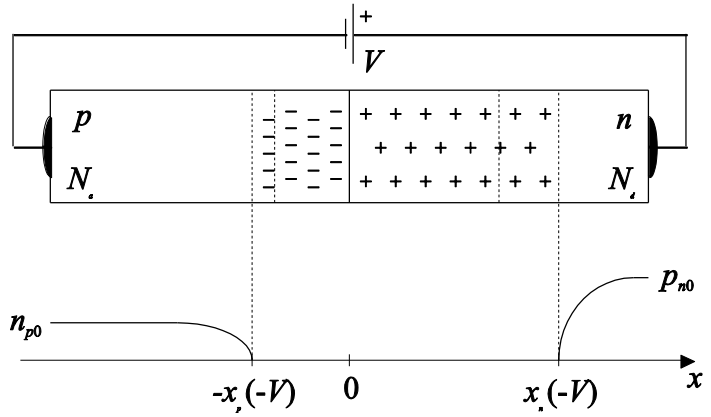


Figure 4.4 *In the case of reverse bias the width of the depletion region increases and the minority carrier concentration at the boundaries of the depletion region decreases to become virtually equal to zero, since the minority carriers are immediately swept over to the opposite side by the strong E-field inside the depletion region.*

The small gradients in the excess carrier concentrations of Figure 4.4 results in a small amount of diffusing minority, carriers flowing towards the depletion region. When these minority carriers reach the boundary of the depletion region they are immediately swept over to the opposite side by the strong E-field inside the depletion region. This results in a small reverse saturation current, with a magnitude essentially independent on the magnitude of the reverse voltage.

Also, the reverse saturation current is partly due to thermally ionized silicon atoms from the depletion region.

Avalanche breakdown

The strong increase of the saturation current for high reverse voltage is termed avalanche breakdown. Avalanche breakdown must be avoided since simultaneously high voltage and current implies high power losses which could destroy the device in short time. Avalanche breakdown is due to a physical phenomenon termed impact ionization. Impact ionization results from highly energized electrons colliding with silicon atoms in such a way that covalent bonds are broken and new electrons become free. If the electric field strength is high enough, both the involved electrons are accelerated and can repeat this process in a multiplicative manner and avalanche breakdown has occurred.

In all types of power electronic semiconductor devices a reverse biased pn-junction accommodates the blocking voltage. To avoid avalanche breakdown, caused by too high electric field strength in the blocking pn-junction of the silicon crystal, one of the sides of the blocking junction must be of low doping according to the expressions derived in this section. Low doping results in a depletion region extending widely into the low doped region, calling for a specific minimum thickness to accommodate the depletion region. This lightly doped region is termed a drift region and is present in all power electronic semiconductor devices.

4.2 The power diode

Power diodes are usually separated into two categories, rectifier and switching diodes. The doping structures for both categories consist of the same layers, at least for the voltage levels discussed here. However, they behave different from each other, due to the fact that the length of the layers (in the direction of the current) and their doping densities are not the same. Therefore, the discussion is applied to power diodes in general, and then the differences between rectifier and switching diodes are explained.

Basic doping structure

The doping structure for a power diode differs, to some extent, from that of a low power and low voltage diode. The main difference originates from the fact that the power diode has to sustain high reverse voltages, implying that the doping structure has to withstand a high electric field, without failure. For this reason, all power electronic semiconductor devices are equipped with a long (in the direction of the electric field) doping layer with low doping density. This

low-doping layer, not found in low voltage and low power devices, is often termed drift layer. In Figure 4.5, the principal doping structure of a power diode is shown. Note that for real devices, the drift layer is considerably thicker than the other. However, in Figure 4.5, the thicknesses of the other layers are exaggerated for visibility.

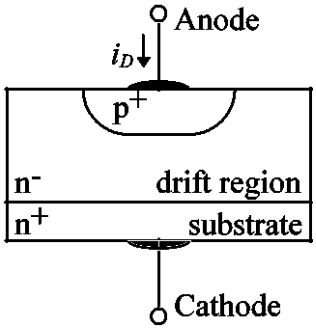


Figure 4.5 Principal power diode doping structure.

From the low power counterpart, the highly doped anode and cathode layers are recognized. The idea of having one thick layer of low doping atom density is that the depletion region formed in the blocking state should mainly be located to this layer. The low doping results in a low peak electric field, and the large thickness means that the entire depletion region is held inside the drift layer, i.e. it does not extend into the cathode layer.

However, for so called punch through (PT) or buffer layer devices a geometrically short but highly doped buffer layer is located in between the drift and cathode layers. In this case, the idea is the same but the drift region has in this case an even lower doping density. Also, the drift region is allowed to extend across the entire drift region to reach the buffer layer. The result of this manipulation is that the drift region can be less thick. According to [14], the thickness can be approximately halved compared to the non-punch through (NPT) case.

The diode is called a natural PT device since no additional drawbacks arises for the PT compared to the NPT device. Also, the highly doped substrate serves as buffer layer for PT-diodes.

For other types of devices, buffer layer devices are often referred to as asymmetrical. The reason is that the PT device cannot block voltage of the opposite polarity, since none of the junctions sustain the high electric field strength resulting from the high doping densities on both sides of the blocking junction. For the diode, no forward blocking capability is provided, implying that this does not matter in this case.

Steady state operation

When a power diode is forward biased, heavy injection of minority carriers into the low-doped drift region occurs. In fact, the density of free carriers can become orders of magnitude higher than depicted by the doping atom density. Consequently, the resistivity of the drift region becomes considerably lower than expected from the doping density. This phenomenon is often referred to as conductivity modulation. Due to conductivity modulation, the forward voltage drop thus becomes rather low.

For conductivity modulation to be established, it is of course important that the diffusion length in the drift region is not too short compared to the thickness (in the direction of the minority carrier movement) of the drift region. Otherwise, recombination effects will cause a low minority carrier density in the bulk of the drift region.

When the diode is reversed biased the anode-drift junction supports the voltage, i.e. a depletion region is formed at the junction. The depletion region extends mostly into the lightly doped drift region, see [14]. In this case a low doping atom density of the drift region is advantageous since the peak electric field strength is reduced. However, the depletion region also expands, implying that the drift region must be thick if it is desired that the depletion region must not extend across the entire drift region. However, since the diode is a natural PT device, the high doping of the cathode layer effectively prevents the depletion region from reaching the cathode contact.

Switching

When a semiconductor device traverses from the conduction state to the blocking, or vice versa, is termed a switching. For a power diode, several phenomena not seen for the low power counterpart appears. In Figure 4.6, typical turn-on and turn-off waveforms are shown for a freewheeling diode in a bridge application. The diode voltage v_D shown in Figure 4.6, is the voltage between the anode and cathode of the diode, i.e. v_D=v_{AC}.

At turn-on, the diode reverse voltage first declines to zero. Then, the rectifying junction becomes forward biased. Consequently, the diode current starts to increase. The diode current rate of change, i.e. the diode current time derivative, is determined by the transistor turning off.

If the current rate of change is extremely high, a diode forward voltage peak is observed [14], depending on the fact that excess carriers are not injected into the drift region at the corresponding rate. Hence, a conductivity modulation lag is observed. This results in a higher forward voltage drop than expected from the

data sheets, due to the higher resistivity of the drift region. To some extent, the voltage peak observed is also due to stray inductance of the bonding wires of the diode housing. As seen in Figure 4.6, a turn-on voltage peak does not show up here.

When the diode current has reached the level determined by the load, the diode voltage returns from its peak level, to eventually reach its steady state. The phenomena where the forward voltage drop becomes considerably higher than what is expected from the data sheets is termed forward recovery.

The steady state forward voltage drop is determined both by the height of the junction potential barrier and the resistivity of the doping layers, mainly the drift region.

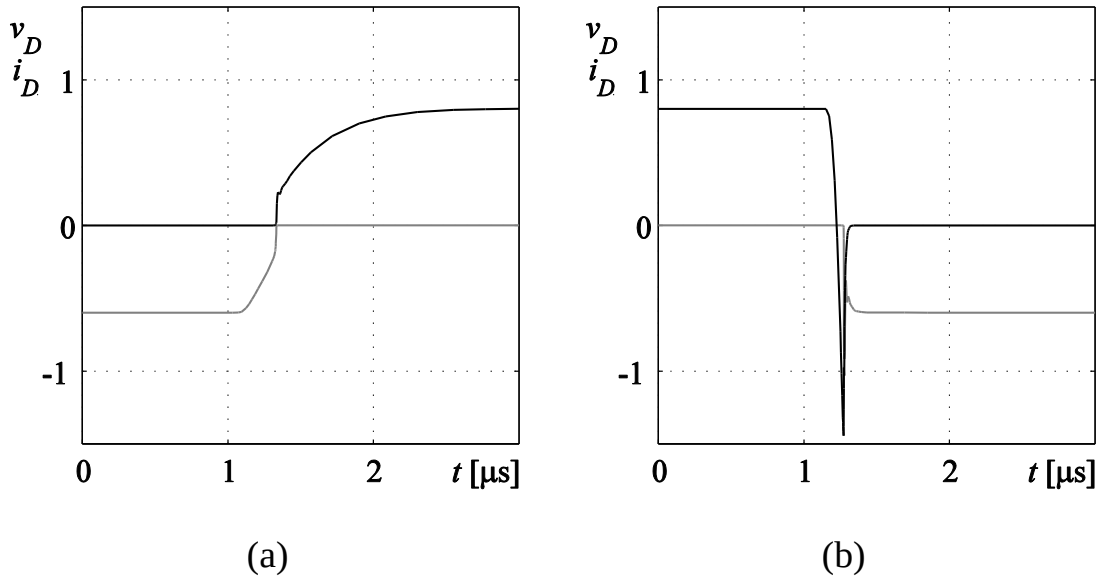


Figure 4.6 *The freewheeling diode current iD (black) and forward voltage vD (grey) during (a) turn-on, and (b) turn-off. Note that the corresponding transistor, an IGBT in this case, is non-ideal, i.e. it affects the diode switching waveforms.*

A freewheeling diode turns off when the complementary transistor is turned on. First, the diode current starts to decrease at a rate determined by the power transistor. On the contrary to what is observed for the low power counterpart, the power diode turn-off current continues to decrease to become negative. This is due to the fact that the excess carriers stored in the drift region forward biases the anode-drift junction.

As soon as these excess carriers are swept out, the anode-drift junction can support the blocking voltage. Thus, the diode voltage becomes negative and the diode current decrease is inhibited. Now, the diode current starts to increase towards zero. The diode current in this part removes the free carriers in the anode and drift layers to create a depletion region. Also, at turn-off a negative voltage peak is observed. The voltage peak in this case is due to the stray inductance of the entire power electronic circuit. Note that the power diode

reverse blocking voltage must be selected to sustain this peak level. The phenomenon in which the diode current becomes negative is termed reverse recovery.

The main difference between freewheeling and rectifier diodes is that conductivity modulation is more pronounced for the latter, i.e. a larger amount of charge is stored in the drift region when a rectifier diode is conducting compared with a freewheeling diode. This implies that the rectifier diode will have a lower forward voltage drop, due to its lower on-state resistance. However, the switching transients will have a longer duration due to the larger amount of charge stored. This implies that the rectifier diode is more sensitive to high diode current derivatives. Both the forward recovery voltage peak and the reverse recovery current peak will thus be higher for a rectifier diode compared to a freewheeling diode of the same rating for a given diode current time derivative.

4.3 The bipolar junction transistor

Similar to the power diode and other types of power semiconductor devices, the power bipolar junction transistor (BJT) has a slightly different doping structure compared to its low power counterpart. The main difference is that for the power BJT a drift region, with large thickness and light doping, is required. Also, the width of the base region is considerably larger for a power BJT compared to the small-signal BJT. These modifications result in differences in the output current-voltage characteristics and the switching behaviour.

The power BJT has in most cases a vertical doping structure since a large cross-sectional area gives a low current density. Furthermore, a large cross-sectional area enhances cooling, since the thermal resistance is inversely proportional to the area of the cooling surface. Figure 4.7 shows typical doping structure for a vertical NPN-BJT. Note that a PNP-BJT would have opposite doping in each layer of the doping structure.

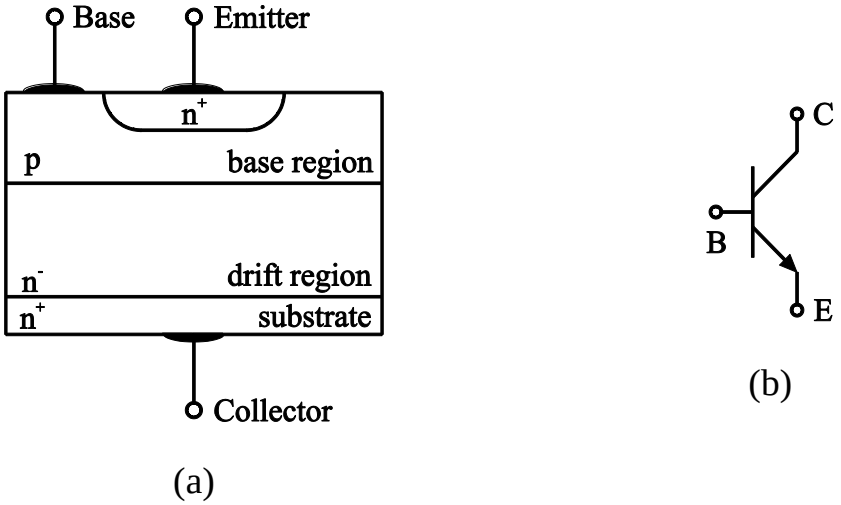


Figure 4.7 Principal doping structure (a) and schematic symbol for a power BJT (b).

The schematic symbol of an NPN-BJT is also shown in Figure 4.7. The power BJT has three terminals denoted base, collector and emitter. In most power application the base serves as input and the collector as output. The emitter is common for both input and output.

The thickness of the drift region determines the breakdown (primary breakdown) voltage of the power BJT, i.e. the maximum voltage the transistor can sustain in blocking state without the occurrence of avalanche breakdown. The thickness of the base affects the current gain to a large extent. A thin base region results in a high current gain but gives a low breakdown voltage, which is discussed in later sections. Commercial BJT:s have their base and emitter regions organized as long and thin fingers, to reduce the risk of current displacement which otherwise could cause locally high losses leading to secondary breakdown. Secondary breakdown is discussed in a later section.

The current gain for power BJT:s is usually rather low in the range 5-20. In most power applications such current gain is too low, since it puts strong requirements on the current handling capabilities of the transistor drivers. Consequently, most power BJT:s are based on monolithic (fabricated on the same wafer) Darlington stages. In such a Darlington stage two or three transistors can be included resulting in a very high current gain.

Output characteristic

The output characteristic consists of a set of curves where the collector current is plotted against the collector-emitter voltage for different levels of the base current. Figure 4.8 shows a typical output characteristic for a power BJT. Several important features are observed from the output characteristic. The collector-emitter breakdown voltage in the case of substantial collector current is denoted BV_{SUS} (Breakdown Voltage, $I_C > 0$ Sustained). The collector-emitter

breakdown voltage for open base circuit is termed BV_{CEO} (Breakdown Voltage, Collector-Emitter, base Open circuited). The breakdown voltage for the collector-base with open emitter circuit is denoted BV_{CBO} (Breakdown Voltage, Collector-Base, emitter Open circuited).

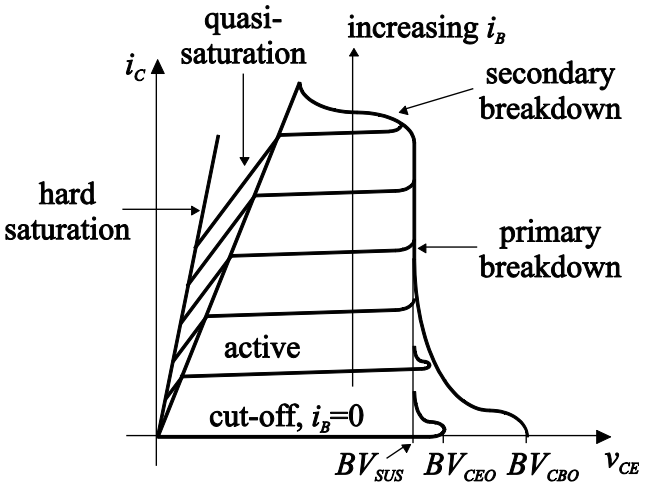


Figure 4.8 Typical output characteristic for a power BJT.

In power application discussed here BV_{CBO} is seldom addressed with one exception; the IGBT. The IGBT can be understood as a Darlington stage consisting of one PNP-BJT and one MOSFET. Primary breakdown denote avalanche breakdown in the collector-base junction. Secondary breakdown is due to locally high power dissipation accompanied by local temperature increase. The underlying physical processes behind secondary breakdown are discussed in a later section.

The main difference between the output characteristic of a power BJT and the one of a small-signal BJT is the region called quasi-saturation. Quasi–saturation is due to the presence of the lightly doped drift region. Since small-signal BJT do not have a drift region the phenomenon is not observed for these transistors.

Current amplification

To understand the mechanisms behind current amplification in power BJT:s the device is investigated in the active or linear region. In the active region the presence of the drift region does not influence the behaviour of the power BJT. Therefore, a BJT with a doping structure equal to the one of a small signal device is investigated first. Then the differences caused by the presence of the drift region are mentioned.

In the active region, the base-emitter junction is forward biased and electrons are injected from the emitter into the base region whereas holes are injected from the base into the emitter region. This minority carrier injection causes minority carrier distributions with very high gradients. Especially, the minority carrier

distribution in the base region exhibits a high gradient due to the fact that the collector-base junction is reverse biased causing the minority carrier density to be virtually zero at the boundary of its depletion region. The large gradient in minority carrier density in the base yields that the base current is caused by diffusion. The current transport mechanism is similar to the one in a power diode but with one important difference; in the BJT the base current consists of both electrons and holes.

The electron current (from the emitter into the base region) is namely larger than the hole current (from the base into the emitter region) for the BJT due to the fact that the transistor structure provides an alternative path for the electrons injected into the base region, e.g. via the collector region. The electrons injected from the emitter into the base region are more likely to leave the base region via the collector region instead of via the base contact for three reasons:

The thickness of the base region (in the direction of the current flow) is small compared to the diffusion length of the electrons which makes it unlikely for the electrons to recombine in the base region.

The area of the collector (perpendicular to the current flow) is much larger than that of the base making it more likely that the diffusing electrons from the base reach the collector region than the base contact.

Since the collector-base is a reverse biased pn-junction, the electron concentration is virtually zero at the boundary of the depletion region. The electric field strength of the depletion region is very high which implies that the electrons reaching its boundary will be swept over to the collector.

Together, these three reasons because the base current to be much lower than the collector current and that the collector and emitter currents are almost equal. A small base current results in a considerably larger collector current, i.e. a strong current amplification takes place. For a high current gain to the following three requirements must be satisfied:

Heavy doping of the emitter region yielding that only a small hole current from the base into the emitter region is required to maintain a high hole concentration.

A long minority carrier lifetime in the base region meaning that only a small amount of the injected electrons recombine in the base.

A thin base (in the direction of the current flow) compared to the diffusion length meaning a large portion of the injected electrons reaches the boundary of the depletion region of the reverse biased collector-base junction.

The third condition of the above listed is complicated to fulfil since a thin base results in a low breakdown voltage for the collector-base junction. Therefore, the thickness of the base is always a compromise between voltage blocking capability and current gain, in most cases in favour for the blocking capability since this cannot be compromised.

For the power BJT, the current gain usually decreases for high collector currents due to conductivity modulation of the base region and also due to emitter current crowding. Conductivity modulation in the base region takes place when the minority carrier concentration equals the doping atom density, i.e. the concentration of injected electrons equals the density of acceptor atoms for an npn-transistor. When this occurs, excess holes of the same amount as the injected electrons must be supplied to the base region in order to maintain space charge neutrality. This implies that the base current must increase because this is the only way to supply the extra holes required. This increase is not accompanied by an increase in the collector current. Consequently, the current gain decreases.

Emitter current crowding is due to a lateral voltage drop, i.e. along the doping layers, in the base region. This voltage drop is subtracted from the applied base-emitter voltage in the interior of the device, resulting in a higher current density at the periphery than at the centre of the emitter region. This variation in current density implies that conductivity modulation of the base region takes place at lower total collector currents than what would be expected if the current density was uniform along the base-emitter junction. Therefore, the base and emitter regions are manufactured as long and thin fingers so that the cross-sectional area becomes large, providing a low current density, and the distance from the emitter periphery to its centre becomes short.

Quasi-saturation

Since quasi-saturation only occurs in the presence of the drift region this must be included in the further analysis. If the collector current increases due to an increase in the base current, the collector-emitter voltage decreases even though the resistive voltage drop across the drift region increases. This implies that the reverse bias across the collector-base junction also decreases. Eventually, the collector-base junction will be forward biased if the base current is high enough. When this occurs, holes from the base region are injected into the drift region. Due to space charge neutrality, electrons of the same amount must also be injected into the drift region. These electrons are supplied from the emitter and diffuse across the base to reach the drift region. This injection results in build-up of excess carriers in the drift region termed quasi-saturation.

In quasi-saturation the drift region is subject to double-injection, but the excess carrier distribution only grows from one side of the drift region since a considerably larger amount of electrons are injected from the emitter via the base into the drift region than from the collector region. As the collector current gradually increases, the excess carrier distribution covers an increasing part of the drift region, which thus becomes more and more conductive. The increased conductivity corresponds to a reduction of the resistance and consequently the collector-emitter voltage drop decreases. Since the excess carrier distribution covers not only the actual base but also a part of the drift region, the effective base becomes virtually thicker. This is termed base width modulation. Base width modulation results in a reduction of the current gain, since the increase in effective base width implies that recombination in the virtual base is more likely.

Hard saturation occurs when the distribution of excess carriers reaches the boundary between the drift and collector regions. It is obvious that hard saturation requires a certain minimum amount of stored charge. If even more charge is stored, by increasing the base current further, the BJT is just driven harder into saturation. This will result in longer switching times and should therefore be avoided.

The base transport factor

To analyse more complicated power semiconductor devices containing npn- and/or pnp-structures, e.g. thyristors, MOSFET:s and IGBT:s, it is feasible to introduce the base transport factor, α. For an npn-BJT, the base transport factor is defined as the ratio between the electrons currents injected from the base into the collector (I_{nc}) and from the emitter into the base region (I_{ne}), respectively. Mathematically this is expressed as

$$\alpha = \frac{I_{nc}}{I_{ne}} \tag{4.19}$$

The current gain β is related to the base transport factor α according to

$$\frac{1}{\beta} = \frac{I_B}{I_C} = \frac{I_{ne} - I_{nc}}{I_{nc}} + \frac{I_{pe}}{I_{nc}} \approx \frac{I_{ne} - I_{nc}}{I_{nc}} = \frac{1}{\alpha} - 1 = \frac{1-\alpha}{\alpha} \;\Rightarrow\; \alpha = \frac{\beta}{\beta+1} \tag{4.20}$$

As shown in the expression above the hole current injected from the base into the emitter region (I_{pe}) is much less than I_{nc}. According to the previous discussions, a small thickness of the base region and a long minority carrier lifetime in the base will result in I_{ne}-$I_{nc} \approx 0$. When a drift region is present the effective base width (thickness) will vary dependent on the extent of the excess

carrier distribution. At high injection levels, I_{ne}-$I_{nc} \approx 0$ no longer holds since substantial recombination occurs in the effective base.

The limit between the active and quasi-saturation operating modes is when the excess carrier distribution just begins to extend into the drift region. As mentioned before, the limit between quasi- and hard saturation is when the excess carrier distribution extends across the entire drift region and just reaches the collector region. The effective base width (thickness) is in the latter case the sum of the thickness of the actual base region and the thickness of the drift region. The drift region is usually considerably thicker than the base (approximately 10 times). Therefore, the current gain decreases when the operating mode changes from active towards hard saturation. If the diffusion length is less than the effective base width, the base transport factor α, and thereby current gain β, decreases significantly.

Switching

To study the behaviour of the BJT in switching transitions the device is assumed to be operating in simple step-down converter (Figure 4.9). Furthermore, it is assumed that the freewheeling diode (FWD) is ideal and that the load is purely inductive and therefore can be modelled as a constant current source with current I_0. In a realistic application the surrounding circuit determines the current but for investigation of the switching waveforms it is sufficient to assume that the load is a current source. The actual collector current together with the minority carrier lifetime of the base determines the amount of stored charge required to maintain the BJT in hard saturation. The actual current gain determines the minimum level of the base current corresponding to this collector current.

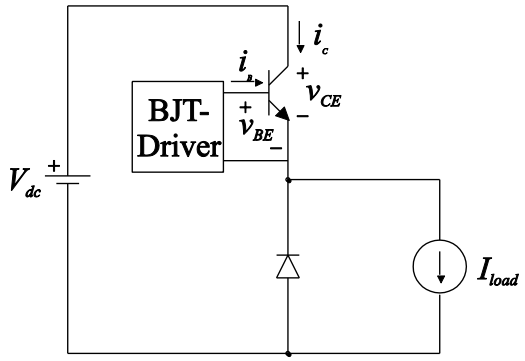


Figure 4.9 Step-down converter circuit used in the analysis of the switching behaviour of the power BJT.

A base current larger than this minimum will establish the required excess carrier distribution in shorter time, and therefore decreases the turn-on time. The larger base current also results in that a larger amount of excess carriers are stored in the interior of the device than required to just maintain hard saturation,

which prolongs the turn-off time to remove the stored charge. The resulting switching waveforms are shown in Figure 4.10.

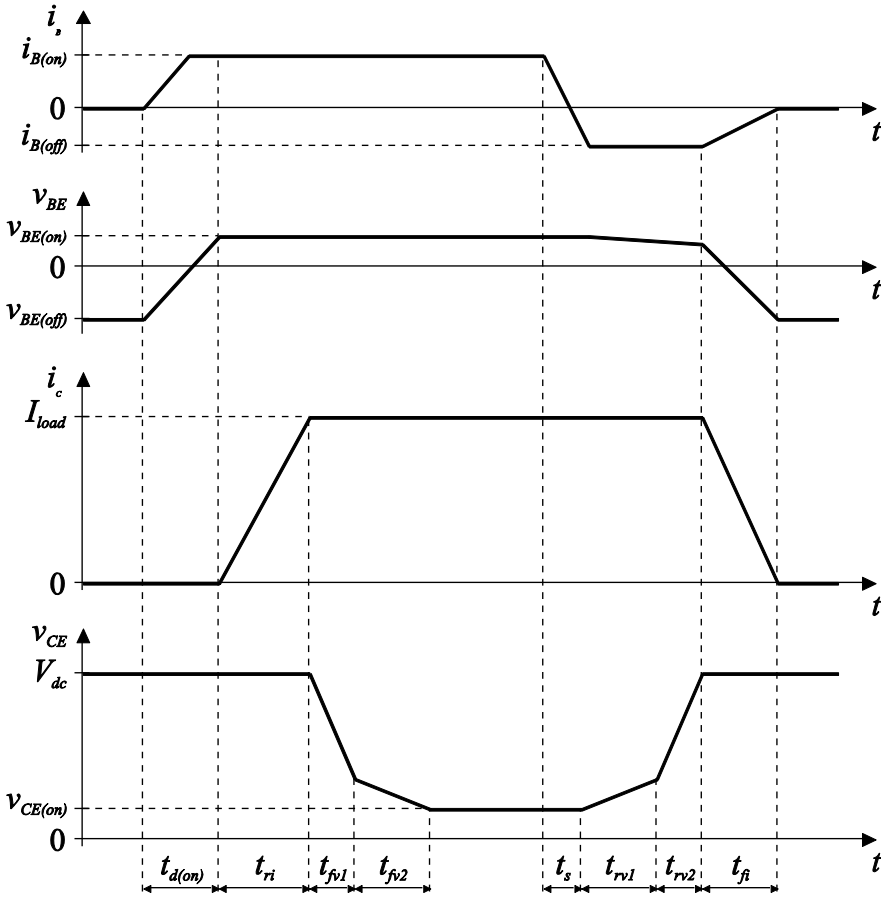


Figure 4.10 Principle switching waveforms for a power BJT. The first half shows turn-on and the second shows turn-off.

Turn-on

Several distinct time intervals are observed for a power BJT at turn-on:

Turn-on delay time ($t_{d(on)}$): No build-up of excess carriers since the depletion region of the base-emitter junction must be removed first. The depletion region is often understood as a space charge capacitance, which in this case must be recharged.

Current rise time (t_{ri}): When the turn-on delay is over the base-emitter junction is forward biased and injection of minority carriers commences at the same time as the collector current begins to increase. Note that the collector-emitter voltage remains high ($v_{CE} \approx V_{dc}$) since the freewheeling diode carries current. The reason why the diode carries current is that since the load is inductive, i.e. acts as a current source, the sum of the BJT and the FWD currents must equal the load current. Since the diode is conducting in this interval, the BJT will not enter saturation but remain in the active region of operation.

Voltage fall time (t_{fv}): This voltage fall time interval is divided into two parts, t_{fv1} and t_{fv2}. During t_{fv1} the collector-emitter voltage falls rapidly. The time interval t_{fv1} ends when the BJT leaves the active region of operation and enters quasi-saturation. During t_{fv2} the collector-emitter voltage falls more slowly than during t_{fv1}, depending on that the current gain decreases due to the increasing thickness of the effective base. When t_{fv2} finishes the BJT enters hard saturation.

Turn-off

Turn-off for a bipolar device like the BJT is primarily a matter of removing the stored charge in the interior of the transistor. It should be clear that it is not possible to just decrease the base current to zero and rely on internal recombination, since recombination is a slow process. Instead, the base current is forced negative so that the excess carriers are swept out. Mainly two different ways of applying this negative base current are utilized; either by abruptly changing the base current from positive to negative or to decrease the base current from positive to negative with a controlled time derivative d_{iB}/dt. It will turn out from the following discussion that the latter with controlled d_{iB}/dt is preferable and therefore this is investigated first. The method with an abrupt change is however not neglected since this is of principal interest due to the fact this is what happens inside several complicated structures such as the Darlington BJT and the IGBT. For turn-off with a controlled base current derivative the following distinct intervals are observed:

Storage time (t_s): The base current is changed from a positive to a negative level but neither the collector current or the collector-emitter voltage is affected. However, the charge stored in excess to the limit of hard saturation is removed. When t_s is over the BJT operates at the boundary between the hard saturation and quasi-saturation regions of operation.

Voltage rise time (t_{rv}): Collector-emitter voltage increase. Note that the freewheeling diode cannot start to conduct until it is forward biased, i.e. the collector-emitter voltage must rise to $v_{CE} \approx V_{dc}$ first. Also the voltage rise interval is divided into two intervals t_{rv1} and t_{rv2}. During t_{rv1} the BJT operates in quasi-saturation, and hence, the collector-emitter voltage increases only slowly (compare with t_{fv2} at turn-on). When the BJT approaches the border between the quasi-saturation and the active regions of operation the interval t_{rv1} is left and the time interval t_{rv2} is entered. During t_{rv2} the collector-emitter voltage increases more rapidly than during t_{rv1} (compare with t_{fv1} during turn-on).

Current fall time (t_{fi}): When the collector-emitter voltage has increased so high that the load current can commutate to the freewheeling diode ($v_{CE} \approx V_{dc}$), t_{rv2} is finished and the collector current can start to decrease to zero.

In the case of an abrupt transition from positive to negative base current, the time intervals t_s and t_{rv} will be shorter and a considerably larger amount of the excess carriers stored in the base have been removed after these time intervals, compared to the case with a controlled base current time derivative. On the other hand, the amount of excess carriers stored in the drift region is in principle unaltered since these are essentially removed by the collector current.

The problem is that it takes much shorter time to reverse bias the base-emitter junction but during this time only a fraction of the excess carriers stored in the drift region are removed. This implies that the only way to remove the remaining excess carriers is by internal recombination, which is a comparably slow process, and by the negative base current equal to the level of the positive collector current, which is also a slow process. The result is a long collector current tail, which could increase the turn-off losses.

Breakdown voltages

When the BJT operates in the forward blocking state, the collector-base junction must sustain the applied voltage. Note that the BJT cannot block collector-emitter voltages of negative polarity, i.e. it has no reverse blocking capability, since the base-emitter junction cannot sustain substantial voltage since the doping levels on both sides are moderate and high, respectively. The low doping of the drift region implies that in forward blocking the depletion region mainly extends into the drift region.

The depletion region in the forward blocking state must not extend so it reaches the emitter region, which is termed reach through. If it does, electrons from the emitter are swept into the base by the strong electric field in the depletion region and thereby causes avalanche breakdown in the collector-base junction. To make sure that this does not occur, the base region is rather heavily doped compared to the drift region and the thickness of the base region is made somewhat larger than the minimum required. This of course reduces the current gain slightly.

Secondary breakdown

Many of the failures of bipolar devices, especially BJT:s, are due to a failure mode termed secondary breakdown, which is not due to impact ionization as avalanche breakdown. The failure appears as a sudden reduction in the collector-emitter voltage at operation in saturation with high collector current. As the collector-emitter voltage decreases the collector current increases, often by large amounts, thereby causing higher conduction losses. What makes the situation extra complicated is that secondary breakdown often occurs in local parts of the power BJT, resulting in a very high temperature locally in the device.

There are several reasons why the BJT is prone to secondary breakdown. Firstly, the drift region of bipolar devices, i.e. semiconductor components where minority carriers transport the current through the doping layers, exhibits a negative temperature coefficient of the conduction resistivity, which means that the resistance decreases as the temperature increases since the probability for recombination decreases with increasing temperature and since the thermal equilibrium concentration of electron-hole-pairs increases.

If the power dissipation increases more than linear with the temperature, an unstable condition has been reached. This condition is termed thermal runaway. The risk of thermal runaway of course increases if the current density in the interior of the device is un-uniform, for example due to emitter current crowding.

Safe operating area (SOA)

To visualize simultaneous maximum voltage and current levels for a power semiconductor in a feasible way a sketch called Safe Operating Area (SOA) is often used. Usually the SOA is specified in forward biased conditions, FBSOA (Forward Bias SOA), and sometimes also in reverse biased conditions, RBSOA (Reverse Bias SOA). FBSOA and RBSOA for a BJT define SOA with the base-emitter junction forward and reverse biased, respectively. For a BJT it is simultaneously applied maximum levels of collector current and collector-emitter voltage that defines the SOA. I_{CM} is the maximum collector current the BJT is allowed to carry (including pulses of short duration), and is limited by bonding wires and metallization's on the wafer. The thermal limit of the transistor is determined by the thermal resistance from the wafer to the copper base-plate (the plate which is mounted on the heat sink) and defines the amount of losses allowed that can be removed without over-heating the device (usually <125 °C for silicon semiconductors). The boundaries for primary and secondary breakdown, provides information on stress levels that should not be exceeded to avoid avalanche breakdown in the collector-base junction and thermal runaway in the base-emitter junction, respectively. Switching transients are assumed to be fast enough so that the main concern is to avoid exceeding I_{CM} and BV_{CEO}, resulting in a rectangular SOA.

4.4 The thyristor

The thyristor is in its design the most simple of the controllable power electronic semiconductor devices. The thyristor is also the most simple in operation of the controllable power electronic semiconductor devices, since the thyristor only exhibit controlled turn-on. Turn-off is not controlled and only occurs when the current through the device is forced to decrease to zero by the external circuitry.

In spite of this obvious disadvantage, the thyristor is a valuable device and perhaps also the power electronic device that is most difficult to replace with another type. The straightforward and un-complicated design namely gives the possibility to manufacture thyristors withstanding very high current (several kA) and sustaining very high blocking voltages (5-10 kV) even though the forward voltage drop in the conducting state is only low (a few Volts). There are other types of controllable devices with such high ratings, for example high power IGBT:s (6 kA, 6 kV). However, it should be kept in mind that such IGBT:s consist of several IGBT chips with lower ratings, connected in series and/or parallel but packaged in a single module. A thyristor of such rating is fabricated from a single wafer. For a high power IGBT packaging is important from a reliability point of view. If the preferred failure mode is open circuit, e.g. in the case of parallel connection, then the IGBT chips in the module should be connected by bonding wires. If short circuit is the preferred failure mode, e.g. in the case of series connection, the chips should be connected in a press-pack housing.

Doping structure

A typical doping structure of a thyristors is shown in Figure 4.11. The doping structure is vertically oriented, similar to most power electronic semiconductor devices, so that the current density is acceptable even though the anode current is high, which otherwise could result in a very high silicon temperature. A common power thyristor is 100 mm in diameter but they can be as large as 150 – 200 mm, which means that an entire wafer is used to manufacture a single device. Note that the drift region (Figure 4.11) should absorb the depletion region in both forward and reverse blocking states of operation. Figure 4.11 also shows a schematic symbol for the thyristor.

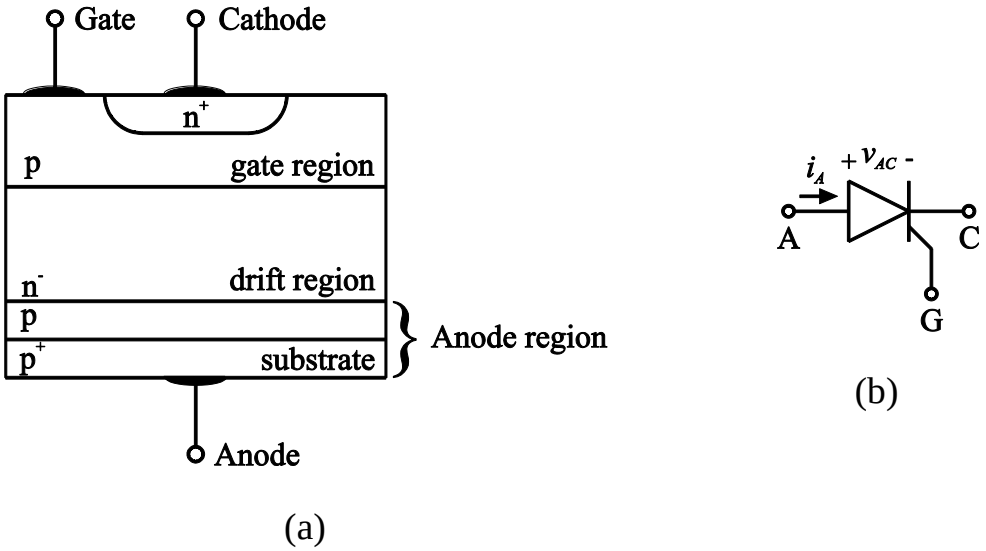


Figure 4.11 Typical doping structure and circuit symbol of a thyristor (b).

The thyristor is a three terminal device (Figure 4.11) where the terminals are termed anode (A), cathode (C) and gate (G). As explained in a later section the thyristor is turned on (trigged), by applying a short current pulse to the gate. This current pulse is the result of a voltage applied between gate and cathode. Still, the device is often termed as current controlled like the BJT, whereas the MOSFET and the IGBT are referred to as charge controlled. The gate metallization is designed in different ways depending on the geometrical size of the thyristor. The reason for this is that it is desirable that an as large portion as possible of the thyristor is forced to trig directly at turn-on, otherwise the local current density in the vicinity of the gate metallization could be very high with high losses occurring.

Stationary characteristic

The current-voltage characteristic of a typical thyristor, shown in Figure 4.12, is defined by the anode current as a function of the anode-cathode voltage. In the reverse blocking state the thyristor characteristic is similar to the one of a diode, with a low reverse saturation current to the voltage level where avalanche breakdown occurs. The voltage level where this occurs is termed BV_{BD} in some literature. Another term, more frequently used is V_{RWM} (Reverse Working Maximum). In the forward direction, the thyristor has two stable operating conditions, forward blocking and forward conducting. Between these two stable states there is one unstable, which appears as a negative resistance in the current-voltage characteristic (since the slope is negative in this region). An important parameter for the thyristor is the break-over voltage V_{BO}, which is defined for an open circuited gate, i.e. with the gate current equal to zero. If the anode-cathode voltage becomes higher than V_{BO} in the forward blocking state then the thyrsitor will by itself perform a transition to the conducting state, even if no gate current is applied. If a gate current pulse is applied, the transition occurs at anode-cathode voltage levels lower than V_{BO}, which is also show in the current- voltage characteristic (Figure 4.12). If only the gate current pulse is high enough this transition will take place for very low forward voltage. Other important parameters are the holding current (I_H) and the holding anode-cathode voltage (V_H). The holding current is the lowest anode current that maintains the thyristor in the conducting state, and the holding voltage is the corresponding anode-cathode voltage. As mentioned earlier, the gate-current do not need to be a continuous DC current; a short pulse so that the thyristor is trigged is sufficient. This is due to the fact that the thyristor is a latching device, a fact to be discussed in a later section. However, it is important to understand that the gate cannot be used to turn off the device. The reason is that if a negative gate current is applied, the thyristor will only turn-off locally, i.e. in the vicinity of the gate metallization. For traction applications, e.g. trains, commutation or turn-

off circuits have been developed, so that the anode current is redirected and the anode decreases to zero and the device turns off.

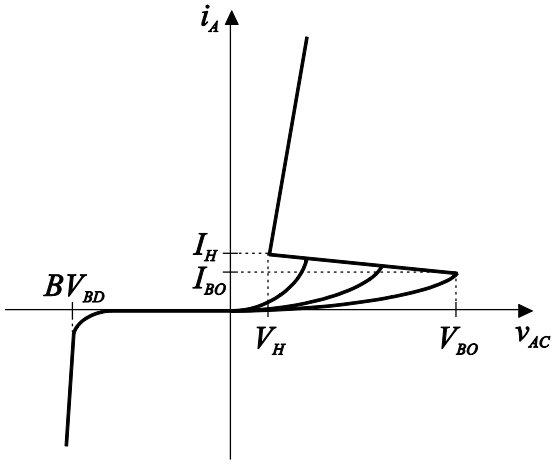


Figure 4.12 Typical output characteristic (anode current versus anode-cathode voltage for different gate currents) of a thyristor.

Operation

To understand the operation of the thyristor, a simplified doping structure is assumed, see Figure 4.13. This simplified doping structure clearly shows that the thyristor can be seen as combination between one pnp- and one npn-structure, which is used to create an electronic equivalent consisting of two BJT:s, see Figure 4.13. The three pn-junctions J_1, J_2 and J_3 are marked in Figure 4.13. Note that junction J_2 is common for the pnp- and npn-structures.

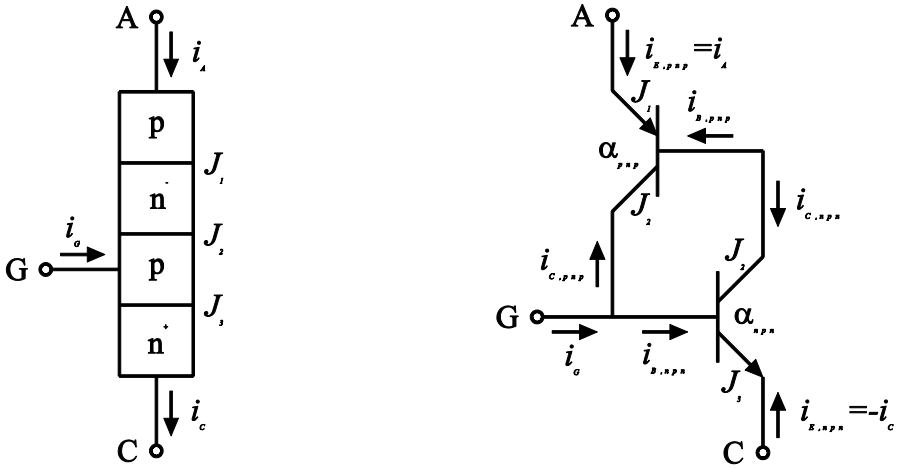


Figure 4.13 Simplified doping structure (a) and equivalent circuit, (b) of a thyristor.

When the thyristor operates in reverse blocking condition, the junctions J_1 and J_3 are both reverse biased. The low-doped drift region absorbs most of the depletion region from junction J_1. Since the p-region at J_3 has lower doping than

the n-region, the p-region absorbs most of the depletion region of the reverse biased junction J_3. However, most of the blocking voltage appears across J_1, so most important is that the drift region is of low doping and thick enough to accommodate the depletion region and withstand the reverse voltage without the occurrence of avalanche breakdown.

When the thyristor operates in the forward blocking state junctions J_1 and J_3 are forward biased and junction J_2 is reverse biased. Therefore, the drift region accommodates the depletion region also in this case. Figure 4.13 is used understand the transition from the forward blocking state to forward conducting state for the thyristor. According to Figure 4.13 the collector currents are written

$$\begin{cases} I_{C,pnp} = -\alpha_{pnp} \cdot I_{E,pnp} + I_{C0,pnp} \\ I_{C,npn} = -\alpha_{npn} \cdot I_{E,npn} + I_{C0,npn} \end{cases} \tag{4.21}$$

where $I_{C0,\,pnp}$ and $I_{C0,\,npn}$ are the collector leakage currents and α_{pnp} and α_{npn} are the base transport factors for the pnp- and npn-structures, respectively. Kirchoff's current law for the thyristor and the npn-structure gives

$$\begin{cases} I_{E,pnp} + I_{E,npn} + I_G = I_A + I_K + I_G = 0 \\ I_{E,npn} + I_{C,npn} + I_{B,npn} = 0 \end{cases} \tag{4.22}$$

Together, the expressions above give the collector currents

$$\begin{cases} I_{C,pnp} = -\alpha_{pnp} \cdot I_A + I_{C0,pnp} = I_G - I_{B,npn} \\ I_{C,npn} = \alpha_{npn} \cdot (I_G + I_A) + I_{C0,npn} = I_G + I_A - I_{B,npn} \end{cases} \tag{4.23}$$

Combining these two expressions gives

$$\begin{aligned} &-\alpha_{pnp} \cdot I_A + I_{C0,pnp} - \alpha_{npn} \cdot I_G - \alpha_{npn} \cdot I_A - I_{C0,npn} = -I_A \quad \Rightarrow \\ &(1 - \alpha_{pnp} - \alpha_{npn}) \cdot I_A = \alpha_{npn} \cdot I_G - I_{C0,pnp} + I_{C0,npn} \end{aligned} \tag{4.24}$$

The anode current I_A is thus written

$$I_A = \frac{\alpha_{npn}}{(1 - \alpha_{pnp} - \alpha_{npn})} \cdot I_G + \frac{I_{C0,npn} - I_{C0,pnp}}{(1 - \alpha_{pnp} - \alpha_{npn})} \tag{4.25}$$

This means that the trig condition for the thyristor in the case of zero gate current $I_G = 0$, i.e. break over, is given by

$$\alpha_{pnp} + \alpha_{npn} = 1 \tag{4.26}$$

From the section on bipolar transistors it is known that the base transport factor is dependent on the effective base width (thickness). For the thyristor operating in forward blocking state the blocking voltage appears across junction J_2 and the depletion region extends mainly into the low doped drift region, i.e. the base of the pnp-structure.

An increasing blocking voltage implies that the depletion region also expands and thus the effective base width of the pnp-structure decreases thereby increasing its base transport factor α_{pnp}. Note that the depletion region expands also into the gate region (p-doping), i.e. the base of the npn-structure. This means that also the base transport factor of the npn-structure (α_{npn}) increases, with increasing forward blocking voltage. When the blocking voltage equals the break over voltage V_{BO} the base transport factors has increased so that their sum equals one and the thyristor passes transits into the forward conducting state.

If a positive gate current is applied, electrons are injected across the forward biased junction J_3 into the base region of the npn-structure termed, i.e. the p-base. The electrons diffuse across the junction J_2 into the base region of the pnp-structure (the drift region), i.e. the n-base. The electrons in the drift region affect the trig condition of the thyristor in two ways. Firstly, the depletion region in the drift region will grow thicker since space charge neutrality requires that an increased number of donor atoms are ionized to compensate the increased amount of negative charge imposed by the electrons in the p-base. Secondly, the electrons injected into the drift region attract holes from the anode region, which in turn attract more electrons from the cathode region, etc.

This means that the thyristor is trigged at anode-cathode voltage lower than the break-over level, V_{BO}, when $I_G > 0$. The regenerative process also provides the latching capability of the thyristor, i.e. only a gate current pulse is needed to trig the thyristor then the minority carrier injection is sufficiently strong in all layers to maintain the device in conducting state. Turn-off occurs, as previously mentioned, when the anode current decreases to zero due to the external circuit. The minority carriers leave the layers both due to recombination and carrier sweep out as a result of the negative current resulting from the reverse bias of the thyristor. This negative current is not blocked initially since the pn-junctions are internally forward biased, and thus have no blocking ability, as long as there are excess minority carriers on either side of the pn-junctions.

Switching

Since the thyristor often is a device of large geometrical size, the duration of the turn-on interval is long. There is a potential risk that the region in the vicinity of the gate latches rapidly as a consequence of the gate current pulse but with a delay in the turn-on of the remote areas of the thyristor. The minority carrier

injection farther away from the gate metallization is delayed since the regenerative process must first be initiated before it can spread to the periphery of the device. In some literature the minority carrier distribution is referred to as plasma, spreading from the region between the gate and anode just beneath the gate metallization. The time delay is consequently termed plasma spreading time, t_{ps}.

Due to the limited plasma spreading time geometrically large thyristors should have a gate that turn-on an as large portion of the thyristor as possible, i.e. a distributed gate. Furthermore, it is often important to limit the time derivative of the anode current at turn-on for all thyristor types. This could be achieved with a snubber circuit limiting the current derivative, i.e. with an inductive snubber. The principal switching waveforms of a thyristor are shown in Figure 4.14, where it is assumed that the external circuit limits the anode current to the level I_0. It is also assumed that the blocking voltage in the forward direction is V_F (v_{AC} positive) and in the reverse direction V_R (v_{AC} negative).

Turn-on

Assume that the thyristor operates in the forward blocking state. The turn-on transient is sketched in the first part of Figure 4.14, where the three exaggeratedly distinct time intervals are shown.

Turn on delay time ($t_{d(on)}$): The turn-on transient starts when a gate current pulse is applied but the anode current and anode-cathode voltage remain unaltered. This due to the fact that excess carriers are being injected into the gate region during this time interval, which results in that the sum of the base transport factors (n- and p-base), $\alpha_{pnp}+ \alpha_{npn}$, equals one when a sufficiently large amount of minority carriers have been injected and the device transits into the conduction state. The time it takes to inject these excess carriers determines the turn-on delay time.

Current rise time (t_{ri}): The duration of this time interval is completely determined by the external circuit in practical applications. The current rise time of the thyristor itself is so short that if its time derivative is not limited only a fraction of the device is conducting when the device carries the full load current. This local turn-on at full load current implies that the current density is very high and results in high losses. In Figure 4.14 it is also shown that the anode-cathode voltage decreases in this time interval. This is due to the larger portion of the cross-section of the thyristor gradually turned on reduces the resistance of the device.

Plasma spreading time (t_{ps}): When the anode-current has reached its stationary level, corresponding to the load current, the excess carrier distribution continues

to spread to the parts of the thyristor that is not yet in the conducting state. This may take substantial time, which implies that forward voltage drop decreases slowly. This time interval is finished when the entire device is operating in the conducting state, and thus, the forward voltage drop has decreased to its stationary level.

Remark: To partly circumvent the problem that the thyristor is only trigged in the vicinity of the gate-metallization, a gate current pulse of an initially elevated level should be supplied (Figure 4.15). The initial peak causes that a larger part of the thyristor is turned on during the current rise interval, resulting in reduced device stress in terms of local current density. Note that the gate current pulse must have a profile that ensures a complete turn-on sequence, i.e. it is important that the gate current does not decrease to a too low level too early.

Turn-off

The turn-off waveform of a thyristor, Figure 4.14, is principally similar to the one of a power diode. The distinct intervals appearing for thyristor turn-off are introduced in the following.

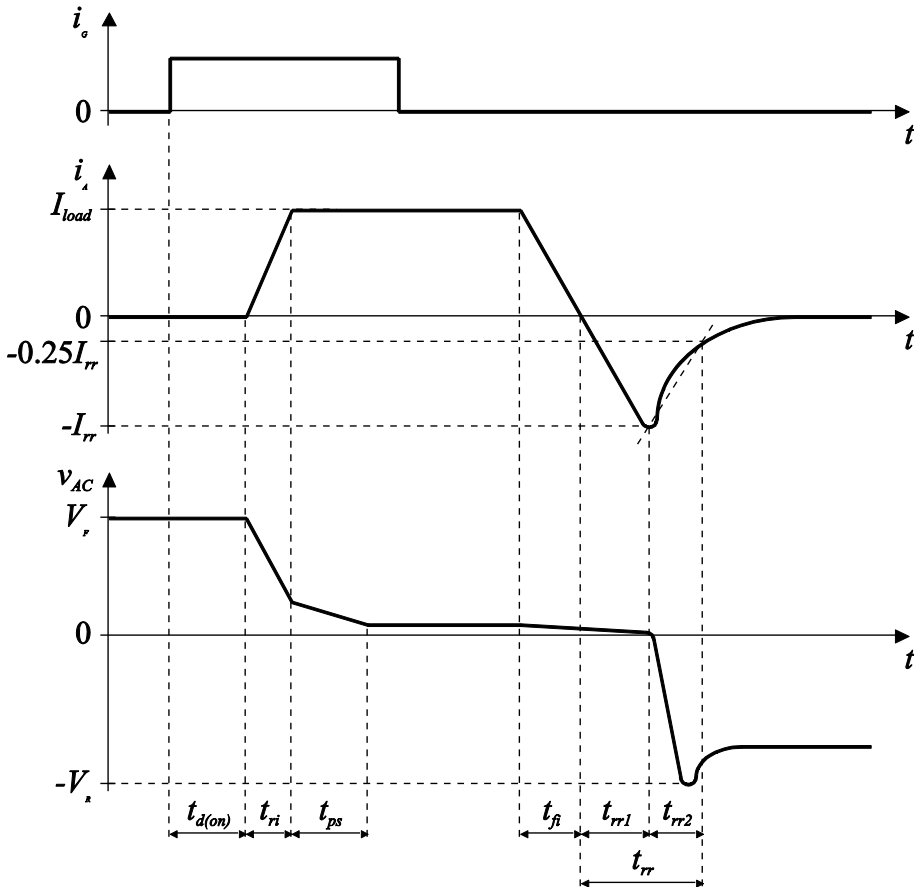


Figure 4.14 *Switching waveforms for a typical thyristor. The left part shows turn-on and the right part turn-off. Note that the relative duration of the time intervals is not consistent with the ones presented by real devices.*

Current fall time (t_{fi}): The anode current decreases with a time derivative determined by the external circuit. The excess carrier concentration decreases during this interval. However, note that neither pn-junctions J_1 nor J_3 are reverse biased.

Reverse recovery time (t_{rr}). During t_{rr} the anode current is negative, first with a negative time derivative (t_{rr1}) and thereafter with a positive time derivative (t_{rr2}). The negative anode current is not blocked during t_{rr1} since neither J_1 nor J_3 are reversing biased. Because of the same reason, the anode-cathode voltage does not decrease. As soon as one of the pn-junctions J_1 or J_3 become reverse biased, which occurs when the excess minority carriers on both sides of a junction are removed, the junction starts to block and the negative anode current starts to increase towards zero. Since one of the pn-junctions is reverse biased and can accommodate the depletion region, the anode-cathode voltage can decrease to a negative level.

As a consequence of the magnitude of the time derivative of the anode current, the anode-cathode voltage will exhibit a voltage overshoot in the same way as a power diode. When the anode-cathode voltage has reached its stationary level and the reverse recovery current has decreased to a level lower 25% of its maximum, I_{rr}, the thyristor is regarded as operating in the blocking state. This is not entirely true, since it is not sure that both pn-junctions J_1 and J_3 are reverse biased. Either junction may have excess carriers on both the p- and n-sides. As long as the thyristor is reversing biased this does not cause any problem.

However, if the thyristor becomes forward biased too early or with a too high voltage derivative, the remaining excess carriers are swept out and redistributed in such a way that the trig condition becomes satisfied. This results in that the thyristor turns on unintentionally. Therefore, a certain minimum off-time (turn-off time t_q) is specified for thyristors. Furthermore, the maximum allowable voltage time derivative for the transition from the reverse to the forward blocking state is also specified.

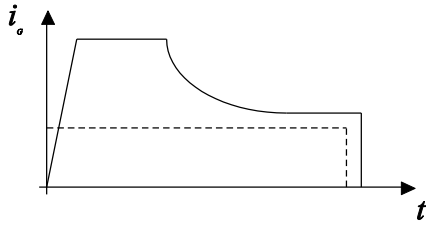


Figure 4.15 Typical gate current turn-on pulse for practical applications (solid) and minimum gate current pulse (dashed).

4.5 The gate turn-off (GTO) thyristor

The thyristor can block the highest voltages and carry the highest current among all controllable power electronic semiconductor devices. Even though the anode current may be as high as several kilo Amperes the forward voltage drop is only a few Volts, which means that in many applications the thyristor is the only controllable semiconductor that can be used. The major drawback of the thyristor is of course that it does not have turn-off capability. Therefore, it cannot be used, as is, in DC-DC conversion applications for example in traction, since the current do not pass through zero at any time instant.

The thyristor can still be used in such application but commutation circuits must be included. The commutation circuit provides an alternative path for the anode current so it can decrease to zero and a commutation is achieved. Commutation circuits are rather complex and are not discussed further. Sufficient to say is that the commutation circuit is a resonant circuit where the oscillation is initiated by other power semiconductors. The oscillation generates a current opposing the anode-current thereby forcing the thyristor into turn-off and eventually into the reverse blocking state.

However, it should be clear that the thyristors main application is in DC-AC and AC-DC conversion applications, where commutation circuits are not needed, for example in HVDC-applications. Note that HVDC is not a DC-DC application but merely a AC-DC-AC application.

In order to not have to rely on the complicated commutation circuits in DC-DC applications, and still have a thyristor with turn-off capability, a few minor changes are added to the original thyristor structure. In the following, the modifications made for three types of turn-off thyristors are described. These are; the GTO thyristor, the MCT and GCT thyristor. It will turn out that at least one of the modifications is generic, e.g. the gate-cathode structure is made much finer than in the thyristor case. Most turn-off thyristors do need snubbers to assist turn-on to limit the time derivative of the anode-current. In some cases a turn-off snubber (voltage clamp) is also required for satisfactory operation. The turn-on waveform is affected by the turn-on snubber operation and therefore the switching waveforms are not discussed in this brief introduction of turn-off thyristors.

The gate turn-off (GTO) thyristor

The Gate Turn-Off (GTO) thyristor was the first thyristor with turn-off capability introduced. The doping structure of the GTO thyristor is essentially the same as for a regular thyristor except for the fact that the gate-cathode structure is distributed across the entire wafer, to minimize the lateral voltage

drop in the gate region occurring at turn-off. This lateral voltage drop results in limitations for the regular thyristor in terms of maximum allowable anode current derivative and problems with current displacement. The distributed gate-cathode structure reduces these problems, which is advantageous also for the turn-on transient. The most important result of the refined gate-cathode structure is however that it facilitates that the gate-cathode junction can be reverse biased throughout the entire wafer and not only locally as for the regular thyristor. A principal doping structure for a GTO thyristor is sketched in Figure 4.16.

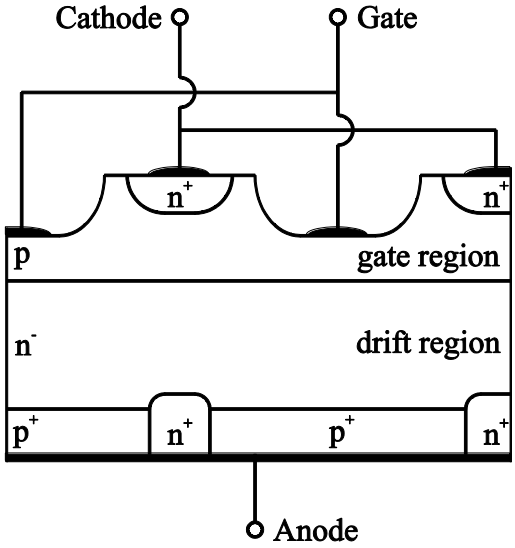


Figure 4.16 Sketch of the doping structure of a GTO thyristor. Note the distributed gate-cathode structure and the thin n-regions at the anode, referred to as anode-shorts.

The same two-transistor equivalent as used for the thyristor can be used for the GTO thyristor to calculate the turn-off current gain β_{off}. Note that this equivalent does not take that the fact that the gate-cathode structure is distributed into account, which is not needed either.

The turn-off transient starts when the gate-cathode junction is reverse biased, causing that the npn-structure of the two-transistor equivalent starts to turn-off, forcing this transistor into the active region of operation. This, in turn, causes a decrease of the base current of the pnp-structure, and therefore also a decrease of the corresponding collector current. This means that the npn-transistor is forced further into the active region of operation. This process continues until the thyristor anode-current has decreased to zero, which implies that also the turn-off of the GTO thyristor is regenerative. Note that the gate-cathode voltage must remain negative during the entire turn-off transition.

The gate-current initiating the turn-off must be negative and its magnitude is given by the turn-off current gain β_{off}. Utilizing the two-transistor equivalent from the previous section on thyristors, the magnitude of the base-current of the npn-transistor for operation in the active region is given by

$$I_{B,npn} < \frac{I_{C,npn}}{\beta_{npn}}, \beta_{npn} = \frac{\alpha_{npn}}{1-\alpha_{npn}} \tag{4.27}$$

where

$$\begin{cases} I_{B,npn} = \alpha_{pnp} \cdot I_A + I_{G,off} \\ I_{C,npn} = I_{B,pnp} = I_{E,pnp} - I_{C,pnp} = \left(1-\alpha_{pnp}\right) \cdot I_A \end{cases} \tag{4.28}$$

Together, this gives

$$I_{G,off} < -\frac{1}{\beta_{off}} \cdot I_A \quad \text{where} \quad \beta_{off} = \frac{\alpha_{npn}}{\alpha_{npn} + \alpha_{pnp} - 1} \tag{4.29}$$

From the last expression above it is obvious that the turn-off current gain becomes high if α_{npn} is close to unity and α_{pnp} is as low as possible. From the section on bipolar transistors it is clear that the p-base region (i.e. the gate-region) should be as thin as possible and that the cathode region should be as highly doped as possible, to give α_{npn} close to unity. Furthermore, the n-base (i.e. the drift region) thickness should be large and the minority carrier lifetime short to result in a low α_{pnp}.

The requirements on the gate and cathode regions are the same as on a regular BJT or thyristor. The drift region must also be thick to host the depletion region. The only conflicting requirement appearing is that the minority carrier lifetime of the drift region should be low since this implies that the resistance of the drift region will increase and that the forward voltage drop of the GTO thyristor will be higher than for a regular thyristor of the same ratings.

By adding anode-shorts, the required reduction of the minority carrier lifetime is decreased causing only a moderate increase of the forward voltage drop in the on-state. At turn-off the excess carriers (holes) stored in the drift region must be removed. In the doping structure of a regular thyristor there are p-layers on both sides of the drift region whose doping atom density I higher than the excess carrier density (density of free holes) in the drift region. This implies that the excess holes do not leave the drift region by diffusion at turn-off. Instead the excess holes are neutralized only by internal recombination, which is a comparably slow process.

If anode shorts are included in the doping structure a negative voltage sweeping out the excess holes from the drift region cannot be applied since there is no pn-junction that can block a negative anode-cathode voltage. A regular thyristor becomes reverse biased at turn-off and therefore the excess carriers are swept out by the accompanying negative current. For the GTO thyristor on the other hand, the drift-region has a highly doped n-region next to it. The excess holes

diffuse towards these anode shorts. This method is very effective in removing the excess holes, that the GTO thyristor exhibits significantly shorter forward recovery and turn-off times than a regular thyristor.

Control signal

Similar to the thyristor, the gate pulse for a GTO thyristor in terms of time derivative and amplitude is of great importance for satisfactory operation, i.e. that a large portion of the device should turn-on when trigged so that the anode current density do not become excessive. For a GTO thyristor, a low amplitude gate current must flow also after the turn-on transient to ensure that the entire device remains in the on-state also for low anode currents. This current is called termed back porch current. Otherwise, if the anode current suddenly increases, the anode current may not be shared by all cathode elements. This could cause thermal runaway and eventually device failure.

For the turn-off it is important to ensure that the negative gate current has sufficiently high amplitude so that the GTO thyristor turns off completely. A typical turn-off current gain is in the range 3-4, i.e. a substantial gate current is required. Current sharing among the cathode elements is also important in order to avoid thermal problems. Since un-uniform current density is mainly present during turn-on and turn-off transients, the manufacturer often specifies minimum on- and off-times.

The mos controlled thyristor (MCT)

Purely bipolar devices have the drawback that they are current controlled, putting strong requirements on the drive circuits. For power semiconductor devices the problem is worsened since the current gain is usually relatively low. However, the bipolar devices have advantageous current handling capabilities since they can carry high currents and still exhibit very low forward on-state voltage and consequently the on-state losses are low. Therefore, there has been a trend over the years to combine the MOS input, which is voltage controlled and therefore have moderate driver requirements, with a bipolar output monolithically, i.e. on the same wafer.

For transistors this has led to the insulated gate bipolar transistor (IGBT), which was introduced in 1984. For thyristors, the same ambition has led to the MOS controlled thyristor (MCT), which has been commercially available since 1994. The MCT has, at least to some extent, the output characteristics of a thyristor, i.e. low conduction losses and high current handling capability, and the input characteristics of a MOSFET, i.e. it is voltage controlled and exhibits short (for a thyristor) switching times. However, it should be kept in mind that the device is still a turn-off thyristor, and should only be used in applications where it

replaces a turn-off thyristor. Usually, a MCT cannot replace a transistor since the switching characteristics are not as good as the ones of a transistor.

So, the MCT is a turn-off thyristor. From a functional point of view it is equipped with one on-FET for turn-on and one off-FET for turn-off. In reality, the MCT has hundreds of cells connected in parallel connected cells; where each cell contains an off-FET but only one cell of 15-20 cells contain an on-FET. Two types of MCTs are manufactured; P-MCTs and N-MCTs. The P-MCT has an on-FET p-channel type and an off-FET of n-channel type. The N-MCT has an on-FET of n-channel type and an OFF-FET of p-channel type.

The first MCTs that were commercially available were P-MCTs, and the description of the operation of the MCT will mainly treat this type. Figure 4.17 shows the doping structure of a P-MCT cell containing both on-FET and off-FET, together with an electrical equivalent used to simplify the analysis.

The channel of the off-FET (n-channel, i.e. p-doped region) is located most nearby the anode. Farther away from the cell periphery the channel of the on-FET (p-channel, i.e. n-doped region) is located. Since the off-FET is of n-channel type, there must be an n-type region closest to the anode so that the n-channel connects two n-type regions. Of course, the typical thyristor pnpn-structure is also found in the doping structure of the MCT. An equivalent circuit corresponding to the two-transistor equivalent is also shown in Figure 4.17. The lightly doped p-region next to the cathode serves as drift region and should as such host the depletion region in the blocking state.

For a P-MCT the drift region (p-type) is the base of the npn-structure of the thyristor. Since this layer has substantial thickness, the npn-structure will possess a low base transport coefficient, compared to the one of the pnp-structure. To turn off the MCT the sum of these two base transport coefficients must be considerably less than unity, which the sum equals when latching.

Since the npn-structure already has a low base transport factor, the most significant reduction of the sum of the base transport factors is achieved by reducing the base transport factor of the pnp-structure. Therefore, the off-FET is located on the anode side of the thyristor so that the base-emitter junction of the npn-structure is short circuited for thyristor turn-off. This yields that the anode current commutates to the off-FET and the pnp-structure starts to turn off. The current gain of the pnp-structure falls to such a low level that the latching condition is not fulfilled and thereby the thyristor turns off in the same manner as the GTO.

Note that the discussion above holds only if the forward voltage drop across the off-FET is below the forward voltage drop of the base-emitter of the pnp-structure operating in the active region for the particular anode current level.

This is one of the reasons why a p-MCT was introduced first. For a p-MCT the off-FET is of n-channel type and the mobility of electrons is approximately three times higher than that of holes and therefore n-channel off-FET will have on-state resistance equal to one third of that of a p-channel off-FET for an n-MCT. This also implies that a p-MCT can turn off a current that is three times that of an n-MCT of comparable geometrical size.

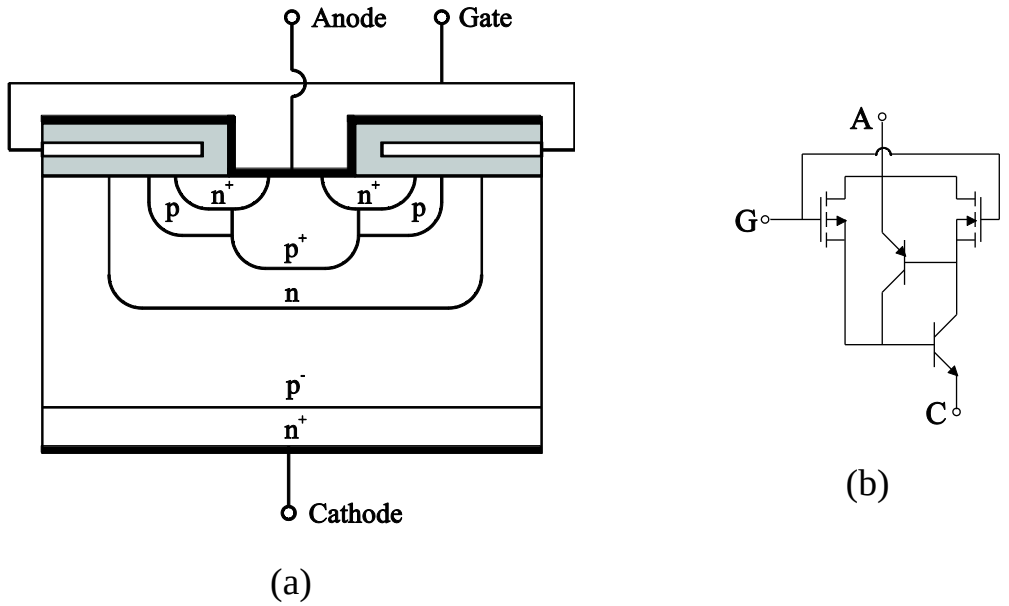


Figure 4.17 The principal doping structure of a p-MCT (a) and electrical equivalent of a p-MCT corresponds to the two-transistor equivalent of a thyristor (b).

The gate commutated turn-off (GCT) thyristor

The GCT is a rather new type off thyristor with turn-off capability introduced in 1997. When first designed the intention was to develop a device that could be used in application where high power GTOs were used previously. After research and development efforts it turned out to be suitable for medium power applications where both GTOs and IGBTs are also used today. The main problem of GTOs in such applications are the need of turn-off snubbers and that low switching frequency must be applied since displacement currents appear in the interior of the device, resulting in local heating. The main problem associated with IGBTs in such applications is that geometrically large IGBT chips cannot be manufactured due to the complicated gate-emitter structures. Medium and high power IGBTs are instead manufactured as modules where several IGBT chips are series and parallel connected through bonding wires to obtain a certain forward blocking voltage level and current handling capability. The numerous bonding wires causes problems in intermittent operation like traction applications with repeated acceleration and braking (for example in trains) causing a higher silicon temperature than expected due to the thermal capacity between the IGBT chips and the heat sink. The thermal cycling caused by the repeated acceleration and braking sequences causes stress on the

soldering connecting bonding wires to the semiconductor chips that could lead to malfunction, like bonding heel-crack or lift-off. This implies that the expected life-time of the IGBTs in such applications is reduced in some cases to a large extent. The GCT is better suited than the IGBT for thermal cycling since the device is fabricated from a single chip and mounted in a press-pack which implies that it can be manufactured without need of bonding wires. The GCT is expected to replace the IGBT in traction applications for voltages exceeding 1.5 kV.

The main difference between the GCT and the regular thyristor is that the GCT has a distributed gate-cathode structure like the GTO thyristor and the MCT, which means that the gate-cathode structure is distributed across the entire surface on one of the sides of the wafer. Even if the GCT thyristor usually is not manufactured with anode-shorts, it can be treated as a GTO thyristor in the analysis, as indicated below. From experiments on the GTO thyristor it was found that if the anode current was forced to commutate from the cathode to the gate before a depletion region was established between the gate and drift-regions, the thyristor could be turned off in a way similar to a bipolar transistor. This implies that by turning off the GTO thyristor like a three-layer structure (BJT) instead of a four-layer structure (thyristor), without affecting the excess carrier distribution in the three remaining regions, the current displacement at turn-off could be reduced significantly.

The results from these experiments were successful also at rated current and voltage for the GTO thyristor. Therefore, turn-off snubber is not needed for GTO thyristors controlled this way. The turn-off current gain is however compromised since the entire anode current commutates to the gate at turn-off, yielding a current gain equal to 1. From the discussion above it is obvious that it is important that the anode current commutates to the gate rapidly for the concept to be well functioning and therefore either a high gate-cathode voltage must be applied or the gate inductance must be low, to obtain a gate current time derivative that is high enough.

Regular GTO thyristors are in most cases controlled via coaxial cable and the gate contact is in most cases geometrically small, both resulting in a high inductance of the gate circuit. To avoid this problem the coaxial cable was exchanged for a flat cable, which has a considerably lower inductance in this case, but anyway, required a control voltage as high as 200 V.

The next step in the transformation of a GTO into a GCT thyristor was to investigate in what way the housing should be changed to further decrease the inductance of the gate circuit. The result was a packaging technique with a coaxial housing which means that the gate terminal extent throughout the entire periphery of the housing. Therefore, the next modification was to replace the flat

cable with a printed circuit board connecting to the gate terminal around the housing. The GCT thyristor driver was mounted on this printed circuit board. These improvements were so advantageous that a 15 V supply voltage is sufficient. This device, with the GCT and the diver integrated is termed IGCT (Integrated Gate Commutated Turn-off) thyristor.

The improvements given above were sufficient to create a new device, the IGCT, but when completed the doping structure was also to be further optimized. The GCT thyristor is intended for bridge applications, and as such, it does not need to have reversed blocking capability due to the anti-parallel freewheeling diode. Therefore, a buffer region was introduced in the doping structure, which in effect means that the thickness of the drift region could be halved. This also means that a freewheeling diode can be easily integrated monolithically since the thicknesses required for the different doping regions of the GCT thyristor agree well with the ones suitable for a freewheeling diode. The diode is a natural punch-through device, due to the p-i-n structure applicable for devices that only block voltage of negative polarity. The next step in device optimization was to introduce a transparent anode structure instead of anode shorts. A transparent anode is a pn-junction with current dependent emitter efficiency. For low currents, the emitter efficiency is very high, implying that the gate turn-on current and the hold current (i.e. the back-porch current) are low. For high current densities, the emitter efficiency is low which implies that the excess carriers are removed as efficiently as in the case of anode shorts, for turn-off.

The principal doping structure of the GCT with an integrated freewheeling diode is shown in Figure 4.18.

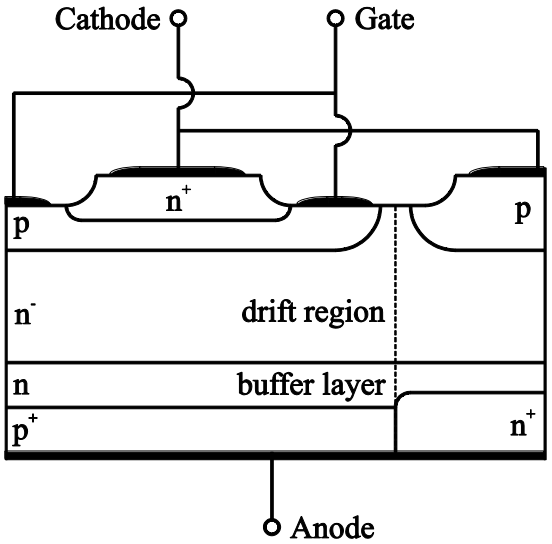


Figure 4.18 The principal doping structure of a GCT thyristor with integrated freewheeling diode.

4.6 The MOSFET

The MOSFET for power electronic applications was introduced in the early 1980s, and was expected a glorious future due to the short switching times. However, also in present power electronic applications the MOSFET is only employed in low power application or at least for low blocking voltage levels (< 600 V usually).

Doping structure

The power MOSFET is in most cases based on a vertically oriented doping structure consisting of four layers. Power MOSFETs always consists of several thousand parallel connected cells to reduce the on-state resistance. In Figure 4.19 one such cell is shown. Note that the relative thickness of the doping layers is not correct in Figure 4.19 since the drift-region is considerably thicker than the other regions of a practical MOSFET. With n-doped drain and source regions, the MOSFET is termed to be of n-channel type and with these regions p-doped instead the device is of p-channel type. Small-signal MOSFETs are available as both enhancement-mode (normally-on) and depletion-mode (normally-off) devices. The power MOSFET is however in principal always fabricated to be operating in depletion-mode since the device should turn-off in the case of driver failure.

The p-doped region adjacent to the source-region is termed body-region and it is in this region the channel is formed when the MOSFET is operating in the on-state. The drift-region, which is not included for a small-signal MOSFET, determines the blocking voltage capability as for other types of power semiconductors. From the principal doping structure it is evident that the gate metallization is separated from the body region by a thin oxide layer so that there is a region between the gate- and the source-regions with very low conductivity. It is also evident that independent of the polarity of the gate-source voltage polarity at least one pn-junction is always reverse biased, which in effect means that current transport in a MOSFET cannot be due to minority carrier injection. Instead, a positive gate-source voltage affects the body-region just beneath the gate-oxide in such a way that an n-channel is created connecting the source- and drain-regions.

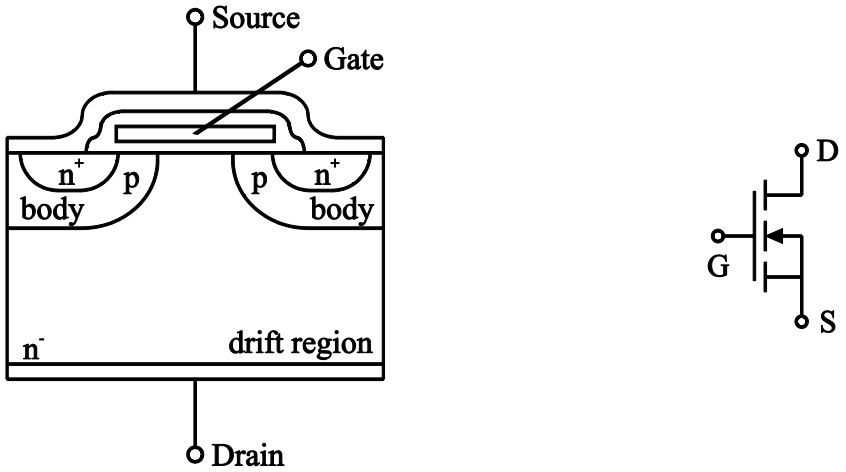


Figure 4.19 Principal doping structure of one cell of a power MOSFET (Left) and the circuit symbol of an n-channel MOSFET (right).

The thickness of the gate-oxide, the width of the gate and the number of gate-source regions (cells) that are parallel connected are all important factors determining the magnitude of the drain-current for a certain gate-source voltage.

The source consists of several (thousands of) circular or polygon shaped cells connected in parallel. The geometrical shape determines, to some extent, the on-state resistance and thereby the on-state losses of the MOSFET. The geometrical shape of the source regions is even used in the marketing of some devices for example the International Rectifier MOSFET family HEXFET with hexagon shaped cells.

The main reason for having many parallel connected source-regions is that it minimizes the current density and the resistance of the channel, since the width-to-length ratio becomes large. The requirement of low on-state resistance also results in that n-channel devices are much more frequently used than p-channel MOSFETs since the mobility of electrons is three times higher than that of holes, which implies that the resistance of an n-channel is lower than the resistance of a comparable p-channel at the same doping level of the silicon.

An important observation from the principal doping structure is that there is a parasitic npn-structure between the source and drain with the body serving as base. This parasitic BJT-structure must not be turned on since the base is not available from any terminal and therefore the structure is not possible to turn off. To counteract that the npn-structure is biased the source-metallization is extended to cover also the body-region to short circuit the base-emitter junction of the parasitic BJT. Instead, a parasitic pn-junction is formed (Figure 4.20) which may be used as a freewheeling diode in bridge applications.

The gate-metallization extends over the drift-region for two reasons. Firstly, the gate-metallization increases the conductivity of the drift-region in the on-state

since an accumulation layer of free carriers is formed beneath the gate. Secondly, the gate-metallization acts as a field plate in the blocking-state, which means that the radius of the depletion layer in the drift-region becomes smoother, decreasing the maximum electric field strength and thus increases the breakdown voltage.

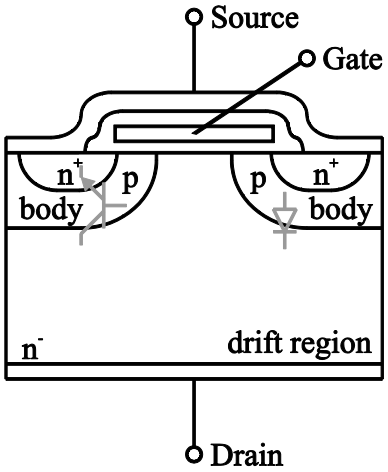


Figure 4.20 Parasitic elements in the doping structure of a power MOSFET.

Current-voltage ccharacteristic

The MOSFET is a semiconductor with three terminals where the terminal termed source usually is common for the input and output at least in power applications. The terminal named gate is the input and the one termed drain is the output. In the active region, the voltage applied over gate-source determines the magnitude of the drain current. Therefore, the output characteristic consists of curves showing the drain current and the drain-source voltage for different gate-source voltages. The output characteristic is divided in three distinct regions as shown in Figure 4.21. The regions are:

Cut-off. The gate-source voltage is lower than the threshold level $V_{GS(th)}$, which is typically 3-5 V. This means that the output forms an open circuit and that the transistor has to block the applied drain-source voltage which must be lower than the breakdown voltage $V_{DS(max)}$. A higher drain-source voltage results in avalanche breakdown in the drain-body junction.

The active region. The gate-source voltage is higher than the threshold level. The drain-current is determined by the applied gate-source voltage and independent of the drain-source voltage. In the active region the drain current is proportional to the square of the difference between the applied gate-source voltage and its threshold level. Note that in the literature the active region is sometimes referred to as the saturated region which in this case means that the drain current does not increase with the drain-source voltage. This should be

avoided since it may result in misunderstanding since it does not correspond to the region where a BJT is said to operate in saturation.

The ohmic region. When such a high gate-source voltage is applied so that the drain current does not saturate the MOSFET is said to operate in the ohmic region. Comparing to a power BJT this region corresponds to the on-state (or saturation). The voltage drop across drain-source is proportional to the drain current, i.e. resistive. The resistance is determined by the geometrical size and the conductivity of the drift region. The resistance is termed $R_{DS(on)}$. The border between the active region and the ohmic region is given by V_{GS}-$V_{GS(th)}$=V_{DS}.

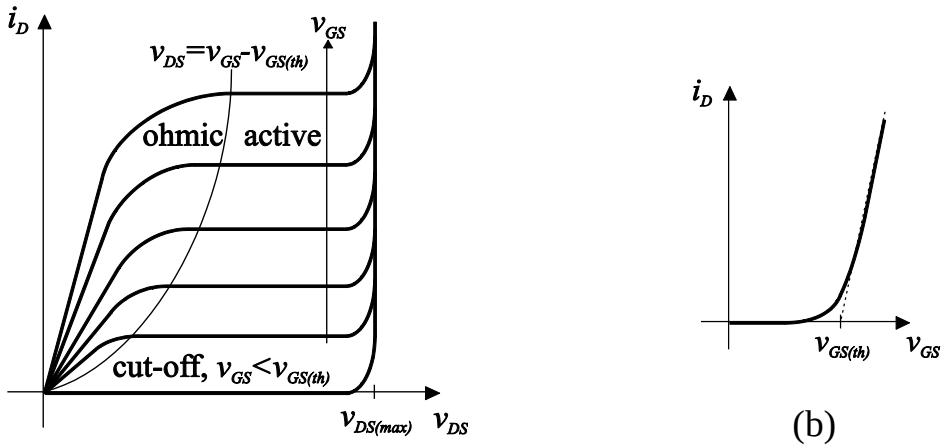
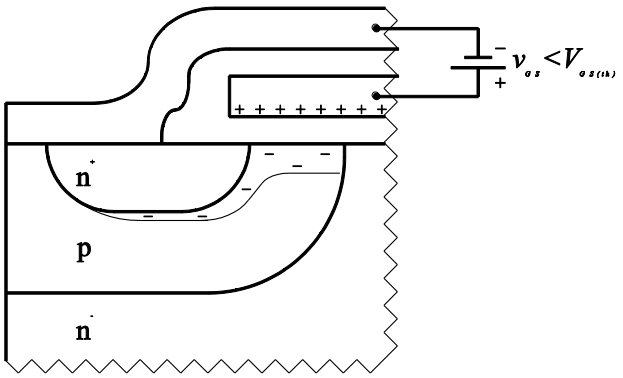


Figure 4.21 Typical output characteristic (a) and transfer characteristic (b) of a power MOSFET.

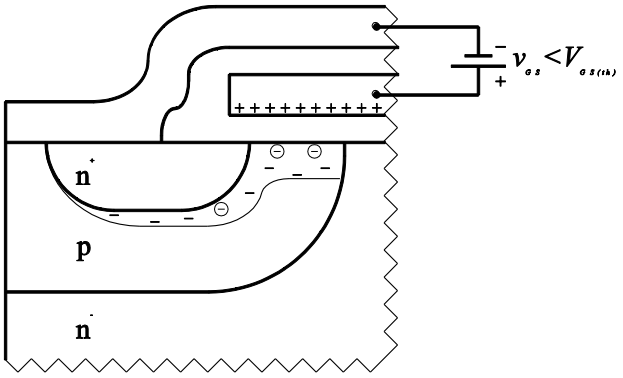
The typical transfer characteristic of a power MOSFET is shown in Figure 4.21. The transfer characteristic describes the minimum level of the drain current as a function of the gate-source voltage for operation in the active region. Drain current, lower than the one depicted by the transfer characteristic, will drive the MOSFET into the ohmic region.

The channel

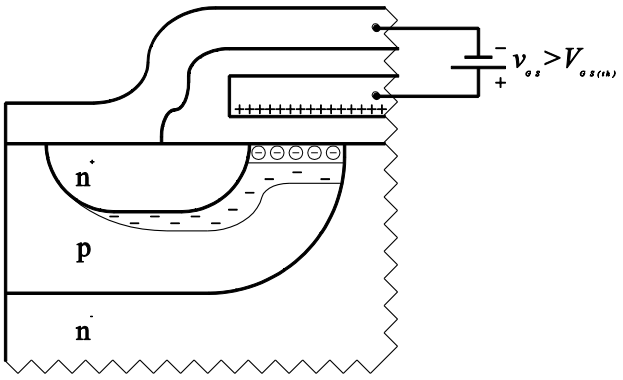
Several of the doping regions of the MOSFET shown in Figure 4.19 is of great importance to understand the basic principles of operation of the power MOSFET. The area in the vicinity of the gate containing gate-metallization, gate-oxide (silicon-dioxide, SiO_2, which is an insulator) and the body-region beneath the gate-oxide together forms the MOS capacitance. If a positive gate-source voltage, lower than the threshold level, is applied it will appear across this capacitance. Furthermore, a depletion region will form in the p-body region just beneath the gate-oxide since a positive charge on the gate-metallization requires an equal amount of charge with opposite sign on the other side of the gate-oxide, i.e. in the p-body, in the same manner as for a capacitor. This is shown in Figure 4.22.



(a)



(b)



(c)

Figure 4.22 The growth of the inversion layer for increasing gate-source voltage. The gate-source voltage increases from (a) to (c). In (a) and (b) the gate-source voltage is lower than the threshold level and in (c) it is higher.

The electrical field that is created by the positive charge on the gate-metallization repels the holes in the p-body, i.e. the majority carriers, just beneath the gate-oxide and thereby the negatively charged acceptor atoms are exposed and a depletion region is formed. If the gate-source voltage increases the thickness of the depletion region must also increase in order to expose more acceptor atoms to increase the amount of negative charge in the depletion region so that it equals the positive charge on the gate-metallization.

When the gate-source voltage is further increased, free electrons will be attracted by the electric field across the oxide-silicon junction. Also, holes will be repelled deeper into the body. This is shown in Figure 4.22, where free electrons are shown as minus signs enclosed in a circle. These free electrons originate from thermal ionization generating free electron-hole pairs. The free holes originating from this process are repelled deeper into the body where they are neutralized by electrons provided from the n-doped source and attracted by the positive charge of the free holes. When the density of free electrons in the oxide-silicon junction has grown equal to the doping atom density of the body region the layer of free electrons is termed inversion layer.

The conductivity of the inversion layer is very high and has the same properties as an n-doped semiconductor (hence the term inversion). The gate-source voltage required to obtain an inversion layer equals the threshold level $V_{GS(th)}$. The inversion layer connects the heavily doped drain and source-regions by providing a path between these two regions. The ability to affect the conductivity of a semiconductor by application of an electric field is termed field-effect.

If the gate-source voltage is further increased, beyond the threshold level, the inversion layer will grow slightly in thickness whereas its conductivity increases strongly so that a channel is created between drain and source according to Figure 4.22. The inversion layer now screens the depletion region so that an increasing gate-source voltage does not result in a further increase of the thickness of the depletion region.

Now consider a MOSFET connected to an electric circuit in such a way that a non-zero drain current i_D flows. This drain current will affect the geometry of the channel. Initially the MOSFET operates in the ohmic region with a gate-source voltage exceeding the threshold level $V_{GS(th)}$ and a relatively low dc link voltage V_{dc}. In this operational mode, the drain current is relatively low and the inversion layer has a equal width in the direction of the current flow.

If the dc link voltage is increased slightly for a constant gate-source voltage V_{GS}, the drain current will increase proportionally since in the ohmic region the inversion layer behaves as a constant resistance connected between the drain and source regions. The increasing drain current results in a voltage drop along the channel, which implies that the channel-to-source voltage is not uniform along the channel, and thereby, the inversion layer will not have a uniform thickness along the channel. Note that this implies that the current density will not be uniform along the channel carrying the drain current i_D.

At the drain end of the channel, channel-to-source voltage equals the drain-source voltage, and therefore, the oxide voltage equal to the difference between

the gate-source and drain-source voltages. Since the inversion layer is thinner at the drain end, the resistance increases when the current increases which is also the reason why the current-voltage characteristic tends to flatten out as the drain-source voltage increase for constant gate-source voltage.

From the discussion above, the conclusion that for high drain currents, the channel would disappear if the oxide voltage at the drain end decrease below the threshold level, and as a result, the device enter cut-off. However, this is not the case, which is explained in the following. Since no minority carrier injection takes place, the entire drain-current flows by drift and consequently depends on an electric field in the direction of the channel. The current density is proportional to the electric field where the constant of proportionality is the conductivity. This implies that the electric field is strongest at the drain end since the current density is highest due to the fact that the inversion layer is thinnest at the drain end. The high current density is important for two reasons. Firstly, as the oxide voltage at the drain end decreases below the threshold level and the electric field in the oxide becomes too small to maintain the inversion layer, the electric field in the channel adopts the responsibility to maintain the inversion layer. This solves the previously announced apprehension that the channel could disappear for high drain current. Secondly, the average carrier velocity for free electrons is dependent on the electric field strength. In the case of high electric field strength in the channel the carrier velocity will saturate. Roughly, saturation occurs when the difference between the gate-source and drain-source voltages equals the gate-source threshold voltage and the MOSFET is about to enter the active region. A further increase of the drain-source voltage increases the electric field strength in the narrowest part of the inversion layer, i.e. at the drain end, which results in a lengthening of this narrow section. In the current-voltage characteristic this is shown as an essentially constant drain current for drain-source voltages exceeding the difference between the actual gate-source voltage and its threshold level. If the gate-source voltage is increased, the inversion layer increases in thickness which implies that a higher drain current is required to reach carrier velocity saturation.

Note that in the active region, the drain current is proportional to the square of the difference between the gate-source voltage and its threshold level. In the case of high drain current (and consequently high carrier density) the mobility of the free carriers decrease both due to collisions between carriers (carrier-carrier scattering) but also due to carrier velocity saturation resulting from the high electric field caused by the high current density.

Switching

The power MOSFET is inherently faster than the power BJT since for the MOSFET excess carrier distribution does not have to be established and

removed at turn-on and turn-off. The only carriers that have to be transported are the ones in the stray capacitances and depletion regions. In most cases the MOSFET is modelled as being capacitive between all three terminals (Figure 4.23). The drain source capacitance is in most cases not needed in calculations but can be used as a part of the snubber capacitance. In the active region a current source is used to model the drain current. The drain current is given by $i_D = g_m \cdot (V_{GS} - V_{GS(th)})$ where g_m is the trans conductance.

In the ohmic region a resistor is used to model the forward voltage drop (drain-source voltage) of the MOSFET and to calculate the on-state losses. The resistance $R_{DS(on)}$ thus corresponds to the resistive losses of the drift region for operation in the on-state. For MOSFETs, the on-state losses are usually substantial since there no conductivity modulation takes place in the drift region as a result of the lack of minority carrier injection.

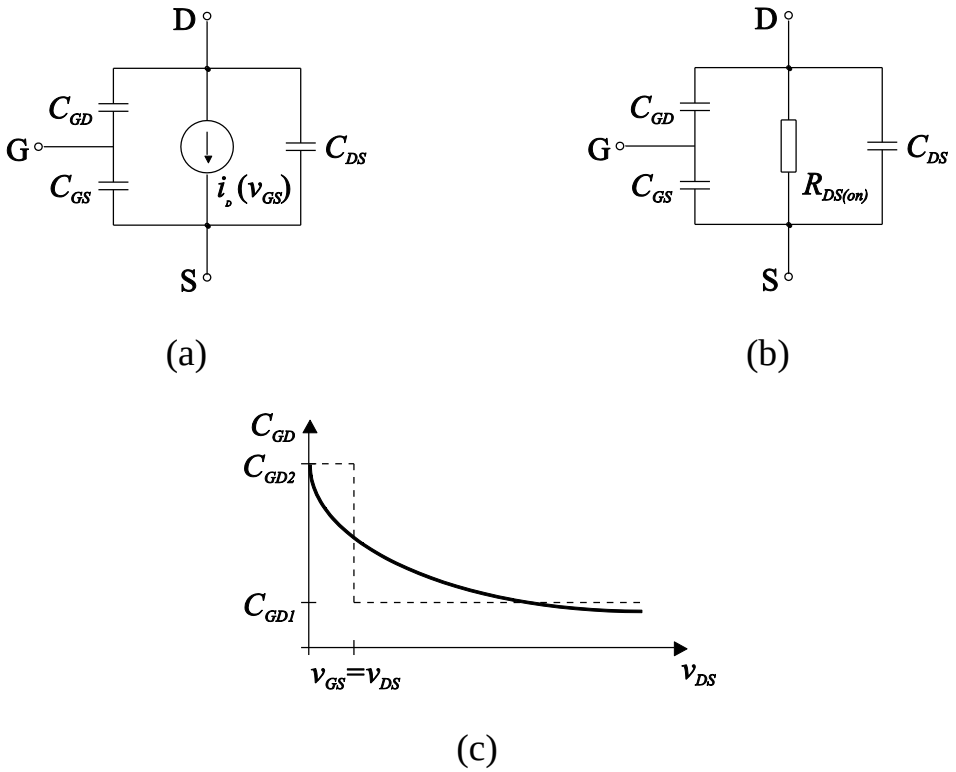


Figure 4.23 Equivalent circuits of the MOSFET for cut off and the active region (a), and the ohmic region (b). The gate-drain capacitance dependency upon the drain-source voltage is shown in (c).

A complication exists in hand calculation of the switching waveforms, and therefore also the switching losses, resulting from the fact that the stray capacitances of the MOSFET are known. Furthermore, it is important to realize that the stray capacitances, especially the gate-drain capacitance, are dependent on the actual drain-source voltage. In cut-off, the gate-drain capacitance is the sum of capacitance of the gate-oxide and the capacitance of the depleted part of the drift region. However, in the ohmic region the gate-drain capacitance is essentially equal to the capacitance of the gate-oxide. As a consequence, the

gate-drain capacitance changes substantially between cut-off and on-state modes of operation (typically a factor 10-100). The dependency of the gate-drain capacitance as a function of the drain-source voltage is of course continuous, but for hand calculation it is sufficient to use two levels one for operation in the off-state and one for operation in the on-state. The gate-source capacitance varies in a similar manner but dependent on the presence of a depletion region beneath the gate-oxide, i.e. on the gate-source capacitance. However, the gate-source capacitance changes only slightly for the possible gate-source voltages and therefore it is considered as being constant for hand calculations.

Similar to the analysis of the switching waveforms of the BJT, MOSFET switching is investigated for a step-down converter with inductively clamped load and ideal freewheeling diode. Note the gate-resistor of the gate drive circuit connected to the gate of the MOSFET. The supply voltage of the gate driver, V_{CC}, is assumed to exceed the threshold gate-source voltage of the MOSFET, $V_{GS(th)}$. Figure 4.24 shows generic switching waveforms of a typical MOSFET.

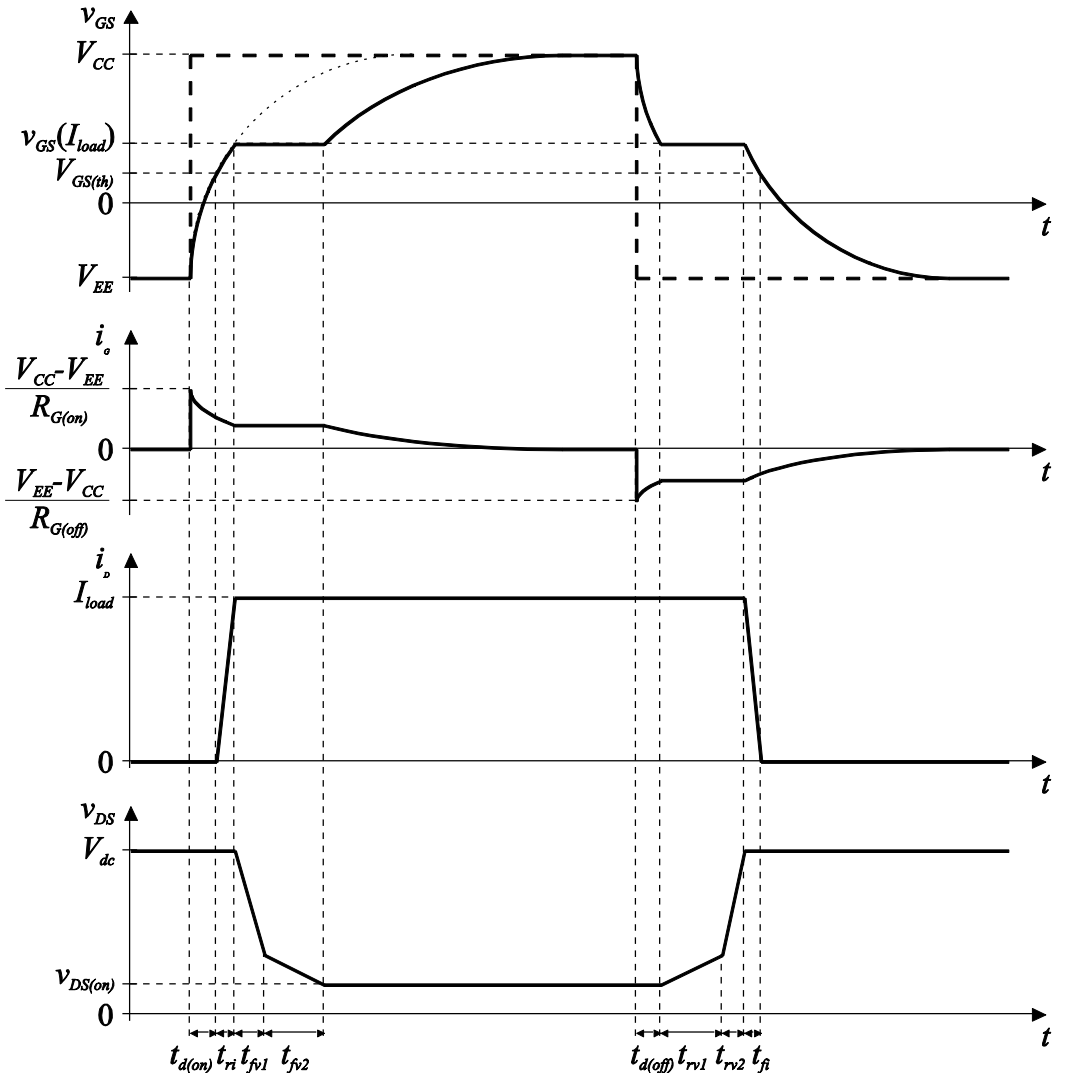


Figure 4.24 *Principle switching waveforms of a typical MOSFET. In the left half the turn-on process is visualized. In right half turn-off is shown.*

Turn-on

Turn-on delay time ($t_{d(on)}$): The output voltage of the driver is assumed to increase momentarily or stepwise. Consequently, the gate-source voltage increases from the initial towards the threshold level $V_{GS(th)}$. Since this voltage change affects both the gate-source and the gate-drain voltages, the gate-source and gate-drain capacitances appear to be in parallel from the driver point of view. The time constant of the gate-source voltage response is determined by the gate-source and gate-drain capacitances (in parallel, i.e. the sum), and the gate resistor. Note that since the MOSFET is blocking, the drift region is partly depleted, and consequently, the gate-drain capacitance is low. Since the turn-on delay time usually is short compared to the time constant, the response of the gate-source voltage appears to be linear.

Current rise time (t_{ri}): When the gate-source voltage reaches the threshold level $V_{GS(th)}$ the inductive load current starts to commutate to the MOSFET, and consequently, the drain current i_D starts to increase. The gate-source voltage continues to increase with the same time constant as during $t_{d(on)}$ interval, since the drain-source voltage is unaffected also during the t_{ri} interval. This is due to the fact that as long as the freewheeling diode carries current, its forward voltage drop is essentially constant low. As soon as the MOSFET carries the entire load current, the freewheeling diode starts to block and the drain-source voltage of the MOSFET decreases.

Voltage fall time (t_{fv}): During this interval the gate-source voltage is clamped to the level determined by the transfer characteristic for the particular drain current $i_D = I_{load}$ since the MOSFET is operating in the active region. Consequently, the entire gate-current is flowing through the gate-drain capacitance, which is recharged so that the drain-source voltage decreases. Note that as the drain-source voltage decreases the depletion layer of the drift region narrows. Therefore, this time-interval is divided in two intervals for basic understanding; one with gate-drain capacitance given by $v_{DS} > v_{GS(th)}$ (C_{GD1}) and one for $v_{DS} < v_{GS(th)}$ (C_{GD2}) according to Figure 4.23. The first part of the voltage fall interval t_{fv1} is valid for $v_{DS} > v_{GS(th)}$, i.e. operation in the active region, and the time constant is given by C_{GD1}. The second part of the voltage fall interval t_{fv2} is valid for $v_{DS} < v_{GS(th)}$, i.e. operation in the ohmic region, and the time constant is given by C_{GD2}. The difference in the derivative of the falling voltage with respect to time is due to the fact that the gate-current is constant, due to the constant gate-source voltage and the change in gate-drain capacitance (recall that a capacitive current is equal to the capacitance times the voltage derivative which gives that a higher capacitance gives a lower voltage derivative with respect to time for a constant capacitor current). The voltage fall interval is sometimes termed the Miller interval and the constant gate-source voltage level during the voltage fall interval is referred to as the Miller plateau.

When the drain-source voltage reaches its steady state level $V_{DS(on)}$ (for the particular drain current) the clamping of the gate-source voltage is lost and the gate-source voltage continues to increase towards the supply voltage level of the driver V_{CC}. Similarly to the $t_{d(on)}$ and t_{ri} intervals, the voltage increase is exponential and the time constant is given by the gate-drain and the gate-source capacitances (in parallel, i.e. the sum) and the gate resistor. Note, however, that the gate-drain capacitance in the on-state is the one valid for low drain-source voltage, i.e. C_{GD2}. Consequently, the time constant is longer than for the $t_{d(on)}$ and t_{ri} intervals.

If the freewheeling diode is no longer considered as being ideal and reverse recovery is included in the analysis another sub-interval appears in the MOSFET voltage fall interval. This sub-interval is located first in the voltage fall interval and is characterised by a very high (negative) derivative in the drain-source voltage with respect to time. The high derivative is due to the fact that the drain current i_D at the start of the voltage fall interval is equal to the sum of the load I_{load} current and the reverse recovery current I_{rr}. However, the reverse recovery current decreases to zero rapidly, which leaves the MOSFET with a gate-source voltage that is higher than the one given by the transfer characteristic. Consequently, the gate-source voltage is decreased rapidly by discharging the charge of the gate-source capacitance to the gate-drain capacitance, which implies that the gate-drain voltage and therefore also the drain-source voltage changes rapidly. As soon as the gate-source voltage has decreased to the level determined by the transfer characteristic and the actual load current, the Miller interval commences and the voltage falls as indicated by t_{fv1} earlier. Note that it is not unlikely, but rather common, that the reverse recovery current is about equal to the load current in magnitude.

Turn-off

The turn-off process of a MOSFET is in principle a reverse turn-on process but the problem associated with diode reverse recovery does not show up. Therefore, the diode is treated as being ideal.

Turn-Off delay time ($t_{d(off)}$): This delay is due to the fact that the gate-source voltage must decrease to the level determined by the transfer characteristic and the actual drain current, i.e. the load current. When this gate-source voltage level is reached the MOSFET enters the active region and the drain-source voltage can rise.

Voltage rise time (t_{rv}): In the same manners as for turn-on, this interval is split into two sub-intervals, depending on whether the drain-source voltage is lower than the gate-source voltage (ohmic region of operation) or higher than the gate-source voltage (active region of operation). Note that also in this case the gate-

source voltage is clamped to a constant level (i.e. a Miller plateau appears), which implies that the entire gate current is charging the gate-drain capacitance. Also in this case the gate-drain capacitance is not constant but depends on the drain-source voltage so that C_{gd2} is valid for t_{rv1} and C_{gd1} for t_{rv2}. When the drain-source voltage reaches a level slightly (about 0.7-1.0 V) higher than the dc link voltage (V_{dc}) the freewheeling diode becomes forward biased and the load current starts to commutate.

Current fall time (t_{fi}): The drain current i_D falls with a speed determined by the rate of fall of the gate-source voltage as specified by the transfer characteristic. Since the rate of fall of the gate-source voltage is determined by the gate current, i.e. the gate resistor and the supply voltage, the duration of t_{fi} can be affected by altering the gate resistor R_G. The gate-source voltage decreases and the gate current increases (negatively) exponentially with a time constant given by the sum of the gate-drain and gate-source capacitances and the gate resistor value. In this time interval the gate-drain capacitance C_{gd1} is valid since the drain-source voltage is higher than the gate-source voltage. Note that the drain current equals zero at the time-instant when the gate-source voltage has fallen to a level equal to the threshold level $V_{GS(th)}$, as expected from the transfer characteristic. Since the duration of t_{fi} is short compared to the time constant, the drain current seems to fall with a constant time derivative.

When the drain current has reached zero the gate-source voltage continues to decrease, exponentially with the same time constant, towards the voltage level supplied from the driver output which is negative in most cases but for some devices zero voltage is sufficient.

An interesting feature of MOSFET switching characteristics is that the device capacitances are not temperature dependent, which implies that the switching losses are in principle constant with respect to device temperature. For bipolar devices, the switching times are strongly temperature dependent since the minority carriers are more energetic at elevated temperatures and therefore the probability of recombination is less at high temperatures.

On-state losses

The MOSFET has due to its short switching times very low switching losses compared to the BJT. On the other hand, the on-state forward voltage drop and thereby the on-state losses, are higher for the MOSFET than for the BJT, especially for high blocking voltage levels. The reason for the high on-state voltage drop is that there is no conductivity modulation taking place in the drift region due to the lack of minority carrier injection. The lack of minority carrier injection is also the main reason for the short switching times.

To be able to block high voltages in the off-state the drift region must have a certain thickness to accommodate the depletion layer. The drift region contributes with most of the resistive voltage drop due to its length in the direction of the current flow and the low doping level. The resistance of the channel and the drain and source regions are considerably lower due shorter geometrical distances and also due to the higher doping levels. The forward voltage drop of a MOSFET is assumed to be purely resistive and given by the on-state resistance $R_{DS(on)}$ and the drain current i_D.

Sometimes it is claimed that the MOSFET can be used as a synchronous rectifier which means that the devices has the ability to control drain currents flowing in both directions, i.e. both positive and negative drain currents. This should be treated with care since for the MOSFET itself it is true but if the parasitic freewheeling diode in the interior of the devices is taken into consideration it is true only for very low drain current densities resulting MOSFET on-state voltage drop lower than freewheeling diode on-state voltage drop. For drain current densities resulting in a MOSFET on-state voltage drop higher than the diode forward voltage drop, the drain current will commutate to the parasitic freewheeling diode and controllability of the current is lost. This is in particular true for MOSFETs with high blocking voltage capability, which are equipped with a thick drift region and consequently has a high $R_{S(on)}$. However, the parasitic freewheeling diode still has a low forward voltage drop since conductivity modulation takes place in this structure of the MOSFET.

Blocking voltage

In forward blocking state, the drift-body junction accommodates the blocking voltage since the drift region is lightly doped compared to the body region which implies that the depletion layer is located in the drift region. As mentioned earlier, there is parasitic npn BJT contained in the doping structure of a power MOSFET which is turned off by extending the source-metallization to the body region to short-circuit the junction between these regions. In this way, the base-emitter junction of the parasitic npn BJT is short circuited, at least in steady-state operation.

In dynamical processes, like switching, there can still be problems with undesired turn-on of the parasitic npn BJT through current displacement and capacitive current. If the drain-source voltage increases very rapidly at MOSFET turn-on this voltage time derivative will connect to the n-base through the gate-drain capacitance and the resistance formed body and source regions for lateral currents, i.e. currents perpendicular to the direction of normal current flow. This limits the maximum allowed drain-source voltage-time derivative which must not be exceeded for practical converter designs. The maximum drain-source voltage derivative with respect to time is in most cases not

explicitly specified in manufacturer's datasheets but in most cases the minimum gate resistor value R_G is specified. Recall from the section on MOSFET switching that R_G can be used to control the drain-source voltage derivative with respect to time.

It is important to understand why the parasitic npn BJT must not be trigged. First the breakdown voltage is reduced if the n-base is an open circuit, compared to the case with the base and the emitter short circuited (recall from the section about the BJT that $BV_{CBO} > BV_{CEO}$). This motivates that the source-metallization is extended to the body region to short circuit the source-body junction. The extended source metallization also contributes to give a smoother profile of the depletion layer boundary, by increasing the minimum radius, which in turn reduces the maximum electric field strength in the depletion layer.

The most severe problem associated with the parasitic npn BJT is if it is turned on and reaches saturated mode of operation, a condition termed latch up. If latch up occurs the power dissipation increases and what is worse the parasitic npn BJT cannot be externally turned off since its base is not connected to any of the three terminals. Instead, the only way to turn the parasitic npn BJT off is by force commutation by an external component.

SOA for the MOSFET

Safe operating area for MOSFETs is determined by the maximum allowed drain current I_{DM} and the breakdown voltage $V_{DS(max)}$ for the specified maximum silicon temperature T_j. The SOA of MOSFETs is in most cases square, since the switching times are so short.

4.7 The IGBT

The IGBT [1][9][14][16][18][19][24], was introduced to be a compromise between the bipolar junction transistor (BJT) and the metal oxide semiconductor field effect transistor (MOSFET). However, the IGBT has become more than just a compromise, due to its high ruggedness, moderate voltage drop and fairly high switching speed. Actually, in many modern designs it is the only reasonable device choice. The main reasons for its popularity is the high blocking voltage, up to 6 kV, its moderate driving requirements and its limited need for protective devices, i.e. snubbers.

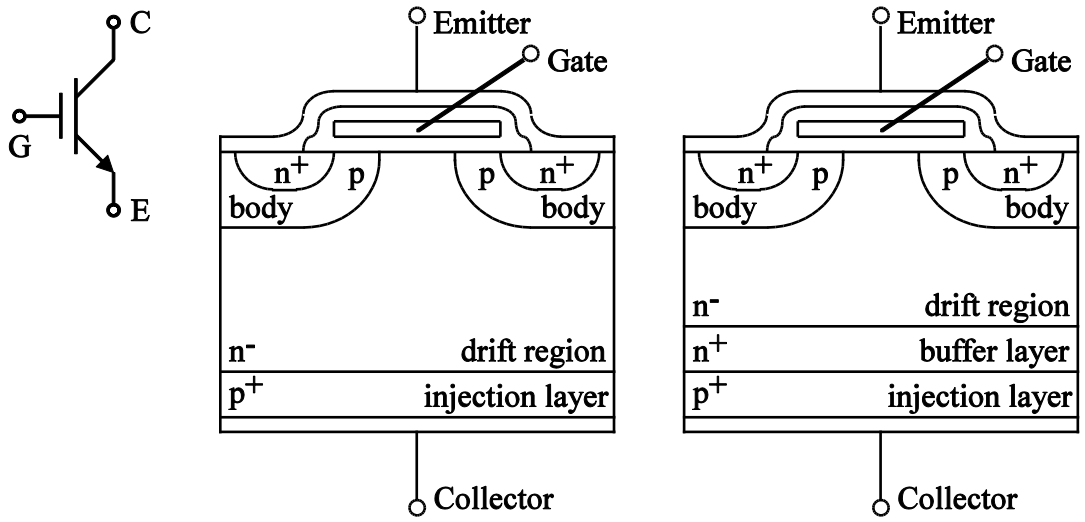
Basic doping structure

The IGBT is a three terminal semiconductor device (Figure 4.25). The terminal named emitter is common to both input and output. The device is controlled by applying a voltage between the gate and emitter terminals. The output current

flows between the collector and emitter terminals. The gate-emitter region is split into thousands of cells. In Figure 4.25, one such cell of a NPT-IGBT and also one of a PT-IGBT are shown. The only structural difference between these two is that the PT device has an additional doping layer, termed buffer.

The description presented here, focuses on the NPT device. The most commonly used n-channel IGBT circuit symbol is used throughout the entire text and is therefore not shown here. Other circuit symbols for the IGBT do exist, but this is the most frequently used.

The reason for dividing the gate-emitter structure into cells is to keep the channel resistance low by increasing its cross-sectional area and decreasing its length. Further on, the emitter metallisation extends over the body region, which is done to short circuit the body-emitter junction. This is used to reduce the risk of entering latch up, a fault condition discussed later on.



(a) (b) (c)

Figure 4.25: IGBT circuit symbol (a) and principal doping structure for (b) the NPT-IGBT, and (c) the PT-IGBT.

Steady state operation

From Figure 4.25 it is seen that the doping structure of an IGBT looks similar to the one of a MOSFET. For the NPT device the only difference is the presence of a heavily doped injection layer at the IGBT collector metallisation. The reason for having this layer is that holes should be injected into the drift region to obtain conductivity modulation when the device operates in the conduction state. The lack of conductivity modulation is the main drawback of the MOSFET, being a majority carrier device.

In the on state, similar to the MOSFET, a channel is created in the body region underneath the gate oxide, connecting the n-doped emitter and drift regions. The

channel supports an electron current, flowing from the emitter into the drift region. The negative charge of the electrons attracts holes from the injection layer, which in turn causes conductivity modulation of the drift region.

When the IGBT operates in the forward blocking state, the drift region supports the voltage. Similar to the power diode case, a depletion region is formed in the drift region. For a NPT device, this is also true for the reverse blocking state, i.e. when the IGBT is blocking a collector-emitter voltage of reverse polarity. However, the PT-IGBT cannot block a collector-emitter voltage of negative polarity, since the injection-buffer junction should support the voltage in this case. Both these regions are highly doped, resulting in high electric field strength even at low blocking voltages. Consequently, avalanche breakdown will occur for low reverse blocking voltage. This is the reason why the PT-IGBT is often termed asymmetric.

In Figure 4.26 the current-voltage characteristics of a typical IGBT is shown. The output characteristic of an IGBT shows the collector current, i_C, as a function of the collector-emitter voltage v_{CE}, for different gate-emitter voltages v_{GE}.

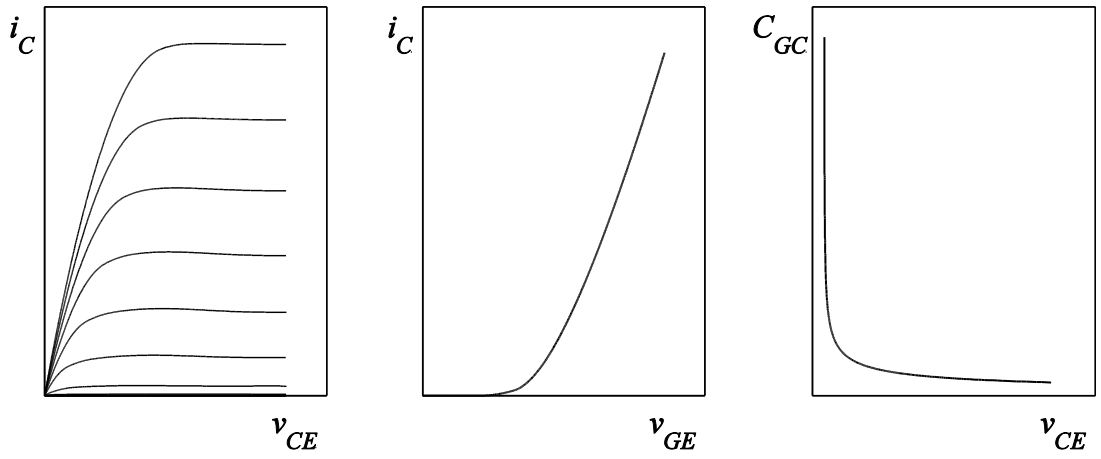


Figure 4.26: Output characteristic for low collector-emitter voltages (a), transfer characteristic (b), and gate-collector capacitance as a function on the collector-emitter voltage (c).

The transfer characteristic shows the collector current as a function of the gate-emitter voltage, when the IGBT is in the active region. The transfer characteristic is thus mainly interesting for the switching transients, since otherwise the device should operate in the on-state or off-state regions.

Switching at inductively clamped load

Figure 4.27 shows typical switching waveforms for an IGBT operating at inductively clamped load in a bridge application. The corresponding

freewheeling diode is regarded as being ideal. At turn-on of the IGBT, the gate-emitter voltage v_{GE} must first reach its threshold level, $v_{GE(th)}$, before the collector current i_C starts to increase. The time needed to do this is termed the turn-on delay time $t_{d(on)}$.

The gate-emitter threshold voltage is seen in the transfer characteristic of Figure 4.26 as a sharp bend where the collector current becomes non-zero.

Note that since the IGBT has a capacitive input, the gate resistor determines the delay time. Also note that the internal capacitance is varying due to depletion layer thickness, i.e. collector-emitter voltage. The gate-collector capacitance is the most affected. It's typical dependence on collector-emitter voltage is shown in Figure 4.26, where the fairly sharp knee approximately corresponds to a collector-emitter voltage equal to the applied gate-emitter voltage.

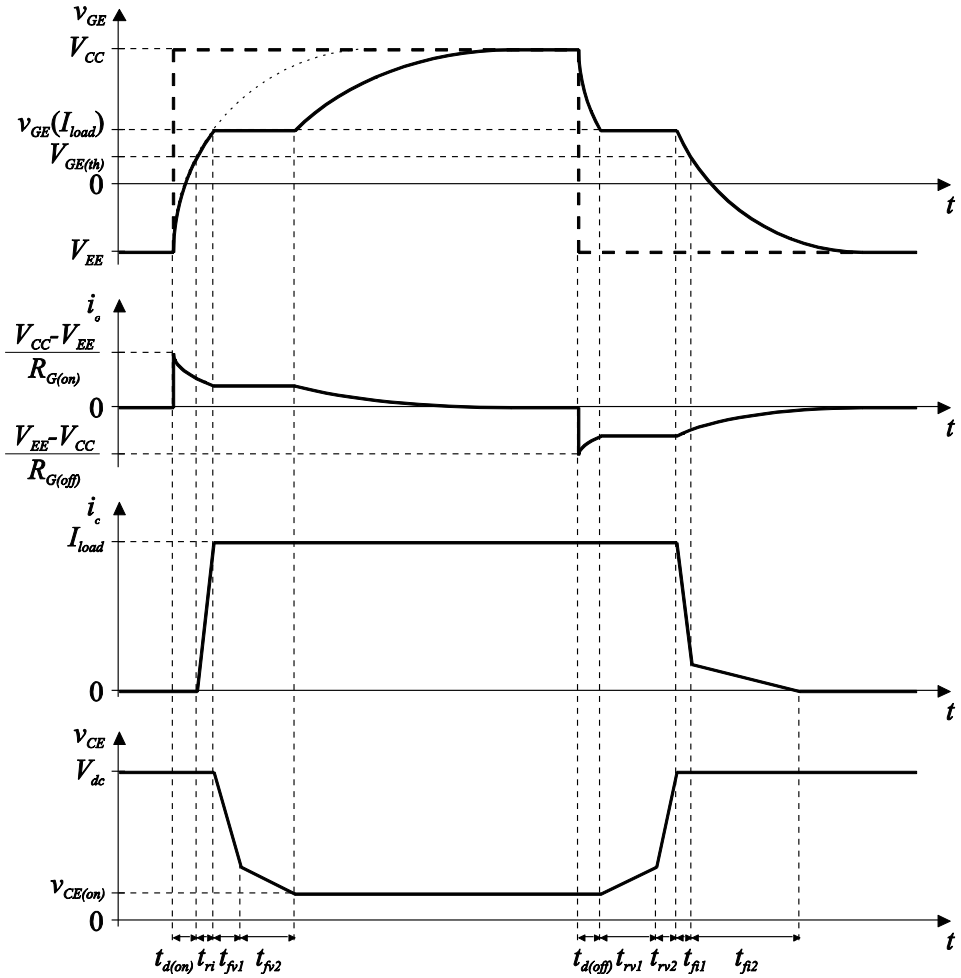


Figure 4.27: Principle switching waveforms of a typical IGBT. To the left the turn-on process is visualized and to the right turn-off is shown. Note that in this figure the terminals collector and emitter are termed drain and source.

After the turn-on delay time, the gate-emitter voltage continues to increase and the collector current starts to increase. The collector current time derivative is determined by the transfer characteristics, see Figure 4.26, which implies that it

is determined by the size of the gate resistor. The duration of this current rise interval is termed current rise time t_{ri}.

Following the current rise interval, the voltage fall interval commences. The duration of this interval is termed voltage fall time t_{fv}. The voltage fall interval can be subdivided into two portions. The first part is similar to the voltage fall of a MOSFET, i.e. the Miller plateau. The second part has considerably lower fall rate mainly due to conductivity modulation lag.

IGBT turn-off also starts with a delay time, $t_{d(off)}$, due to the fact that the gate-emitter voltage has to decrease to the level determined by the transfer characteristic before the collector-emitter voltage starts to increase. During the voltage rise time t_{rv} the collector-emitter voltage increases and the gate-emitter voltage is constant due to the Miller effect.

When the collector-emitter voltage has reached the level of the DC link voltage, the freewheeling diode becomes forward biased. Hence, the IGBT collector current starts to decrease. The collector current fall rate is also determined by the gate-emitter voltage through the transfer characteristic, i.e. the fall rate is essentially determined by the gate resistor value. When the gate-emitter voltage has decreased to its threshold level $v_{GE(th)}$, the channel in the body region is removed, corresponding to that the MOSFET part of the IGBT is turned off. In IGBT data sheets this part of the collector current fall is termed t_{fi}.

However, there are still excess carriers stored in the IGBT drift region. Since the channel is not present, these carriers cannot be removed through the channel. Instead, there will be a long collector current tail where the current falls at a very low rate. This low rate is due to the fact that the excess carriers are removed by internal recombination, which is a very slow process [14]. Note that the current tail fall rate is not affected by the gate resistor value. However, the gate resistor value has some influence since the initial magnitude of the current tail is dependent on the amount of excess carriers removed from the drift region during t_{fi} [1]. For a low gate resistor value, t_{fi} will be short and hence the density of excess carriers is almost unaltered during t_{fi}. For a high gate resistor value on the other hand, a large amount of the excess carriers are removed during the comparably long t_{fi}. Thus, the initial magnitude of the collector current tail will be lower in the second case.

Latchup

The doping structure of the IGBT contains a parasitic thyristor (pnpn) structure, see Figure 4.25. The injection layer and the emitter region forms the anode and cathode, respectively, of the parasitic thyristor structure, see Figure 4.28. The body region forms the gate (p-base) and the drift region the n-base of this

structure. The parasitic thyristor structure must not be triggered, a state termed latchup, since the thyristor cannot be turned off unless its anode current is forced to zero by surrounding circuit elements.

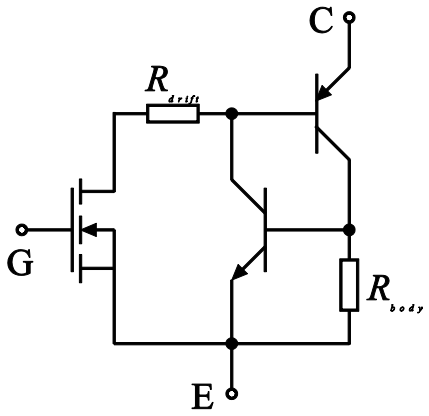


Figure 4.28: Simple equivalent circuit of an IGBT.

Though not obvious, the thyristor structure can be trigged by a lateral current, forward biasing the body-emitter junction. The lateral current flow arises from a part of the hole current, injected from the collector, first being attracted by the negative charge of the channel. When the holes enter the body region, electrons from the emitter metallisation covering the body are attracted, which the injected holes recombine with. Hence, a lateral current flow in the body region results. If the resistivity of the body is high enough, the lateral current can forward bias the body-emitter junction. Note that even though the emitter metallisation also covers the body, i.e. short circuits the emitter and body regions, the body-emitter junction can be forward biased in the interior of the structure.

Actually, latch up is distinguished into two different cases, static which occurs in the on-state and dynamic which occurs during turn-off. Static latch up is due to a too high collector current, resulting in a lateral voltage drop forward biasing the body-emitter junction. Hence, static latch up is avoided by not allowing a higher collector current than specified by the manufacturer.

Dynamic latchup occurs at turn-off, after the channel has been removed. Since the collector-emitter voltage already is at the blocking level, see **Figure 7.6**, the depletion region is being established in the drift region. As a consequence, the emitter efficiency for the remaining pnp-structure increase since the effective base length decreases. The effective base length is the portion of the drift region remaining un-depleted. The emitter efficiency increases due to the fact that the possibility of recombination decreases if the base length is short.

The problem is that an increased emitter efficiency results in the fact that a larger part of the total current is collector current of the pnp-structure, which also is the lateral current in this case. Hence, if a gate resistor of low resistance is used, the lateral current might increase when the channel is removed,

compared to the on-state case. Then, dynamic latch up will occur for a lower IGBT collector current than static would. Therefore, it is also important not to use a gate resistor value lower than specified by the manufacturer.

Switching under zero voltage conditions

One of the main problems related with soft switching appears due to poor understanding of power semiconductor physics, since it is assumed that data sheet information is still valid at soft switching. However, data sheet information for IGBTs is in most cases given for inductively clamped load, i.e. constant load current during the switching transients [10]. Also, the information is only valid for a certain constant DC link voltage. In the literature [10][12][18], a lot of problems appearing due to soft switching are discussed.

One of the most discussed phenomena observed, is the current tail bump occurring at IGBT zero voltage turn-off. In **Figure 7.13**, the current tail bump is clearly seen. According to [18] the reason for this bump is that during ZVS turn-off the excess carriers stored in the drift region are not forced out by the expanding depletion region as for hard switched turn-off. Consequently, after the channel is removed the collector current continues its decrease and no current tail is observed until the IGBT collector-emitter voltage begins to increase. The current tail bump results in higher losses at ZVS turn-off than expected from data sheet information. Nevertheless, the turn-off losses are lower for ZVS than for hard switching [10].

Another problem observed, is due to the conductivity modulation lag [10][12], appearing for high collector current time derivatives. For an IGBT, the origin of this lag is the same as for the power diode discussed earlier. This means that the collector current increases at a rate higher than the rate which excess carriers are injected into the drift region, needed to establish conductivity modulation.

In [10] problems arising from device packaging are also discussed. Here, the problem is that bonding wire inductance not only gives an increased forward voltage drop but also can result in uneven current distribution among several IGBT chips packaged in the same module, at high current time derivatives. Uneven current distribution leads to increased losses and also increased device stress.

The soft switching problems discussed above are investigated in the simulations presented in Chapter **Error! Reference source not found.**, where at least the current tail bump is clearly seen.

Safe operating area for the IGBT

The most important parameters of the SOA for an IGBT are the absolute maximum collector current I_{CM}, the absolute maximum $V_{CE(max)}$ and the absolute maximum gate-emitter voltage $V_{GE(max)}$. The breakdown voltage BV_{CBO} (section 4.3) of the pnp-structure essentially determines the breakdown voltage of the IGBT. IGBTs are typically manufactured to be able to withstand a collector current 6 to 10 times higher than for I_{CM} for time periods less than 10 μs. Note that IGBT drivers are often designed to be able to turn off such a current level without device failure occurring.

Most IGBTs are designed for a maximum operation temperature of at 150 °C. Furthermore, IGBTs can be designed so that the losses are in principle not temperature dependent since the MOSFET part of the IGBT corresponds to a resistance with positive temperature coefficient and the BJT part corresponds to a resistance with negative temperature coefficient. In most cases IGBTs are designed to exhibit slight negative temperature coefficient, to simplify parallel operation of devices.

The safe operating area of an IGBT is essentially square in switch-mode. The reverse blocking SOA (RBSOA) of some IGBTs includes a limitation of the turn-off voltage derivative to avoid dynamic latch up. The collector-emitter voltage derivative with respect to time is limited by setting a lower limit of the gate resistor. Note that increasing the resistance of the gate resistor increases the switching losses. To avoid failures due to dynamic latch up, gate resistors lower than the minimum specified by the manufacturer should never be used.

4.8 Silicon Carbide

High temperature capability

Silicon-Carbide is a material, which is a mixture of silicon and carbide. Both materials have four valence electrons in its outer electron shell. Silicon Carbide is neither n-doped nor p-doped, so there is a significant wide energy gap between the electron valence band and the electron conduction band, see figure i.

In a metal, e.g. copper, a many electrons are found in the conduction band, and therefore a metal is a good electric conductor. Contrary to the situation in a metal, in a semiconductor the electrons are normally found in the valence band. By means of an external signal, e.g. a voltage, the electrons can get sufficient energy to ”jump” from the valence band to the conduction band, where it remains for a certain time and thereafter spontaneously falls back to the valence

band, and another electron can jump up to the conduction band. This is how semiconductors are controlled.

A serious problem is high junction temperature. Due to the high temperature the electrons in the valence band get sufficient energy to jump up to the conduction band, but as the electrons now continuously get energy they are not falling back to the valence band. We call this effect thermal ionisation. As the electrons are not falling back, control of the semiconductor is lost. The wider the energy gap between the valence band and the conduction band, the higher temperature is requested for thermal ionisation.

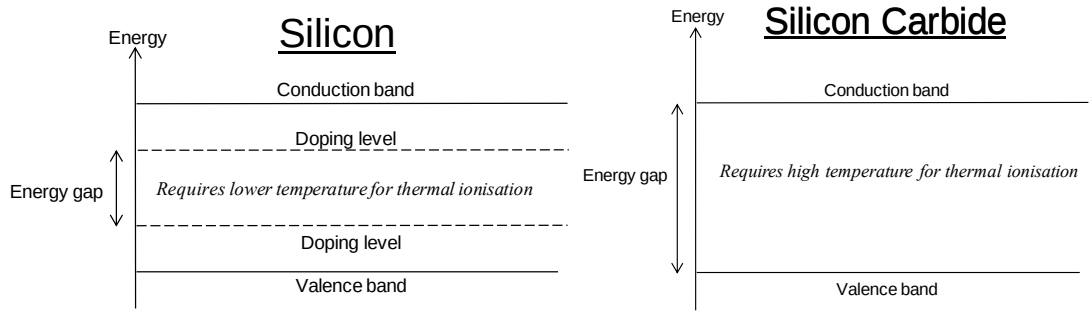


Figure 4.29 The energy gap in silicon carbide is much wider than in doped silicon, which reduces the risk for thermal ionisation.

As the risk for thermal ionisation in silicon carbide is reduced, the silicon carbide components can operate at much higher temperature than silicon component. Maximum operating temperature for silicon carbide components is several hundred centigrade, while maximum operating temperature for silicon is limited to between 100 and 200 centigrade.

However, the high temperature capability relates to the silicon carbide material itself. Today, the real silicon carbide component uses more or less the same casing as the silicon component. So far the casing limits the possibility to fully utilize the high temperature property of silicon carbide.

Important properties of silicon carbide components

- As already mentioned, the silicon carbide can operated at high junction temperature
- Low losses, especially switch losses.
- Both the high temperature capability and the low losses reduce the size of the cooling system.
- E.g. in hybrid electric vehicles, the power electronics and the combustion motor can use the same cooling system, see figure ii

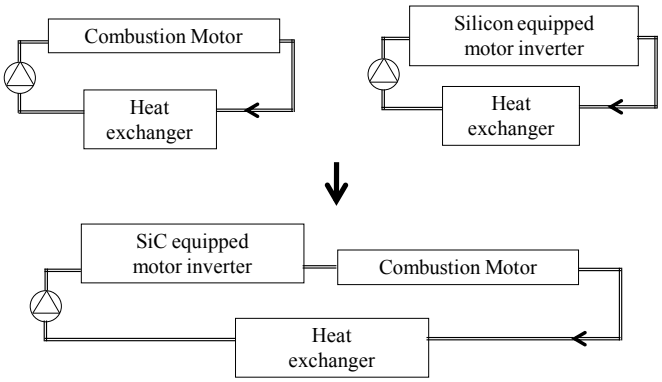


Figure 4.30 A motor converter, equipped with silicon carbide components, can share the same cooling system with the combustion motor

- High voltage capability.
- The high voltage capability reduces the number of series connected components in high voltage applications.
- Parallel connection of components.
- The silicon carbide resistance, even if it is low, has a positive temperature coefficient, which balances the current in the parallel connected components.

Different silicon carbide component types

MOSFET
- Voltage controlled –simple gate drive unit
- Low on-state losses
- Lower voltage capability than BJT
- An anti-parallel diode can be included in the MOSFET diffusion pattern.
- The MOSFET seems now (2018) to be the standard silicon carbide component for motor drive applications

BJT Bipolar Junction Transistor
- Current controlled.
- Low on-state losses
- The current control requires a more complicated base drive unit. The losses in the base drive unit must be included in the component total loss losses.
- High voltage capability

JFET Field effect transistor
- Voltage controlled –simple gate drive unit
- "Normally-on" or "Normally-off".
- "Normally-on" means that the component goes into conducting state, when the control signal on the gate equals zero or disappears. Thus, the gate drive units of a "Normally-on" component must always have voltage supply, otherwise a short circuit of the dc-link can occur.

- A “Normally-on” component has lower on state losses than a “Normally-off” component.
- Lower voltage capability than BJT

5 Passive Components

In this chapter the basics in design of inductive elements, i.e. inductors and transformers are introduced. A brief description of the capacitor types most widely used in power electronic design is also given.

5.1 Inductive components

Inductive elements consist of one or several copper wire windings which in most cases are wound upon an iron core. In this section the different core materials used in most cases are discussed. Core materials and inductor design are thoroughly discussed in for example [11][14][23][24].

Magnetic components for power electronic applications, such as inductors and transformers, are often designed only for the intended application, due to the fact that it is impossible to maintain storage of the wide variety of components needed. This means that for most power electronic applications, inductor design is a natural part of the development, since almost every power electronic circuit contains this kind of components.

Magnetic materials

Selection of appropriate core material is an important issue in magnetic component design. The material choice is to a large extent determined by the frequency spectra of the magnetising current. The frequency spectra of interest in power electronics range from 50 Hz (thyristor bridge commutation inductors) to several MHz (pulse transformers in for example transistor drive circuitry). There are mainly three types of core materials used to cover these frequency spectra.

For the low frequency region (50 Hz to some tens of kHz), alloys of iron and for example chrome, silicon or cobalt are used. These alloys are characterised by high electric conductivity and high saturation flux density (up to 1.8 T). The high conductivity makes this material subject to eddy current losses, i.e. the applied magnetics flux easily induces current in the core. The induced currents results in resistive losses, referred to as eddy current losses. To partly overcome this problem these cores are laminated, or in some cases manufactured as steel tape, to decrease the length of the current paths and the magnitude of the currents induced.

For the mid frequency region (1 kHz to 100 kHz), cores made from powder of the same mixtures as in the previous case, are used. The iron powder particles

are covered with an insulating layer, and forced together with pressure. Since the powder particles are electrically insulated from each other, the conductivity and thus the eddy current losses of the core become low. However, another loss component referred to as hysteresis loss, becomes more significant compared to the case for the previously discussed laminated cores.

In the upper frequency region (30 kHz to 10 MHz), soft ferrite cores are commonly used. Ferrites are ceramic compounds, consisting of iron oxide together with usually zinc or nickel. Other compound materials are also used, and also mixtures of several compound materials are possible. The crystals of the ferrite are typically 10–20 μm in dimension, thus limiting the eddy current path length and the magnitude of the current induced.

Ferrites, like other magnetic materials have losses associated with the hysteresis of the *B*-*H* loop. Hysteresis occurs due to friction associated with magnetic domain wall movement and magnetic domain rotation when an external magnetic field *H* is applied, causing the change in magnetic flux density to lag the change in magnetic field. This lag is known as hysteresis. For an AC magnetic field, a loop is formed in the *B*-*H* plane, forming the hysteresis curve. The area enclosed by such a loop corresponds to the hysteresis loss.

In most cases the manufacturer specifies the core losses as curves showing the loss per volume or mass unit, depending on the flux density and the frequency. In some cases, empirically derived expressions are given instead of curves. There are however also analytically derived loss models like the Steinmetz formula [14] where the specific losses are given by

$$p_{Fe} = k_h f^{a_1} \hat{B}_{ac}^{a_2} + k_{ec} f^2 \hat{B}_{ac}^2 \tag{5.1}$$

where a_1, a_2, k_h, and k_{ec} are material dependent constants. The first term of the Steinmetz formula above, is due to hysteresis losses and the second term due to eddy current losses. Note that the AC peak magnetic flux density is needed to calculate the iron losses. Several other forms of Steinmetz formula are often used. For example, the hysteresis loop is travelled once each period, implying that a_1=1. Also, if the hysteresis loop is approximated as being rectangular, a_2=2 can be assumed.

For the magnetic core materials in [13], the manufacturer gives the specific iron losses according to the empirical relationship

$$p_{Fe} = \frac{f}{\dfrac{a}{\hat{B}_{ac}^3} + \dfrac{b}{\hat{B}_{ac}^{2.3}} + \dfrac{c}{\hat{B}_{ac}^{1.65}}} + \left(d f^2 \hat{B}_{ac}^2 \right) \tag{5.2}$$

where *a*, *b*, *c* and *d* are material dependent constants. Specific iron losses in this case mean that the losses are given in the unit mW/cm^3.

5.2 Inductor design

In this appendix, inductor design is reviewed. The design method used, is basically the same as the one described in [11]. Nevertheless, some of the steps are also found in [14].

Inductance

First, the inductance of a gapped iron core inductor, see Figure 5.1, is derived.

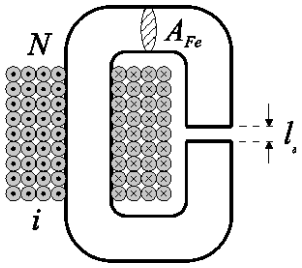


Figure 5.1 Basic gapped iron core inductor.

The inductance L is defined as

$$L = \frac{d\psi}{di} \tag{5.3}$$

where ψ is the flux linkage resulting from the current i. For a linear media, i.e. disregarding magnetic saturation and hysteresis in the case of an iron core, this is equivalent to

$$L = \frac{\psi}{i} \tag{5.4}$$

For the gapped iron core of Figure 5.1, Ampère's circuital law gives

$$N \cdot i = H_{Fe} \cdot l_{Fe} + H_\delta \cdot l_\delta \tag{5.5}$$

Here, N denotes the number of winding turns, H_{Fe} and H_δ the magnetic field intensity in the iron core and air gap, respectively. The magnetic flux mean path length is denoted l_{Fe} in the iron and l_δ in the air gap. The magnetic flux densities, B_{Fe} and B_δ, is defined from

$$\begin{cases} B_{Fe} = \mu_0 \mu_{Fe} H_{Fe} \\ B_\delta = \mu_0 H_\delta \end{cases} \tag{5.6}$$

where μ_0 is the permeability of air and μ_{Fe} is the relative permeability of the iron core material. Substitution into Ampère's circuital law gives

$$N \cdot i = \frac{B_{Fe}}{\mu_0 \mu_{Fe}} \cdot l_{Fe} + \frac{B_\delta}{\mu_0} \cdot l_\delta \tag{5.7}$$

Assuming uniform flux densities, the flux linkage is expressed as

$$\psi = N \quad B_{Fe} A_{Fe} = N \quad B_\delta A_\delta \tag{5.8}$$

where A_{Fe} and A_δ are the cross-sectional areas of the iron core and the air gap, respectively. Substituting the flux linkage into Ampère's circuital law gives

$$N \cdot i = \frac{\psi}{\mu_0 N} \cdot \left(\frac{l_{Fe}}{\mu_{Fe} A_{Fe}} + \frac{l_\delta}{A_\delta} \right) \tag{5.9}$$

Fringing flux in the vicinity of the air gap is neglected, which is equivalent to

$$B_\delta = B_{Fe} = B \Leftrightarrow A_\delta = A_{Fe} \tag{5.10}$$

This gives

$$N \cdot i = \frac{\psi}{\mu_0 A_{Fe} N} \cdot \left(\frac{l_{Fe}}{\mu_{Fe}} + l_\delta \right) \tag{5.11}$$

Rearranging this, gives an expression for the inductance according to

$$L = \frac{\psi}{i} = \frac{\mu_0 A_{Fe} N^2}{\frac{l_{Fe}}{\mu_{Fe}} + l_\delta} \tag{5.12}$$

In most cases the relative permeability of the iron core is high and the air gap long, i.e.

$$l_\delta >> \frac{l_{Fe}}{\mu_{Fe}} \tag{5.13}$$

which implies that the inductance is approximately given by

$$L = \frac{\mu_0 A_{Fe} N^2}{l_\delta} \tag{5.14}$$

Inductor core size selection

One of the first steps in the design of an inductor is to select an appropriate core size. The problem of selecting the core size arises from the fact the geometrical properties of a core are composed in a wide variety of ways. It is thus desirable to select the core size in a formalised way, by assigning a general quantity to a core that is not depending on its actual geometrical shape, i.e. the ratio between the geometrical measures of the core. A common way to accomplish is by selecting the core based upon the desired area product. Here, this method is reviewed from [14].

The peak flux linkage is expressed as

$$\hat{\psi} = L\hat{i}_m = N \quad A_{Fe}\hat{B}_m \tag{5.15}$$

where i_m is the magnetising current. The winding window A_w of a core is in the case of a single inductor coil expressed as

$$A_w = \frac{N A_{Cu}}{k_{Cu}} \tag{5.16}$$

where A_{Cu} is the winding copper conductor cross-sectional area and k_{Cu} is the copper fill factor, expressing how tightly wound the inductor coil is. The copper conductor RMS current I_{Cu} is expressed as

$$I_{Cu} = A_{Cu} J_{Cu} \tag{5.17}$$

where J_{Cu} is the corresponding current density. Substitution gives

$$L\hat{i}_m = k_{Cu} \frac{J_{Cu}}{I_{Cu}} \hat{B}_m A_w A_{Fe} \tag{5.18}$$

The area product, *AP*, of a core is defined according to

$$AP = A_w A_{Fe} \tag{5.19}$$

Substitution and rearranging the terms gives

$$AP = \frac{L\hat{i}_m I_{Cu}}{k_{Cu}\hat{B}_m J_{Cu}} \tag{5.20}$$

For an inductor, the copper conductor current and the magnetising currents are equal from a design point of view, i.e.

$$i_m(t) = i_{Cu}(t) \tag{5.21}$$

Thus, in this case the area product is written

$$AP = \frac{L\hat{i}_{Cu} I_{Cu}}{k_{Cu} \hat{B} J_{Cu}} \tag{5.22}$$

The last result is also found in [14]. If several inductor coils are wound upon a single core, like for a regenerative snubber [24], this expression to be slightly modified. This is done by assuming equal current density of each of the inductor coils. This means that the winding window area A_w is equally divided among the winding coils. Thus, the winding window area available for one coil is written

$$A_{w1} = \frac{A_w}{N_w} = \frac{N \quad A_{Cu}}{N_w k_{Cu}} \tag{5.23}$$

where N_w is the number of windings, i.e. the number inductors, wound upon the core. Note that, similar to the case of a transformer, the inductance and both the magnetising and copper conductor currents must be referred to the same coil. The *AP* value for a core with several inductor coils thus becomes

$$AP = N_w \frac{L_k \hat{i}_{m,k} I_{Cu,k}}{k_{Cu} \hat{B}_m J_{Cu}} \tag{5.24}$$

where the inductance and currents are referred to coil *k*.

Transformer core size selection

For transformers the required AP-value is determined in a similar way. In the case of transformers acting as inductors, for example in flyback converters and regenerative snubbers, the design is actually exactly the same as for inductors but with several windings. For regular transformers, the apparent power *S* is used as a design parameter instead of the inductance for core size selection. Furthermore, the number of windings N_w is always higher than one. For transformers the necessary AP value for the core is determined in a similar way but starting from the apparent power, S. In other aspects the method of core size selection derived below is similar to the one utilized for inductors.

For sinusoidal voltage and current, the apparent power *S* is determined by

$$S = V \cdot I \tag{5.25}$$

where the rms voltage *V* and the rms current *I* are referred to the same winding. The flux-linkage derivative with respect to time is

$$\frac{d\Psi}{dt} = e = v - R_{Cu} \cdot i \approx v \tag{5.26}$$

where e is the back-emf and R_{Cu} is the copper winding resistance. Since flux-linkage is the product of core flux and number of winding turns (still for the same physical winding) and core flux is the product of flux density and core cross-sectional area, the flux-linkage time derivative can be expressed as

$$\frac{d\Psi}{dt} = N \cdot A_{Fe} \cdot \frac{dB}{dt} \approx v \tag{5.27}$$

If the voltage is sinusoidal, the flux-density will also be sinusoidal which means that

$$V = N \cdot A_{Fe} \cdot \omega \cdot \frac{\hat{B}}{\sqrt{2}} \tag{5.28}$$

where $\omega = 2 \cdot \pi \cdot f$ is the angular frequency of the sinusoidal quantities. The winding current is written as

$$I = I_{RMS} = A_{Cu} \cdot J_{RMS} \tag{5.29}$$

where A_{Cu} is the cross-sectional area and J_{RMS} the rms current density of the winding conductor. Since there are N_w windings, the winding window is split between each winding. The winding windows for each winding $A_{w,i}$, i=1,..,N_w are equal in size, which implies that the current density of each winding conductor is also the same.

$$\begin{aligned} &A_{w,1} = A_{w,2} = \ldots = A_{w,N_w} = \frac{A_w}{N_w} \quad \Leftrightarrow \\ &J_{RMS,1} = J_{RMS,2} = \ldots = J_{RMS,N_w} = J_{RMS} \end{aligned} \tag{5.30}$$

Therefore, the total winding window, i.e. the one of the core, is given by

$$A_w = N_w \cdot \frac{N \cdot A_{Cu}}{k_{Cu}} \tag{5.31}$$

This means that the winding current of a particular winding with N turns is expressed as

$$I = \frac{k_{Cu} \cdot A_w}{N \cdot N_w} \cdot J_{RMS} \tag{5.32}$$

The apparent power is thus expressed as

$$S = \frac{1}{N_w} \cdot k_{Cu} \cdot J_{RMS} \cdot A_w \cdot A_{Fe} \cdot \frac{dB}{dt} = \frac{1}{N_w} \cdot k_{Cu} \cdot J_{RMS} \cdot AP \cdot \omega \cdot \frac{\hat{B}}{\sqrt{2}} \tag{5.33}$$

which means that for sinusoidal voltage and current, the area product of a transformer core should be selected from

$$AP = \frac{\sqrt{2} \cdot S \cdot N_w}{k_{Cu} \cdot J_{RMS} \cdot \omega \cdot \hat{B}} \tag{5.34}$$

From the expression above it is obvious that the higher the frequency, the smaller the required core size is. However, note that if the frequency is increased, the magnetizing losses also increases which implies that the larger core might still be needed to obtain a sufficient area of the cooling surface. On the other hand, a different type of material can be used instead but the saturation flux density of a high frequency material is in most cases lower than for a low frequency material.

If the current and voltage is not sinusoidal, for example in a switch-mode application, the relation between magnetic flux density and voltage has to be identified based on the application. For example if the winding voltage is square with amplitude V_{dc} and duty cycle equal to D, the peak magnetic flux density is obtained by integration

$$N \cdot A_{Fe} \cdot \frac{dB}{dt} \approx V_{dc} \quad \Leftrightarrow \quad \hat{B} = B_0 + \frac{V_{dc} \cdot D \cdot T_{sw}}{N \cdot A_{Fe}} \tag{5.35}$$

where B_0 is the initial magnetization and T_{sw} the switching period (note that the switching frequency is $f_{sw} = 1/T_{sw}$). The power transferred in, for example, a forward converter is

$$P = V_{dc} \cdot D \cdot I \tag{5.36}$$

where the average load current I is referred to the primary, i.e. the actual load current is $N \cdot I$. For a forward converter B_0 is equal to zero which gives

$$\begin{aligned} P &= \frac{k_{Cu} \cdot A_w}{N \cdot N_w} \cdot J_{RMS} \cdot V_{dc} \cdot D = \frac{k_{Cu} \cdot A_w}{N \cdot N_w} \cdot J_{RMS} \cdot \frac{N \cdot A_{Fe} \cdot \hat{B}}{T_{sw}} = \\ &= \frac{1}{N_w} \cdot k_{Cu} \cdot J_{RMS} \cdot AP \cdot f_{sw} \cdot \hat{B} \end{aligned} \tag{5.37}$$

This means that

$$AP = \frac{P \cdot N_w}{k_{Cu} \cdot J_{RMS} \cdot f_{sw} \cdot \hat{B}} \tag{5.38}$$

Again it is obvious that if the switching frequency is increased, i.e. the switching period reduced, the core size may be reduced. However, the same precautions as for sinusoidal voltage and current are still valid.

Core configurations of inductive elements

As described in the previous section, the area product *AP* is utilized to select appropriate core size. The most straightforward to find a suitable core is to calculate *AP* for one and then decide if the needed core should be larger or smaller, and continue until a core with almost the desired *AP* is found. If design of inductive components is made regularly, then *AP* should be calculated for a large number of cores and stored in tabular form since this makes core selection easier. Some core manufacturers also provide *AP* values for their cores. In [11] it is pointed out that a core with a somewhat lower *AP* value than required should be selected but to be sure a core with a somewhat larger *AP* value should be selected. The reason for this is that the peak flux density is included in the calculation of the *AP* value, which complicates the design process. Since not only the saturation flux density but also the allowable temperature increase affects the maximum peak flux density due to the correlation between core losses and peak flux density and frequency. Therefore, design of inductive components is in most cases an iterative process.

Once a required initial value of the area product is obtained, the number of winding turns *N* should be decided. In the case of cores with air gap the number of winding turn is also only a starting value since including an air gap implies that the number of winding turns must be reduced to obtain a certain inductance. Note that a reduction of the number of winding turns results in an increase of the peak flux density, which is the reason why a core with a somewhat larger area product than required should be selected. To select the starting value of the number winding turns the following expression is used

$$\hat{\Psi} = L \cdot \hat{i} = N \cdot A_{Fe} \cdot \hat{B} \tag{5.39}$$

$$N = \frac{L \cdot \hat{i}}{A_{Fe} \cdot \hat{B}} \tag{5.40}$$

As just mentioned, the introduction of an air gap will affect the number of winding turns. The reason for this is that the flux fringing in the vicinity of the air gap causes an increase of the cross-sectional area of the air gap, i.e. perpendicular to the direction of the magnetic flux which means that $A_\delta > A_{Fe}$. In [11] an expression for calculating the ratio between the cross-sectional areas of the effective air gap and the core is provided. This ratio is referred to as the flux fringing factor k_{FF} and is different for different core geometries.

Laminated EE and EI cores

Laminated cores are in most cases manufacture as EE or EI cores shown in Figure 5.2.

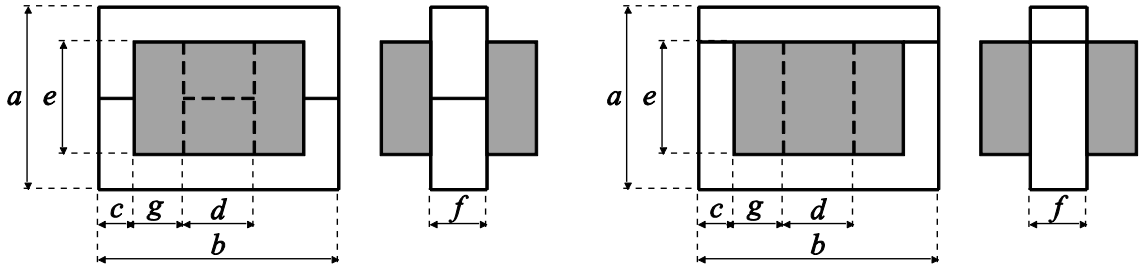


Figure 5.2: EE core (left) and EI core (right) with windings (grey) and geometrical dimensions.

Laminated cores occupy an exceptional position among cores in the sense that there is a wide variety in the geometrical dimensions depending on application and manufacturer. This is due to the fact that laminated cores are often designed for their intended application and manufactured in coil winder workshops. Consequently, it is sometimes difficult for the uninitiated to translate the area product to suitable geometrical proportions. One guideline to help the uninitiated is that bobbins (winding formers) for smaller cores are standardized.

In [14] normalized relative dimensions for EE cores, optimized for minimum volume, are presented. These normalized dimensions are provided in Table 5.1 where the dimensions refer to the ones in Figure 5.2. The effective geometrical dimensions for ferrite EE and EI cores are specified in manufacturer data sheets and therefore *AP* is easy to calculate in case of ferrite cores.

Table 5.1: Optimum dimensions for EE and EI cores according to [14]. The area product is *AP=3.75d0*. If the fact that the bobbin decreases the effective area of the winding window the area product is *AP*=2.1 *d0* according to [14]. A reduction of the copper fill factor *kCu* is in most cases used to compensate for the bobbin

a	*3.5d0*
b	*4d0*
c	*0.5d0*
d	*d0*
e	*2.5d0*
f	*1.5d0*
g	*d0*

For any EE or EI core the following holds

$$A_{Fe} = k_{Fe} \cdot d \cdot f \tag{5.41}$$

$$A_{Bobbin} = k_{Bobbin} \cdot e \cdot g \tag{5.42}$$

where *d*, *e*, *f* and *g* are dimension according to Figure 5.2 and k_{Fe} and k_{Bobbin} are constants less than unity. Since the laminations are covered with shell lack to reduce the electrical conductivity, the effective core cross-sectional area is less

than the one given by the geometrical dimensions. This reduction is considered as constant except that it depends on the thickness of the laminations. The reduction is quantified as the stacking factor k_{Fe} and typical values, from are provided in Table 5.2. The other scaling factor k_{Bobbin} is included to point out that the effective winding window is reduced due to the bobbin. This scaling factor is not included in any literature but should be either determined based on experience or be treated as included in the copper fill factor k_{Cu} of the winding.

Table 5.2: Typical stacking factor *kFe* for EE and EI cores as a function of the lamination thickness.

Lamination thickness [mm]	k_{Fe}
0.013	0.50
0.025	0.75
0.05	0.85
0.10 - 0.25	0.90
0.27 - 0.36	0.95

To utilize the core dimensions effectively the dimensions are in most cases selected as $d = 2c = a - e$ as in Table 5.1 above. This implies that the flux density is equal in all parts of the core, neglecting flux displacements in corners etc. This choice of dimensions also results in that the air gap is the sum of the centre limb air gap (width d) and one of the outer limb (width c) air gaps. This requires that both the outer limb air gaps are equally long but it is not necessary that they are equally long as the centre limb air gap. For ferrite EE cores and other similar core configurations the entire air gap is usually located in the centre limb of the core.

The required length of the air gap is given by

$$l_\delta = \frac{\mu_0 \cdot A_{Fe} \cdot N^2}{L} \tag{5.43}$$

if $l_\delta >> l_{Fe}/\mu_{Fe}$. The flux fringing factor for EE and EI cores is

$$k_{FF} = 1 + \frac{l_\delta}{\sqrt{A_{Fe}}} \cdot \ln\left(\frac{2 \cdot e}{l_\delta}\right) \tag{5.44}$$

according to [11]. As discussed earlier, the number of winding turns should be adjusted after introducing an air gap. The number of winding turns is now given by

$$N = \sqrt{\frac{l_\delta \cdot L}{\mu_0 \cdot A_{Fe} \cdot k_{FF}}} \tag{5.45}$$

As a result of the reduction in number of winding turns, the peak flux density will increase

$$\hat{B} = \frac{L \cdot \hat{i}}{A_{Fe} \cdot N} \tag{5.46}$$

Note that the peak flux density must be below (in most applications well below) the saturation flux density. Also the core losses increase due to the increase in the peak flux density.

The losses are of inductive elements are often divided into three separate parts; iron losses (hysteresis and eddy current losses) P_{Fe}, air gap losses (in the core close to the air gap) P_{Gap} and copper losses (in the winding) P_{Cu}. The iron losses are obtained from loss curves or empirical expressions provided by core and core material manufacturers. To calculate the iron and air gap losses, the current spectrum should be known since in the method presented here the loss components for each frequency are added to obtain the total losses. Thus

$$P_{Fe} = \sum_i P_{Fe,i} \tag{5.47}$$

where i=1,2,3, … is the harmonic order. Note that this method is not entirely accurate, since superposition is only valid for linear systems whereas the core material exhibits both hysteresis and saturation effects. For magnetic materials, current harmonics results in minor hysteresis loops on top of the fundamental and to calculate the losses of the harmonics more advanced methods of loss calculations have to be utilized [20]. Nevertheless, some manufacturers claim that their expressions for loss calculations hold also for superposition of loss contributions of harmonics [13].

The air gap losses can be estimated by approximates expressions [11]. The approximate air gap losses for EE and EI cores given in [11] are

$$P_{Gap,i} = 1550 \cdot d \cdot l_\delta \cdot f_i \cdot \hat{B}_i^2 \tag{5.48}$$

for each harmonic component i. The total air gap losses are thus

$$P_{Gap} = \sum_i P_{Gap,i} \tag{5.49}$$

The copper losses are resistive and to calculate these, the total length of the winding conductor must be known. To simplify the calculation, the mean length

per winding turn, *MLT*, is estimated. In the expressions for the *MLT* given here a full bobbin is assumed. Also, the total RMS current is needed to calculate the copper losses

$$I_{RMS} = \sqrt{\sum_i I_{RMS,i}^2} \tag{5.50}$$

which gives the copper losses

$$P_{Cu} = R_{Cu} \cdot I_{RMS}^2 \tag{5.51}$$

The copper resistance is

$$R_{Cu} = \rho_{Cu} \cdot \frac{N \cdot MLT}{A_{Cu}} \tag{5.52}$$

where ρ_{Cu} is the resistivity and A_{Cu} the cross-sectional area of the copper conductor. The mean length per winding turn is given by

$$MLT = 2 \cdot d + 2 \cdot f + 4 \cdot g \tag{5.53}$$

with reference to the dimensions in Figure 5.2.

The enclosing surface of an inductive component is used for calculation of the temperature rise provided that the losses are known. This calculation is not very accurate which will be commented later and therefore there is no meaning in calculating the enclosing surface with a high degree of accuracy either. Therefore the corners of the winding and core are assumed to be sharp and not having any smoother curvature. Furthermore it is assumed that the bobbin is full, i.e. that the winding covers the entire winding window. This is not a conservative restriction since we assume that we have selected a core with an appropriate *AP* value.

For EE och EI cores the enclosing surface of the core (not covered by the winding) and the winding are given by

$$A_{\Psi,Fe} = 4 \cdot c \cdot (b + e) + 2 \cdot f \cdot (a + b) \tag{5.54}$$

$$A_{\Psi,Cu} = 4 \cdot g \cdot (d + e + 2 \cdot g) + 2 \cdot e \cdot (d + 2 \cdot g) \tag{5.55}$$

To estimate the temperature rise, the power flow (unit W/m^2) through the core and winding surfaces is determined. In this estimation it is important to divide the core losses in a reasonable way since a part of the heat created in the core will be transported through the winding. Also it is important to decide if the winding should cover the air gap or not. It is good practice to not let the winding

cover the air gap in case of laminated cores since the air gap losses, which are local in nature, is far easier to cool then. The expressions for the power flow through the iron-to-air and the copper-to-air surfaces in the case that the winding covers the air gap are

$$\Psi_{T,Fe} = \left(\frac{b}{a+b-d} \cdot P_{Fe} \right) \cdot \frac{1}{A_{\Psi,Fe}} \tag{5.56}$$

$$\Psi_{T,Cu} = \left(\frac{e}{a+b-d} \cdot P_{Fe} + P_{Gap} + P_{Cu} \right) \cdot \frac{1}{A_{\Psi,Cu}} \tag{5.57}$$

These power flows through the enclosing surface of the inductor together with temperature curves shown in a later example in this chapter gives the temperature rise. Note the difficulty in determining the surface which the air gap losses are cooled through, due to the local nature of the air gap losses. In some designs, the air gap losses are very high and as a consequence of this difficulty the temperature rise is very complicated to estimate. Most likely the enclosing surface of the copper winding should be used for calculation of temperature rise instead of the small part of the core in the vicinity air gap not covered by the winding, since otherwise the estimated temperature rise would be huge.

Tape wound C cores

Tape wound C cores are used in applications where a high *AP* value is required together with a fine lamination, i.e. in high power and high frequency applications. The lamination thickness is determined by the frequency in such a way that the iron losses can be kept on an acceptable level. The lamination thickness for tape wound C cores may be a low as 0.025 mm.

Iron powder or ferrite core materials are often optimized for high frequency. However, these cores are complicated to manufacture for large geometrical dimensions, i.e. high AP-values. Laminated EE and EI cores are easy to manufacture for low frequency but for high frequency the laminations must be thin to keep the iron losses at acceptable levels. Laminated cores with thin laminations would result in complicated manufacturing since many laminations are required and low stacking factor (Table 5.2) resulting in a poor effective cross-sectional area of the core. In this type of applications tape wound C cores are used instead. For tape wound C cores the laminations are actually a steel band or tape (to indicate that it is very thin) that is wound upon a rectangular (with smooth corners) core former so that a rectangular (with rounded corners) core is formed. This core is then cut so that two C shaped halves are obtained. Bobbins and air gap spacers can easily be mounted. A metal or plastic band keeps the C cores together. For C cores the stacking factor is much higher than for laminated cores since the steel tape of the core is wound with high tension. A

sketch, including geometrical dimensions, of a inductive component manufactured from C cores is shown in Figure 5.3.

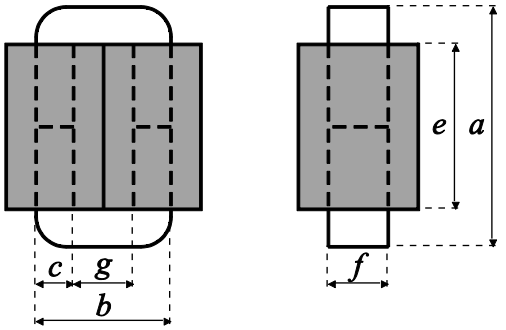


Figure 5.3: Inductive component based on two C core halves with windings (grey) and geometrical dimensions. Note that two windings are used.

The main difficulty with inductive components made from laminations or steel tape is that long air gaps may result in high air gap losses. The air gap losses results from the component of the fringing flux in the vicinity of the air gap that is perpendicular to the laminations or steel tape (Figure 5.7). This flux component causes very high eddy current losses locally. The problem could be overcome by cutting the laminations close to the air gap or by introducing another material for example iron powder in the air gap. The fringing flux also results in eddy current losses in the copper conductors of the winding, which is often handled by splitting the winding in such a way that it does not cover the air gap. As already mentioned, this also has the advantage that the air gap is easier to cool.

In the same way as for EE and EI cores the stacking factor k_{Fe} of tape wound C cores depends on the tape thickness. Typical values of k_{Fe} are found in [15], and reproduced in Table 5.3.

Table 5.3 Typical stacking factor *kFe* for tape wound C cores as a function of the steel tape thickness [15].

Lamination thickness [mm]	k_{Fe}
0.025	0.83
0.05	0.89
0.1	0.90
0.18	0.92
0.22 - 0.3	0.95

The design of inductive components based on tape wound C cores is similar to the one based on EE and EI cores. Therefore, the comments are given for EE and EI cores are valid also in this case.

The effective cross-sectional area of the core is

$$A_{Fe} = k_{Fe} \cdot c \cdot f \tag{5.58}$$

The length of the air gap is (if $l_\delta >> l_{Fe}/\mu_{Fe}$):

$$l_\delta = \frac{\mu_0 \cdot A_{Fe} \cdot N^2}{L} \tag{5.59}$$

The flux fringing factor in the case of a C core is

$$k_{FF} = 1 + \frac{l_\delta}{\sqrt{A_{Fe}}} \cdot \ln\left(\frac{2 \cdot e}{l_\delta}\right) \tag{5.60}$$

As earlier pointed out the number of winding turns must be adjusted after introducing an air gap to obtain the correct inductance

$$N = \sqrt{\frac{l_\delta \cdot L}{\mu_0 \cdot A_{Fe} \cdot k_{FF}}} \tag{5.61}$$

The peak flux density increases due to the decresed number of winding turns, according to

$$\hat{B} = \frac{L \cdot \hat{i}}{A_{Fe} \cdot N} \tag{5.62}$$

The flux density of each current harmonic is calculated from the previous expression and from the flux density harmonics the iron and air gap losses are calculated

$$P_{Fe} = \sum_i P_{Fe,i} \tag{5.63}$$

$$P_{Gap,i} = 388 \cdot f \cdot l_\delta \cdot f_i \cdot \hat{B}_i^2 \tag{5.64}$$

$$P_{Gap} = \sum_i P_{Gap,i} \tag{5.65}$$

Note that the empirically derived constant (from [11]) in the expression for calculating the air gap losses of C cores is much less than the corresponding for EE and EI cores.

The copper losses are calculated in the same way as for EE and EI cores. However, note that the calculation of *MLT* given here is only valid for the case with two windings and full bobbins.

$$P_{Cu} = R_{Cu} \cdot I_{RMS}^2 \tag{5.66}$$

where

$$I_{RMS} = \sqrt{\sum_i I_{RMS,i}^2} \tag{5.67}$$

and

$$R_{Cu} = \rho_{Cu} \cdot \frac{N \cdot MLT}{A_{Cu}} \tag{5.68}$$

$$MLT = 2 \cdot c + 2 \cdot f + 2 \cdot g \tag{5.69}$$

For C cores with two windings and full bobbins the following enclosing surfaces can also be derived

$$A_{\Psi,Cu} = 4 \cdot e \cdot (c+g) + 2 \cdot e \cdot (f+g) + 4 \cdot g \cdot (c+g) + 2 \cdot f \cdot g \tag{5.70}$$

$$A_{\Psi,Fe} = 4 \cdot b \cdot c + 2 \cdot b \cdot f + 4 \cdot c \cdot f \tag{5.71}$$

The power flow through the core-to-air and winding-to-air interfaces is given by

$$\Psi_{T,Fe} = \left(\frac{b}{b+e} \cdot P_{Fe} \right) \cdot \frac{1}{A_{\Psi,Fe}} \tag{5.72}$$

$$\Psi_{T,Cu} = \left(\frac{e}{b+e} \cdot P_{Fe} + P_{Gap} + P_{Cu} \right) \cdot \frac{1}{A_{\Psi,Cu}} \tag{5.73}$$

provided that the core losses that are generated in a part of the core that is covered by the winding is also transferred through the winding to become a part of the power flow through the winding-to-air interface.

As for EE and EI cores the power flow in the vicinity of the air gap is complicated to calculate in the case that the winding is split to not cover the air gap.

Iron powder toroid cores

Commercially available iron powder toroid cores are only manufactured in smaller sizes, since the press tools required to manufacture larger cores are costly. As a consequence of the oxide layer covering the iron powder grains, iron powder cores possess a distributed air gap. In design of inductive components based on iron powder toroid cores this distributed air gap is the only air gap present, i.e. the core is not split up in parts to introduce an air gap.

Instead, an appropriate core material with an explicit permeability is selected so that the inductance equals the required. At first this seems a bit unfeasible since the larger the grains the higher the permeability and at the same time the smaller the grains the better the high frequency characteristics for all iron powder materials. However, if a certain high frequency inductor design requires a high permeability resulting in a material with poor high frequency characteristics the losses will be high calling for a core with larger *AP* since the peak flux density must be reduced. This results in a larger core requiring a lower permeability and therefore a material with more favourable high frequency characteristics is selected. This means that even if seems to be a complicated way of selecting appropriate iron powder material the method will converge towards a few materials and core sizes to select from. For the experienced inductive component designer it is not so complicated since he or she knows the suitable peak flux density for certain frequency and therefore also suitable *AP* values.

Since there is no additional air gap there can be no winding former or bobbin present, which of course makes manufacturing of the winding more complicated and costly than for other geometries. A sketch of a toroid core with geometrical dimensions is shown in Figure 5.4.

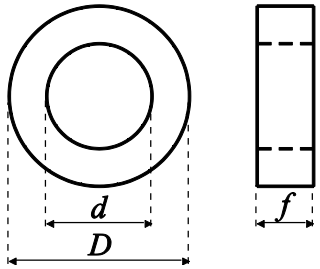


Figure 5.4: Toroid core and its geometrical dimensions. Note that the winding is not included.

The required number of winding turns is determined by

$$N = \frac{L \cdot \hat{i}}{A_e \cdot \hat{B}} \tag{5.74}$$

as for the configurations presented earlier. The expression for calculation of inductance (neglecting hysteresis and saturation effects)

$$L = \frac{\mu_0 \cdot \mu_e \cdot A_e \cdot N^2}{l_e} \tag{5.75}$$

where μ_e, A_e and l_e are the effective relative permeability, cross-sectional area of the core and core length, respectively. Rearranging gives the desired effective relative permeability according to

$$\mu_e = \frac{l_e \cdot L}{\mu_0 \cdot A_e \cdot N^2} \tag{5.76}$$

Inductive components based on toroid cores are often designed so that only a fraction of the usable part of the winding window is used, for example in single layer windings. Note that in such a case it is rather complicated base the design on the area product since the copper fill factor is used to determine the required area product. Furthermore, the enclosing surface of the inductive component must also be calculated from the specific design.

Ferrite cores

For the electro-magnetic design of inductive components based on ferrite cores equivalent properties normalized as if the core was a toroid, i.e. the electro-magnetic calculations are made as if the core was a toroid even if it is not. This method should for example be used for the core data given in [13]. This means that the actual air gap of the ferrite core is included in the equivalent relative permeability μ_e of the particular core for the particular material. The actual permeability of the ferrite (without any explicit air gap) is termed initial permeability μ_i. Also, for both iron powder cores and ferrite cores the core manufacturer often provides a quantity called the A_L value, which denotes the inductance per winding turn squared. Design of inductive components based on ferrite cores is essentially the same as for other materials and geometries and therefore it is not addressed again.

The inductance is calculated from the A_L value according to

$$L = \frac{\mu_0 \cdot A_e \cdot N^2}{\frac{l_e}{\mu_i} + l_\delta} = A_L \cdot N^2 \tag{5.77}$$

which means that the A_L value is equal to

$$A_L = \frac{\mu_0 \cdot \mu_e \cdot A_e}{l_e} \tag{5.78}$$

The equivalent permeability is equal to

$$\mu_e = \frac{\mu_i}{1 + \mu_i \cdot \frac{l_\delta}{l_e}} \tag{5.79}$$

Small deviations between the A_L value given in manufacturers data sheets and the ones that are calculated with the expression above are likely, which is due to

the fact that the A_L value given in data sheets is found from experiments and therefore takes flux fringing into account.

As for iron powder toroid cores the winding window is often only partly utilized, which results in the same shortcomings as in the case of iron powder cores. Ferrite cores are manufactured in a wide variety of different configurations including the ones presented earlier. Therefore, the enclosing surface for temperature rise calculations are not given here but should of course be derived in the design process.

Inductor design example

In this section one way of designing a certain type of inductors is given through an example. The inductor designed is actually used in the line side LCL-filter of the battery charger presented in [8]. The design is based on a tape wound C-core. An advantage of tape wound cores is that a high stacking factor is obtained even though a thin steel tape (0.025-0.3 mm) is used.

The specification of the inductor is given in Table 5.4.

Table 5.4 Inductor specification.

L	0.3 mH
I_{RMS} @ 1 kHz	120 A
I_{PEAK} @ 5 kHz	10 A
I_{PEAK} @ 10 kHz	5 A

Thus, the total RMS current is approximately given by the 1 kHz component, i.e.

$$I_{Cu} \approx 120\,\text{A} \tag{5.80}$$

By using four parallel conductors of rectangular cross-section, 3×5 mm, a RMS current density equal to 2 A/mm^2 is obtained. The absolute maximum magnetising, and thus copper conductor, current is calculated according to

$$\hat{i}_{Cu} = 120\sqrt{2} + 10 + 5\,\text{A} = 185\,\text{A} \tag{5.81}$$

To select core size, an initial guess of a suitable peak magnetic flux density has to be done. Here

$$\hat{B} = 0.35\,\text{T} \tag{5.82}$$

is selected. Also, a reasonable guess on the copper fill factor k_{Cu} must be made. A common value is 0.4. If this is used it is found that

$$AP = \frac{L\hat{i}_{Cu} I_{Cu}}{k_{Cu} \hat{B} J_{Cu}} = 2379\,\mathrm{cm}^4 \tag{5.83}$$

The C-core TELMAG Su 150b (Figure 5.5), have geometrical properties according Table 5.5.

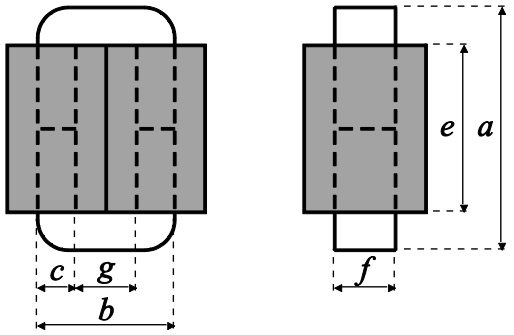


Figure 5.5: Inductor based on a C-core. The winding (grey) is split into two parallel connected windings.

Table 5.5: Geometry of the core Su 150b.

a	255.6 mm
b	150.2 mm
c	49.4 mm
d	76.2 mm
e	154.0 mm
g	50.0 mm

By including the stacking factor of the core with 0.1 mm tape, the iron core area is found

$$A_{Fe} = 0.968 \cdot 33.9\,\mathrm{cm}^2 = 32.8\,\mathrm{cm}^2 \tag{5.84}$$

The winding window area is approximated as

$$A_w \approx 15.4 \cdot 5.0\,\mathrm{cm}^2 = 77.0\,\mathrm{cm}^2 \tag{5.85}$$

Hence, the area product for this core becomes

$$AP = A_w A_{Fe} = 2527\,\mathrm{cm}^4 \tag{5.86}$$

An initial guess, i.e. without air gap, of the number of winding turns required, is given by

$$N = \frac{L\hat{i}_{Cu}}{A_{Fe}\hat{B}} = 48\,\text{turns} \tag{5.87}$$

The total air gap length is

$$l_\delta = \frac{\mu_0 A_{Fe} N^2}{L} = 32\,\text{mm} = 2 \cdot 16\,\text{mm} \tag{5.88}$$

The fringing flux factor, k_{FF}, is calculated as

$$k_{FF} = 1 + \frac{l_\delta}{\sqrt{A_{Fe}}} \cdot \ln\left(\frac{2e}{l_\delta}\right) = 2.265 \tag{5.89}$$

The fringing flux factor is then used to adjust the number of winding turns, to compensate for the fringing flux introduced in the vicinity of the air gap.

$$N = \sqrt{\frac{l_\delta \quad L}{\mu_0 A_{Fe} k_{FF}}} = 32\,\text{turns} \tag{5.90}$$

The reduced turns ratio implies that the peak magnetic flux density becomes higher than expected

$$\hat{B} = \frac{L\hat{i}_{Cu}}{A_{Fe}N} = 0.53\,\text{T} \tag{5.91}$$

The peak flux density is calculated for each frequency component and then the iron core loss for each component is found from the manufacturer data sheets, which for 0.1 mm tape gives

$$\begin{cases} \hat{B}_{1\,\text{kHz}} = 0.486\,\text{T} \\ \hat{B}_{5\,\text{kHz}} = 0.029\,\text{T} \\ \hat{B}_{10\,\text{kHz}} = 0.014\,\text{T} \end{cases} \Rightarrow \begin{cases} P_{Fe,1\,\text{kHz}} = 141\,\text{W} \\ P_{Fe,5\,\text{kHz}} = 10\,\text{W} \\ P_{Fe,10\,\text{kHz}} = 6\,\text{W} \end{cases} \tag{5.92}$$

If minor loops are neglected, the total iron core loss is found directly summing the loss associated with each current component, i.e.

$$P_{Fe} = \sum_i P_{Fe,i} = 157\,\text{W} \tag{5.93}$$

Another iron core loss component, air gap loss due to fringing flux components being perpendicular to the tape surface, introduces additional eddy current losses. In [11], the air gap losses are calculated from the empirically derived expression

$$P_{\delta,i} = 388 d l_\delta \quad f_i \quad \hat{B}_i^2 \tag{5.94}$$

The corresponding air gap losses, for the different current components, is found to be

$$\begin{cases} P_{\delta,1\,\mathrm{kHz}} = 212\ \mathrm{W} \\ P_{\delta,5\,\mathrm{kHz}} = 4\ \mathrm{W} \\ P_{\delta,10\,\mathrm{kHz}} = 2\ \mathrm{W} \end{cases} \tag{5.95}$$

Again, the total air gap losses are found by summing the components

$$P_\delta = \sum_i P_{\delta,i} = 218\ \mathrm{W} \tag{5.96}$$

which seems to be fairly high. To calculate the losses in the winding, referred to as copper losses, the mean length per turn, *MLT*, is calculated. For a C-core with a winding geometry according to Figure 5.5, i.e. two coils, *MLT* is given by

$$MLT = 2c + 2d + 2g = 351\,\mathrm{mm} \tag{5.97}$$

The total winding resistance of the inductor is calculated at a winding temperature of 90 °C. At this temperature the resistivity of copper, ρ_{Cu}, equals $2.156 \cdot 10^{-2}$ Ωmm^2/m. The winding resistance is thus

$$R_{Cu} = \rho_{Cu} \frac{N \cdot MLT}{A_{Cu}} = 4.04\,\mathrm{m\Omega} \tag{5.98}$$

The copper losses is calculated from

$$P_{Cu} = R_{Cu} I_{Cu}^2 = 58\ \mathrm{W} \tag{5.99}$$

In order to estimate the temperature rise of the winding, its surface area is calculated according to

$$A_{T,Cu} = 4e(c+g) + 2e(d+g) + 4g(c+g) + 2dg = 0.1276\,\mathrm{m}^2 \tag{5.100}$$

Also, the iron core surface area is calculated

$$A_{T,Fe} = 4bc + 2bd + 4cf = 0.0676\,\mathrm{m}^2 \tag{5.101}$$

The total thermal flux density through the copper winding, is calculated as

$$\Psi_{T,Cu} = \frac{1}{A_{T,Cu}} \left(\frac{e}{b+e} P_{Fe} + P_\delta + P_{Cu} \right) = 2782\ \mathrm{W/m}^2 \tag{5.102}$$

In the same manner, the thermal flux through the iron core surface not covered by the copper winding, is calculated

$$\Psi_{T,Fe} = \frac{1}{A_{T,Cu}}\left(\frac{b}{b+e}P_{Fe}\right) = 1154\ \mathrm{W/m^2} \tag{5.103}$$

According to the literature, for example [11][14], heat transfer is due to two physical processes, radiation and convection. Radiation follows the expression

$$\Psi_{T,rad} = 5.70\cdot 10^{-8}\varepsilon\left(T_s^4 - T_a^4\right) \tag{5.104}$$

where ε is the emissivity of the surface, and T_s and T_a are the surface and ambient temperatures, respectively. Heat transfer by convection is according to [11], expressed as

$$\Psi_{T,conv} = 2.17F\left(T_s - T_a\right)^{\eta}\sqrt{p} \tag{5.105}$$

where F is an air friction factor, η is a factor depending on the shape and orientation of the surface and p is the relative pressure. Note that other textbooks present different methods for calculating the convection heat transfer. The total heat transfer is given by the sum of the radiation and convection components, i.e.

$$\Psi_T = \Psi_{T,rad} + \Psi_{T,conv} \tag{5.106}$$

However, since the heat transfer by convection is complicated to model, heat transfer is instead calculated from experience. In [11] it is stated that the heat transfer is usually 55 % radiation and 45 % convection, which means that the total heat transfer can be approximated directly from the radiation component. The temperature rise at an ambient temperature of 40 °C, using this approximation, is shown in Figure 5.6. In Figure 5.6, another approximation is also shown where the temperature rise is calculated according to

$$T_s = T_a + \frac{\Psi_T}{\alpha_0 + k_\alpha T_a} \tag{5.107}$$

where the two constants are selected as α_0=12 W/m^2·°C and k_α=0.1 W/m^2·°C^2. As seen in Figure 5.6, both methods give similar results.

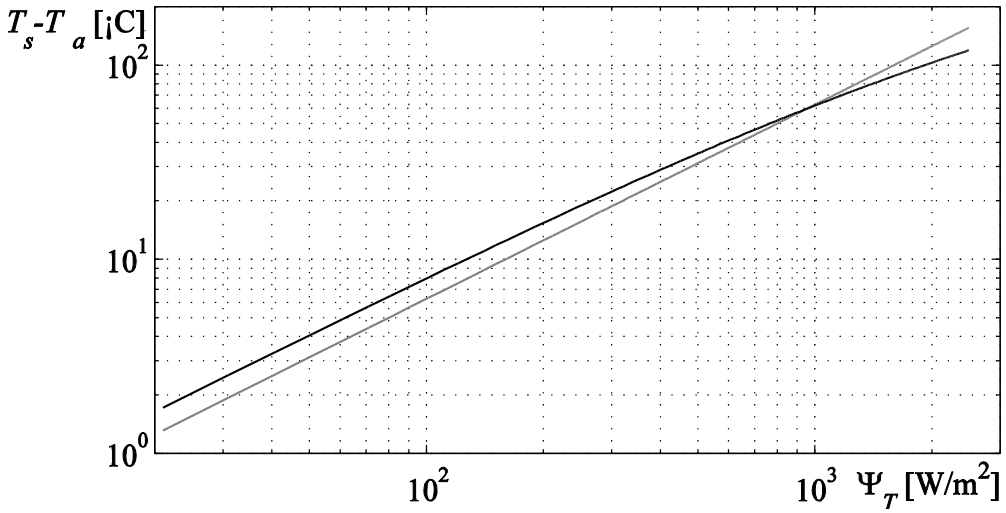


Figure 5.6: Calculated temperature rise at an ambient temperature of 40 ℃, based on radiated heat (black) and an approximate method (grey).

If simplified approximation is used, the temperature rise of the copper winding surface becomes

$$T_{s,Cu} - T_a = 174\ ^\circ\mathrm{C} \tag{5.108}$$

which is by far too high. The temperature rise of the iron core surface becomes

$$T_{s,Fe} - T_a = 72\ ^\circ\mathrm{C} \tag{5.109}$$

The calculated temperature rise of the copper winding surface is high, mainly due to the air gap losses. Often, the air gap losses are not included in the temperature calculation, which in this case gives

$$\Psi_{T,Cu} = \frac{1}{A_{T,Cu}}\left(\frac{e}{b+e}P_{Fe} + P_{Cu}\right) = 1074\ \mathrm{W/m^2} \tag{5.110}$$

$$T_{s,Cu} - T_a = 67\ ^\circ\mathrm{C} \tag{5.111}$$

Another question to be asked is how good the empirical air gap loss model is. However, when the inductor was implemented and tested, the copper winding surface close to the air gap was by far too hot, i.e. more than 130 °C.

To partially overcome this problem, the winding was split close to the air gap, in order to give a more efficient cooling of the iron core in the vicinity of the air gap, see Figure 5.7. Another advantage gained by splitting the winding, is that the eddy currents induced in the copper winding, are reduced. To some extent this solved the problem, but fan cooling was also needed.

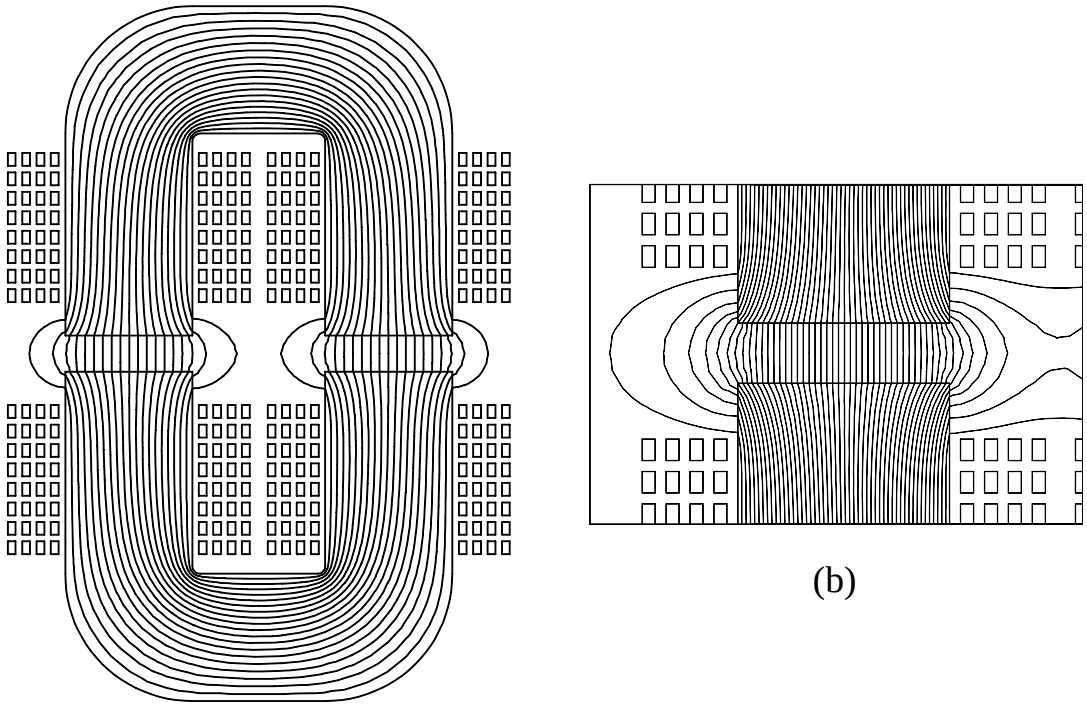


Figure 5.7: Magnetic flux lines (a) of the entire inductor, and (b) of the region around one of the air gaps. Note the component of the fringing flux that is perpendicular to the surface of the steel tape.

5.3 Capacitors

Different kinds of capacitors are used in power electronic circuits. Here, only two are discussed, the non-polarised metallized film polypropylene capacitor and the polarised wet aluminium electrolytic capacitor. Non-polarised capacitors are used in output filters and for applications with high capacitor voltage time derivative, like commutation circuits. Polarised capacitors are used when a high capacitance is needed, for example for DC link capacitors.

Design

Metallized film polypropylene capacitors have a thin plastic film to support the metal layer of the electrodes. The plastic used for the film can for example be polyester. If the plastic film has electrodes (of the same polarity) on both sides it is referred to as double metallized film. The dielectric consists of a polypropylene film. To avoid air pockets resulting in locally high electric field strength, the polypropylene film should be somewhat porous to be able to absorb oil, according to [23].

Wet aluminium electrolytic capacitors contain a fluid, the electrolyte, between the aluminium electrodes. The electrolyte is absorbed by paper in between the aluminium electrodes, in order to avoid air pockets. Since the electrolyte is

conductive, the aluminium electrodes are electrically close together, only separated by the dielectric of the capacitor. The dielectric constitutes of a thin aluminium oxide layer on the positive electrode.

Simulation model

Both the metallized polypropylene and the wet aluminium electrolytic capacitors are modelled by the capacitor equivalent shown in Figure 5.8.

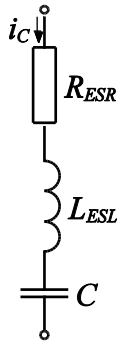


Figure 5.8: Capacitor simulation model.

In Figure 5.8 two parasitic elements are shown, L_{ESL} and R_{ESR}. The equivalent series inductance L_{ESL} is due to stray inductance of the leads and the metal layers forming the electrodes. A numerical value for L_{ESL} is often specified in the manufacturer data sheets. The equivalent series resistance R_{ESR} is due to the resistance of the leads and also dielectric losses. Therefore, R_{ESR} is frequency dependent according to

$$R_{ESR}(f) = R_s + \frac{\tan\delta_0}{2\pi f C} \tag{5.112}$$

where the term tan δ_0 is referred to as the dielectric dissipation factor. As the name hints, the dielectric dissipation factor is due to losses within the dielectric itself. This term is regarded as constant for the possible operating frequencies. However, since the voltage across the dielectric decreases with increasing frequency for a constant RMS value of the capacitor current i_C, the dielectric losses also decrease with increasing frequency as concluded from (5.112). The constant term R_s contained in the expression for $R_{ESR,}$ is due to the resistance of the leads and the electrodes.

However, since the equivalent series resistance is frequency dependent, it can not be used directly in an online simulation model. Instead, spectral analysis is used to calculate the RMS capacitor current for each frequency. Then the loss of the corresponding frequency is calculated from

$$P_{ESR}(f) = R_{ESR}(f) \cdot I_C^2(f) \tag{5.113}$$

The loss contribution for the different frequency components are summed, which gives the total losses P_{ESR}. Finally, an equivalent series resistance without frequency dependency is calculated according to

$$R_{ESR} = \frac{P_{ESR}}{I_C^2} \tag{5.114}$$

where I_C is the RMS value of the total capacitor current. Also in this case, the simulation model R_{ESR} is only used to influence the rest of the power electronic circuit. To calculate the capacitor losses, each frequency component is treated separately, as in expression (5.113).

6 Converter Loss Estimation and Thermal Design

Power loss in semiconductor devices is, from the user point of view, divided into two separate types, conduction and switching loss. Different types of semiconductors, e.g. thyristors, MOSFETs, IGBTs etc., have different on-state, blocking-state and switching characteristics and are therefore suitable for different applications. The user has to select appropriate semiconductors for a given application and converter specification. The thermal design is of course of great importance for reliable operation of the converter. Therefore, at least a rough estimation of the semiconductor losses has to be done, which is then verified in tests.

6.1 Loss estimation

Here, thermal design based on data sheet information, i.e. inductively clamped load current, is discussed. If the power semiconductors do not operate under inductively clamped current, the data sheet information is no longer valid. This could be the case for resonant converters or if the converter is equipped with RCD charge-discharge snubbers. In such a case, simulation has to be used to calculate the semiconductor losses. Figure 6.1 shows a step down converter with inductively clamped load current and without snubbers. This circuit is used to highlight the steps taken to perform a loss estimation based on data sheet information.

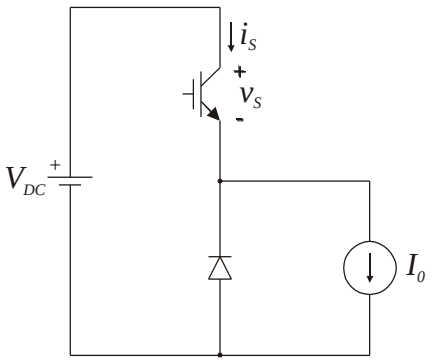


Figure 6.1: Step down converter used to illustrate loss estimation.

Figure 6.2 shows approximate switching waveforms for a generic switch *S*. The instantaneous losses are obtained by multiplication of the instantaneous voltage across the switch, v_s, and the instantaneous current, i_s, through the switch.

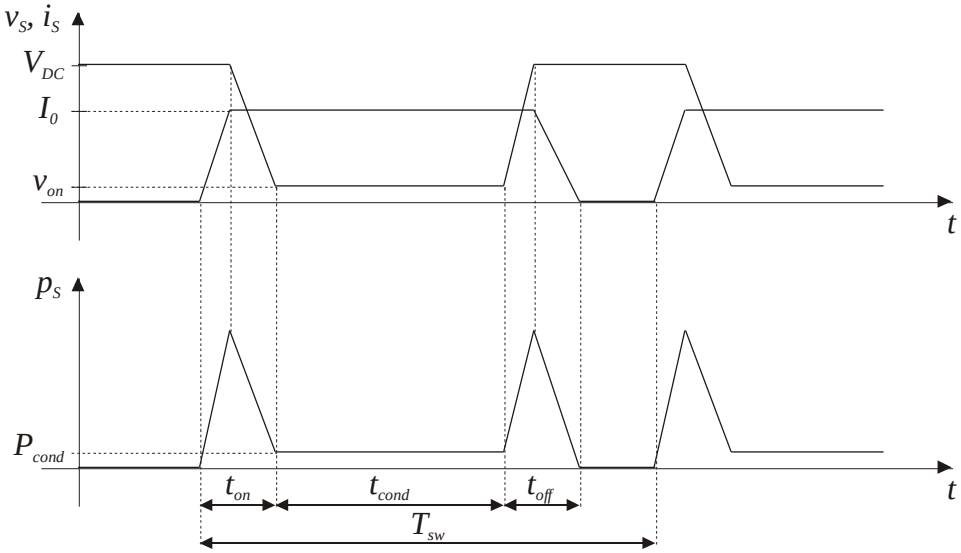


Figure 6.2: Approximate switching waveforms for the switch S.

$$p_S(t) = v_S(t) \cdot i_S(t)$$

The instantaneous loss is integrated for one switching period T_{sw}, yielding the energy loss for one switching period. The energy is divided into three parts; turn-on (index *on*), on-state or conduction (index *cond*) and turn-off (index *off*) losses.

$$E_S(T_{sw}) = \int_{T_{sw}} p_S(\tau) d\tau = E_{S,on}(T_{sw}) + E_{S,cond}(T_{sw}) + E_{S,off}(T_{sw}) \tag{6.1}$$

$$E_{S,on}(T_{sw}) = \int_{t_{on}} p_S(\tau) d\tau = V_{DC} \cdot I_0 \cdot \frac{t_{on}}{2} \tag{6.2}$$

$$E_{S,cond}(T_{sw}) = \int_{t_{cond}} p_S(\tau) d\tau = V_{S(on)} \cdot I_0 \cdot t_{cond} \tag{6.3}$$

$$E_{S,off}(T_{sw}) = \int_{t_{off}} p_S(\tau) d\tau = V_{DC} \cdot I_0 \cdot \frac{t_{off}}{2} \tag{6.4}$$

Note that $V_{S(on)}$ is current dependent, according to

$$V_{S(on)} = V_{S0} + R_S \cdot I_0 \tag{6.5}$$

where V_{S0} is the forward voltage at zero current and R_S is the resistance, both from the linearized *IV* characteristic. Note that R_S is often temperature dependent.

This amount of energy is transferred to heat and must be removed by means of cooling. Otherwise the silicon chip temperature (often referred to a as the junction temperature) will be too high. The typical maximum junction

temperature specified in data sheets is 150 °C, whereas 125 °C is often used for practical converter designs. The next step is to divide the energy losses with the switching time period T_{sw}, to give the average power losses.

$$P_S(T_{sw}) = \frac{E_S(T_{sw})}{T_{sw}} = P_{S,on}(T_{sw}) + P_{S,cond}(T_{sw}) + P_{S,off}(T_{sw}) \qquad (6.6)$$

$$P_{S,on}(T_{sw}) = \frac{E_{S,on}(T_{sw})}{T_{sw}} = E_{S,on}(T_{sw}) \cdot f_{sw} = \frac{V_{DC} \cdot I_0 \cdot t_{on}}{2} \cdot f_{sw} \qquad (6.7)$$

$$P_{S,cond}(T_{sw}) = \frac{E_{S,cond}(T_{sw})}{T_{sw}} = V_{S(on)} \cdot I_0 \cdot \frac{t_{cond}}{T_{sw}} = V_{S(on)} \cdot I_0 \cdot D_S \qquad (6.8)$$

$$P_{S,off}(T_{sw}) = \frac{E_{S,off}(T_{sw})}{T_{sw}} = E_{S,off}(T_{sw}) \cdot f_{sw} = \frac{V_{DC} \cdot I_0 \cdot t_{off}}{2} \cdot f_{sw} \qquad (6.9)$$

In the expression for the on-state losses a parameter denoted D_S shows up, which corresponds to the duty-cycle of the device in question. Note that this not have to be the same as the duty cycle for the output or load voltage, even though it is for the simple step down converter. In the case of a full bridge DC-DC converter the output and device duty cycles are not the same. By adding the turn-on and turn-off losses, the switching loss is calculated.

$$P_{S,sw}(T_{sw}) = P_{S,on}(T_{sw}) + P_{S,off}(T_{sw}) \qquad (6.10)$$

This means that an approximate loss calculation can be done. To do this the turn-on (t_{on}) and turn-off (t_{off}) times are needed. These usually vary with the DC link voltage and load current. Unfortunately they are in most cases not given in the data sheets. Usually only the current rise (t_{ri}) and fall (t_{fi}) times are given for a specified load condition. On the other hand the turn-on ($E_{on,n}$) and turn-off ($E_{off,n}$) energies are often specified in data sheets. These are given for a specific DC link voltage ($V_{DC,n}$) and load condition ($I_{0,n}$). These energies also include voltage rise and fall times. Therefore, to get a more accurate estimation of the switching losses, the expressions

$$E_{S,on}(T_{sw}) = \frac{E_{on,n}}{V_{DC,n} \cdot I_{0,n}} \cdot V_{DC} \cdot I_0 \qquad (6.11)$$

$$E_{S,off}(T_{sw}) = \frac{E_{off,n}}{V_{DC,n} \cdot I_{0,n}} \cdot V_{DC} \cdot I_0 \qquad (6.12)$$

should be used, if possible.

For the freewheeling diode in Figure 6.1 will also have some power loss. The most obvious are the conduction losses, which can be written:

$$P_{D,cond}(T_{sw}) = V_{D(on)} \cdot I_0 \cdot D_D \tag{6.13}$$

where

$$V_{D(on)} = V_{D0} + R_D \cdot I_0 \tag{6.14}$$

V_{D0} and R_D are the diode knee voltage and resistance from the linearized *IV* characteristic. Note that R_D is often temperature dependent. D_D is the duty-cycle for the freewheeling diode, which can be approximated as:

$$D_D \approx 1 - D_S \tag{6.15}$$

Except the conduction losses there will also be loss associated with reverse recovery. Since the voltage is high across the freewheeling diode (=V_{DC}) for the latter half of reverse recovery and the diode current at the same time high, there will be high loss. The reason for this reverse current is that there are excess carriers stored in the interior of the diode, mainly in the drift region, that must be swept out. To calculate the amount of the losses, the charge Q_f has to be estimated. Sometimes, the manufacturers of power semiconductor devices give this charge but most often only the total reverse recovery charge Q_{rr} is given. In such a case a reasonable estimation of the two times t_{rr1} and t_{rr2} (t_{rr1}+t_{rr2} = t_{rr}) is made, which gives

$$Q_f \approx \frac{1}{S+1} \cdot Q_{rr} \quad \text{where} \quad S = \frac{t_{rr1}}{t_{rr2}} \tag{6.16}$$

The estimate of the relation between t_{rr1} and t_{rr2} can often be based on what is known for other freewheeling diodes. As previously discussed, the reverse voltage across the diode will equal the DC link voltage under the time interval t_{rr2}, that is for the time that Q_f is swept out from its drift region. Reverse recovery takes place once each switch period, which gives

$$P_{D,rr} = V_{DC} \cdot Q_f \cdot f_{sw} \tag{6.17}$$

Sometimes the manufacturers gives the turn-off energy also for the freewheeling diodes, which in that case includes the losses associated with reverse recovery. The expression for the diode turn-off losses becomes

$$P_{D,off} = E_{D,off}(T_{sw}) \cdot f_{sw} \quad , \quad E_{D,off}(T_{sw}) = \frac{E_{off,n}}{V_{DC,n} \cdot I_{0,n}} \cdot V_{DC} \cdot I_0 \tag{6.18}$$

and

$$Q_f = \frac{Q_{f,n}}{I_{0,n}} \cdot I_0 \tag{6.19}$$

Remark: In the calculations above the loss contribution from the reverse recovery current superimposed on the load current for the switch *S* during turn-on is not taken into account. In most cases, transistors and freewheeling diodes are integrated in the same housing or module and therefore these losses are included in the specified $E_{S,\,on}$.

Now an analytical method for estimation of the semiconductor losses of a self-commutated three-phase VSC is described. The calculations are made for a single half-bridge, see Figure 6.3. Since the output from a three-phase VSC is a line-to-line voltage, the corresponding line-to-neutral voltage has to be calculated if only a single half-bridge is considered. The original sinusoidal voltage references are used even though they are phase potentials. This is applicable since the average losses are calculated only for the fundamental current component, i.e. neglecting the ripple.

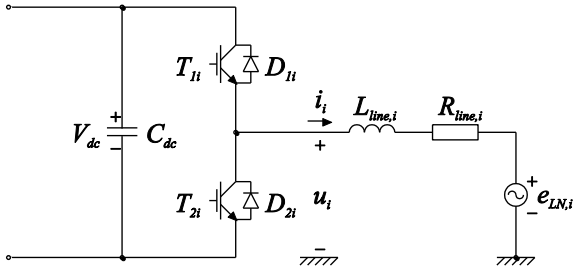


Figure 6.3: One half-bridge of a three-phase voltage source converter.

The calculations are based on data sheet information, which is usually only valid for inductive load currents. For a VSC utilising hard-switching, this is not a limitation since the AC side filters are inductive in this case. Furthermore, the calculations only take the fundamental current into account, i.e. the current ripple is not considered.

Figure 6.4 shows the fundamental components of the output voltage and current, for one half-bridge of the converter. The converter output current lags the voltage with an angle φ.

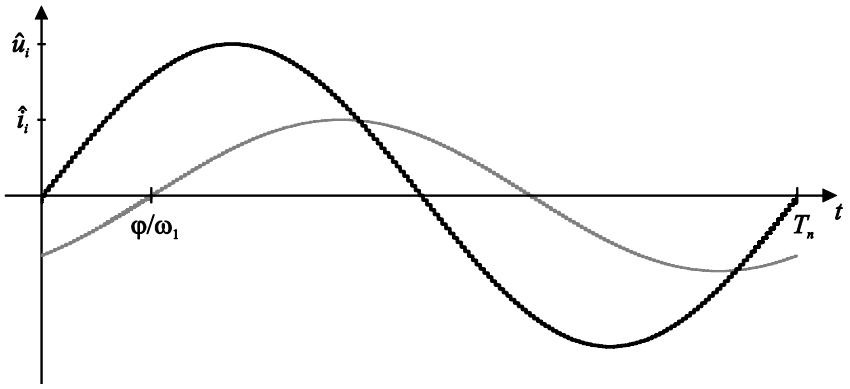


Figure 6.4: Converter output voltage and current. The current is displaced by an angle φ relative to the voltage.

For the IGBTs of one half-bridge, the switching losses are estimated from

$$\begin{aligned}\overline{P}_{Ti,sw} &= \frac{1}{T_n}\int_{T_n}\left(P_{on}+P_{off}\right)dt = \frac{f_{sw}}{T_n}\int_{T_n}\left(E_{on}+E_{off}\right)dt = \\ &= \frac{E_{on,n}+E_{off,n}}{V_{dc,n}\cdot I_n}\cdot\frac{V_{dc}\,f_{sw}}{T_n}\int_{T_n}\left|\hat{i}_i\sin\left(\omega_1 t-\varphi\right)\right|dt = \\ &= \frac{2\sqrt{2}}{\pi}\cdot\frac{E_{on,n}+E_{off,n}}{V_{dc,n}\cdot I_n}\cdot V_{dc}I_i\,f_{sw}\end{aligned} \tag{6.20}$$

where $E_{on,n}$ and $E_{off,n}$ are the turn-on and turn-off energy loss per switching cycle, specified in data sheets for DC link voltage $V_{dc,n}$ and inductively clamped output current I_n. The switching losses for the freewheeling diodes of a half-bridge are calculated in a similar manner. The turn-on loss for a switching diode is usually negligible whereas the turn-off losses are almost completely due to reverse recovery (E_{Drr}). This gives

$$\begin{aligned}\overline{P}_{Di,sw} &= \frac{1}{T_n}\int_{T_n}\left(P_{on}+P_{off}\right)dt = \frac{f_{sw}}{T_n}\int_{T_n}\left(E_{on}+E_{off}\right)dt = \\ &= \frac{2\sqrt{2}}{\pi}\cdot\frac{E_{off,n}}{V_{dc,n}\cdot I_n}\cdot V_{dc}I_i\,f_{sw} = \frac{2\sqrt{2}}{\pi}\cdot\frac{E_{Drr,n}}{V_{dc,n}\cdot I_n}\cdot V_{dc}I_i\,f_{sw}\end{aligned} \tag{6.21}$$

The conduction losses are more complicated to calculate. The forward voltage drop of an IGBT and a freewheeling diode are given by

$$\begin{cases}V_{T(on)} = V_{T0} + R_{T(on)}i_T \\ V_{D(on)} = V_{D0} + R_{D(on)}i_D\end{cases} \tag{6.22}$$

For the half period where the current is positive, i.e. in the interval $\omega_1 t = [\varphi..\pi+\varphi]$ of Figure 6.4, the transistor T_{1i} and the freewheeling diode D_{2i} are conducting. The duty cycle d_{T1i} for T_{1i} in this interval is given by

$$d_{T1i} = \frac{1}{2} + \frac{u_i}{V_{dc}} = \frac{1}{2} + \frac{\hat{u}_i \sin(\omega_1 t)}{V_{dc}} \tag{6.23}$$

The corresponding duty cycle for D_{2i} is

$$d_{D2i} = 1 - d_{T1i} = \frac{1}{2} - \frac{\hat{u}_i \sin(\omega_1 t)}{V_{dc}} \tag{6.24}$$

The conduction losses for one IGBT and freewheeling diode are written

$$\begin{cases} P_{T1i,cond} = V_{T(on)} i_i d_{T1i} \\ P_{D2i,cond} = V_{D(on)} i_i d_{D2i} \end{cases} \tag{6.25}$$

The average losses for one period of the output fundamental are found by integration. Note that T_{1i} and D_{2i} are only conducting when the output current is positive, i.e. in the interval $\omega_1 t = [\varphi .. \pi + \varphi]$. This gives

$$\begin{aligned} \overline{P}_{T1i,cond} = &\left(\frac{1}{2\pi} \cdot V_{T0} \hat{i}_i + \frac{1}{8} \cdot R_{T(on)} \hat{i}_i^2 \right) + \\ &+ \left(\frac{1}{4} \cdot V_{T0} \hat{i}_i + \frac{1}{3\pi} \cdot R_{T(on)} \hat{i}_i^2 \right) \cdot \frac{\hat{u}_i \cos(\varphi)}{V_{dc}} \end{aligned} \tag{6.26}$$

and

$$\begin{aligned} \overline{P}_{D2i,cond} = &\left(\frac{1}{2\pi} \cdot V_{D0} \hat{i}_i + \frac{1}{8} \cdot R_{D(on)} \hat{i}_i^2 \right) - \\ &- \left(\frac{1}{4} \cdot V_{D0} \hat{i}_i + \frac{1}{3\pi} \cdot R_{D(on)} \hat{i}_i^2 \right) \cdot \frac{\hat{u}_i \cos(\varphi)}{V_{dc}} \end{aligned} \tag{6.27}$$

For one half-bridge, the conduction losses are twice the losses of a single IGBT and diode, i.e.

$$\begin{aligned} \overline{P}_{Ti,cond} = &\left(\frac{\sqrt{2}}{\pi} \cdot V_{T0} I_i + \frac{1}{2} \cdot R_{T(on)} I_i^2 \right) + \\ &+ \left(V_{T0} I_i + \frac{4\sqrt{2}}{3\pi} \cdot R_{T(on)} I_i^2 \right) \cdot \frac{U_i \cos(\varphi)}{V_{dc}} \end{aligned} \tag{6.28}$$

and

$$\overline{P}_{Di,cond} = \left(\frac{\sqrt{2}}{\pi} \cdot V_{D0} I_i + \frac{1}{2} \cdot R_{D(on)} I_i^2 \right) - \\ - \left(V_{D0} I_i + \frac{4\sqrt{2}}{3\pi} \cdot R_{D(on)} I_i^2 \right) \cdot \frac{U_i \cos(\varphi)}{V_{dc}} \tag{6.29}$$

Note that the converter output RMS voltage U and current I for phase i are used in the two last expressions.

In stationary conditions, the RMS single line equation

$$U - j\omega_1 LI - E_{LN} = 0 \tag{6.30}$$

applies. By separation of the real and imaginary parts

$$\begin{cases} \omega_1 LI \cos(\varphi) = E_{LN} \sin(\delta_{load}) \\ U - \omega_1 LI \sin(\varphi) = E_{LN} \cos(\delta_{load}) \end{cases} \tag{6.31}$$

is obtained. By putting the load angle equal to the phase lag, i.e.

$$\delta_{load} = \varphi \tag{6.32}$$

equivalent to unity power factor for the source or load, it is found that

$$\varphi = \arctan\left(\frac{\omega_1 LI}{E_{LN}} \right) \tag{6.33}$$

for inverter operation. For rectifier operation the same expression is valid but with negative line current. The converter output voltage is given by

$$U = E_{LN} \cos(\delta_{load}) + \omega_1 LI \sin(\varphi) = \frac{E_{LN}^2 + (\omega_1 LI)^2}{E_{LN}} \cos(\varphi) \tag{6.34}$$

which means that the RMS output voltage of the converter is equal in both cases. The phase displacement is also equal but with opposite sign.

This is compared to measurement results of [8], where the losses of a battery charger consisting of four half-bridges are measured. The fourth half-bridge operates as a step-down converter with the output connected to the batteries via an inductive filter. The losses of the battery side half-bridge are given by

$$\overline{P}_T = \left(V_{T0} I_{batt} + R_{T(on)} I_{batt}^2 \right) \cdot \frac{V_{batt}}{V_{dc}} + \frac{E_{on,n} + E_{off,n}}{V_{dc,n} \cdot I_n} \cdot V_{dc} I_{batt} f_{sw} \tag{6.35}$$

for the IGBT and

$$\overline{P}_D = \left(V_{D0} I_{batt} + R_{D(on)} I_{batt}^2\right) \cdot \left(1 - \frac{V_{batt}}{V_{dc}}\right) + \frac{E_{Drr,n}}{V_{dc,n} \cdot I_n} \cdot V_{dc} I_{batt} f_{sw} \qquad (6.36)$$

for the freewheeling diode. All the half-bridges of the battery charger are Semikron SKM50GB123D modules, with approximate data according to Table 6.1.

Table 6.1: Semikron SKM50GB123D data.

V_{T0}	2.0 V	
$R_{T(on)}$	27 mΩ	Data valid for:
$E_{on,n}+E_{off,n}$	11.5 mWs	$T_{junction}$=125 °C $V_{dc,n}$=600 V
V_{D0}	1.2 V	I_n=40 A
$R_{D(on)}$	11.7 mΩ	R_{Gate}=22 Ω
$E_{Drr,n}$	1.9 mWs	

The battery charger operates at V_{dc}=650 V. The input and output power, neglecting the filters, are P_{in}=5765 W and P_{out}=5581 W. The input and output power are measured simultaneously with a four-channel Norma 6100 wide band power analyser. The three-phase boost rectifier is connected to a 400 V, 50 Hz grid. Consequently, U=233 V. The step down converter is connected to a battery with V_{batt}=243 V. The calculated losses equal 159 W, whereas the measured are 184 W. Note that the losses of DC link bleeder resistors, with a total resistance of 22 kΩ, and the resistors in series with the LEM LV50P DC bus voltage transducer, with a resistance of 75 kΩ, are not included. The total power loss in these resistors equals 25 W, which is also equal to the difference between the calculated and measured losses. Note that the losses in the DC bus capacitors are not included in the calculation. In [8], the simulated DC link capacitor losses for a 10 kW hard-switched battery charger are found to be approximately 5 W, or 0.05% of the output power. In this case this corresponds to approximately 3 W. Thus, the estimated converter losses equal 187 W.

6.2 Thermal design

The theoretical upper limit for the temperature of a semiconductor is the so-called intrinsic temperature, i.e. the temperature where the intrinsic carrier density equals the doping atom concentration in the region with the lowest doping atom density. This means that the rectifying properties are removed since

the depletion region is short circuited by the intrinsic carriers. In other words, the semiconductor becomes a conductor at the intrinsic temperature.

For commercial semiconductor devices, the maximum allowed temperature is considerably lower than the intrinsic temperature, since an increased silicon temperature will give increased resistive losses. A common maximum silicon temperature is 150 °C. Below this temperature, the manufacturer guarantees the specified on-state voltage drop, blocking voltage etc. In practical designs, the silicon (or junction) temperature for the device with the lowest temperature margin mounted on the heat sink, is 20-40 °C below the rated maximum. It is very important to do a good thermal design both for active and passive components, to get an acceptable life-time of the equipment.

Heat transfer is accomplished through three principally different processes; heat conduction, radiation and convection. For heat conduction and radiation there exists accurate models, at least for simple geometries. For convection it is more complicated to get reasonably accurate models. One dimensional heat conduction is calculated from

$$\begin{aligned} P_{cond} &= R_{cond} \cdot (T_2 - T_1) \\ R_{cond} &= \frac{\lambda \cdot A}{l} \end{aligned} \tag{6.37}$$

Where T_1 and T_2 are the temperatures in the end surfaces of the body or object, A is the cross sectional area perpendicular to the power flow, l the length of the object and λ ([W/m·K]) the thermal conductivity. Heat transfer through radiation is given by Stefan-Boltzmanns law

$$\begin{aligned} P_{rad} &= 5.6705 \cdot 10^{-8} \cdot \varepsilon \cdot A \cdot \left(T_s^4 - T_a^4\right) \\ R_{rad} &= \frac{T_s - T_a}{P_{rad}} \end{aligned} \tag{6.38}$$

Where T_s and T_a are the heat sink and the ambient temperatures, respectively, A is the area of the surface which the power is flowing through and ε the emissivity for the surface. To get a high level of heat transfer through radiation the emissivity has to be high. The emissivity for black enamelled aluminium, which is the common surface of commercial heat sinks, is close to 0.9. From the expressions valid for heat transfer through heat conduction and radiation it is seen that these will not be affected by forced cooling. Convection is on the other hand strongly enhanced by forced cooling.

For natural cooling (no fan or coolant other than the ambient air) the heat sink should be mounted with the fins vertically oriented, since this maximizes the

heat transfer by means of convection. Heat conduction and radiation are not affected by the orientation of the heat sink.

The designers of power electronic equipment do not have to deal with heat transfer other than by means of thermal resistance, which the manufacturer specifies. The thermal resistance is often specified for different air flows since forced cooling enhances the heat transfer, i.e. lowers the thermal resistance. For power electronic applications thermal grease (often called silicon grease) is often inserted between the semiconductor modules and the heat sink to remove undesired air, which could degrade the heat transfer. Figure 6.5 shows the heat transfer through the module and the heat sink.

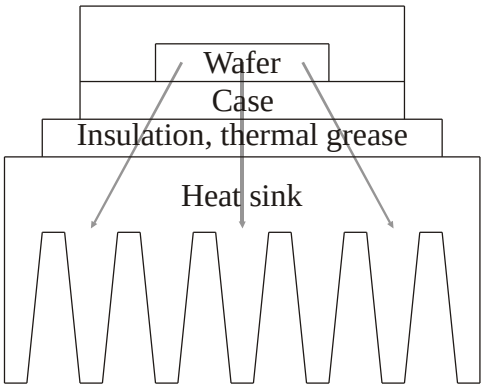


Figure 6.5: Heat transfer from the silicon wafer to the ambience, via a heat sink.

The thermal resistance for one component in a module (with several components) is written as

$$R_{thja,comp} = R_{thjc,comp} + R_{thch,\text{module}} + R_{thha} \tag{6.39}$$

Index *j* denotes silicon (junction), *c* denotes case, *h* means heat sink and *a* stands for ambience. Furthermore, index *comp* means a single component and *module* means for the entire module. Note that the thermal resistance from silicon grease (or mica) is added to $R_{thch,\ module}$. The silicon (junction) temperature for the component is calculated from

$$T_{j,comp} = P_{d,comp} \cdot R_{thja,comp} + T_a \tag{6.40}$$

where $P_{d,comp}$ is the power loss for the component in question.

In power electronic applications a single heat sink is often used to remove heat from several loss sources, i.e. several components or modules, which means that the method of calculation has to be modified somewhat. The calculations above can be understood from an electrical equivalent where the losses are modeled as a current source, the thermal resistances as resistances and the temperatures as voltages, see Figure 6.6.

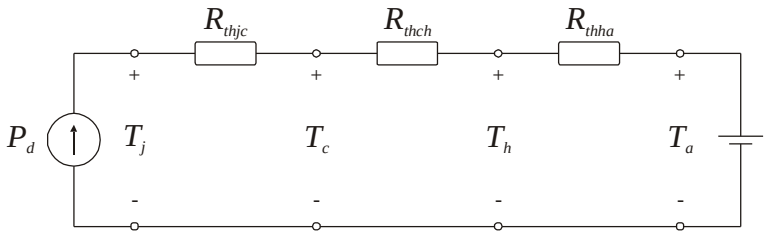


Figure 6.6: Electrical equivalent where the heat sink is used for one single component.

The electrical equivalent for *n* parallel components mounted on the same heat sink is shown in Figure 6.7 below.

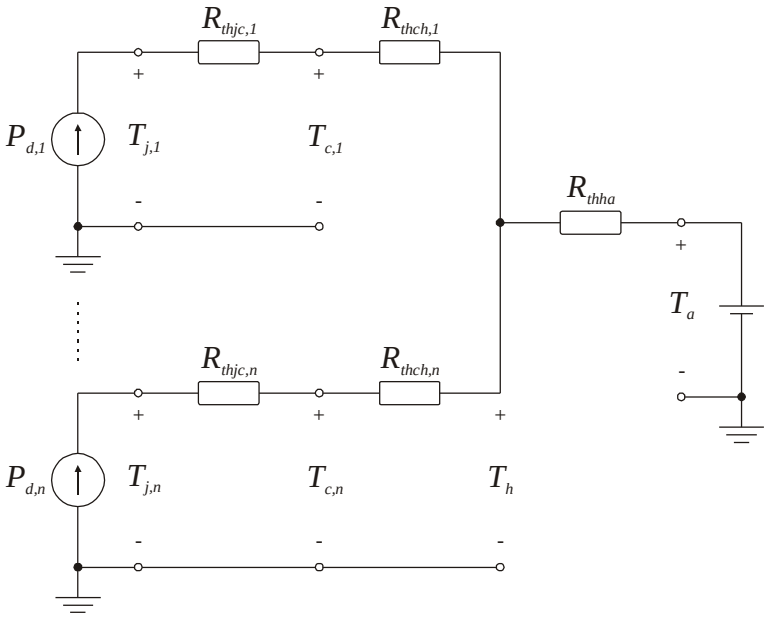


Figure 6.7: Electrical equivalent for the case where several components are mounted on a single heat sink. Note that several components can be mounted in the same module.

The steps for thermal design in this case are given below. The method assumes that the power losses and thermal resistances for all the components are known. The result of the calculation is a maximum value of the thermal resistance for the heat sink.

3. Determine the needed heat sink temperature for all the components assuming a certain maximum silicon temperature, i.e. for each component calculate

$$T_{h,i} = T_{j,i} - P_{d,i} \cdot \left(R_{thjc,i} + R_{thch,i} \right) \tag{6.41}$$

4. The component requiring the lowest heat sink temperature will determine the allowed maximum temperature. Use this as the heat sink temperature, i.e. set

$$T_h = \min\left(T_{h,i} \right) \tag{6.42}$$

5. This selection of heat sink temperature means that some components will have a silicon temperature lower than the maximum allowed. The maximum allowable thermal resistance for the heat sink is calculated from

$$T_h = T_a + R_{thha} \cdot \sum_{i=1}^{n} P_{d,i} \quad \Rightarrow \quad R_{thha} = \frac{T_h - T_a}{\sum_{i=1}^{n} P_{d,i}} \tag{6.43}$$

Forced cooling can reduce the thermal resistance to a quite large extent, often 5-10 times. Note also that no precautions have been made regarding transient thermal impedance. The transient thermal impedance often has a large impact in traction applications where accelerations and retardation are usual resulting in large variations in the silicon temperature of the semiconductors even though the heat sink temperature only varies a couple of degrees. In practical applications this temperature variation results in problems like bonding lift-off and bonding heel-crack.

7 Snubbers

Snubbers are a collective term for protective devices utilized to reduce stress imposed on power electronic devices, especially semiconductors. The main stress factors that are attempted to reduce with snubbers are over-voltage and excessive voltage and current derivatives with respect to time. In earlier chapters, the sensitivity of power semiconductors regarding over-voltage has been discussed. The sensitivity of thyristor devices against high current derivatives has also been highlighted. The risk of dynamic latch-up of IGBTs limits the turn-off collector-emitter voltage derivative with respect to time. Also, from an Electromagnetic compatibility (EMC) point of view, the voltage derivative with respect to time needs to be limited due to capacitive coupling between conductors.

For transistors, snubbers are mainly active at device turn-off. Snubbers protecting against over-voltage are clamping (limiting) the voltage at power semiconductor turn-off. These are needed since the circuit stray inductance in combination with the rapid current decrease at turn-off otherwise would cause substantial over-voltage at device turn-off. Snubbers reducing the voltage derivative with respect to time at power semiconductor turn-off aims to avoid simultaneously high current through and voltage across the device, i.e. provide soft-switching, thereby reducing the losses.

The need of snubbers is in most cases dependent on power semiconductor device family. Power diodes do not require snubbers in most applications, but may require an over-voltage snubber (non-polarized *RC* circuit), in the case of snappy reverse recovery current causing high voltage drop due to stray and leakage inductance. Thyristors with large wafer diameters need turn-on snubbers to avoid excessive local current density and modern GTO thyristors in bridge applications need turn-on snubbers to reduce the anode current derivative with respect to time since the freewheeling diode is subject to reverse recovery at GTO thyristor turn-on. Both thyristors and transistors do need over-voltage snubbers to reduce the impact of inductive parasitic circuit elements at device turn-off. Also, transistors may be equipped with snubbers controlling the voltage derivative with respect to time at turn-off, i.e. snubbers providing soft turn-off.

Power semiconductors with capacitive input (often referred to as charge controlled) for example the MOSFET and the IGBT can be operated without snubbers due to the fact that the output (drain or collector) current derivative

with respect to time is controlled by the gate resistor. Increasing the resistance of the gate resistor, results in a reduction of the output current derivative and therefore less voltage stress caused by circuit stray inductance. However, since the switching times increase, the switching losses also increase. The switching losses are dissipated in the device silicon, which typically has a temperature limit of 150 °C. A properly designed snubber is often a by far more suitable solution. Of course, there are losses in resistive elements of the snubber. However, the snubber is not mounted on the heat sink and therefore the size of the heat sink does not have to be increased. Also, resistors operating at a temperature of 300 °C are available.

The main component of over-voltage snubbers and snubbers controlling the voltage derivative with respect to time is a capacitor. Therefore, the initial analyses of these snubbers are based on purely capacitive snubbers. Later on in the investigations, the other elements of the snubber, i.e. a resistor and sometimes also a diode, are included in the analyses. The main component of a snubber controlling the current derivative with respect to time is an inductor. Therefore, a purely inductive snubber is analysed first for this type of snubber. For the snubber that can be used for single devices, this case is analysed first and then extended to cover also half bridges. There are also snubbers developed for half bridges, which are of great interest since a large amount of power semiconductor are manufactured as modules forming half bridges.

7.1 Diode snubbers

In the analysis of this snubber it is assumed that the device operates as a freewheeling diode in a step down converter, where the controlled switch is considered as ideal (Figure 7.1).

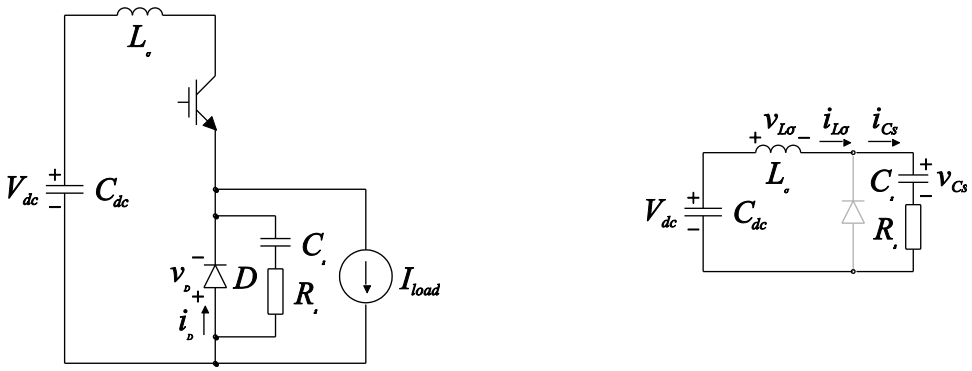


Figure 7.1: Turn-off snubber for diodes (left) and equivalent circuit at reverse recovery (right).

The reverse recovery current is assumed to be extremely snappy, so that the anode current returns from the reverse recovery peak $-I_{rr}$ to zero instantly, which means that the time interval t_{rr2} has zero duration (equivalent circuit shown in Figure 7.1). This abrupt return of the reverse recovery current is the worst case

scenario with respect to over-voltage caused by the equivalent stray inductance of the circuit L_σ.

The snubber is intended to protect against this over-voltage and therefore the snubber design is somewhat conservative since the reverse recovery current is not this nappy for real diodes. The snubber is formed by the *RC* circuit connected in parallel with the device to be protected, i.e. the freewheeling diode. To simplify the initial analysis the snubber resistor is neglected, i.e. $R_s = 0$. In the case of *LC* circuits the stored energy can be used to calculate the maximum voltage appearing across the protected device. However, in the method presented here, the differential equations of the system are used together with the initial conditions, to express the over-voltage. For the equivalent circuit of Figure 7.1 the differential equations are written

$$\begin{cases} V_{dc} - v_{L\sigma} - R_s \cdot i_s - v_{Cs} = 0 \\ i_s = i_{Cs} = C_s \cdot \dfrac{dv_{Cs}}{dt} \\ v_{L\sigma} = L_\sigma \cdot \dfrac{di_{L\sigma}}{dt} = L_\sigma \cdot \dfrac{di_{Cs}}{dt} = L_\sigma C_s \cdot \dfrac{d^2 v_{Cs}}{dt^2} \end{cases} \tag{7.1}$$

which is expressed as a single second order differential equation

$$L_\sigma C_s \cdot \frac{d^2 v_{Cs}}{dt^2} + R_s C_s \cdot \frac{dv_{Cs}}{dt} + v_{Cs} = V_{dc} \tag{7.2}$$

For the purely capacitive snubber, i.e. with $R_s = 0$, the differential equation below is obtained. The initial conditions are also stated below. Note that it is assumed that the current flowing in the equivalent stray inductance commutates to the snubber capacitor abruptly.

$$\begin{cases} L_\sigma C_s \cdot \dfrac{d^2 v_{Cs}}{dt^2} + v_{Cs} = V_{dc} \\ v_{Cs}(t_0) = 0 \\ i_{Cs}(t_0) = I_{rr} \end{cases} \tag{7.3}$$

The solution to this differential equations is given by

$$v_{Cs}(t) = V_{dc} \cdot (1 - \cos(\omega_0 (t - t_0))) + I_{rr} \cdot \sqrt{\frac{L_\sigma}{C_s}} \cdot \sin(\omega_0 (t - t_0)) \tag{7.4}$$

where the characteristic frequency is

$$\omega_0 = \frac{1}{\sqrt{L_\sigma \cdot C_s}} \tag{7.5}$$

To simplify the analysis a normalisation or base value of the capacitance

$$C_{base} = L_\sigma \cdot \left(\frac{I_{rr}}{V_{dc}} \right)^2 \tag{7.6}$$

is introduced. Inserting into the solution of the differential equation gives

$$v_{Cs}(t) = V_{dc} \cdot \left(1 - \cos(\omega_0 (t - t_0)) + \sqrt{\frac{C_{base}}{C_s}} \cdot \sin(\omega_0 (t - t_0)) \right) \tag{7.7}$$

The voltage the freewheeling has to sustain is thus given by

$$V_{D,max} = V_{dc} \cdot \left(1 + \sqrt{1 + \frac{C_{base}}{C_s}} \right) > 2V_{dc} \tag{7.8}$$

From the last expression it is obvious that no matter how high the snubber capacitance is, the diode must be able to block at least twice the dc link voltage. This holds for the purely capacitive snubber in case of infinitely snappy reverse recovery current. Moreover, when no snubber resistor is present, the diode blocking voltage exhibits an un-damped oscillation, since the oscillating energy is not dissipated. In a real circuit there is of course parasitic resistance, but this only yields a low damping.

If a snubber resistor is employed, a feasible damping can be obtained. Moreover, the snubber resistor may also contribute to reduce the maximum voltage the freewheeling diode has to sustain so that it is lower than twice the dc link voltage also for snappy reverse recovery current. However, the mathematical analysis results in expressions that is more complicated to interpret, so the design should be carried out with aid of simulation software. Still, it should be emphasized that simulations do not give a more accurate result than the device simulation models provide. For example, a reverse recovery current returning abruptly to zero is not realistic and the equivalent stray inductance of the circuit is unknown. Therefore, in snubber design, suitable component values are derived by simple hand calculations or circuit simulations and then tested in the application. During the test phase, the snubber component values are tuned to obtain a feasible behavior. A rule of thumb for the diode over-voltage snubber is to select

$$\begin{cases} C_s = C_{base} \\ R_s = 1.3 \cdot R_{base} \text{ where } R_{base} = \dfrac{V_{dc}}{I_{rr}} \end{cases} \tag{7.9}$$

According to [14] these snubber component values minimizes the voltage stress imposed on the freewheeling diode so that the maximum blocking voltage the diode has to sustain is less than twice the dc link voltage.

Diode snubbers are often not utilized for freewheeling diodes in bridge applications, since the snubbers protecting the anti-parallel GTO thyristors or transistors protects also the diodes from over-voltage.

7.2 Snubbers for thyristors

Like diodes, thyristors and GTO thyristors usually require over-voltage (voltage clamp) snubbers. For regular thyristors the need of snubbers is mainly due to snappy reverse recovery as for diodes. Therefore *RC* snubbers are used also for thyristors. Snubbers are also used to ensure that the thyristors is not exposed for excessive anode-cathode voltage derivative when traversing from reverse to forward blocking state, which otherwise may cause unintended triggering of the thyristor. To prevent this, the RC snubber is also suitable.

For GTO thyristors in bridge applications, an over-voltage snubber termed RCD voltage clamp snubber is sometimes used. This snubber is also used for transistors, which are discussed in later sections.

As earlier stated, the anode current derivative with respect to time at turn-on of GTO thyristors should not exceed the level specified by the device manufacturer, since this may result in un-uniform current distribution and in bridge applications also a very high reverse recovery current of freewheeling diodes. This type of snubber utilizes a deliberately inserted inductor connected between the actual GTO thyristor bridge (i.e. the half bridges) and the DC link capacitor. Of course, such an inductor will result in over-voltage at GTO thyristor turn-off if a voltage clamp snubber is not included. Therefore, the inductive turn-on snubber is always accompanied by a voltage clamp snubber to avoid excessive voltage across the power semiconductor devices. This snubber is also discussed in the section on snubbers for transistors.

7.3 Transistor snubbers

Two types of transistor snubbers are addressed in this section, the RCD (resistor-capacitor-diode) voltage clamp snubber and the RCD charge-discharge

snubber. Both these snubbers may also be used for GTO thyristors in bridge applications. Inductive turn-on snubbers, limiting the turn-on current derivative is usually not used for charge controlled transistors (e.g. the MOSFET and the IGBT) but may be used for current controlled transistors (e.g. the BJT).

The RCD voltage clamp snubber

The RCD voltage clamp snubber (Figure 7.2) is intended to limit (clamp) the over-voltage across the power semiconductor device at turn-off by providing an alternative current path for the current flowing in the circuit stray inductance so that the corresponding energy is transferred to the snubber capacitor. Note that for an inductor the stored energy is expressed by

$$W_L = \frac{1}{2} L i_L^2 \tag{7.10}$$

whereas for a capacitor the stored energy is written

$$W_C = \frac{1}{2} C v_C^2 \tag{7.11}$$

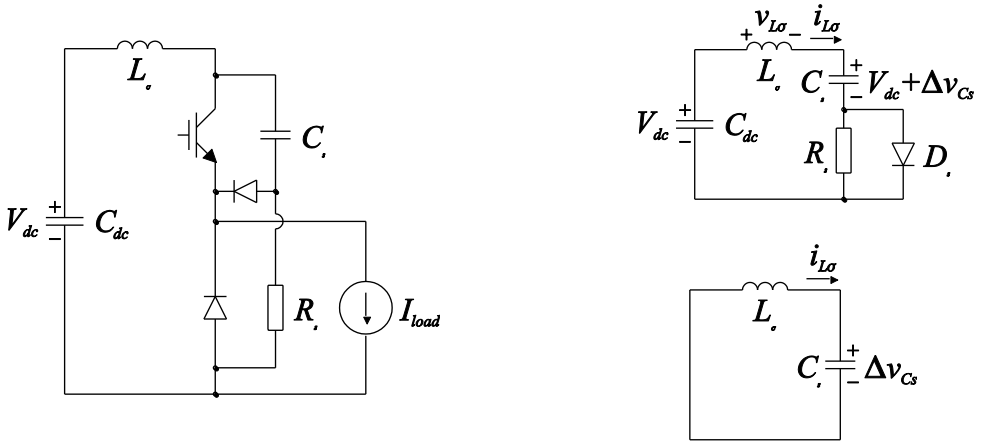


Figure 7.2: RCD voltage clamp snubber for transistors and GTO thyristors in bridge applications (right) and its equivalent circuits at transistor turn-off (left).

This means that a certain stray inductance, current and snubber capacitance will result in a certain snubber capacitor voltage at turn-on, i.e. the snubber capacitor voltage is clamped to the level determined by these parameters. When the transistor operates in stationary blocking condition, i.e. between turn-off and turn-on instants, the blocking voltage it needs to sustain should be equal to the DC link voltage. The resistor of the RCD clamp snubber provides a path to discharge the snubber capacitor against the dc link capacitor where the resistance is controlling the time constant of the discharge. Charging the snubber capacitor, i.e. at transistor turn-off, should take place with an as short time constant as possible, and therefore the snubber resistor is bypassed with a diode.

To analyse the circuit, it is simplified by determining an equivalent circuit valid at the time instant of transistor turn-off (Figure 7.2). In this analysis, it is assumed that the transistor turns-off instantly, i.e. the transistor current falls with infinite (negative) derivative with respect to time. If the snubber is not present, the voltage across the equivalent stray inductance would be infinite (or at least very high in real applications) since

$$v_{L\sigma} = L_\sigma \frac{di_{L\sigma}}{dt} \tag{7.12}$$

This voltage would appear as an over-voltage across the transistor now operating in the blocking state, since the dc link act as a voltage source and the on-state voltage of the freewheeling diode is determined by the load current. Therefore, the only device where the over-voltage can show up is across the blocking transistor. The snubber provides a path for the current flowing in the stray inductance in such a way that charge is added to the snubber capacitor. The first circuit equivalent (Figure 7.2) show this. The second circuit equivalent is derived in the following. Since the DC link capacitor has a capacitance that is considerably higher than the one of the snubber capacitor and these appear to be connected in series at transistor turn-off, the equivalent capacitance is given by

$$\frac{1}{C_{eq}} = \frac{1}{C_{dc}} + \frac{1}{C_s} \;\Leftrightarrow\; C_{eq} \approx C_s \quad \text{since } C_{dc} >> C_s \tag{7.13}$$

The snubber capacitor voltage and the DC link capacitor voltage have opposite polarities therefore the voltage across the equivalent capacitor is given by (Note that the initial snubber capacitor voltage is equal to the dc link voltage V_{dc})

$$v_{Ceq} = v_{Cs} - V_{dc} = V_{dc} + \Delta v_{Cs} - V_{dc} = \Delta v_{Cs} \tag{7.14}$$

This means that the voltage across the equivalent snubber capacitance is equal to the over-voltage across the transistor, i.e. in excess of V_{dc}. The initial current flowing through the equivalent stray inductance is equal to the load current, since the load current is provided by the DC link capacitor prior to transistor turn-off. This information is sufficient to establish an energy relationship saying that all the energy stored in the stray inductance is transferred to the equivalent snubber capacitor

$$\frac{1}{2} \cdot C_s \cdot \Delta V_{Cs,max}^2 = \frac{1}{2} \cdot L_\sigma \cdot I_{load}^2 \;\Rightarrow\; \Delta V_{Cs,max} = I_{load} \cdot \sqrt{\frac{L_\sigma}{C_s}} \tag{7.15}$$

The snubber resistor is selected in such a way that the snubber capacitor is discharged to a certain level in one switching period. Transistor turn-on is not

critical for the operation of this snubber since the snubber diode is blocking during transistor on-state operation, which is also the reason why the snubber capacitor voltage never decreases below the voltage of the DC link capacitor. A typical choice of snubber resistor is

$$R_s < \frac{1}{6 \cdot C_s \cdot f_{sw}} \tag{7.16}$$

This resistance value corresponds to that the snubber capacitor is discharged to 0.25% of the DC link voltage on the average, i.e. assuming a duty cycle of 0.5, at transistor turn-on.

Figure 7.3 shows a time-domain simulation of an RCD voltage clamp snubber protecting an IGBT (note the collector current tail) operating in a step-down converter according to Figure 7.2.

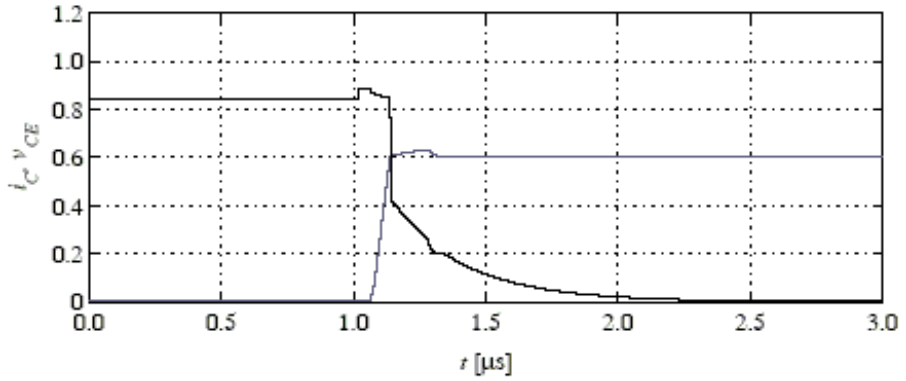


Figure 7.3: Time-signals showing normalized transistor current (black) and voltage (grey) at turn-off of the IGBT in a step down converter. The IGBT is protected by an RCD voltage clamp snubber.

The IGBT is rated 23 A and 600 V, which is also the normalization base used in Figure 7.3. The IGBT is thus operating approximately at 0.6 per unit (pu) of the rated voltage and 0.8 pu of the rated current. The equivalent stray inductance is 100 nH and the snubber capacitance is C_s = 68 nF. The snubber is designed to allow an over-voltage of 5 %, i.e. 0.05 pu at IGBT turn-off in the case of rated current (23 A).

Remark**:** For GTO thyristor bridges (composed of one or several half bridges) this snubber is used as a turn-on snubber. In this case a snubber inductor L_s is placed on purpose at the location where the equivalent stray inductance L_σ is placed in the previous analysis. The other snubber components are kept to avoid over-voltage at GTO thyristor turn-off. If the snubber inductance is low, the current rise time is determined by the device turning on, but in the case of substantial snubber inductance, the current rise time is determined by

$$\frac{di}{dt} = \frac{V_{dc}}{L_s} \;\Rightarrow\; t_{ri} = \frac{L_s I_{load}}{V_{dc}} \tag{7.17}$$

The RCD voltage clamp snubber presented above is intended for limiting the voltage across a single transistor. However, the snubber can be used also for transistors operating in half bridges. For discrete transistors, one RCD voltage clamp snubber is used for each transistor (Figure 7.4). For transistor modules, one RCD voltage clamp snubber is often used to protect both transistors of the module (Figure 7.4). Note that the RCD voltage clamp snubber intended for modules resembles the snubber equivalent of Figure 7.2. Actually, the snubbers operate in similar way, so the parameter selection is the same and gives the same behaviour of the snubbers.

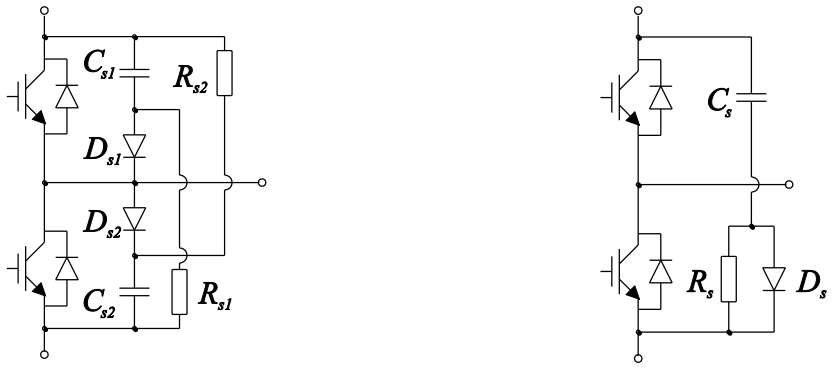


Figure 7.4: RCD voltage clamp snubbers for half bridges based on discrete transistors (left) and transistor modules (right).

The RCD charge-discharge snubber

Now the second type of transistor snubber, the RCD charge-discharge snubber is investigated. The investigation starts with a presentation of the switching behaviour of the hard switched step down converter. Then, a capacitive snubber is introduced. This is followed by the introduction of the full RCD charge-discharge snubber applied both for the step down converter and a transistor half bridge.

One of the most basic transistor bridge configurations for power electronic applications is the step down converter. It consists of a voltage source (DC link capacitor), a power transistor (IGBT in this case) and a freewheeling diode, see Figure 7.5. Since this is a voltage source converter, the load is a current source, i.e. inductive. When the state of the switch is changed from on to off (turn-off) or from off to on (turn-on), the transition will take a finite time in the non-ideal case. Furthermore, for non-ideal circuits there are parasitic components, for example stray inductance that can cause over-voltage across the semiconductor devices at turn-off.

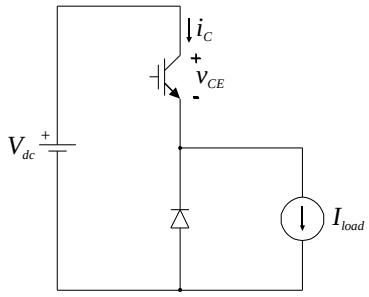
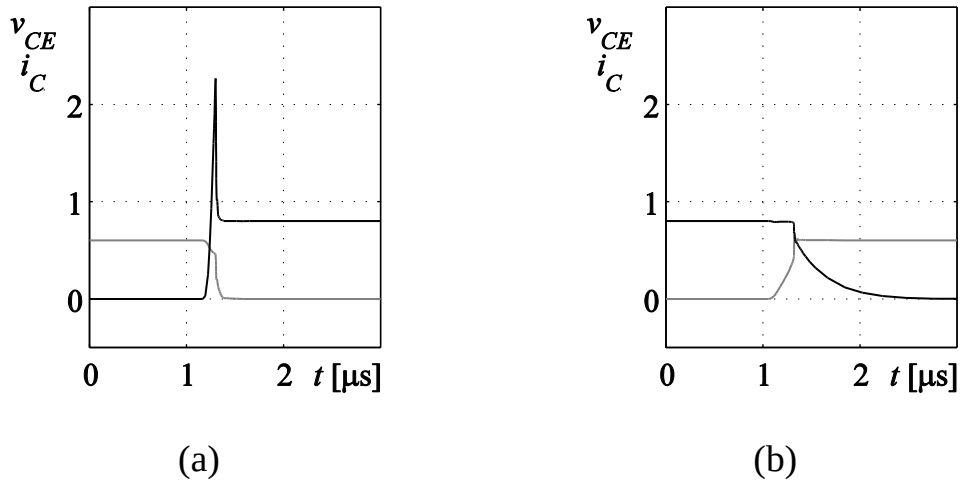


Figure 7.5: Basic step down converter used in the analysis.

Figure 7.6 shows typical collector current and collector-emitter voltage for a power transistor (IGBT in this case), when used in the step down converter above. Note that the collector current and collector-emitter voltage is expressed in p.u., where the normalization values are selected as the rated maximum continuous collector current, I_C, and maximum collector-emitter voltage that can be sustained across the device, V_{CES}. In the simulation of the step down converter the DC link voltage V_{dc} and load current I_{load} are selected as

$$\begin{cases} I_{load} = 0.8 I_C \\ V_{dc} = 0.6 V_{CES} \end{cases} \tag{7.18}$$



(a) (b)

Figure 7.6: Time-signals showing normalized transistor current (black) and voltage (grey) at (a) turn-on and (b) turn-off of the power transistor in the step down converter.

In Figure 7.6, the power transistor is exposed to a current spike at turn-on which is due to reverse recovery of the freewheeling diode. Furthermore, it is clearly seen that the IGBT is exposed to simultaneously high current and voltage at the switching instants. This leads to high switching losses, especially at turn-off since the IGBT exhibits a collector current tail here. The physics of reverse recovery of a power diode and the IGBT current tail are discussed in Section 4.7.

A common way to visualize the stress levels imposed on a power electronic semiconductor device is to plot the switching trajectories, i.e. current versus voltage, on top of the safe operating area (SOA) of the device. The power

semiconductor manufacturers specify the SOA on absolute maximum values that must not be exceeded. However, in some cases there are two SOAs, the second valid only for very short pulses (transients). Figure 7.7 shows the switching trajectories corresponding to the time signals in Figure 7.6 on top of the SOA for the particular IGBT used in the simulation of the step down converter above.

The switching trajectories in Figure 7.7 shows that there are no problems with over-voltage or over-current extending outside the SOA. The duration of the time intervals where the power transistor is exposed to high current and voltage simultaneously, causing high losses, is however not seen. This can only be seen in Figure 7.6. Both these plots are valuable since either of them are delimiting to what extent, by means of transferred power, the power converter can be used.

In some cases, the circuit stray inductance can be high causing a high over-voltage at turn-off, which means that there must be a large margin between the DC link voltage used and the rated voltage. In other cases the switching losses can be high causing a low value of rated converter current in order to keep the junction temperature of the power transistor at an acceptable level.

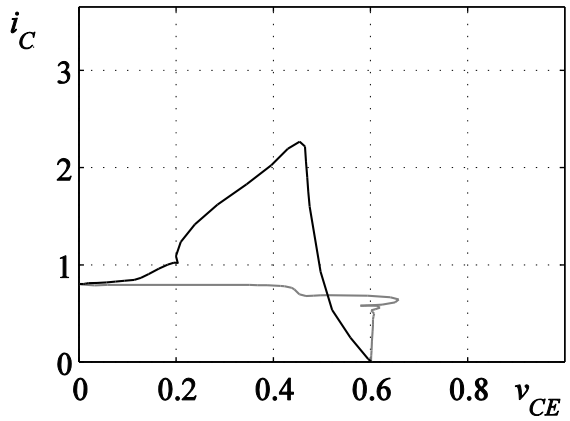


Figure 7.7: Switching trajectories (turn-on black and turn-off grey) of an IGBT in a step down converter application. The total area corresponds to the safe operating area valid for short pulses.

To partly overcome this problem and be able to use the semiconductor devices in a more efficient way, snubber circuits are introduced [14][23][24]. There are several different snubber circuits used for different purposes, for example to reduce the overvoltage caused by stray inductance, at turn-off. There are also snubber circuits where the aim is to move the switching loci further into the SOA, which means lower losses, at least if the duration of the switching intervals are not prolonged. One such snubber is the RCD (resistor, capacitor and diode) charge-discharge snubber.

From now on the RCD charge-discharge-snubber is referred to only as RCD snubber, even though there are several types of RCD snubbers for different

purposes. At turn-off, this snubber behaves as a pure capacitor. To simplify the preliminary analysis, the snubber is considered as consisting only of a single capacitor denoted C_s in Figure 7.8.

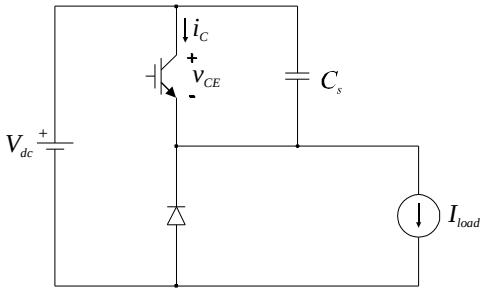
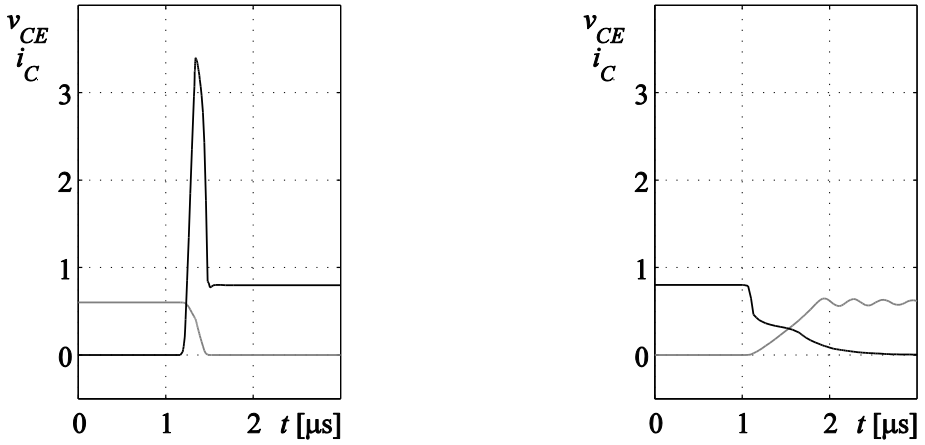


Figure 7.8: The step down converter with a purely capacitive turn-off snubber.

By introducing a capacitor across the power transistor output terminals, the voltage derivative with respect to time can be controlled, see for example [14]. This is possible since at turn-off, a part of the transistor current finds an alternative path through the capacitor, which means that the transistor collector current falls and the collector-emitter voltage rises simultaneously, see Figure 7.9. This is not possible in the previous case, i.e. without a snubber, since here the only alternative way for the load current is through the freewheeling diode.



(a) (b)

Figure 7.9: Time-signals showing transistor current (black) and voltage (grey) at (a) turn-on and (b) turn-off of the power transistor in the step down converter with a capacitive snubber across the power transistor. Note the high collector current peak at turn-on, and that the collector current and collector-emitter voltage changes simultaneously at turn-off.

The freewheeling diode only carries substantial current when forward biased, which means that the transistor collector-emitter voltage must be approximately equal to the DC link voltage, V_{dc}, before the collector current can begin to decline at turn-off for the case without a snubber.

It is sometimes stated that this type of snubber circuit provides turn-off under zero voltage switching (ZVS) conditions, or soft-switched conditions [18]. This refers to the fact that when the turn-off sequence is initiated, the collector-

emitter voltage is approximately zero. However, at the end of the current fall interval the collector-emitter voltage is high, ideally equal to the DC link voltage V_{dc}, according to [14]. This means that the snubber capacitor should be selected according to

$$i_{Cs} = I_{load} - i_C = I_{load} - I_{load} \cdot \left(1 - \frac{t}{t_{fi}}\right) = I_{load} \cdot \frac{t}{t_{fi}} \quad 0 < t < t_{fi} \tag{7.19}$$

$$v_{Cs} = \frac{1}{C_s} \int_0^t i_{Cs} dt = \frac{I_{load} \cdot t^2}{2 \cdot C_s \cdot t_{fi}} \tag{7.20}$$

$$v_{Cs}(t_{fi}) = V_{dc} \;\Rightarrow\; C_s = C_{base} = \frac{I_{load} \cdot t_{fi}}{2 \cdot V_{dc}} \;\Rightarrow\; v_{Cs} = V_{dc} \cdot \left(\frac{t}{t_{fi}}\right)^2 \tag{7.21}$$

Converters not using snubbers providing soft transitions neither at turn-on or turn-off, are often termed hard switched [18]. Another characterization of soft switching is also used in this context, zero current switching (ZCS), which refers to that either or both the turn-on and turn-off transitions take place at virtually zero current [14].

From Figure 7.9 it is seen that the snubber capacitor is affecting the turn-off waveforms in such a way that the turn-off losses do decrease. Also, the collector-emitter voltage derivative is controlled which can be an important aspect by means of electromagnetic compatibility (EMC) [17]. On the other hand, it is also seen that the turn-on waveforms becomes less favourable.

For the hard switched step down converter there was a short collector current spike due to reverse recovery of the freewheeling diode at turn-on of the power transistor. For the case with a capacitive snubber the current spike is even higher, thus adding stress to the transistor at turn-on. Prior to turn-on, the snubber capacitor is fully charged, i.e. the capacitor voltage equals V_{dc}. The voltage across the freewheeling diode remains close to zero as long as its junction is forward biased, which implies that the capacitor voltage, and thereby the collector-emitter voltage, cannot decrease before the pn-junction of the freewheeling diode becomes reverse biased. The freewheeling diode becomes reverse biased exactly when the reverse recovery current reaches its peak value.

Reverse recovery is due to stored minority charge carriers close to the pn-junction of the diode, causing the junction to be forward biased even though the diode current is negative. The negative current sweeps out the minority carriers and eventually the junction becomes reverse biased. As mentioned earlier,

power electronic diodes and IGBTs are investigated further in Chapter **Error! Reference source not found.**.

The discussion above implies that the capacitor discharge starts when the reverse recovery current reaches its peak. Even if the diode recovers fast, the transistor has to carry an excessive current for quite some time since the only discharge path for the capacitor is through the power transistor.

The switching waveforms in Figure 7.9 are obtained with $C_s = C_{base}$ (from the design expressions) and it can be shown that the higher the snubber capacitance is the further into the SOA of the device the turn-off switching trajectory will show up. On the other hand, the current peak in Figure 7.9 at turn-on will also be both higher and wider in the case of higher snubber capacitance. Therefore, $C_s = C_{base}$ is the preferable value.

To cope with the problem of the capacitive snubber, a resistor is added. The intention with the resistor is to limit the capacitor discharge current at transistor turn-on. A diode is placed in parallel with the resistor since at turn-off, the capacitor gives the desired behaviour and the resistor would cause increased losses. The complete RCD snubber is shown in Figure 7.10.

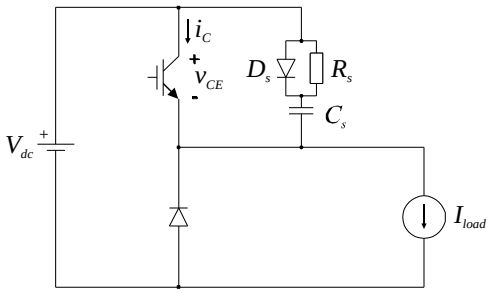
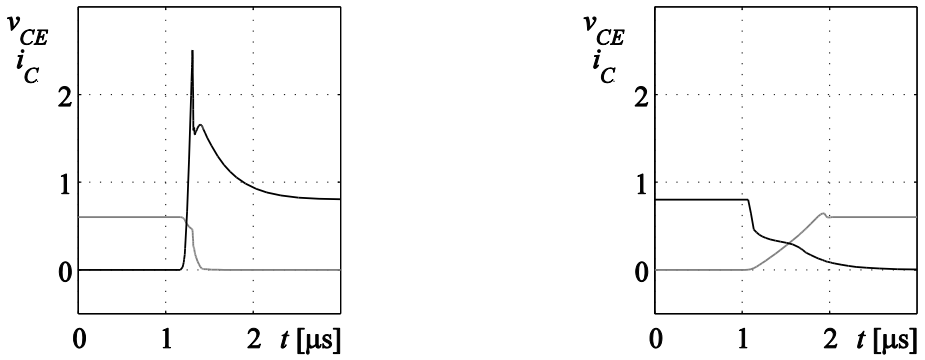


Figure 7.10: The step down converter with the full RCD charge-discharge snubber placed across the output terminals of the device.

The resistor value is chosen in such a way that the capacitor peak discharge current do not exceed the peak reverse recovery current of the freewheeling diode [14], i.e.

$$R_s > R_{base} = \frac{V_{dc}}{I_{rr}} \tag{7.22}$$

Figure 7.11 shows the simulated transistor collector current and collector-emitter voltage.



(a) (b)

Figure 7.11: Time-signals showing transistor current (black) and voltage (grey) at (a) turn-on and (b) turn-off of the power transistor in the step down converter with a full RCD charge-discharge snubber across the power transistor.

It is clearly seen that the turn-off sequence is similar to the one for the purely capacitive snubber, but the stress levels at turn-on is only slightly higher than it was for the hard switched step down converter.

In Figure 7.12 a half bridge with a RCD snubber across each transistor is shown. The upper RCD snubber is intended to provide soft turn-off for the upper IGBT. The lower RCD snubber is intended to provide soft turn-off for the lower IGBT. The simulated collector current and collector-emitter voltage are shown in Figure 7.13. Note the very high current peak at transistor turn-on.

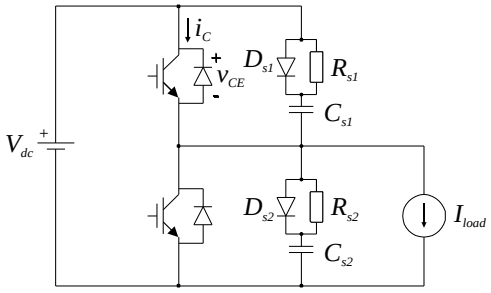


Figure 7.12: A half bridge with one RCD snubber connected across each transistor.

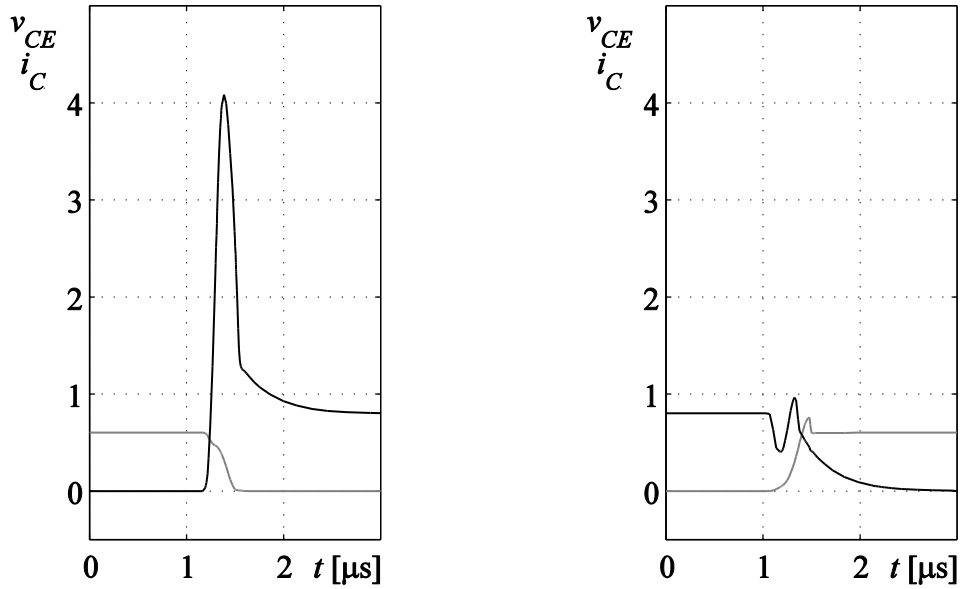


Figure 7.13: Time-signals showing transistor collector current (black) and collector-emitter voltage (grey) at (a) turn-on and (b) turn-off of the upper IGBT in the converter consisting of a bridge leg. One RCD snubber is used across each IGBT. Note the high collector current peak at turn-on. Also note the poor behaviour at turn-off.

The occurrence of this current peak is due to the capacitive current path seen from the IGBT output terminals, [14]. At turn-on of the upper IGBT, the collector-emitter voltage should decrease for the upper IGBT and increase for the lower. This means that the upper snubber capacitor in Figure 7.12, denoted C_{s1}, should be discharged and the lower, denoted C_{s2}, should be charged.

The discharge current of C_{s1} is limited by the snubber resistor R_{s1} as previous but the charging current of C_{s2} is not limited by any other component. Furthermore, the only path possible for the charging current is through the upper IGBT. In this way, the charging current of the lower snubber capacitor gives a large contribution to the collector current of the upper IGBT, at turn-on. The same problem appears for the lower transistor.

Another problem seen in Figure 7.13 is that the turn-off is not soft, i.e. the collector current fall and the collector-emitter voltage rise do not occur simultaneously. Instead, the collector current falls somewhat, then increases and eventually the current falls again. This is referred to as a current tail bump, which is discussed in Section 2.

The switching loci for the converter consisting of a bridge leg with RCD snubbers across the transistors are shown in Figure 7.14. Note that the switching trajectory is outside the SOA valid for very short pulses, which is very dangerous, since this can lead to device failure.

Small variations of the snubber circuit can be used to partly solve this problem. First, the snubber diodes in parallel with the resistors can be removed. In this way the capacitive path is broken but soft turn-off is also lost. However, EMC related problems can still be reduced [14].

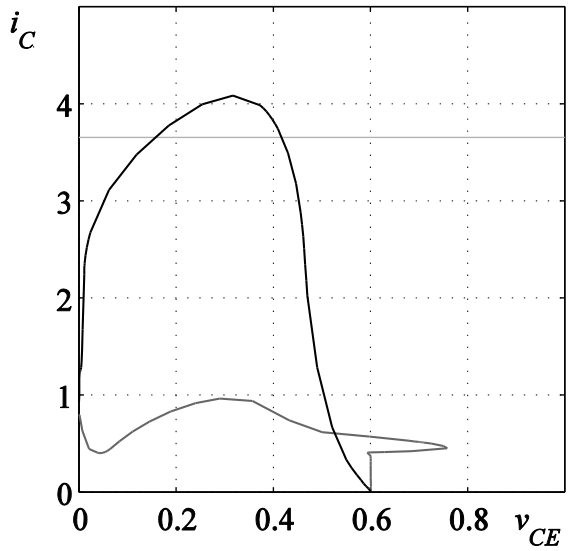


Figure 7.14: Switching loci (turn-on black and turn-off dark grey) of the bridge leg with RCD snubbers. Note that the turn-on collector current is beyond the current limit of the dynamic SOA (light grey), valid for short pulses.

Another way to solve this problem is by introducing inductors in series with the snubbers, thus reducing the current derivative with respect to time [24]. This works, but gives a bulky snubber. This is not desirable, especially not since eight such snubbers are needed for the battery charger considered.

8 Control of Electromechanical Energy Converters

8.1 Introduction

The term electromechanical energy converter is a very general concept and usually guides the thoughts towards rotating machines. In a general interpretation every device linking in terms of energy a mechanical and an electrical system is a converter of this type. Energy transportation can occur in both directions. If energy is transported from the mechanical to the electrical system the converter operates in generation mode. On the other hand motor operation means that energy is transported from the electrical system to the mechanical system. The energy transforming process is in principle reversible, so that all motors may as well be used as generators and vice versa.

An electromechanical converter operates based on phenomena related to electrical charge transport, in this way using electric and magnetic fields. The most important exploited physical phenomena are:

1. A current carrying conductor located in a magnetic field experiences force acting on it. Forces are also experienced between two current carrying conductors due to the interaction of their magnetic fields. The energy conversion process is reversible: an electromotive force is induced in a conductor moving in a magnetic field.

2. A ferromagnetic material located in a magnetic field is experiencing mechanical forces. If a coil generates the magnetic field the energy transformation process is reversible: when moving the ferromagnetic material the flux linkage will vary and a voltage is induced in the coil.
3. The plates of a charged capacitor and dielectric material within an electrical field experience mechanical forces. As opposed to this phenomenon, a relative movement of the capacitor plates or the dielectric material results in a variation of the electric charge on the plates, or voltage between them, or both.
4. Several crystalline structures show small deformation if voltage gradients are applied on certain directions. If in turn the crystalline material is stressed mechanically a voltage can be measured along the crystal. This phenomenon is known as the *piezoelectric effect*. Very strong mechanical forces can occur, even if the deformation in the presence of an applied voltage is small.

5. Most of the ferromagnetic materials experience very small deformations under the influence of magnetic fields and their magnetic properties change when they are mechanically stressed. The phenomenon is known as *magnetostriction* and, similar to the piezoelectric phenomenon; mechanical forces can become large, even if deformations are very small.

The choice of the energy conversion principle used in practice depends on the desired motion and on the magnitude of the forces to be developed by the converter. Two basic types of motion are implemented in practical electromechanical energy converters: rotation and translator motion or vibration.

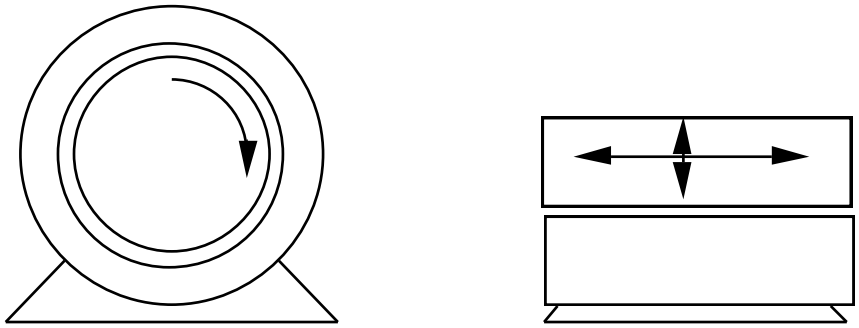


Figure 8.1 Rotation (left) and translation or vibration (right).

Rotating electromechanical energy converters are almost exclusively based on the first two principles described above. Electrical machines of his type are built for power ranges reaching from fractions of watts up to approximately 1 GigaWatt (largest generators in operation today).

To the category of translatory or vibrating converters belong for example: electro-mechanical relays, loudspeakers, microphones, different types of measurement devices, etc. All principles mentioned earlier are used in practical design for this category of converters. If large mechanical forces and movements of the active parts of the converter are desired, the first two principles are commonly employed.

The reason why specifically the magnetic field is more interesting from the point of view of energy transmission becomes clear if the energy densities are compared for the electric and magnetic fields in air or vacuum.

$$\begin{aligned} &\text{Magnetic energy density:} && \frac{B^2}{2\mu_0} \; {}^{J}\!/_{m^3} \\ &\text{Electric energy density:} && \frac{E^2}{2\varepsilon_0} \; {}^{J}\!/_{m^3} \end{aligned} \qquad (8.1)$$

In the air gap of a rotating electric machine, magnetic flux densities in the order of magnitude of about 1 Tesla (Wb/m^2) can easily be obtained yielding an energy density:

$$\text{Magnetic energy density: } \frac{1}{2}\cdot\frac{1^2}{4\cdot\pi\cdot 10^{-7}} \approx 4\cdot 10^5 \text{ J/m}^3$$

Electrical field intensities on the order of magnitude of $3\cdot10^6$ V/m may be reached in air under normal atmospheric pressure, resulting in energy densities of:

$$\text{Elektric energy density: } \frac{1}{2}\cdot\frac{10^{-9}}{36\cdot\pi}\cdot 9\cdot 10^{12} \approx 4\cdot 10^{1} \text{ J/m}^3$$

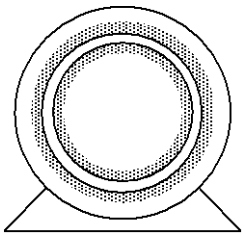
The magnetic field allows therefore reaching approximately 10^4 times higher energy densities than the electric field under the same conditions. It is therefore more attractive as an energy transmission media, especially for applications where large amounts of energy are handled.

The following presentation refers to rotating electromechanical energy converters using magnetic fields as the coupling media. The emphasis lies on showing how the electrical and mechanical system parts determine mechanical torque and, finally, how this mechanical torque may be controlled.

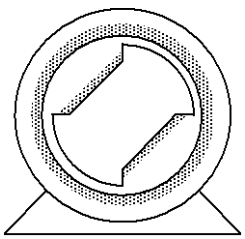
8.2 Mechanical Construction

Electrical machines are produced in a large variety of construction types, each of them with its specific characteristics and special properties. What at first glance seems to be an entire set of operating concepts which cannot be easily surveyed, at a closer examination turns out to contain only a few basic constructions with respect to mechanical and electrical aspects. Looking at the mechanical construction they can be divided into the following categories listed below. The figures show cross-sections in the rotation plane where shaded areas indicate the electrical winding locations.

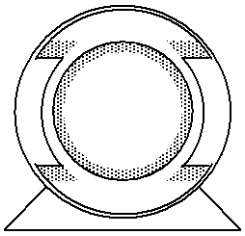
1 Cylindrical stator
Cylindrical rotor
Both constructions are used for synchronous machines (field winding on the rotor) as well as for asynchronous machines.



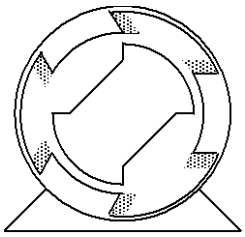
2 Cylindrical stator
Rotor with salient poles
Construction predominantly used for synchronous machines



3 Stator with salient poles
Cylindrical rotor
Construction used in general for direct current - *dc* - machine



4 Stator with salient poles
Rotor with salient poles
Construction used only for reluctance machines (special type of synchronous machine). Will not be treated in the following.



In this presentation the inner, rotating part of an electrical machine is called *rotor* whereas the outer, stationary part of the machine is the *stator*. Some constructions, for example fans, use the reversed arrangement: the rotor is the outer part of the machine, rotating around the stator.

We distinguish further between *field winding* and *armature winding*. The field winding is often realized as a compact winding and serves to generate the main magnetic flux of the machine. For the synchronous machine above (construction 1 and 2) the field winding is placed on the rotor whereas for the dc-machine (construction 3) it is situated on the stator. In many cases the field winding is replaced by permanent magnets. The armature winding is always the winding where emf's are induced either due to the rotating magnetic field (synchronous machine) or due to the rotation of the armature windings in the stationary magnetic field (dc-machine). The armature winding is usually not a compact winding. In the synchronous machine (construction 1 and 2) it is placed in the stator whereas in the dc-machine (construction 3) it is placed on the rotor.

8.3 Number of poles, mechanical and electrical angle

The machines in the figures in section 8.2 have two poles, except the reluctance machine where stator and rotor have a different number of poles for several reasons not discussed here. Electrical machines do often in practice have a larger number of poles than just two. Figure 8.2 below shows a two pole and a four pole synchronous machine with salient poles and a one-turn stator winding placed above a pole pitch.

The stator winding consists of two conductors' perpendicular to the drawing plane and with the back ends connected together. If the rotor is rotating an ac voltage is induced in the winding. For the two pole machine one period of the

induced voltage corresponds to 360 mechanical degrees that is a full turn of the rotor.

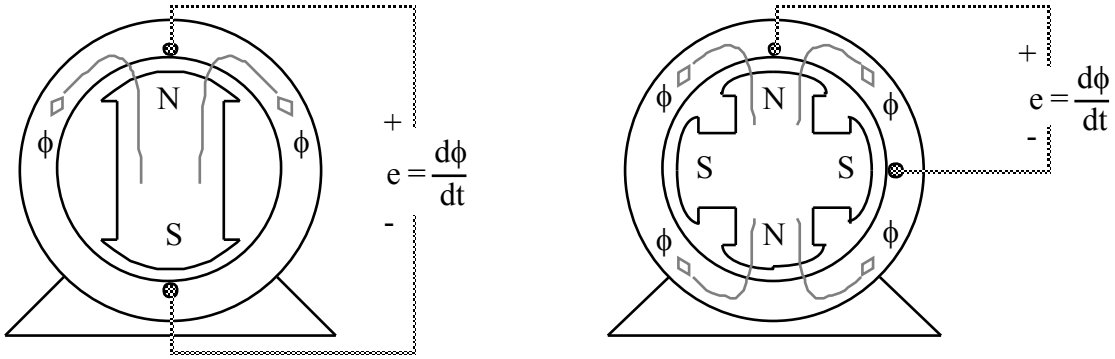


Figure 8.2 Two poles (left) and four poles (right).

For the four pole machines only a half mechanical turn is necessary to obtain a full electric period. It directly follows that:

$$\theta_{el} = \frac{p}{2} \cdot \theta_{mec} \tag{8.2}$$

The number of poles p indicates the number of magnetic poles and may be considered as the transmission factor of a changing device between the electrical and mechanical angle. The same relation holds for the derivatives of equation (8.2), i.e. for the angular velocities:

$$\omega_{el} = \frac{p}{2} \cdot \omega_{mec} \tag{8.3}$$

This relation between the angles influences the expression for mechanical torque and power. Normally the electro dynamical torque of an electrical machine (T_{el}) is calculated per pair of poles. If the machine has more than one pair of poles the resulting mechanical torque (T_{mech}) is calculated:

$$T_{mec} = \frac{p}{2} \cdot T_{el} \tag{8.4}$$

The value of the total transmitted power must be the same no matter if mechanical or electrical parameters enter its expression. The instantaneous active power *p(t)* is calculated:

$$p(t) = \omega_{el} \cdot T_{el} = \frac{p}{2} \cdot \omega_{mec} \cdot \frac{2}{p} \cdot T_{mec} = \omega_{mec} \cdot T_{mec} \tag{8.5}$$

In the following only two pole machines are considered. Angles/angular velocities and torque are thereby angles/angular velocities and torque per pole

pair. Equation (8.3) and (8.4) are used to calculate mechanical rotation speed and torque.

8.4 Torque production

The main task of an electrical machine is to generate a mechanical torque. The torque, which depends on rotation speed and rotor position, may be used as it is generated, for example when laminated metal sheets or paper have to be rolled up on a roll maintaining a given tensional stress in the material. The generated torque can be used to indirectly control the rotation speed of a shaft. In other more advanced applications rotation speed is used to control position. An example for this latter case is robots. All these applications rely basically on torque control. It is therefore crucial to have detailed knowledge about how torque is generated in different types of electrical machines and the possibilities to optimally control the torque required for a specific application.

The following section treats torque generation with a very simple machine model. Then the basic principles for torque generation in more complex dc and ac machines are detailed. Finally each machine type is treated separately including specific methods for practical torque control.

8.5 Sinusoidal mmf and flux waves

It is generally of interest to obtain mmfs (*m*agneto *m*otoric *f*orces) and flux density waves with sinusoidal spatial distribution in electrical machines.

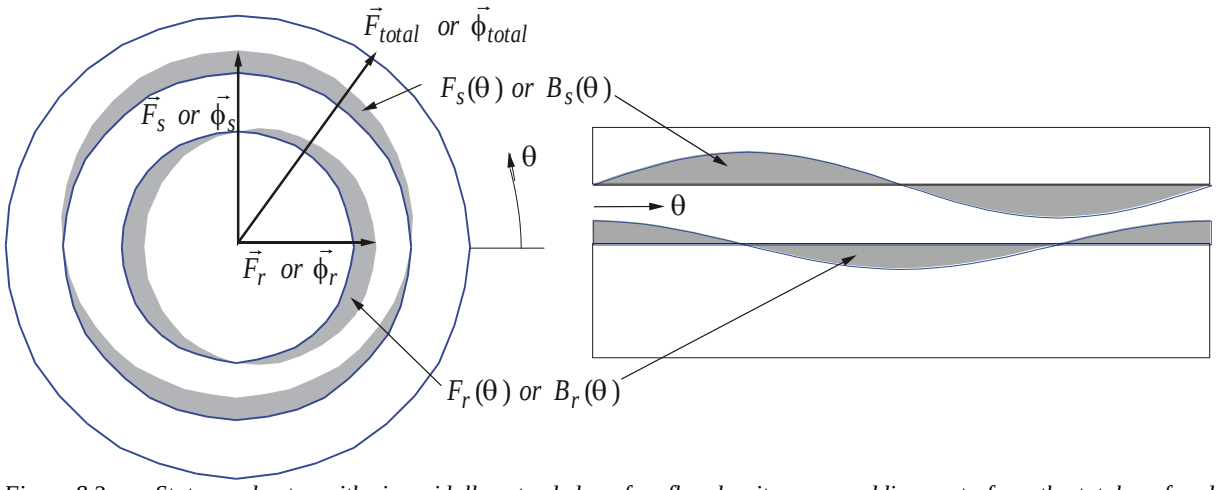


Figure 8.3 *Stator and rotor with sinusoidally extended mmf or flux density waves adding up to form the total mmf and flux density waves*

Sinusoidal distribution means that the mmf [$F(\square)$] or the flux densities [$B(\square)$] show a sinusoidal dependence on the angle as Figure 8.3 shows. This becomes more transparent if the circular cross section of the machine is viewed as linear. (See right hand drawing in Figure 8.3 above). The shaded area indicates the

mmf contribution or flux density, from the stator's inner surface mmf or flux density and the rotor's mmf or flux density.

The positive reference directions are chosen to the right and upper half plane in Figure 8.3, causing the mmf or flux density to be directed out of the stator inner surface for the angle range $[0,\pi]$ and into the stator surface in the interval $[\pi,2\pi]$. The reasons why the mmf or flux waves should be sinusoidal distributed are two:

6. If several windings contribute to the resultant mmf or flux wave the wave form of the resultant is the same as of the component waves.
7. The waves can be described by complex vectors, which in turn may be added with superposition. The vectors are represented in a complex reference frame with $\theta=0$ on the real axis and $\theta=\pi/2$ on the imaginary axis. The vectors are directed along the peak of the quantity they describe (e.g. the peak of the flux density wave) and have a magnitude proportional to the maximum of the quantity they describe. As an example the orientation of the flux vector is given by the main flux density orientation and the peak value is determined by the total flux.

For the example in Figure 8.3 the total flux is:

$$\phi_s = \int_0^{\pi} B_s(\theta)\cdot l\cdot r\cdot d\theta \quad \text{where} \begin{cases} l = \text{stator length} \\ r = \text{stator radius} \end{cases} \tag{8.6}$$

The magnetic flux is driven by the magneto motive force in exactly the same manner as the electric current is driven by the electromotive force (voltage). The scaling factor between the driving force, the mmf, and the corresponding magnetic flux is called *reluctance*. In a given direction the reluctance is defined as the ratio between the average mmf in that direction and the resulting flux in the same direction. The reluctance for the resulting flux in Figure 8.3 is therefore:

$$R = \frac{F_{total,average}}{\phi_{total}} = \frac{\frac{2}{\pi}\cdot\hat{F}_{total}}{\phi_{total}} \tag{8.7}$$

Normally the materials chosen for electrical machines have high magnetic "conductivity" (μ_r has very large values) and they can be considered as magnetic "short circuits". The reluctance is therefore almost exclusively determined by the air gap between the stator and the rotor. If the air gap is constant the reluctance will be the same for all directions (see Figure 8.3). For machines with salient poles the reluctance will be different for different directions, as the following section will describe further.

8.6 The permanent magnet machine

DC machines and synchronous machines where the field windings have been replaced with permanent magnets are generally called permanent magnet machines. For the following we go even a step further and assume the armature windings replaced with permanent magnets as well. The permanent magnet synchronous machine (PMSM) will serve as an example. The rotor contains permanent magnets and is therefore permanently magnetized. The generated mmf wave has approximately sinusoidal spatial distribution.

We choose a rotor with salient poles and thus the reluctances in the main rotor direction (direction of rotor magnetization in Figure 8.4) and the perpendicular directions are different. The purpose of this PM-PM-design is to demonstrate how the magnitude and direction of the mmfs together with the rotor saliency determine the resulting mechanical torque.

Assume that:

1. the machine has salient poles and is manufactured as a two-pole design
2. the cross-section is represented in the complex plane (x,y) and the x axis is given by the main rotor direction (also called the direct axis), which is the orientation of the salient poles
3. the rotor generates a sinusoidal spatially extended mmf wave with its main orientation along the direct rotor axis (see Figure 8.4). The rotor has a sinusoidal spatial mmf represented by the complex vector $\vec{F}_m$ where $\left|\vec{F}_m\right| = \hat{F}_m$
4. the stator is cylindrical and generates also a sinusoidal spatial extended mmf wave with its main orientation leading the rotor itself at an angle γ. The stator mmf wave is described with the complex vector $\vec{F}_s$ where $\left|\vec{F}_s\right| = \hat{F}_s$.
5. the construction determines the reluctances $\mathcal{R}_x$ and $\mathcal{R}_y$ in x- and y-direction respectively
6. stator and rotor materials - often iron - have infinitely large permeability, i.e. no magnetic energy is stored in the iron.

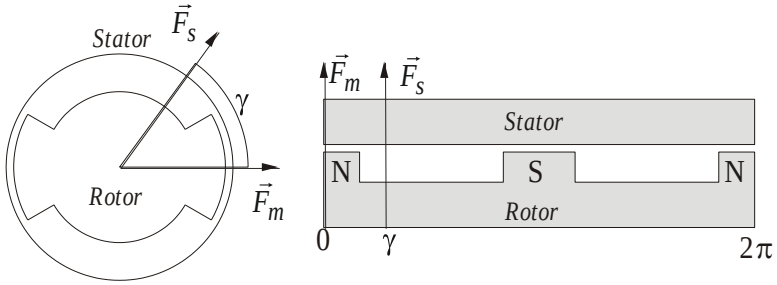


Figure 8.4 Cross section of the PM-PM-machine (left) and linear cross section (right).

The resulting mmf is calculated as:

$$\vec{F}_\delta = \vec{F}_m + \vec{F}_s = F_{\delta x} + F_{\delta y} \tag{8.8}$$

where

$$\begin{aligned} F_{\delta x} &= \hat{F}_s \cdot \cos(\gamma) + \hat{F}_m = \hat{F}_{sx} + \hat{F}_m \\ F_{\delta y} &= \hat{F}_s \cdot \sin(\gamma) = \hat{F}_{sy} \end{aligned} \tag{8.9}$$

In the air gap we have a linear dependency between the mmf and the magnetic flux. It can be demonstrated that the magnetic energy stored in the air gap is expressed as a function of the reluctances R_x and R_y:

$$\begin{aligned} W_{magn} &= \frac{1}{2}\frac{\hat{F}_x^2}{R_x} + \frac{1}{2}\frac{\hat{F}_y^2}{R_y} = \\ &= \frac{1}{2}\cdot\left(\frac{\hat{F}_s^2 \cdot \cos^2\gamma + 2\cdot\hat{F}_s\cdot\hat{F}_m\cdot\cos\gamma + \hat{F}_m^2}{R_x}\right) + \frac{1}{2}\cdot\left(\frac{\hat{F}_s^2\cdot\sin^2\gamma}{R_y}\right) \end{aligned} \tag{8.10}$$

For an angular displacement of the armature mmf dγ the mechanical energy dW_{mec} is transferred with the torque T as:

$$\frac{dW_{mec}}{d\gamma} = T \tag{8.11}$$

If the system has no losses and does not exchange other energy types than mechanical and electrical, the sum of the electrical and mechanical energy of the system must be constant according to the energy conservation principle:

$$W_{magn} + W_{mec} = \text{constant} \tag{8.12}$$

From equations (8.10), (8.11) and (8.12) the torque T is obtained as:

$$T = -\frac{dW_{magn}}{d\gamma} = \ldots = \frac{\hat{F}_{sy}\cdot\hat{F}_m}{R_x} + \hat{F}_{sx}\cdot\hat{F}_{sy}\cdot\left(\frac{1}{R_x} - \frac{1}{R_y}\right) \tag{8.13}$$

If one of the mmf's is generated by electrical windings the torque is called an *electro dynamical* torque. It has basically two components; the first component is determined only by the rotor mmf and stator mmf in the *y* direction together with the permeance in *x* direction. This component causes for example a circular compass needle to adjust itself to the direction of an external magnetic field. The second component is called *reluctance torque* and is given only by F_{sx} and F_{sy} together with the difference in permeance (!) in the *x* and *y* directions respectively. This component does not depend on the rotor magnetization. It is the *reluctance torque* that causes an entirely non magnetized iron nail to adjust its direction to the direction of an external magnetic field. Figure 8.5 shows five different examples. The grey arrows indicate the direction of the magnetization of the rotor and the external field. The thin black arrows indicate the direction of the torque. Note that the two right most cases are equal in geometry, but different in magnetization. The rotor saliency does not have to coincide with the rotor magnetization and the two rightmost cases represent the cases when the rotor is magnetized along the saliency and perpendicular to the saliency respectively. In the rightmost case the torque can only be determined by (8.13) since the two terms in the equation have opposite sign.

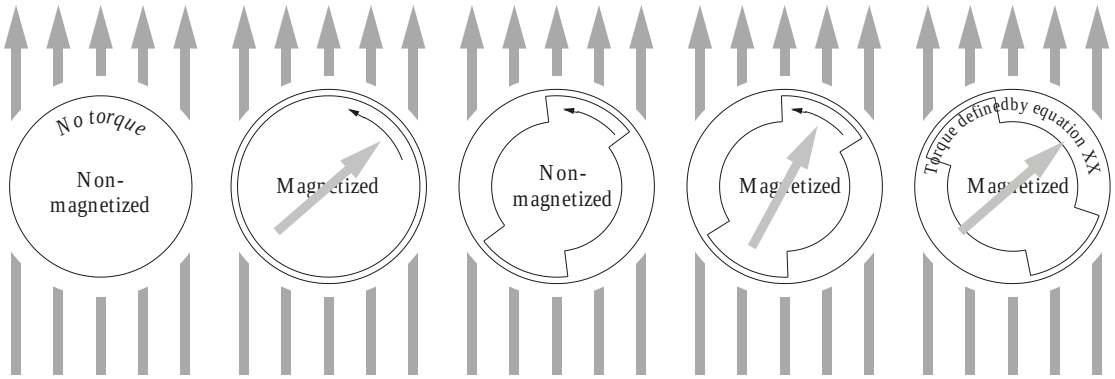
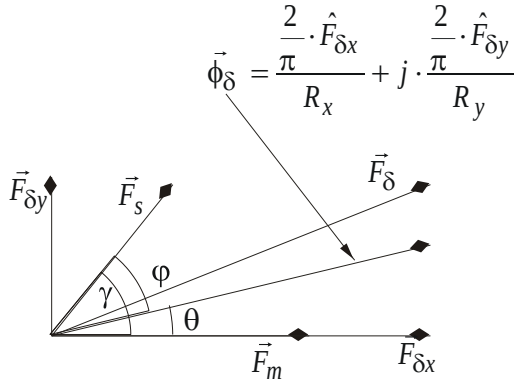


Figure 8.5 Electrodynamic and reluctance torque.

Expression (8.13) can be further simplified using the total magnetic flux in the air gap ϕ_δ and the armature mmf F_s. Since the reluctances in *x* and *y* direction are different, the resulting flux will not coincide with the resulting mmf. The larger reluctance in *y* direction will cause the mmf to generate a smaller flux in *y* direction than it would generate in *x* direction. The *y* component of the resulting flux will therefore be proportionally smaller. This leads to the following equations for the torque *T*:

$$
\begin{aligned}
T &= \frac{1}{R_x}\cdot\left(\hat{F}_m + \hat{F}_{sx}\right)\cdot\hat{F}_{sy} - \frac{1}{R_y}\cdot\hat{F}_{sy}\cdot\hat{F}_{sx} = \\
&= \frac{\pi}{2}\cdot\left(\phi_{\delta x}\cdot\hat{F}_{sx} - \phi_{\delta y}\cdot\hat{F}_{sy}\right) = \\
&= \frac{\pi}{2}\cdot\phi_\delta\cdot\hat{F}_s\cdot\left(\cos\theta\sin\gamma - \sin\theta\cos\gamma\right) = \\
&= \frac{\pi}{2}\cdot\phi_\delta\cdot\hat{F}_s\cdot\sin\left(\gamma-\theta\right) = \frac{\pi}{2}\cdot\phi_\delta\cdot\hat{F}_s\cdot\sin\left(\varphi\right) = \\
&= \frac{\pi}{2}\cdot\vec{\phi}_\delta\cdot\vec{F}_s
\end{aligned}
\qquad (8.14)
$$

Equation (8.14) shows an important relationship valid both for machines with and without salient poles. Torque can always be expressed as the cross product of the air gap flux vector and armature current vector.



8.7 Armature mmf generated by an electric winding

Disregard for a moment the rotor in the examples of the preceding section and let the permanent magnetized stator be replaced by a stator with electric windings like in Figure 8.6. The winding is a so-called diameter coil with N_s turns. The winding is placed in a special slot in the inner stator periphery, as Figure 8.6 shows.

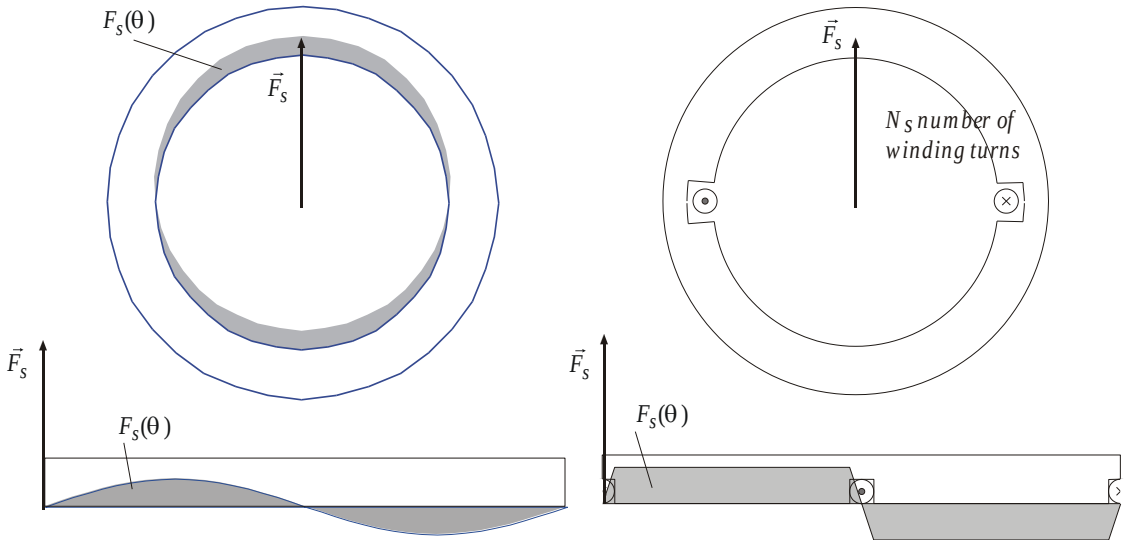


Figure 8.6 Ideal PM-magnetized armature circuit (left) and electrically magnetized circuit (right).

Since the armature winding is a diameter winding the mmf wave will have a rectangular shape, in Figure 8.6 drawn as almost rectangular due to the finite distribution of the conductor slots. The easiest way to demonstrate this fact is by using Ampere's circuit law, which states that the mmf along a closed path equals the total current crossing the surface enclosed by the path.

$$F = \int H \cdot dl \tag{8.15}$$

The mmf along, an arbitrary integration path in Figure 8.6 (right), equals $N_s \cdot i_s$. The assumption of having the reluctance of the magnetic circuit concentrated in the air gap leads to the integral of H to be concentrated in the air gap as well. For symmetry reasons the magnetic field intensity H will have the same value on opposite sides of the rotor but changed sign. This causes the mmf to have the same distribution around the rotor surface. As the mmf has the same value independent of the integration path the mmf wave will be rectangular. This becomes transparent in the stretched view of the winding. The amplitude of the mmf wave becomes:

$$\hat{F}_s = \frac{N_s \cdot i_s}{2} \tag{8.16}$$

The Fourier expansion contains the amplitudes for the fundamental wave and the harmonics:

$$\begin{aligned} \hat{F}_{s1} &= \frac{4}{\pi} \cdot \frac{N_s \cdot i_s}{2} = \frac{2}{\pi} \cdot N_s \cdot i_s \\ \hat{F}_{sn} &= \frac{4}{\pi} \cdot \frac{N_s \cdot i_s}{2 \cdot n} = \frac{2}{\pi} \cdot \frac{N_s \cdot i_s}{n} \quad n = 3,5,7,... \end{aligned} \tag{8.17}$$

For ac machines having several stator windings the mmf's interact and it is desired to reduce the content of harmonics in the armature mmf as much as possible. The harmonics content affects the torque shape causing it to present undesired ripples. For this reason the stator windings are placed in a distributed manner (Figure 8.7) in contrast to diameter winding.

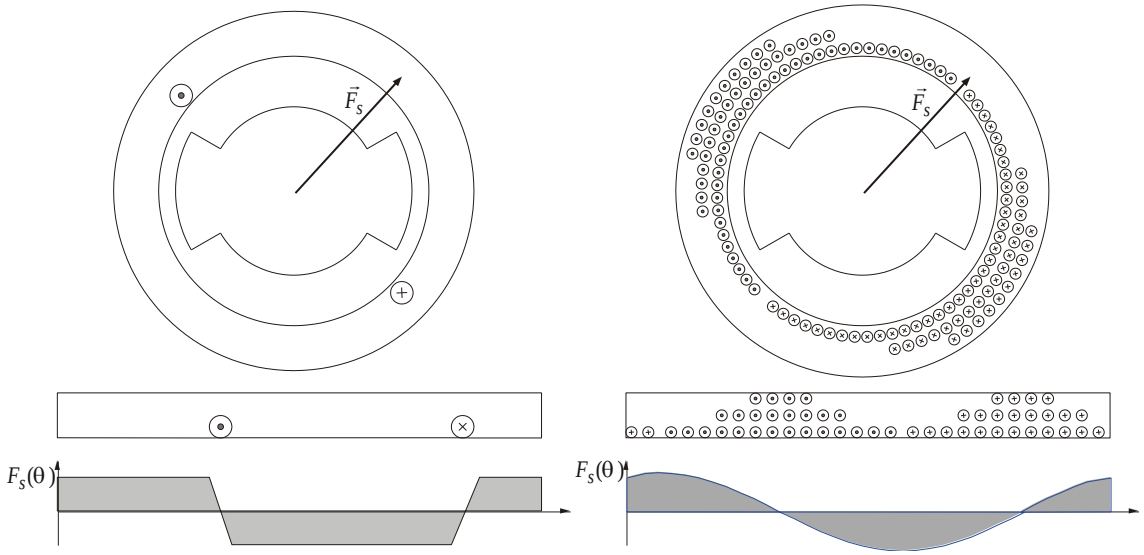


Figure 8.7 Diameter winding (left) and distributed winding (right).

Further design refinements to reduce the harmonics content in the armature mmf wave are often necessary in practice. Since it is not the aim of this text to discuss these design details, the interested reader is referred to specialized literature in the field of electrical machine design. The amplitude of the fundamental and harmonics in the armature mmf wave will be smaller for the distributed winding compared to the diameter winding.

A corrective factor is therefore considered in the expression for the corresponding amplitude. It is commonly called the winding factor k_r.

$$\hat{F}_{s1} = \frac{2}{\pi} \cdot N_s \cdot i_s \cdot k_{r1} = \left(\frac{2}{\pi} \cdot N_s \cdot k_{r1} \right) \cdot i_s$$
$$\hat{F}_{sn} = \frac{2}{\pi} \cdot \frac{N_s \cdot i_s}{n} \cdot k_{r1} = \left(\frac{2}{\pi} \cdot N_s \cdot k_{rn} \right) \cdot \frac{1}{n} \cdot i_s \quad n = 3,5,7,... \tag{8.18}$$

When using a distributed winding the mmf fundamental is reduced (k_{r1}<1) but at the same time the factors k_{rn} for the harmonics are reduced even more, which emphasizes the fundamental component. The following assumptions are now made: k_{rn}=0 and the effective number of turns for the armature winding is with regard to the fundamental wave:

$$N_{s,eff} = N_s \cdot k_{r1} \tag{8.19}$$

Introducing the armature current the complex expression for the armature mmf becomes:

$$\vec{F}_s = \hat{F}_{sx} + j\hat{F}_{sy} = \frac{2}{\pi} \cdot N_{s,eff} \cdot \vec{i}_s = \frac{2}{\pi} \cdot N_{s,eff} \cdot \left(i_{sx} + j i_{sy}\right) \tag{8.20}$$

This expression defines also the armature current vector with its modulus equaling the current amplitude and the orientation of the main direction of the mmf. With (8.13) and (8.20) the torque is calculated as:

$$\begin{aligned} T &= \frac{1}{R_x} \cdot \left(\hat{F}_m + \hat{F}_{sx}\right) \cdot \hat{F}_{sy} - \frac{1}{R_y} \cdot \hat{F}_{sy} \cdot \hat{F}_{sx} = \frac{1}{R_x} \cdot \hat{F}_m \cdot \hat{F}_{sy} + \hat{F}_{sy} \cdot \hat{F}_{sx} \cdot \left(\frac{1}{R_x} - \frac{1}{R_y}\right) = \\ &= \frac{1}{R_x} \cdot \hat{F}_m \cdot \frac{2}{\pi} \cdot N_{s,eff} \cdot i_{sy} + \left(\frac{2}{\pi} \cdot N_{s,eff}\right)^2 i_{sx} \cdot i_{sy} \cdot \left(\frac{1}{R_x} - \frac{1}{R_y}\right) = \\ &= \psi_m \cdot i_{sy} + i_{sx} \cdot i_{sy} \cdot \left(L_{mx} - L_{my}\right) = \left(\psi_m + L_{mx} \cdot i_{sx}\right) \cdot i_{sy} - L_{my} \cdot i_{sx} \cdot i_{sy} = \\ &= \psi_{\delta x} \cdot i_{sy} - \psi_{\delta y} \cdot i_{sx} \end{aligned} \tag{8.21}$$

or, with the lower most row in equation (8.14):

$$T = \frac{\pi}{2} \cdot \vec{\phi}_\delta \times \vec{F}_s = \left(\frac{\pi}{2} \cdot \vec{\phi}_\delta\right) \times \left(\frac{2}{\pi} \cdot N_{s,eff} \cdot \vec{i}_a\right) = \vec{\psi}_\delta \times \vec{i}_a = \psi_\delta \cdot i_a \cdot \sin(\varphi) \tag{8.22}$$

where

ψ_m = *flux from rotor magnets linking the stator windings*
ψ_δ = *air gap flux linking the stator windings*
L_{mx} = *main inductnace in the x – direction*
L_{my} = *main inductnace in the y – direction*

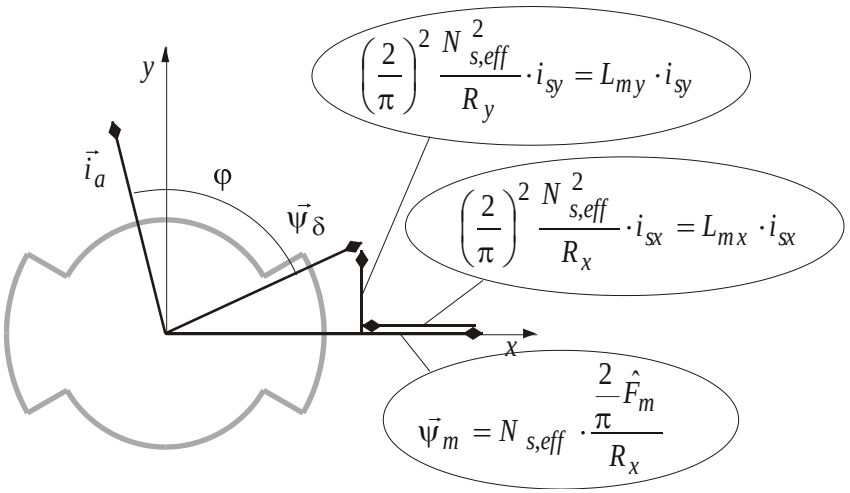


Figure 8.8 Armature current and resulting flux.

The mechanical torque can hereby be expressed in several different ways. However, it can always be written as the cross product of the air gap flux and the armature current vectors no matter what reference frame is chosen. More explicitly stated, the torque is the scalar product of the two vectors weighted with the sine of the angle between them. The above relation is of fundamental importance for understanding torque generation in electrical machines. Therefore the reader is advised to spend some thoughts about this relation. Some questions might be of help in this context:

7. Decompose the stator current vector in components along and across the air gap flux vector. Which one is the component generating torque with the air gap flux ?
8. For a machine with a so-called turbo rotor (L_{mx}=L_{my}), what role does i_{sx} play in the torque production?
9. The mechanical torque can hereby be expressed in several different ways. However, it can always be written as the cross product of the air gap flux and the armature current vectors no matter what reference frame is chosen. More explicitly stated, the torque is the scalar product of the two vectors weighted with the sine of the angle between them. The above relation is of fundamental importance for understanding torque generation in electrical machines. Therefore the reader is advised to spend some thoughts about this relation. Some questions might be of help in this context:
10. Decompose the stator current vector in components along and across the air gap flux vector. Which one is the component generating torque with the air gap flux ?
11. For a machine with a so-called turbo rotor (L_{mx}=L_{my}), what role does i_{sx} play in the torque production?

8.8 Rotation

In the preceding section the mechanical torque generated by an electrical machine was shown to be determined by the armature current, for which a complex vector $\vec{i}_s$ is defined, and the air gap flux, also described by a corresponding vector $\vec{\psi}_\delta$. In general there is an angle between these two vectors in the complex plane. The armature current has up to now only been considered flowing in the distributed stator windings. But it can be shown that the same equations are also valid for the case where permanent magnet is located in the stator and the distributed windings in the rotor, i.e. an “inverted” machine. If stator and rotor are not rigidly fixed in space the torque will give rise to a relative torque between them. To maintain the torque at a required value the angle between $\vec{\psi}_\delta$ and $\vec{i}_s$ is required to be constant, that is the armature winding

has to move relative to the stator, or, for the machine with the armature winding placed on the rotor, relative to the rotor.

One can imagine that displacing the armature conductors relative to the stator where they are mounted, is an impractical solution aiming at assuring a constant angle between the armature current and air gap flux vector when the rotor is rotating. There are several methods to achieve the same result, and the most important ones are discussed below.

8.9 Multiphase armature windings

Two phases

Consider a machine with the armature winding on the stator as in figure Figure 8.9 (left). A fixed reference frame (□□□) with the real axis (□) on the horizontal is assumed, having the origin on the rotational symmetry axis. The armature winding in Figure 8.9 (left) is replaced by an equivalent of two windings localized on the □- and □-axis respectively in Figure 8.9 (right).

For a distributed armature winding we can assume the armature mmf wave in Figure 8.9 (left) to be approximately sinusoidal. Then the armature winding may be replaced by two distributed armature windings where the induced currents are $i_{s\alpha}$ and $i_{s\beta}$ respectively, so that the resulting mmf remains unchanged.

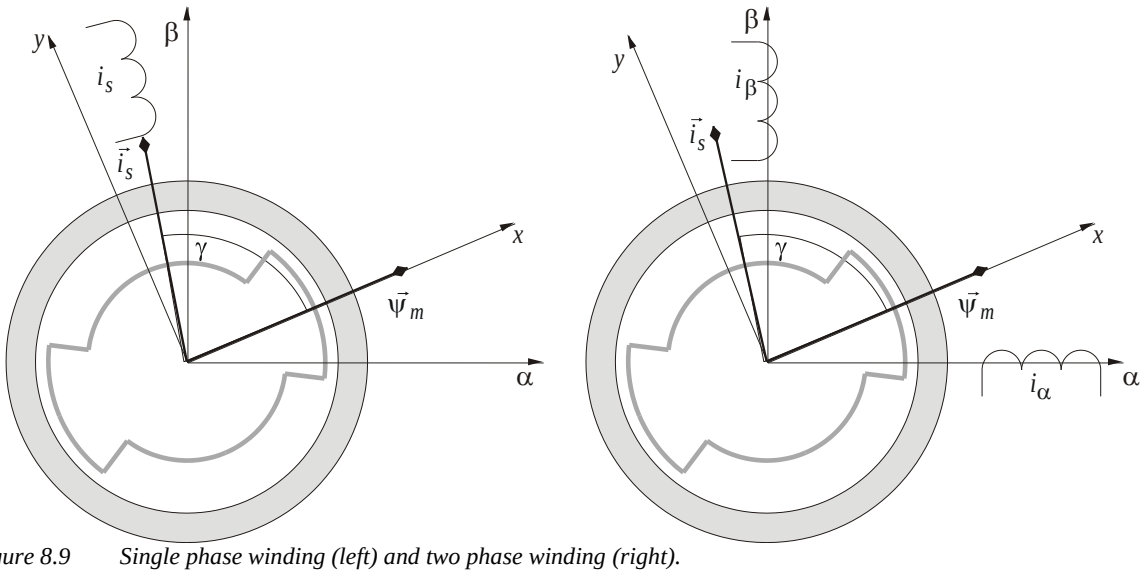


Figure 8.9 Single phase winding (left) and two phase winding (right).

The vector representation of the armature current in the reference frame (*x,y*) has already been used:

$$\vec{i}_s^{xy} = i_{sx} + ji_{sy} = i_s \cdot e^{j\gamma} \tag{8.23}$$

By means of a simple coordinate transformation we can express the armature current vector in the reference frame (α,β).

$$\vec{i}_s^{\alpha\beta} = \vec{i}_s^{xy} \cdot e^{j\theta_r} = i_s \cdot e^{j(\theta_r+\gamma)} = i_{s\alpha} + ji_{s\beta} \tag{8.24}$$

Consider the rotor performing a rotational motion with the angular speed ω_r ($\theta_r = \omega_r \cdot t$) and with constant torque. i_{sx} and i_{sy} are therefore constant, that is, using amplitude and angle, i_s and γ are constant:

$$i_{s\alpha} + ji_{s\beta} = i_s \cdot \cos(\omega_r t + \gamma) + ji_s \cdot \sin(\omega_r t + \gamma) \tag{8.25}$$

Equation (8.25) shows that the new components of the armature current in the reference frame (α,β), that is $i_{s\alpha}$ and $i_{a\beta}$, are sinusoidal alternating currents. Thus, in order to keep the angle between the armatures current and air gap flux vector constant we could replace the dc current carrying armature winding (that would require to be rotated relative to the armature) with two ac current carrying windings.

Three phases

In the most practical applications ac machines are built with three phase stator windings (Figure 8.10 b). Two phase stator windings are only used in a few special cases. Several reasons exist for this, e.g. the sum of all three phase currents is zero in a sinusoidal symmetrical three phase system allowing to operate the system under symmetric three phase load without the return wire (known also as the neutral wire).

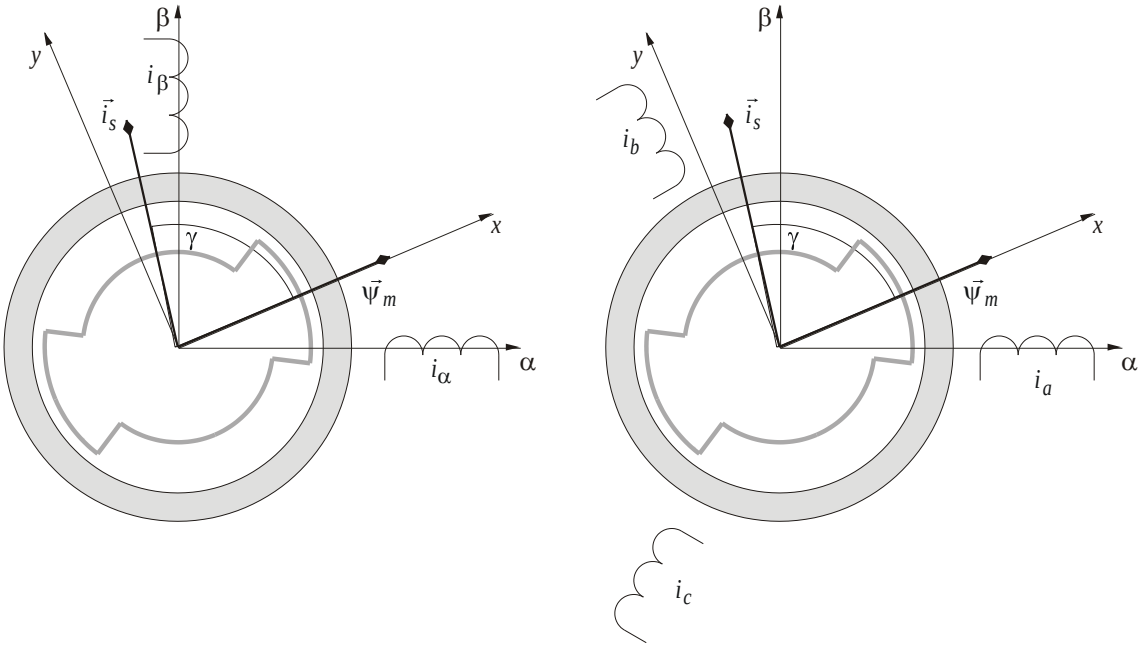


Figure 8.10 Two phase armature winding (left) and three phase armature winding (right).

The windings are symmetrically distributed along the stator periphery forming angles of 120 degrees between them. The ac currents carried by these three windings can be calculated with the earlier introduced $i_{s\alpha}$ and $i_{s\beta}$ using a two-phase-to-three-phase transformation. For details see Chapter **Error! Reference source not found.**.

$$\begin{aligned} i_a &= \sqrt{\frac{2}{3}} \cdot i_{s\alpha} \\ i_b &= \sqrt{\frac{2}{3}} \cdot \left(-\frac{1}{2} i_{s\alpha} + \frac{\sqrt{3}}{2} i_{s\beta} \right) \\ i_c &= \sqrt{\frac{2}{3}} \cdot \left(-\frac{1}{2} i_{s\alpha} - \frac{\sqrt{3}}{2} i_{s\beta} \right) \end{aligned} \tag{8.26}$$

In conclusion we can say that to generate a given torque the necessary currents are calculated first as i_{sx} and i_{sy}. From these values $i_{s\alpha}$ and $i_{s\beta}$ result calculated by means of a coordinate transformation. Finally, the three phase currents i_a, i_b and i_c.result from the two-phase-to-three-phase transformation. Appropriate voltages are further required to drive these currents and to compensate the emfs induced in the stator windings and the resistive voltage drops.

Resistance, leakage inductance and induced emf

Up to now only the inductances L_{mx} and L_{my} mapping the armature current into the stator air gap flux linkage for the *x* and *y* direction respectively have been considered. In practice part of the stator flux will not link the armature but the stator winding only, thus not contributing to the air gap flux, as Figure 8.11 illustrates.

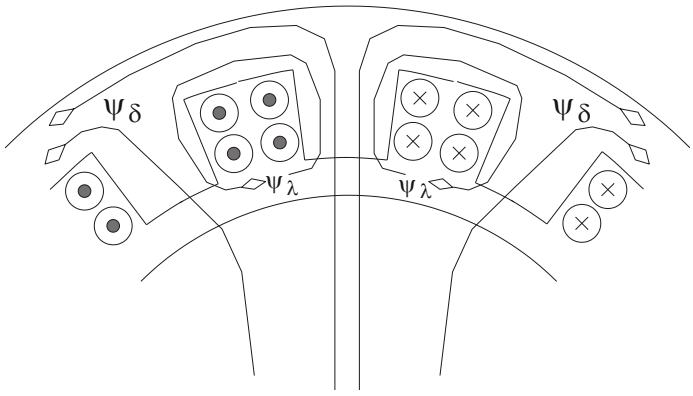


Figure 8.11 Air gap flux (Ψδ) leakage flux (Ψλ).

This part of the stator flux is commonly called leakage flux and is usually indexed with λ. Since the leakage flux does not contribute to the air gap flux it does not contribute to torque generation either. It is described by its equivalent leakage inductance and from the point of view of torque generation this

inductance could be considered as a component not belonging to the machine itself, but which nevertheless influences the voltage equations for the machine. The leakage flux does not contribute to the air gap flux but remains part of the entire stator flux and is calculated as:

$$\begin{aligned}\psi_{a\lambda} &= N_{s,eff} \cdot \phi_{a\lambda} = L_{s\lambda} \cdot i_a \\ \psi_{b\lambda} &= N_{s,eff} \cdot \phi_{b\lambda} = L_{s\lambda} \cdot i_b \\ \psi_{c\lambda} &= N_{s,eff} \cdot \phi_{c\lambda} = L_{s\lambda} \cdot i_c\end{aligned} \tag{8.27}$$

The entire voltage along a winding is the sum of all in the winding induced emfs and all resistive voltage drops along the winding. If we assume that for the three phase stator winding machine all three leakage inductances are constant (linear magnetic behavior) the following equations can be written:

$$\begin{aligned}u_a &= R_s \cdot i_a + \frac{d\psi_a}{dt} = R_s \cdot i_a + \frac{d}{dt}\left(\psi_{\delta a} + L_{s\lambda} \cdot i_a\right) \\ u_b &= R_s \cdot i_b + \frac{d\psi_b}{dt} = R_s \cdot i_b + \frac{d}{dt}\left(\psi_{\delta b} + L_{s\lambda} \cdot i_b\right) \\ u_c &= R_s \cdot i_c + \frac{d\psi_c}{dt} = R_s \cdot i_c + \frac{d}{dt}\left(\psi_{\delta c} + L_{s\lambda} \cdot i_c\right)\end{aligned} \tag{8.28}$$

where

$\psi_{\delta a,b,c}$ = the air gap flux linkages
$\psi_{a,b,c}$ = the total fluxes linking the three windings

A description of the electrical behaviour of the machine with the above equations starts to become difficult to handle when the derivatives of the air gap fluxes must be written explicitly, since they contain the inductances for the *x* and *y* direction which in turn depend on the rotor position. It is thus much more convenient to express the same equations as a vector equation, and then express the vector equation in the (*x*,*y*) reference frame since the machine parameters are constant with respect to the rotor angle in that reference frame. The method is the same as described in chapter 4 – the generic three phase load. In the (α,β) reference frame the three phase vector equation for the armature winding voltage may be written:

$$\vec{u}_s^{\alpha\beta} = R_s \cdot \vec{i}_s^{\alpha\beta} + \frac{d\vec{\psi}_s^{\alpha\beta}}{dt} = R_s \cdot \vec{i}_s^{\alpha\beta} + \frac{d}{dt}\left(\vec{\psi}_\delta^{\alpha\beta} + L_{s\lambda} \cdot \vec{i}_s^{\alpha\beta}\right) \tag{8.29}$$

The vectors in (8.29) are expressed in the stator coordinates (α,β). However, the vectors entering the three phase vector equation above may be expressed in an arbitrary reference frame, e.g. in (*x*,*y*) coordinates. The transformation is identical to the one used in chapter 4 to transfer the three phase vector equation

from the (α,β) reference frame to the (x,y) reference frame and yields the following results:

$$\vec{u}_s^{xy} = R_s \cdot \vec{i}_s^{xy} + \frac{d}{dt}\left(\vec{\psi}_\delta^{xy} + L_{s\lambda} \cdot \vec{i}_s^{xy}\right) + j\omega_r \cdot \left(\vec{\psi}_\delta^{xy} + L_{s\lambda} \cdot \vec{i}_s^{xy}\right) \qquad (8.30)$$

This equation represents the vector equation for the stator winding voltage where the vectors are expressed in rotor coordinates (x,y). As opposed to the similar equation (8.29) with the vectors expressed in stator coordinates, in (8.30) enters a term containing the rotor angular velocity and the resulting total stator flux. This term is the voltage induced in the stator winding due to the rotating stator flux wave. This voltage is of course entering equation (8.29) but there it is hidden in the terms containing the derivative of the stator flux expressed in stator coordinates. In (8.30) the derivative terms is zero except when the stator flux and stator current are changing in terms of modulus or phase angle in rotor coordinates. Separating real and imaginary parts the equations for the fictitious armature windings in the x and y directions are obtained (see equation (8.31).

$$\begin{aligned}
u_{sx} &= R_s \cdot i_{sx} + \frac{d}{dt}\left(\psi_m + L_{mx} \cdot i_{sx} + L_{s\lambda} \cdot i_{sx}\right) - \omega_r \cdot \left(L_{my} \cdot i_{sy} + L_{s\lambda} \cdot i_{sy}\right) = \\
&= R_s \cdot i_{sx} + \frac{d}{dt}\left(\psi_m + L_{sx} \cdot i_{sx}\right) - \omega_r \cdot L_{sy} \cdot i_{sy} \\
u_{sy} &= R_s \cdot i_{sy} + \frac{d}{dt}\left(L_{my} \cdot i_{sy} + L_{s\lambda} \cdot i_{sy}\right) + \omega_r \cdot \left(\psi_m + L_{mx} \cdot i_{sx} + L_{s\lambda} \cdot i_{sx}\right) = \\
&= R_s \cdot i_{sy} + L_{sy} \cdot \frac{di_{sy}}{dt} + \omega_r \cdot \left(\psi_m + L_{sx} \cdot i_{sx}\right)
\end{aligned} \qquad (8.31)$$

We showed before that in order to have constant torque i_{sx}, i_{sy} and ψ_m are required to be constant, that is all variables in (8.31) are constant in the steady state (constant torque) differing from the alternating variables connected to the stator winding and expressed in stator coordinates, which have the same frequency as the rotor angular velocity.

To conclude, the electrical description of the three phase armature winding is given by the differential equations (8.28). They can be written as the compact vector differential equation (8.29) in arbitrary coordinates using appropriate coordinate transformations. If the reference frame is chosen to be fix with respect to the rotating flux wave (that is a reference frame coupled to the rotor (x,y)) the variables determining the torque become constant for the steady state. This result can be used in analysis and control of ac machines and will be applied in the sections where different machine types are treated separately.

9 Speed Control in Electrical Drives

Motor control means that e.g. speed or sometimes position has to be controlled with high accuracy. As in all control problems, disturbance signals are present, which is the reason that closed loop control is needed at all. The most common disturbance is load variations, but parameter variations are also common, and the closed loop control system is used to compensate for such variations. In e.g. an industrial robot, the inertia can change rapidly whereas it varies slowly in a reeling machine. The electrical parameters, e.g. resistances and inductances, do also vary as a function of e.g. temperature.

Closed loop control is important not only to control a certain quantity, but also to limit it. Thus an important function of a current controller in an electrical machine is to limit the maximum possible current in a fault situation.

The control loops in an electrical drive are usually always connected in cascade. This concept does only work if the bandwidth increases from the outer to the inner loops. The innermost loops are thus the fastest and the outermost the slowest. This means that the position control loop only can work well if the speed control loop quickly responds to any change in the reference value. The setting of suitable control parameters can thus be simplified if the inner loops are regarded as infinitely fast, or can be modeled with a simple model.

In this chapter speed control is discussed, without any demand for the type of torque source. The actual motor can thus be of any kind without affecting the design of the speed control loop. The mechanical system is briefly repeated in section 9.1. In section 9.2 the basics of cascade control are repeated. In chapter 9.3 the fundamental characteristics of the PI controller is discussed and in chapter 9.4 this is applied to speed control.

9.1 The mechanical system

The torque produced by the electrical machine, the electromechanical energy converter, is denoted T_{el} whereas the load torque that is a result of friction etc. is denoted T_l. A rotating mass with the inertia J has a speed ω that is determined by the differential equation:

$$\frac{d}{dt}(J \cdot \omega) = T_{el} - T_l \tag{9.1}$$

The inertia J is not necessarily constant and thus the derivative can be written as:

$$\frac{d}{dt}(J\cdot\omega)=J\cdot\frac{d\omega}{dt}+\omega\cdot\frac{dJ}{dt}=T_{el}-T_l \tag{9.2}$$

There are several examples of time varying inertia, e.g. industrial robots, reeling machines for paper and spin-dryers. In this context however we will consider J as constant. In the discussion so far we have assumed that the mechanical system is stiff. That is not true in many larger dives, i.e. the torsion of the shaft must be considered. Such analysis is also left out of this context.

9.2 Cascade control

If it is assumed that the torque production is infinitely fast, i.e. the real torque responds to changes in the reference value without time delay, then a block diagram of a speed and position control system can be described as in figure Figure 9.1.

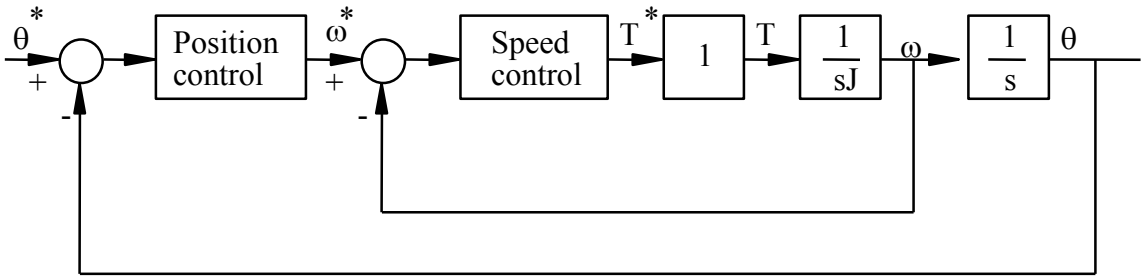


Figure 9.1 Block diagram for speed and position controllers.

The system contains two integrations. This gives a hint about two properties of the system:

12. one must watch out for the stability margins;
13. integrators may help to eliminate remaining errors.

The feedback in the speed loop has a directly stabilizing effect. To illustrate this, assume that both controllers are of proportional type with gain K_p for the position loop and K_ω for the speed loop. The closed system transfer function is then:

$$\frac{\theta(s)}{\theta^*(s)}=\frac{K_p\cdot K_\omega}{J\cdot s^2+K_\omega\cdot s+K_p\cdot K_\omega} \tag{9.3}$$

with the poles:

$$poles=-\frac{K_\omega}{2J}\pm\sqrt{\frac{K_\omega^2}{4J^2}-\frac{K_p\cdot K_\omega}{J}}$$

The poles are always in the left half plane if K_{ϖ} and K_p are both positive. With these two parameters the poles of the position control system can be placed arbitrarily. Cascade control is thus nothing but a state feedback from the two states position and speed.

If we try to control the position with only one controller, i.e. the speed controller feedback is cut off and K_{ϖ}=1, the corresponding closed system transfer function will be:

$$\frac{\theta(s)}{\theta^*(s)} = \frac{K_p}{J \cdot s^2 + K_p} \tag{9.4}$$

The poles of this system are both on the imaginary axis, i.e. the system cannot be stabilized with only a proportional controller, it has to be derivative. However, derivating a position signal is equivalent of measuring the speed.

9.3 The PI controller

The PI controller can be used in most control loops in an electrical drive. There are reasons to take a closer look at this controller. It is available both in analogue and digital forms. At very high sampling frequency the digital controller resembles the analogue, but the implementation of anti-windup functions are different.

Analogue PI controllers

The most commonly used parameterization of a PI controller uses the control error *e(t)*,

$$e(t) = y^*(t) - y(t) \tag{9.5}$$

where $y^*(t)$ is the reference value and *y(t)* is the measured value. The controller output is proportional to the error and the integral of the error,

$$u(t) = K_p \cdot \left(e(t) + \frac{1}{\tau_i} \int e(t)dt \right) \tag{9.6}$$

It can also be written as,

$$u(t) = K_1 \cdot e(t) + K_2 \cdot \int e(t)dt \tag{9.7}$$

which theoretically is equal to equation (9.6), but from practical reasons the difference is big. In the equation (9.6) controller the gain can be changed without changing the time constant. This is important in commissioning.

The transfer functions with the controllers in equations (9.6) and (9.7) differ in another important sense as well. For (9.6) the transfer function can be written

$$u(s)=K_p\frac{1+s\tau_i}{s\tau_i}e(s) \tag{9.8}$$

This means that at a change in K_p the breakpoint in the Bode diagram will be unchanged $1/\tau_i$. The transfer function with controller according to (9.9) is

$$u(s)=\left(K_1+\frac{K_2}{s\tau_i}\right)e(s) \tag{9.9}$$

As can be seen from (9.9) both gain and breakpoint will be affected by both K_1 and K_2.

Analog anti windup

If the control error *e(t)* remains for a longer time it is likely that the integral part becomes too large and saturates the control signal. In that case the integral part of the output signal must be limited. In a digital PI controller this can be done fairly easy by giving the integral part the value of the output (control) signal subtracted with the proportional part. In an analogue PI controller it is done with the use of two zener diodes in the feedback according to Figure 9.2

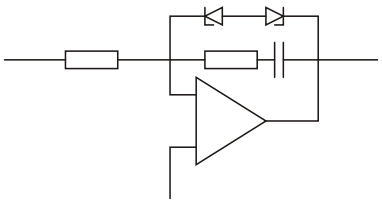


Figure 9.2 Anti windup connection of analogue PI controller.

Digital PI controllers

Digital PI-controllers can just like analog run into saturation of the output signal, i.e. limitation of the reference of the torque set point, and protection against "windup" must be introduced.

The digital PI controller is here assumed to operate at equidistant sampling instants denoted [k,k+1,k+2, …]. The output signal is built up of a proportional part and an integrating part according to:

$$u(k) = K_p \cdot \left(y^*(k) - y(k) + \frac{T_s}{T_i} \sum_{n=0}^{n=k} \left(y^*(n) - y(n) \right) \right) \tag{9.10}$$

where

$$K_p \cdot \frac{T_s}{T_i} \sum_{n=0}^{n=k} \left(y^*(n) - y(n) \right) = u_{\text{int}}(k) \tag{9.11}$$

and T_s is the sampling time of the controller. Thus the digital PI-controller can be expressed as:

$$u(k) = K_p \cdot \left(y^*(k) - y(k) \right) + u_{\text{int}}(k) \tag{9.12}$$

Note that the use of K_p in the calculation of the integral part of the controller has the same function as using equation (9.6) instead of (9.7) for the analog PI-controller, since a change of gain K_p will preserve the integral time constant.

$$u = K_p \frac{1 + \frac{T_s}{T_i} - z^{-1}}{\left(1 - z^{-1}\right)} \cdot e \tag{9.13}$$

Digital anti-windup

If there is a limitation in the control signal *u(k)* that may introduce windup of the integral part of the controller, an "anti-windup" function can easily be introduced with equation (9.14).

$$\begin{aligned}
&u_{\text{int}}(k) = K_p \cdot \frac{T_s}{T_i} \sum_{n=0}^{n=k} \left(y^*(n) - y(n) \right) \\
&u(k) = K_p \cdot \left(y^*(k) - y(k) \right) + u_{\text{int}}(k) \\
&\text{if } u(k) > u_{\max} \text{ then} \\
&\quad u(k) = u_{\max} \\
&\quad u_{\text{int}}(k) = u_{\max} - K_p \cdot \left(y^*(k) - y(k) \right) \\
&\text{elseif } u(k) < u_{\min} \text{ then} \\
&\quad u(k) = u_{\min} \\
&\quad u_{\text{int}}(k) = u_{\min} - K_p \cdot \left(y^*(k) - y(k) \right) \\
&\text{end if}
\end{aligned} \tag{9.14}$$

9.4 Adjustment of the speed controller

Assuming that the torque source is infinitely fast, the speed controller can be sketched as a block diagram in Figure 9.3. With a PI controller, the system will contain two integrators. It is possible to place the poles of that controller arbitrarily. Assume that the speed controller has the gain K and integration time constant T_i.

The closed system transfer function will then be

$$\frac{\omega(s)}{\omega^*(s)} = \frac{K(s \cdot T_i + 1)}{JT_i s^2 + KT_i s + K} \tag{9.15}$$

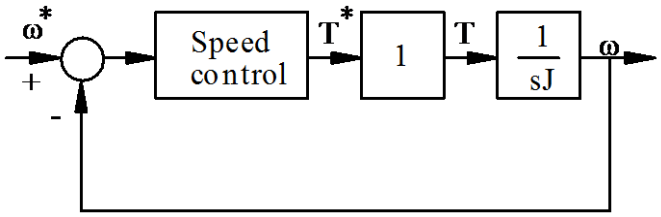


Figure 9.3 The speed control system.

The poles to the system will be

$$poles = -\frac{K}{2J} \pm \sqrt{\frac{K^2}{4J^2} - \frac{K}{JT_i}} \tag{9.16}$$

It can be seen from equation ((9.16) that the poles of the system can be placed arbitrarily with K and T_i. There are of course practical limitations to how fast the control system can be made. If K is to large, the assumption that the torque loop is infinitely fast are no longer valid. Furthermore the measured speed signal has to be filtered. If the toque loop dynamics are includes in the calculations a method called symmetric optimum can be used.

Symmetric optimum

A criterion often used in drives is "symmetric optimum", which is a standard criterion for circuits containing a double integrator. The idea is to select the cross over frequency ω_0 (when $|G(j\omega)| = 1$) at the frequency where the phase margin is maximal. This will give the best damping for the closed system

The actual torque response for a power electronically fed electrical machine is rather complicated to model. A first order approximation that is very useful is to assume that the actual torque responds to a change in the reference as a first

order low pass filter with a time constant equal to the sampling period in the torque control loop.

$$\frac{T(s)}{T^*(s)} = \frac{1}{1+s\cdot T_{tc}} \tag{9.17}$$

where T_{tc} is the band width of the torque control system. The speed control system can now be described as

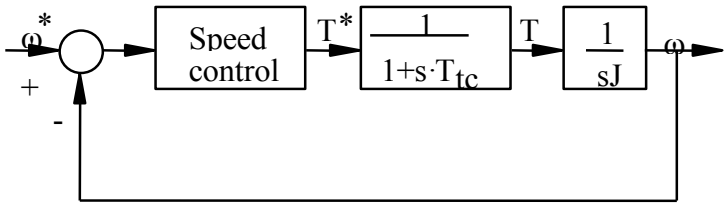


Figure 9.4 *Block diagram for the speed controller, with a first order approximation of the torque loop dynamics included.*

The open loop transfer function for the system in Figure 9.4 is,

$$G(j\omega) = K_\omega \frac{1+s\cdot T_i}{s\cdot T_i}\cdot\frac{1}{1+s\cdot T_{tc}}\cdot\frac{1}{s\cdot J} \tag{9.18}$$

We assume that $T_{tc} < T_i$. Then the Bode diagram for $G(j\omega)$ has the principal behaviour in Figure 9.5.

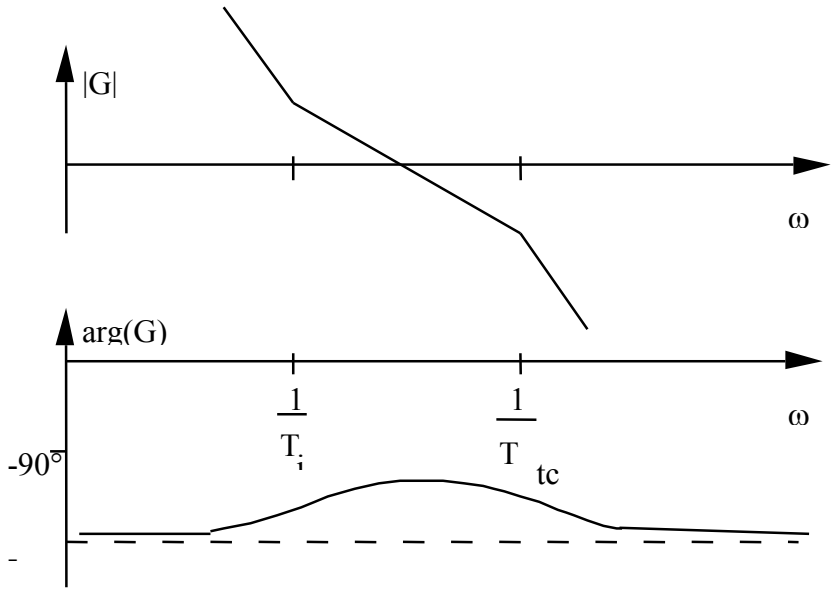


Figure 9.5 *The principal behaviour of G (j□) with PI controller, torque dynamics and mechanics included.*

The maximum phase appears between the two breaking frequencies $1/T_i$ and $1/T_{tc}$

$$\omega_0 = \frac{1}{\sqrt{T_i \cdot T_{tc}}} \tag{9.19}$$

To determine the parameters of the PI controller, two equations are needed. The condition $|G(j\omega)| = 1$ at the frequency ω_0 gives the first equation. Assuming that $T_{tc} < T_i$:

$$T_i = a^2 \cdot T_s, \text{ where} a > 1$$

$$\omega_0 = \frac{1}{\sqrt{T_i \cdot T_{tc}}} = \frac{1}{a \cdot T_{tc}} = \frac{a}{T_i} \tag{9.20}$$

At the cross over frequency it is thus valid that

$$|G(j\omega)| = \left| K_\omega \frac{1 + j \cdot a}{j \cdot a} \cdot \frac{1}{1 + \frac{j}{a}} \cdot \frac{1}{\frac{j \cdot a}{T_i} \cdot J} \right| = K_\omega \cdot \frac{T_i}{a \cdot J} = 1$$

or

$$K_\omega = \frac{a \cdot J}{T_i} \tag{9.21}$$

The closed systems poles can be calculated explicitly. The transfer function for the closed speed loop is:

$$\begin{aligned} \frac{\omega(s)}{\omega^*(s)} &= \frac{K_\omega \frac{1 + s \cdot T_i}{s \cdot T_i} \cdot \frac{1}{1 + s \cdot T_{tc}} \cdot \frac{1}{s \cdot J}}{1 + K_\omega \frac{1 + s \cdot T_i}{s \cdot T_i} \cdot \frac{1}{1 + s \cdot T_{tc}} \cdot \frac{1}{s \cdot J}} = \frac{K_\omega (1 + s \cdot T_i)}{s \cdot T_i \cdot (1 + s \cdot T_{tc}) \cdot s \cdot J + K_\omega (1 + s \cdot T_i)} = \\ &= \frac{K_\omega (1 + s \cdot T_i)}{s^3 \cdot J \cdot T_i \cdot T_{tc} + s^2 \cdot J \cdot T_i + s \cdot K_\omega \cdot T_i + K_\omega} \end{aligned} \tag{9.22}$$

With K_ω (9.21) and the relations (9.19), (9.20) inserted the characteristic equation can be written as

$$s^3 \cdot \frac{T_i^2}{a^2} + s^2 \cdot T_i + s \cdot a + \frac{a}{T_i} = 0 \tag{9.23}$$

One root is

$$s = -\omega_0 = -\frac{a}{T_i} \tag{9.24}$$

Polynomial division of the characteristic polynomial ((9.23) gives

$$s^2 \cdot \frac{T_i^2}{a^2} + s \cdot T_i \cdot \frac{a-1}{a} + 1 = 0 \tag{9.25}$$

and then the other roots can be calculated analytically. We assume that they are complex conjugates,

$$s_{2,3} = -\omega_0 \left(\xi \pm \sqrt{\xi^2 - 1} \right) \tag{9.26}$$

where

$$\zeta = \frac{a-1}{2} \tag{9.27}$$

and ω_0 is given by equation (9.19). We can see that ζ is the relative damping for poles to the system. The parameter T_i can now be selected to place the poles in a suitable way. If the relative damping is selected to $1/\sqrt{2}$, then a= 2.41. This gives directly a value on T_i as a function of T_{tc}. If we on the other hand assume that no complex poles are to exist in the speed control system that means that ζ = 1, corresponding to a = 3. All three poles will thus be equal to -ω_0.

Compensation for sensor noise

In most applications with speed control a speed sensor is needed, and the output signal of this speed sensor is usually subjected to noise that needs to be filtered out to some extent. Thus, the feedback loop involves a filter that affects the dynamics of the closed loop system. As shown above the characteristic equation of the closed loop is of the 3'rd order even without a filter, and with the filter included it would be of the 4'th order, thus rather difficult to find a useful algebraic solution.

A feasible engineering solution to this problem can be found by comparing the filter time constants of the speed sensor filter with the response time constant of the torque loop. The torque loop is usually very fast, expressed as a first order time constant the real torque follows the torque reference within 50…500 microseconds. The speed sensor filter time constant on the other hand is usually at least one and maybe two orders of magnitude longer, i.e. up to 50 milliseconds. To realize the impact of this on the control system parameter selection, a block diagram of the control system needs to be drawn. First it must

be noted that it is not the speed that is controlled, but the speed measurement signal after the filter. Note from Figure 9.6 that from a control point of view the torque dynamics and the filter dynamics are series connected.

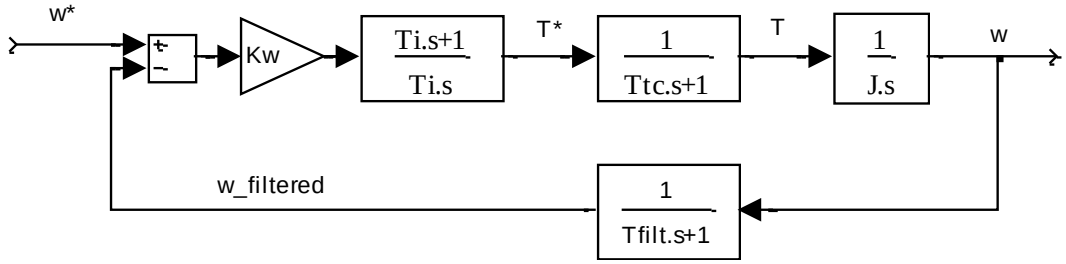


Figure 9.6 Speed control system with speed sensor filter dynamics.

Since the torque dynamics are at least one order of magnitude faster than the filter dynamics it can be omitted. With the torque dynamics omitted and the filter dynamics moved to structurally replace the torque dynamics it is evident that the system can be designed using symmetric optimum in the same way as in the previous section, but with the torque dynamics replaced by the filter dynamics.

10 The DC machine

The DC machine is the most common machine used for control purposes. The reason is that its construction allows simple methods to control the generated torque. The price paid for this advantage is the relatively complicated construction, which requires regular maintenance (especially the commutator with its sliding contacts which wear out and must be repaired/replaced). It is not used for power generation, since the cheaper and less service demanding ac machines, eventually with the aid of power electronics, can do the same job. The market for ac machines has grown continuously even in the field of control applications. Ac machines have the powerful advantage of robust mechanical construction but the disadvantage of needing more complicated torque controllers which however is less relevant as μ-processors become cheaper and faster.

Modern DC machines are technically well developed (they have been used for more than 150 years) but they are still very interesting for control applications. Among common applications for the DC machine we can mention servo drives, electric trucks, railroad traction and, not to forget, electric vehicles in general. A special type of DC machine is the so-called "universal motor" and is used in most house hold devices, as for example vacuum cleaners, food processors, hand drilling machines, electrical cork screws (!), etc., etc.

10.1 Mechanical construction

The mechanical construction of the DC machine is based on an inverted version of the machine discussed in chapter 8.9. The stator contains the magnets, or an electric field winding and the rotor contains the armature winding. With a field winding the stator is usually arranged with salient poles. Since eddy currents substantially deteriorate the dynamic properties of the machine its stator is usually manufactured of laminated metallic sheets. The armature winding in the rotor is an AC winding (!), but not arranged as a three phase winding. The armature winding is a single continuous winding with a finite number of nodes connected to a segment of a commutator, see Figure 10.1. The current distribution is not sinusoidal, but rectangular. This exception from the requirement of sinusoidal windings is compensated by the commutator that causes the rotor mmf distribution to be triangular, independent of the rotor position, see Figure 10.1. Since the requirement of sinusoidal winding distribution partly was made to allow superposition of several windings, and this is overcome by the commutator function, the vector theory can still be applied, keeping in mind that the mmf from the rotor is triangular, not sinusoidal. Two

sliding contacts are connected to two of the commutator segments and thus allowing electrical connection to the armature winding. The commutator segments with their insulation are pressed on the rotor axis.

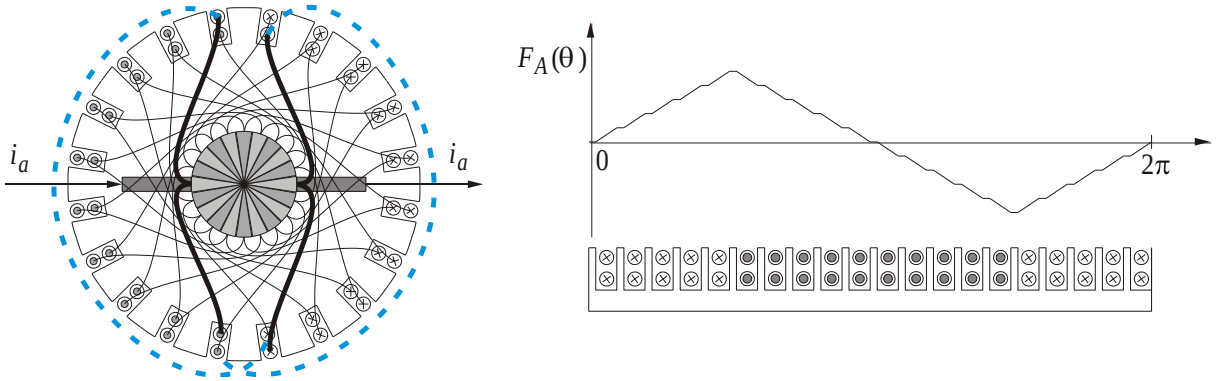


Figure 10.1 Rotor winding arrangement (right). (left) and triangular mmf from rectangular winding distribution

Besides the armature winding placed on the rotor the construction contains further two windings that also carry armature current. One of these additional windings is the so-called commutation winding and is arranged in such a manner that the armature current generates a local mmf interacting with the conductor closest to the sliding contacts. This additional winding does not influence the generated torque but helps to avoid sparking at the sliding contacts when current is switched from one segment to the other. The other additional winding is placed on the poles of the field winding and reduces the quadrature armature mmf to obtain only a direct fundamental of the resulting air gap flux. In other words, the inductance in the *y*-axis is negligible. In Figure 10.2 all conductors having the current direction indicated are carrying the same armature current. Small machines have in general a very simple construction: the stator is usually manufactured of stamped metal sheets and permanent magnets are often used to generate the main machine flux. Furthermore they do not have any commutator and compensation winding.

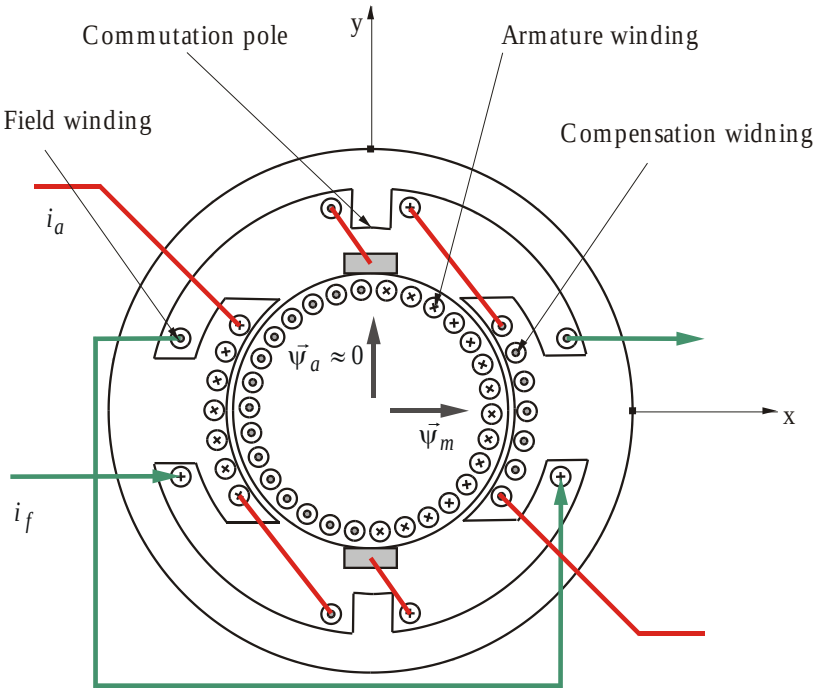


Figure 10.2 *Schematic view of the DC machine with commutator and compensation winding. The conductor dimensions are exaggerated.*

10.2 Mathematical model

The electrical equation for the armature winding and the torque equation can be directly taken from equation (8.31). The commutator connects to the armature winding in such a way that the *y*-axis equation is connected, but not the *x*-axis equation. Another set of commutator contacts placed in between the ones indicated in Figure 10.1 would be connecting to the *x*-axis winding, but since these are never used, the *x*-axis equation is an open circuit. Thus the *x*-axis current in equation (8.31) is always zero (i_{sx}=0). Note that index "*y*" is omitted and that the index "*s*" for stator in equation (8.31) is replaced with "*a*" for armature. The armature flux [ψ_a] is very small because of the compensation winding and the large air gap but cannot be neglected since it determines the armature current dynamics and thus the control options. The armature (rotor) voltage equation can thus derived by directly using the *y*-axis part of equation (8.31).

$$u_a = R_a \cdot i_a + L_a \cdot \frac{di_a}{dt} + \omega_r \cdot \psi_m \qquad (10.1)$$
$$T = \psi_m \cdot i_a$$

If the main machine flux is generated by a field winding and not by permanent magnets the electrical equation for the field winding becomes:

$$\begin{aligned} u_f &= R_f \cdot i_f + L_f \cdot \frac{di_f}{dt} \\ \psi_m &= L_m \cdot i_f \\ L_f &= L_m + L_{f\lambda} \end{aligned} \tag{10.2}$$

A symbolic model for the DC-machine may be sketched as

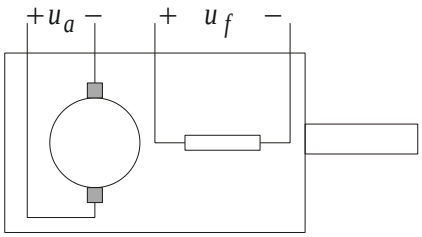


Figure 10.3 Symbolic model of a DC-machine.

In different practical applications armature and field winding circuit are coupled, that is they are connected in series, direct parallel or in a combination of these connections. Such machines are named series; shunt or compound machines are designed to give a certain torque/speed characteristic at constant dc voltage. In modern drive system applications these control means are replaced with power electronic control. Since the feeding voltage usually is variable, torque can be controlled independent of rotation speed within certain limits.

10.3 Torque control

Torque control of a DC machine is obtained by controlling the armature current i_a. The control algorithm for the armature current can be derived by identifying that the voltage equation (10.1) has the same structure as the equation of the generic one phase load equation (3.1). The current control algorithm derived for the generic one phase load can thus be used. With the use of DC machine notation, the equivalent of equation 3.11, for a "fast computer" is given by equation (10.3).

$$u_a^*(k) = \left(\frac{L_a}{T_s} + \frac{R_a}{2}\right) \cdot \left(\left(i_a^*(k) - i_a(k)\right) + \frac{T_s}{\left(\frac{L_a}{R_a} + \frac{T_s}{2}\right)} \cdot \sum_{n=0}^{n=k-1} \left(i_a^*(n) - i_a(n)\right) \right) + e_a(k) \tag{10.3}$$

$$e_a(k) = \omega_r(k) \cdot \psi_m$$

When the machine is intended to be operated at speeds exceeding nominal speed field weakening becomes necessary in order to prevent the induced voltage (which is rotation speed dependent) to exceed the maximum allowable output voltage of the power amplifier connected to the armature. Knowing this

maximum allowable output voltage $e_{a,max}$ a suitable reference value for the magnetizing flux linkage (ψ_m) can be determined by equation (10.4).

$$\psi_m^* = \begin{cases} \psi_{m,nom} \; if \;\; \omega_r < \omega_{r,nom} \\ \psi_{m,nom} \cdot \dfrac{\omega_{r,nom}}{\omega_r} \; if \;\; \omega_r > \omega_{r,nom} \end{cases} \qquad (10.4)$$

Due to magnetic saturation there is usually a non-linear relation between the magnetizing flux and the field current. The saturation curve is usually used to calculate a suitable field current reference once the magnetizing flux linkage reference is known. The input of to that calculation is the speed. Another way to determine the current reference value in the field weakening region is to calculate the induced emf (e_a) using the measured armature voltage and current. A PI controller having the input signal ($e_{a,max}$-e_a) can generate the field current reference as its output for speeds greater than nominal speed. Figure 10.4 indicates the two different ways to calculate the field current reference.

Note that the upper alternative in Figure 10.4 only gives the correct field current reference if the DC link voltage is the one that was the basis for the calculation of the look up tables. If the DC link voltage is increased, for any reason, it is possible to run the machine up to higher speed without field weakening, something that the look up table does not account for. The lower alternative in Figure 10.4 makes it easier to account for a variable DC link voltage since the maximum allowed induced voltage ($e_{a,max}$) can be made a function of the DC link voltage.

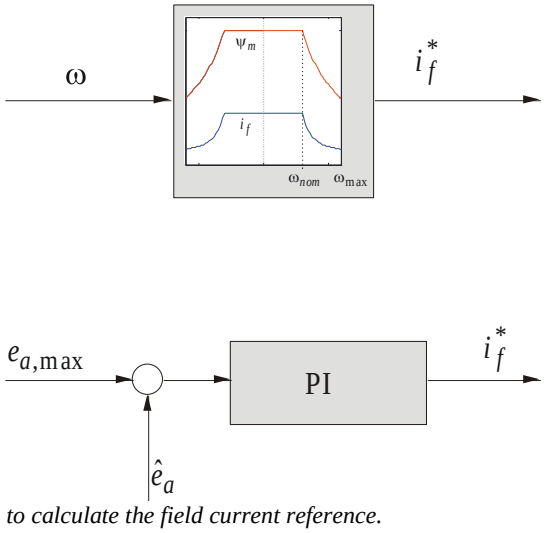


Figure 10.4 Two different ways to calculate the field current reference.

The field current itself requires further control. Since the time constant for the field winding is generally very large the field current cannot be controlled with "dead beat" response in the same time scale as the armature current. Nevertheless, the same method and criteria considered for the armature current

controller can be used to dimension the field current controller. If the field current variations required are large the proportional component of the field current controller presumably requires a larger voltage reference during a sampling interval than the power amplifier in use can supply. An anti-windup function is thus necessary to protect the integral part of the field current controller. The field current controller can be implemented with an algorithm according to equation (10.5).

$$u_f^*(k) = \left(\frac{L_f}{T_s} + \frac{R_f}{2}\right) \cdot \left(\left(i_f^*(k) - i_f(k)\right) + \frac{T_s}{\left(\frac{L_f}{R_f} + \frac{T_s}{2}\right)} \cdot \sum_{n=0}^{n=k-1} \left(i_f^*(n) - i_f(n)\right) \right) = \tag{10.5}$$

$$= \frac{L_f}{T_s} \cdot \left(\left(i_f^*(k) - i_f(k)\right) + \frac{T_s}{\frac{L_f}{R_f}} \cdot \sum_{n=0}^{n=k-1} \left(i_f^*(n) - i_f(n)\right) \right)$$

The closed loop control system with the controller designed according to this section is presented in the block diagram in Figure 10.5.

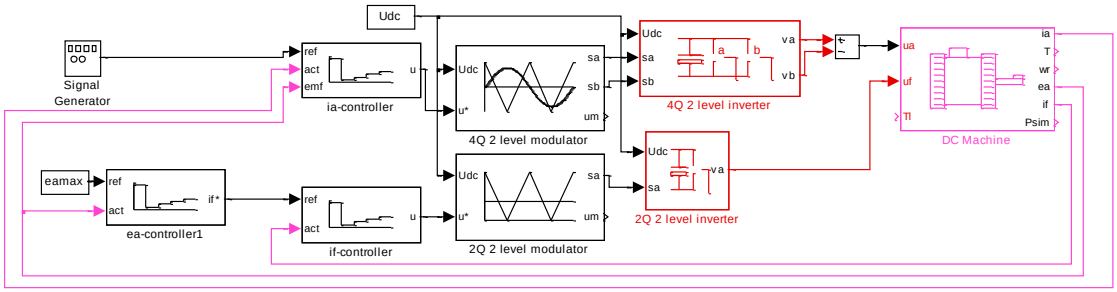


Figure 10.5 Simulink diagram of top level of a control system model for a field wound DC machine.

10.4 Torque control of the dc machine

Figure 10.5 shows a Simulink diagram of a DC machine torque control system with a 4 quadrant converter for the armature current and a 2 quadrant converter for the field current power supply. The emf (e_a) is assumed to be estimated and a PI controller is used to maximize the emf with the only restriction that the field current is limited to its nominal value.

The field circuit saturation curve is given in Figure 10.6. Figure 10.7 shows a free acceleration with a 2 Hz +/- 20 A armature current reference. The drive system parameters are:

```
>> Udc=400;
>> La=0.03;
>> Ra=2.4;
>> Lf=2.5;
>> Rf=292;
>> Ts=0.0005;
```

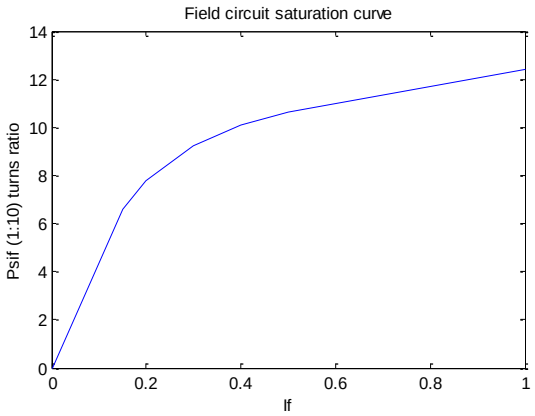


Figure 10.6 Field circuit saturation curve.

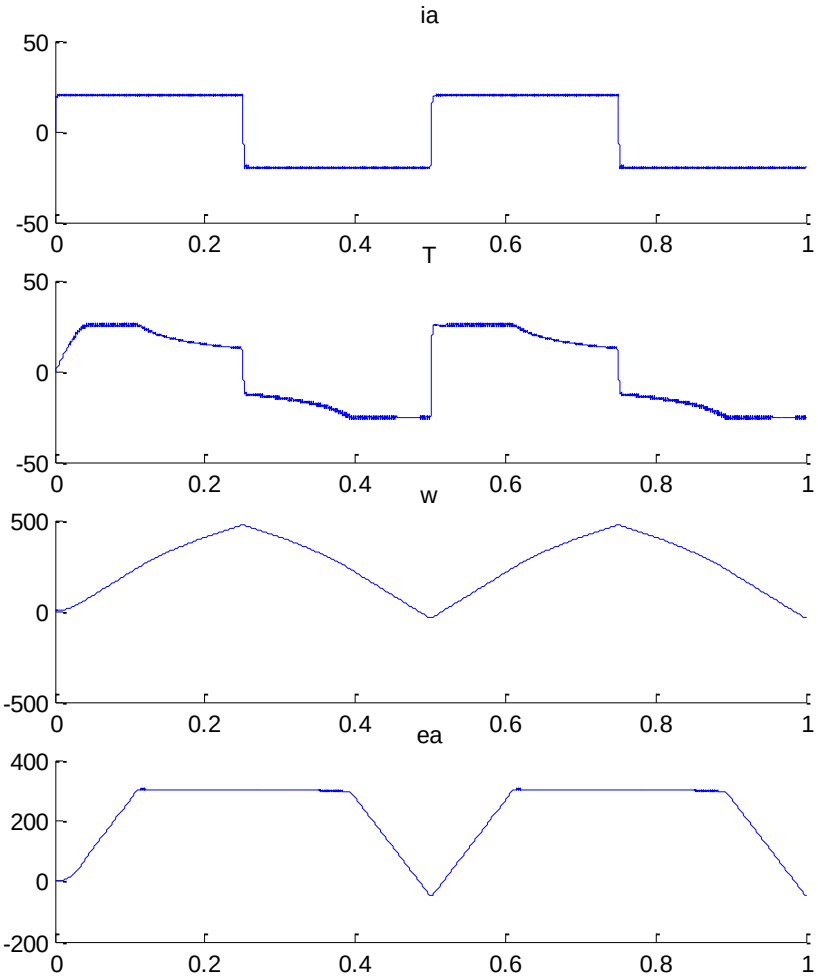


Figure 10.7 Dynamic response of a no-loaded DC machine at a 2 Hz +/- 20 A armature current. Diagrams from top: ia, T, □, ea.

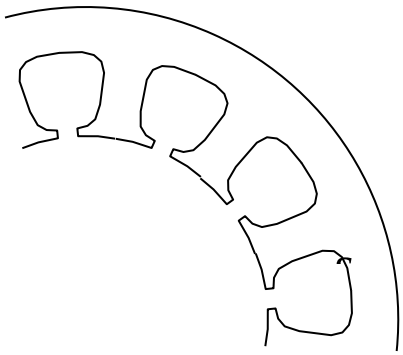
11 The Permanently Magnetized Synchronous Machine

An important application area for the synchronous machine is large scale power generation. In the majority of power stations synchronous machines operate as generators and their design depends on the rotational speed required. Multipole machines with salient poles are used for relatively slow rotation speeds whereas for higher speeds (for example gas turbine driven generators) the machines have lower pole numbers and cylindrical rotors - so-called turbo rotors. But such big machines are also employed in electrical drives for traction, rolling mills, mining etc. A special type is the permanent magnetised synchronous machine (PMSM), which is often used for servo systems up to 100 kW.

In the following section, the mechanical design and control methods for PM synchronous machines will be discussed.

11.1 Mechanical design

The stator has usually cylindrical shape with slots on the inner surface where the stator windings are placed. In general the number of slots is large for distributed windings (typically 6 per phase). The stator is manufactured using laminated metallic sheets to minimize the eddy currents induced by the rotating flux.



The rotor of a permanently magnetized synchronous machine can have the magnets applied either on the rotor surface or buried deep into the rotor. The most common materials for the permanent magnets are Samarium-Cobalt and Neodynium-Boron Iron, which are very durable (resist to vibration and to relatively high temperatures) and allow high magnetic flux densities.

The permanent magnets have a relative permeability close to one $\mu_r \approx 1$. Since the rotor is cylindrical the air gap will be relatively large (approximately equal to the height of the magnets), especially for magnets on the outer rotor surface. As a consequence the reluctance torque will be small (11.5), sometimes negligible.

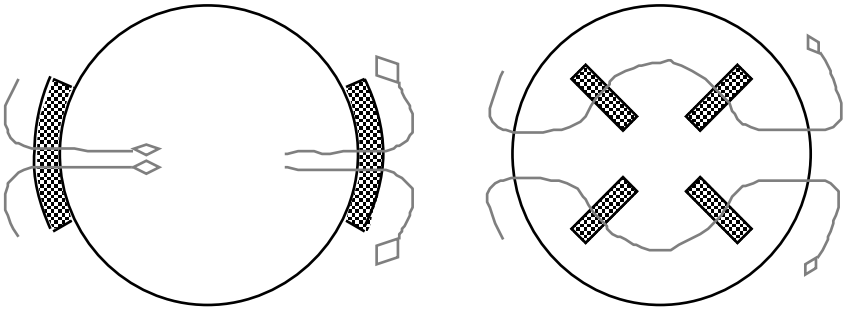


Figure 11.1 A 2-pole rotor with outer magnets (left) and 2-pole rotor with inner magnets (right).

11.2 Mathematical model

The permanently magnetized synchronous machine is structurally identical to the generalized machine described in chapter **Error! Reference source not found.**. Thus, the equations describing the machine can also be described with the same equations as derived in chapter **Error! Reference source not found.**. Figure 11.2 defines the stator (a,b) and rotor (x,y) reference frame used for control of the PMSM.

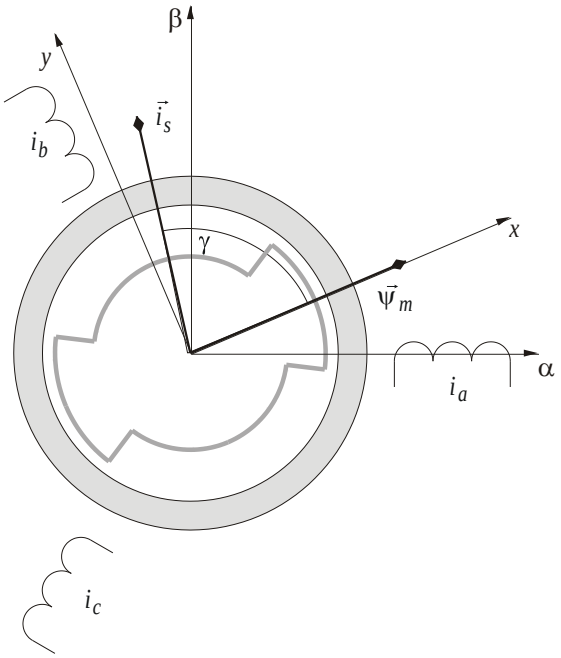


Figure 11.2 The PMSM structure and reference frames.

With reference to Figure 11.2, the electrical equations in the stator reference frame are:

$$\vec{u}_s^{\alpha\beta} = R_s \cdot \vec{i}_s^{\alpha\beta} + \frac{d\vec{\psi}_s^{\alpha\beta}}{dt} = R_s \cdot \vec{i}_s^{\alpha\beta} + \frac{d}{dt}\left(\vec{\psi}_\delta^{\alpha\beta} + L_{s\lambda} \cdot \vec{i}_s^{\alpha\beta}\right) \tag{11.1}$$

The same vector equation expressed in the rotor reference frame is:

$$\vec{u}_s^{xy} = R_s \cdot \vec{i}_s^{xy} + \frac{d}{dt}\left(\vec{\psi}_\delta^{xy} + L_{s\lambda} \cdot \vec{i}_s^{xy}\right) + j\omega_r \cdot \left(\vec{\psi}_\delta^{xy} + L_{s\lambda} \cdot \vec{i}_s^{xy}\right) \tag{11.2}$$

or, alternatively expressed in its components:

$$\begin{aligned} u_{sx} &= R_s \cdot i_{sx} + \frac{d}{dt}\left(\psi_m + L_{mx} \cdot i_{sx} + L_{s\lambda} \cdot i_{sx}\right) - \omega_r \cdot \left(L_{my} \cdot i_{sy} + L_{s\lambda} \cdot i_{sy}\right) = \\ &= R_s \cdot i_{sx} + \frac{d}{dt}\left(\psi_m + L_{sx} \cdot i_{sx}\right) - \omega_r \cdot L_{sy} \cdot i_{sy} \\ u_{sy} &= R_s \cdot i_{sy} + \frac{d}{dt}\left(L_{my} \cdot i_{sy} + L_{s\lambda} \cdot i_{sy}\right) + \omega_r \cdot \left(\psi_m + L_{mx} \cdot i_{sx} + L_{s\lambda} \cdot i_{sx}\right) = \\ &= R_s \cdot i_{sy} + L_{sy} \cdot \frac{di_{sy}}{dt} + \omega_r \cdot \left(\psi_m + L_{sx} \cdot i_{sx}\right) \end{aligned} \tag{11.3}$$

The flux relations are:

$$\begin{aligned} \psi_{sx} &= \psi_m + L_{sx} \cdot i_{sx} = \psi_m + (L_{mx} + L_{s\lambda}) \cdot i_{sx} \\ \psi_{sy} &= L_{sy} \cdot i_{sy} = (L_{my} + L_{s\lambda}) \cdot i_{sy} \end{aligned} \tag{11.4}$$

where $L_{mx,y}$ is the main inductance in the x- and y directions respectively, and $L_{s\lambda}$ is the stator leakage inductance that is equal in all directions since there is no saliency in the stator.

The torque equation can be expressed in the following ways:

$$\begin{aligned} T &= \vec{\psi}_s \times \vec{i}_s = \psi_{sx} \cdot i_{sy} - \psi_{sy} \cdot i_{sx} = \\ &= \left(\psi_m + (L_{mx} + L_{s\lambda}) \cdot i_{sx}\right) \cdot i_{sy} - (L_{my} + L_{s\lambda}) \cdot i_{sy} \cdot i_{sx} = \\ &= \psi_m \cdot i_{sy} + (L_{mx} - L_{my}) \cdot i_{sx} \cdot i_{sy} \end{aligned} \tag{11.5}$$

The torque expression thus contains two components, the first being independent of the rotor shape and the second being independent of the permanent magnetization. The latter term is often called the reluctance torque, since it depends on difference in reluctance in the x- and y-direction. Both these terms can be used in torque control.

11.3 Torque control

The torque generated by a synchronous machine can be controlled in several similar ways. The choice of the control concept depends in general on the requirements on torque quality but tends also to be a matter of taste, since different schools have their own "best" solutions. Among the requirements on torque quality we can mention:

Fast response the torque should respond fast to a change in the reference value. Under normal operating conditions response times of < 0.5 ms are common, independent of the size of the machine.

Accuracy the desired steady state value should be reached as accurately as possible. The accuracy depends essentially on the quality of the machine model that is how god the model describes the real machine. An error source is nonlinear saturation effects which have an impact on the inductances. Thus high quality requirements make it necessary to use adaptive control schemes.

Stability the controlled system must be stable, small parameter variations should not affect system stability.

Simplicity a practical implementation should be easy to achieve and as cheap as possible, often leading also to robustness and simple operation of the controlled system.

Different concepts to control the torque for a synchronous machine are analysed in the following section

Torque controlled by means of quadrature stator current

With this very simple method we can control torque without requiring limits on other variables such as stator flux or power factor. Essentially the method bases on keeping the direct armature current zero (i_{sx}=0). With these assumptions the torque is:

$$T = \psi_m \cdot i_{sy} \tag{11.6}$$

Given its reference value *T*(k)* at the time instant kT_s, the reference value for the quadrature stator current (i_{sy}*) is calculated as:

$$\begin{cases} i_{sx}^* = 0 \\ i_{sy}^* = \dfrac{T^*}{\psi_m} \end{cases} \tag{11.7}$$

This is the simplest possible way to determine the current references of a PM synchronous machine. A draw back with this method is that the power factor, that is the cosine of the angle (φ) between the induced voltage $\vec{e}_s = j\omega \cdot \vec{\psi}_s$, and the stator current $\vec{i}_s$, are free variables, as well as the value for the resulting stator flux that increases with i_{sy}.

The angle α in Figure 11.3 can be rather big, close to p/4. The result is a low power factor as seen from the phase terminals. Figure shows an example where this control method is applied to a 50 kW PMSM with approximately the following data: R_s=0.01 Ω, L_s=0.0002 H, ψ_m=0.067. A 200 A current step in the y-axis current is applied when the machine is at stand still and a free acceleration takes place.

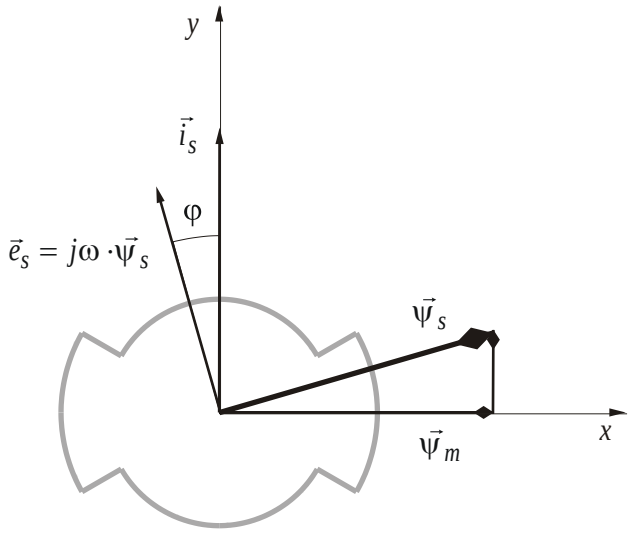


Figure 11.3 If isx=0 the stator flux increases with isy and the power factor is a non-controlled variable.

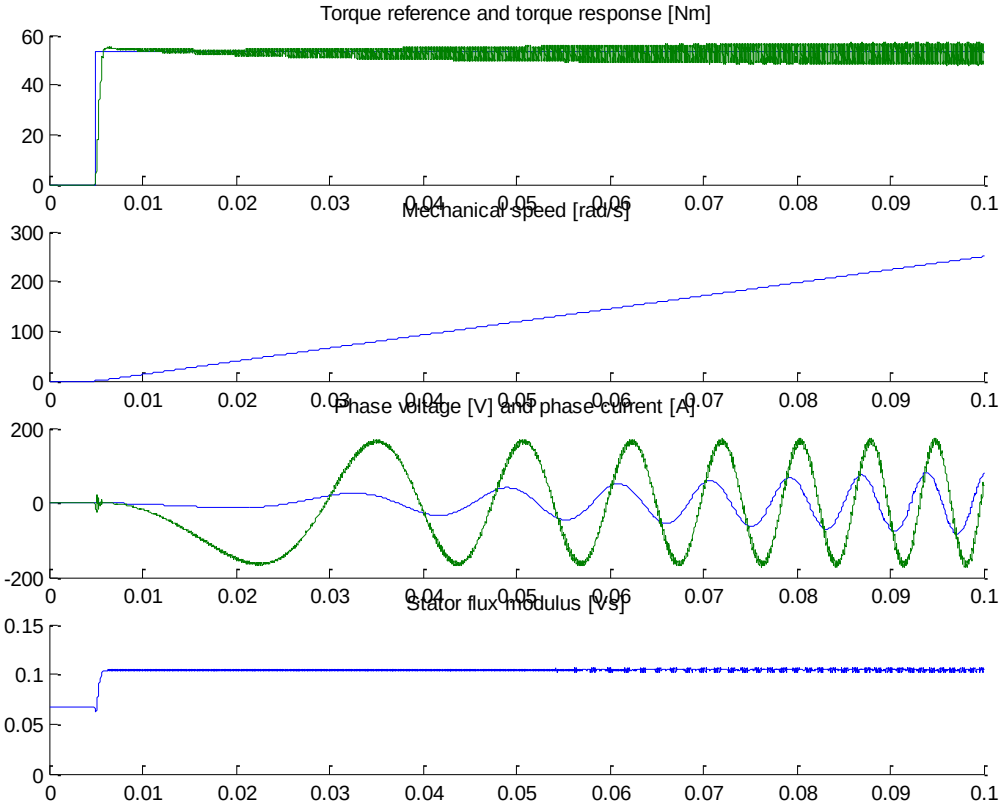


Figure 11.4 Free acceleration with a 50 kW PMSM and only q-axis current control.

Note in Figure 11.4 that there is a phase lag between the voltage and current of a phase. This means that the power converter has to supply reactive power to the stator, an undesired aspect of the control.

Torque controlled with minimal stator current

The method presented in the previous section did not take advantage of the reluctance torque, a possibility that can be used to reduce the need for stator current. Assume the task to control a PMSM with salient poles in such a way that the y-axis inductance is higher than the x-axis inductance. This means a machine with generic proportions of the rotor according to Figure 11.5.

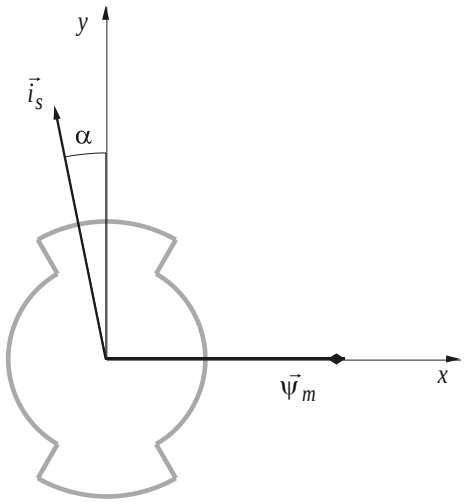


Figure 11.5 Generic rotor geometry for the PMSM.

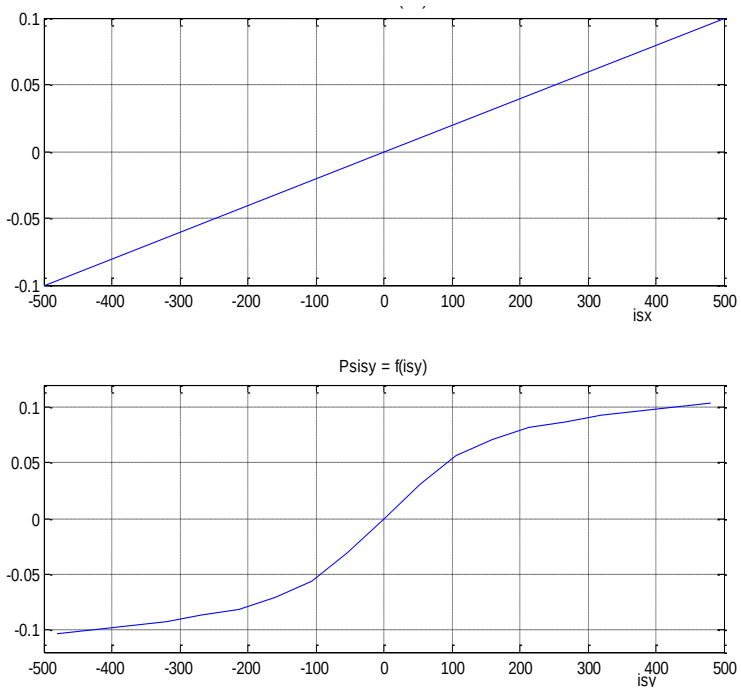


Figure 11.6 Saturation curves for the PMSM along the x- and y-axes.

By letting the stator current vector into the 2'nd quadrant, the reluctance torque will contribute to the total torque production with the same sign as the torque component $\psi_m \cdot i_{sy}$. For each and every torque level, there is a specific angle α in Figure 11.7 that minimizes the stator current. Due to the non-linear saturation effects, it is not possible to resolve an algebraic expression for this angle, but it has to be found numerically. One possible way to do this is as follows.

Assume that the inductance variations, due to the stator currents in the x- and the y-axis, are known. It is then possible to calculate the torque for any location of the stator current vector in the (x,y)-plane. Note that for each location of the stator current vector, a unique inductance for each direction (x,y) is calculated and then used in the torque expression in equation (11.5). If this is done on the motor described in Figure 11.5, the result is a set of "iso"-torque lines according to Figure 11.7.

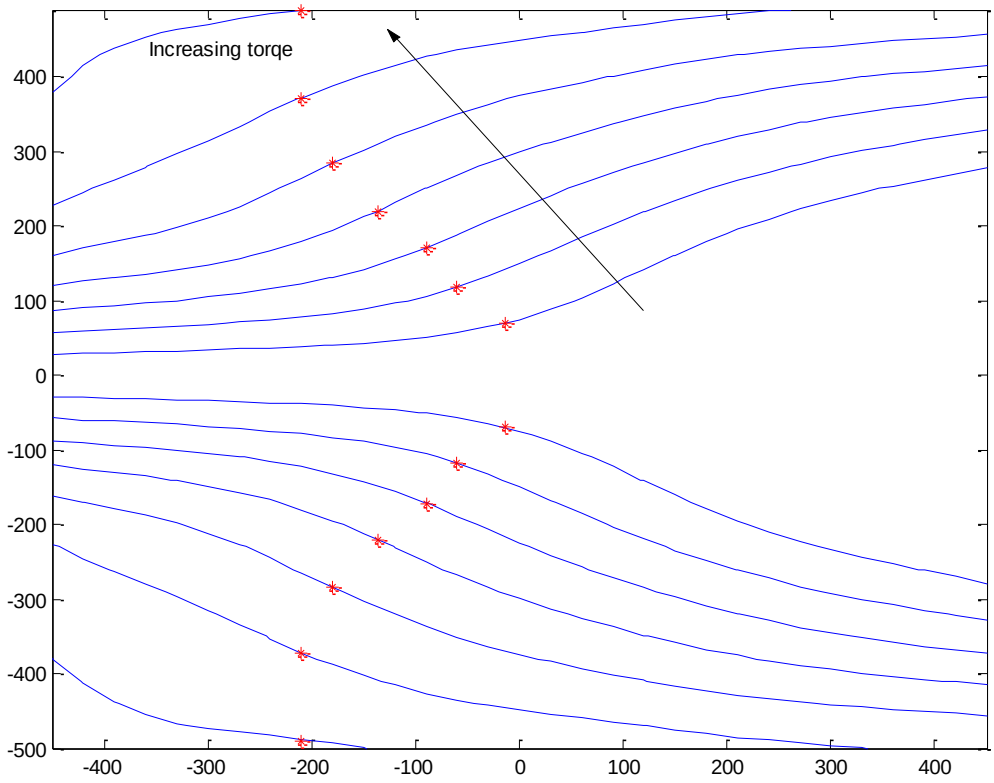


Figure 11.7Torque contribution for all combinations of isx and isy. The lines are "iso"-torque lines, i.e. locations of the stator current vector producing the same torque. The stars are the spots on the lines that are located closest to origo, i.e the location of the stator current vector that gives the highest torque for a given current.

From the lines in Figure 11.7 it is possible to locate the spot on each line that is closest to origo. This spot is the location of a particular stator current vector giving the highest torque per Ampére stator current. A set of the locations can be

made into a look up table and used in controlling the machine. Figure 11.8 contains the data for such a look up table, only for positive torque levels. Implemented in a simulation program, the result is a torque source that uses a minimum of stator current for a specific application. Figure 11.9 shows a simulation with an optimal choice of stator current vector components for minimum stator current, but producing the same torque as in Figure 11.4 where only the y-axis current was used.

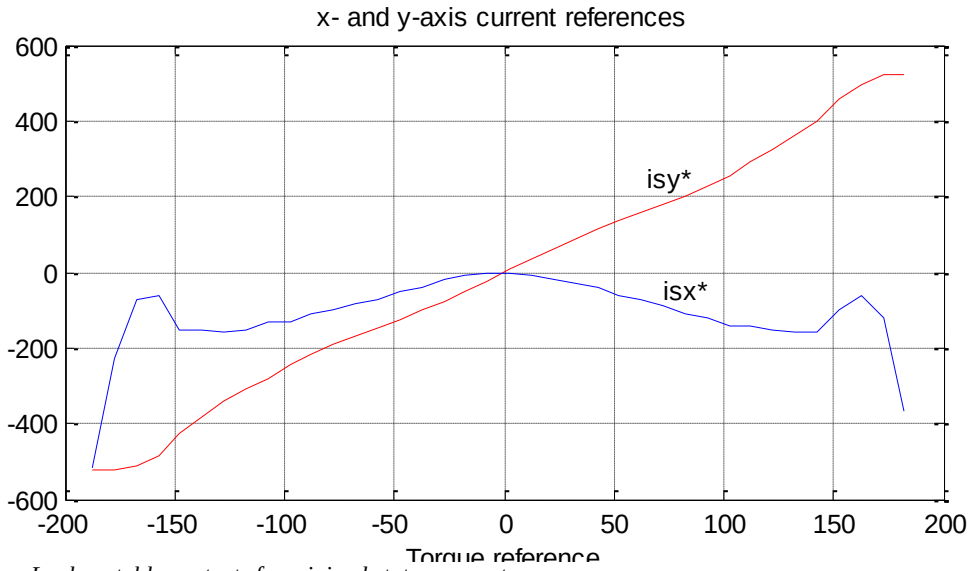


Figure 11.8 Look up table contents for minimal stator current.

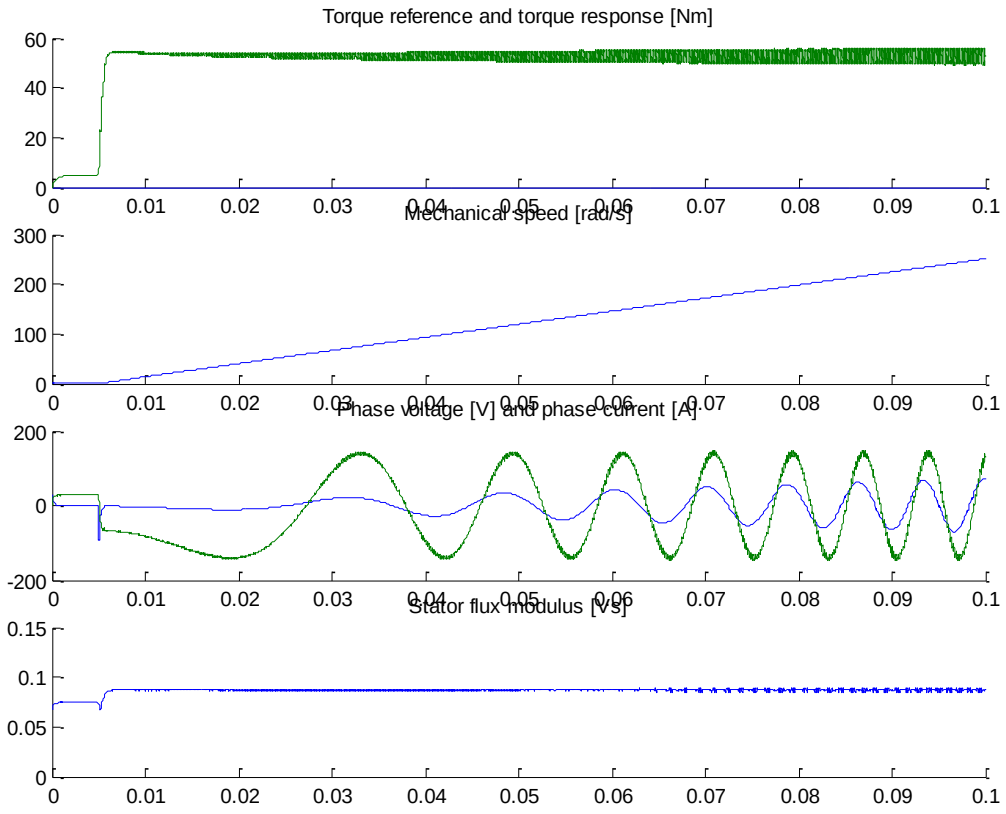


Figure 11.9 Free acceleration with a 50 kW PMSM and current control for minimal stator current.

Note that the phase current amplitude is lightly lower in Figure 11.9 than in Figure 11.4 , as should be expected. Due to the lower y-axis current component, the stator flux level is also less affected by the stator currents.

11.4 Current control

The currents of the PMSM can be controlled in the same way as was derived in chapter 4 for a generic three phase load.

$$u_{sx}^{*}(k)=\left(\frac{L_{sx}}{T_s}+\frac{R_s}{2}\right)\cdot\left(\left(i_x^{*}(k)-i_x(k)\right)+\frac{T_s}{\left(\frac{L_{sx}}{R_s}+\frac{T_s}{2}\right)}\cdot\sum_{n=0}^{n=k-1}\left(i_{sx}^{*}(n)-i_{sx}(n)\right)\right)-\omega_r\cdot L_{sy}\cdot i_{sy}(k)$$

$$u_{sy}^{*}(k)=\left(\frac{L_{sy}}{T_s}+\frac{R_s}{2}\right)\cdot\left(\left(i_y^{*}(k)-i_y(k)\right)+\frac{T_s}{\left(\frac{L_{sy}}{R_s}+\frac{T_s}{2}\right)}\cdot\sum_{n=0}^{n=k-1}\left(i_{sy}^{*}(n)-i_{sy}(n)\right)\right)+\omega_r\cdot\left(\psi_m+L_{sx}\cdot i_{sx}(k)\right) \tag{11.8}$$

Note that some parameters in the current controllers are no longer constant due to saturation of the inductances. When using fixed controller parameter values, the lowest in the range of the real parameters should be used to avoid instability. Note also that servomotors usually operate with a very high switching frequency, and thus there is not a significant need for prediction of the stator current. Deadbeat step response is usually not possible due to the short sampling intervals. With a switching frequency above 10 kHz, and a corresponding possible sampling frequency above 20 kHz, the sampled current control does easily get a bandwidth of several kHz without using prediction.

11.5 Field weakening control

Permanently magnetized synchronous machines are used in many applications where the speed range is so wide that it is necessary to bring the machine into field weakening. The goal in field weakening is to control the stator currents in such a way that the stator voltage is limited, if possible without reducing the torque. This is partly possible to accomplish and to understand how it must be done, it is necessary to study the voltage need at various speeds and currents.

The stator voltage of a PMSM can be expressed in the rotor oriented reference frame as:

$$\left|\vec{u}_s\right|=\sqrt{\left(R_s\cdot i_{sx}-\omega_r\cdot L_{sx}(i_{sx},i_{sy})\cdot i_{sx}\right)^2+\left(R_s\cdot i_{sy}+\omega_r\cdot\left(\psi_m+L_{sx}(i_{sx},i_{sy})\cdot i_{sx}\right)\right)^2}$$

Note in this expression that the inductances in both the x- and the y-directions are functions of the currents in both the x- and the y-directions. The stator voltage is thus a non-linear function of the stator currents.

The stator voltage can be illustrated in the (x,y)-frame with the loci of the stator current vector that result in the same stator voltage, i.e. "iso-voltage" lines. The way to calculate this is similar to calculate the "iso-torque"-lines illustrated in Figure 11.7. The steps are as follows:

1. Scan all combinations of stator current components (i_{sx}, i_{sy}).
2. For each combination, calculate the stator flux linkage. $\vec{\psi}_s = \left(\psi_m + L_{sx}(i_{sx}, i_{sy}) \cdot i_{sx}\right) + j \cdot L_{sy}(i_{sx}, i_{sy}) \cdot i_{sy}$
3. For a finite set of rotor speeds, calculate the induced voltage as $|\vec{u}_s| = \omega_r \cdot \left|\vec{\psi}_s\left(i_{sx}, i_{sy}\right)\right|$

   The stator resistance is omitted, but of very little significance due to the high induced voltage in the field weakening range.
4. Locate and plot combinations of stator current components that give the same stator voltage according to 3.

The result of applying the above mentioned procedure on the 50 kW PMSM discussed in the previous sections is shown in Figure 11.10.

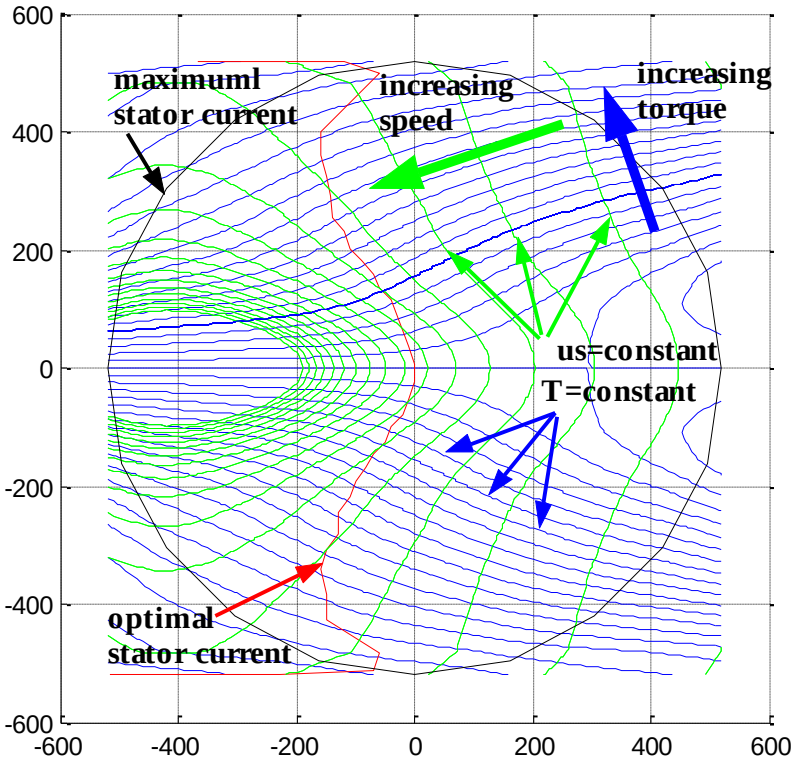


Figure 11.10 "iso"-torque (T=constant), "iso"-voltage (us=constant), optimal stator current and maximal stator current for a PMSM.

The consequence for torque control of a PMSM with properties according to Figure 11.10 is that the stator current must be selected to fulfil the requirements of Table .

Table 11.1 Selection of stator current references according to priority.

Priority	Action
1	Stator current less than maximum stator current (Current limitation to protect machine and power electronics)
2	Stator voltage less than maximum stator voltage (Voltage limitation to facilitate current control)
3	Minimal stator current for a given torque (Stator current on the optimal line)

Figure 11.11 shows the stator current vector loci for a step change in the torque reference from zero to a constant positive torque.

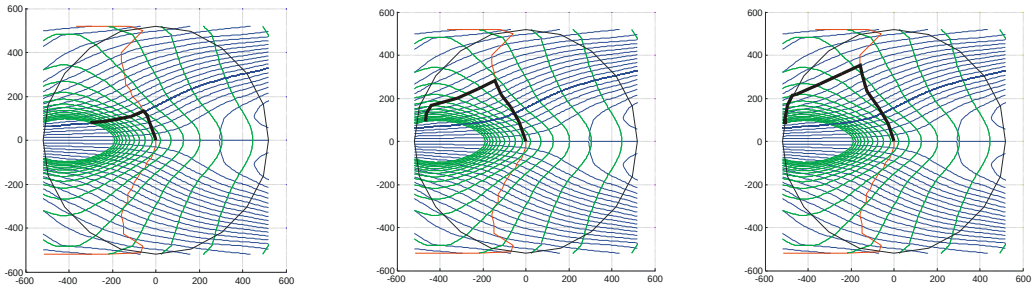


Figure 11.11: Stator current loci when the machine accelerates from zero to higher speeds at constant torque reference of three levels

Note in Figure 11.11 (left) how the current at low speed is selected along the optimum line and, as the speed increases, the current stays within the voltage limitation line at the constant torque line, thus maintaining the torque at the reference level but not with minimum stator current at higher speeds. When the speed increases beyond a limit where the torque level cannot be maintained, then the stator current has to be changed in order to stay within the voltage limitation line but with a decreased torque, see Figure 11.11 (middle). Finally, if the current reaches the current limit, the torque may have to be reduces just to protect the system from over current, see Figure 11.11 (right).

In the control system, this means that there has to be a function that takes care of field weakening and changes the stator current references from the optimal setting to a setting that takes the priority list of Table 11.1 into account Figure 11.12 shows the contents of a current controller in Simulink that has such a function included. Figure 11.13 shows an example of contents of the field weakening block and Figure 11.14 shows a simulated example of a free acceleration into field weakening speed.

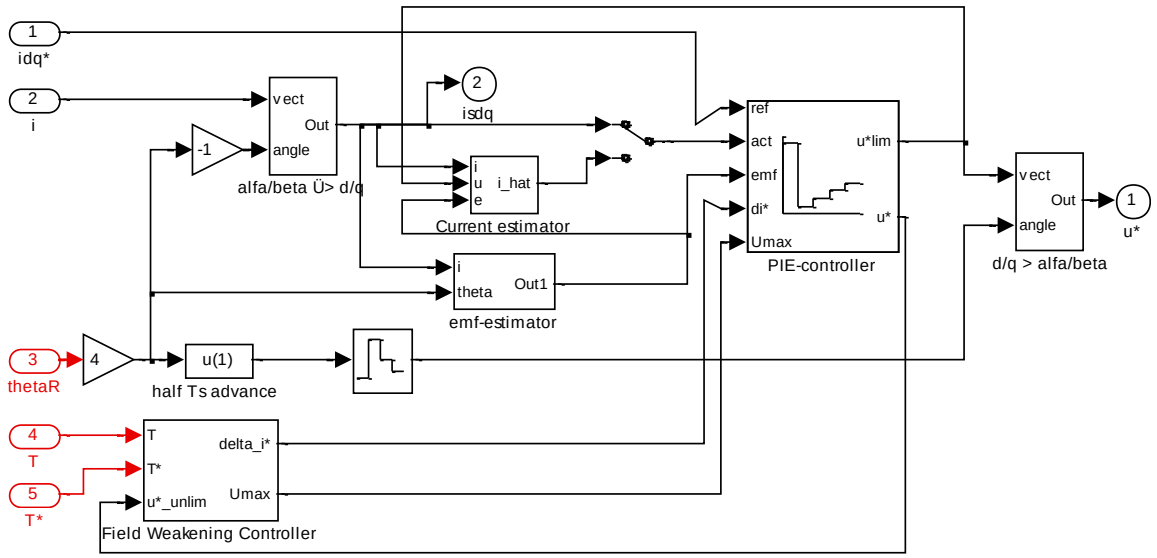


Figure 11.12 Current controller with field weakening.

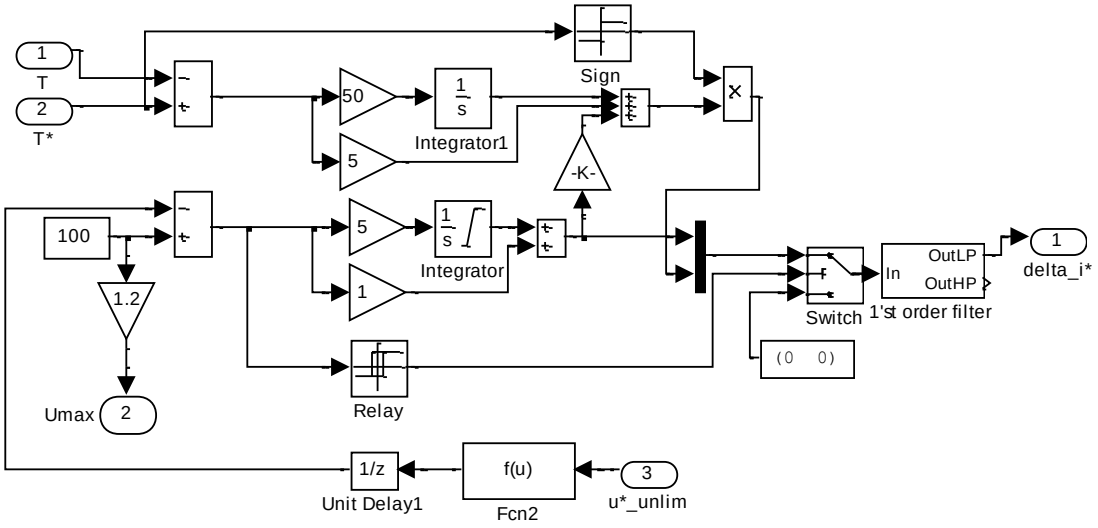


Figure 11.13 Contents of the field weakening block of Figure 11.12.

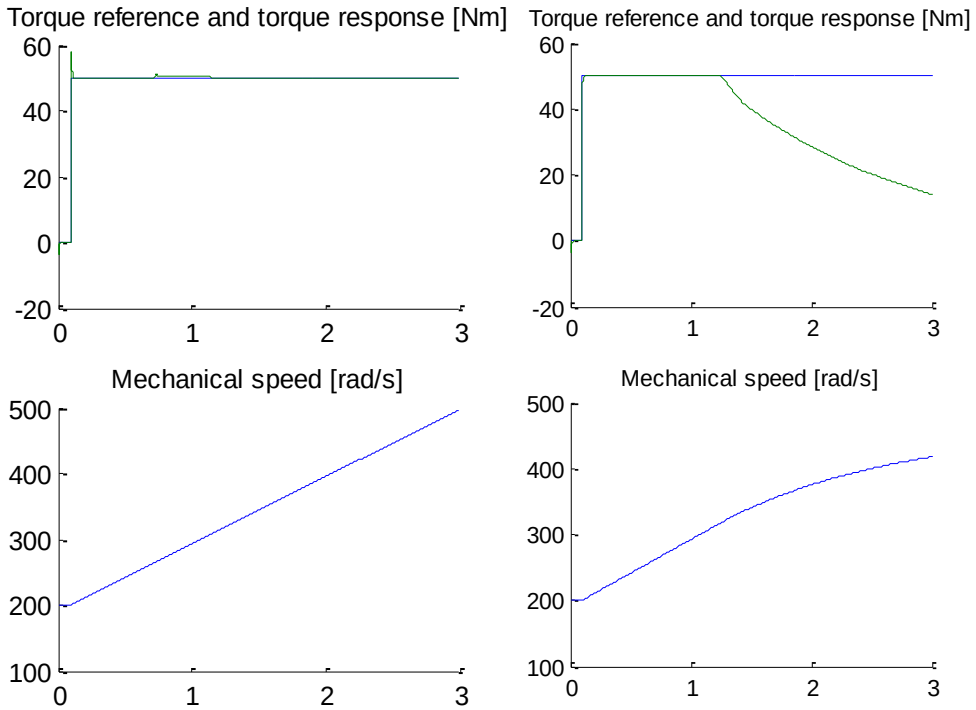


Figure 11.14 Free acceleration with (left), and without (right), the field weakening controller.

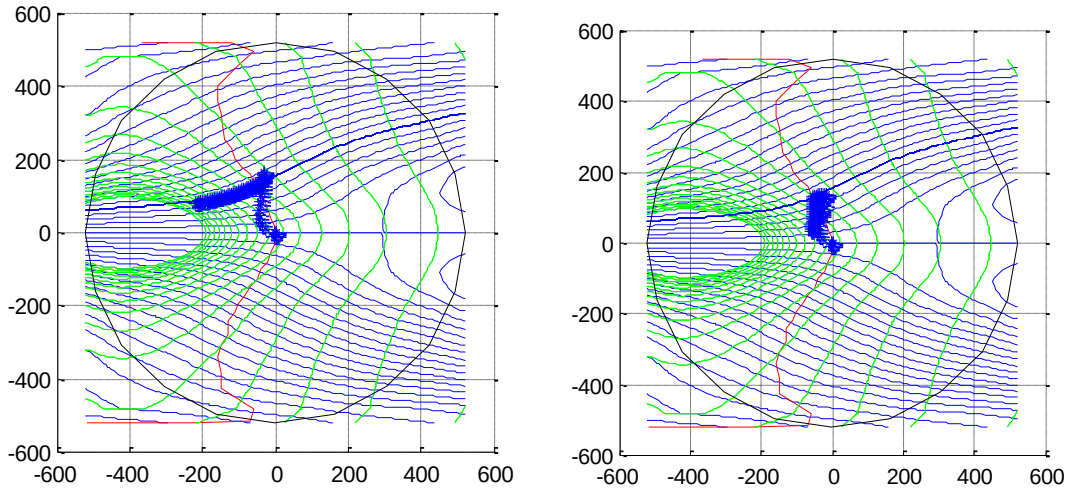


Figure 11.15 Current loci with and without a field weakening controller.

12 The Induction Machine

The most common type of electrical machine used in practice is the induction machine, often called the asynchronous machine (AM). It is almost exclusively used as a motor in applications ranging from tape recorders up to railroad traction, electric vehicles, elevators etc. covering both low power and high power applications. The main reason why the induction machine is so wide spread in practical use is its simple and cheap construction. Since the machine does not have brushes and slip rings it does not require any other maintenance than possibly lubrication of its bearings. For this reason the AM is one of the most rugged electrical machine types.

Compared to the synchronous machine (SM) and the DC machine the AM does not have a field winding connected to an external voltage source. The magnetic flux wave in the machine results as the superposition of the stator and rotor flux similar to all the other machine types. The rotor currents are driven by voltages induced in the rotor winding when the rotor flux rotates relative to the rotor. The flux wave rotates at synchronous rotational speed relative to the stator. In order to maintain the flux wave rotating relative to the rotor as well, the latter has to rotate at different speed (lower or higher) than the synchronous speed. The rotor speed is therefore "asynchronous" and this is where the name for the AM comes from. Since the rotor currents are induced in the rotor the AM is sometimes also called "induction machine".

From the point of view of torque control the AM is somewhat more difficult to handle than the SM. The main problem is to estimate the flux wave position, either stator-, air gap- or rotor-flux. For the SM it can easily be calculated when the rotor position and the currents are known. The rotor position for the AM does not provide any information about the actual flux wave position since the flux wave rotates relative to the rotor. It is tempting to try to get direct information about the flux wave using a flux measuring sensor in the air gap (for example a hall sensor). Such solutions are not used since they have other drawbacks, for example time consuming and expensive shut down, disassembly and assembly operations in case a hall sensor fails to operate. Practical solutions therefore try to employ mathematical methods to calculate the flux wave position using the variables easily accessible for measurements. One practicable and useful additional measurement is rotor position measurement, which is obtained from an extra sensor mounted on the rotor axis.

12.1 Mechanical construction

The stator windings of a three phase AM are located in stator slots in the same way as the stator windings of the SM. From a mechanical and electrical point of view the stator circuits of the two machine types are exactly the same. It is the rotor that essentially differs and may be classified as follows:

5. short-circuited rotor
6. slip ring rotor

Since the magnetic flux wave rotates relative to the rotor, the latter has to be manufactured using laminated iron sheets as usually necessary for the stator. On its periphery the rotor has slots where the rotor-winding conductors are situated, see Figure 12.1.

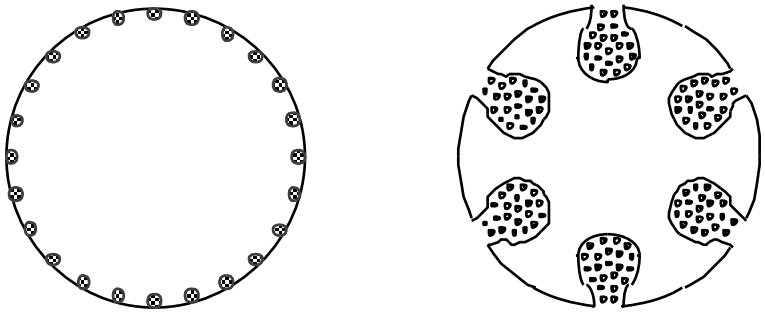
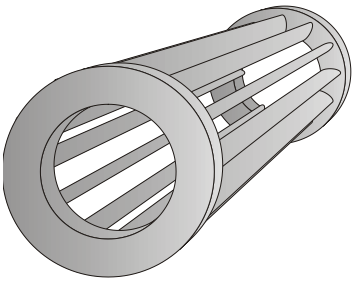
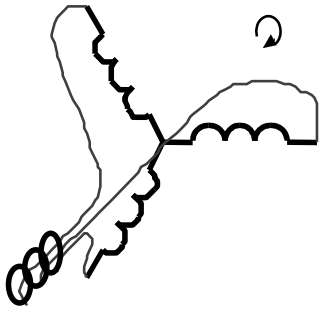


Figure 12.1: Different rotor constructions.

The short circuited rotor winding is made of copper or aluminium bars which are pressed into the rotor slots. On both ends the bars are short circuited by means of two conductor rings as in the right figure above. The bar profiles depend on the properties to be achieved. The slot profile itself influences mainly the start-up characteristics for a direct start of the AM with given supply voltage frequency (for example 50 Hz), and the efficiency of the machine in inverter operation. Improving one characteristic usually makes another worse, every design is a compromise. Frequency converters are often used for torque control purposes, which cause the mechanical torque to have ripples (torque "harmonics") determined, among other factors, by the shape of the rotor-winding conductors. In some applications such ripples cannot be tolerated and a means to reduce them is necessary. Since the majority of the AM are constructed to have best performance when connected to the 50-Hz network, optimizing an AM drive means not only adjusting the control system but also the entire construction of the machine.



The slip ring rotor type contains windings similar to the stator windings. The three phase windings are connected in "Y" with the loose ends coupled to their slip rings as indicated in the figure to the right. Using slip rings has two advantages: firstly the electrical characteristics of the rotor can be modified (for example additional resistances may be introduced in the rotor circuit, allowing to enhance the starting properties of the machine), and secondly having direct access to the rotor windings better possibilities to control torque are opened.



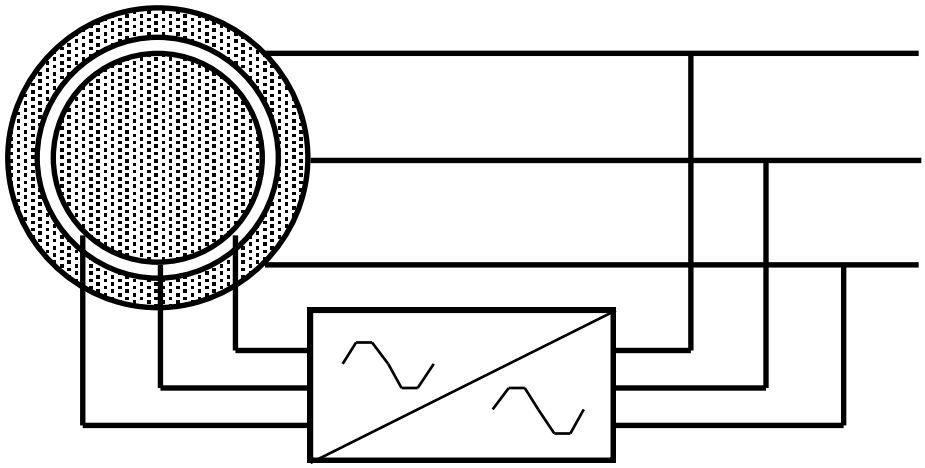


Figure 12.2: AM with cascaded rotor coupling.

If the drive application requires having only moderate deviations from the nominal speed (for example +/- 20 %) the "cascaded rotor" coupling can have advantages. In this case the frequency converter needs only to handle a fraction of the transmitted power (the so-called "slip power") and can therefore be smaller and cheaper. On the other hand a slip ring rotor machine is more expensive than the short-circuited rotor machine.

What type of drive system is chosen for an application finally depends on the specifics of the application itself.

12.2 Mathematical model

Since the stator of an AM is in principle identical with the stator for the SM the same equations can be written for both machines, either written explicitly for each phase or in the more compact matrix form. The stator vector equation for the AM with variables expressed in stator coordinates is:

$$\vec{u}_s^s = R_s \cdot \vec{i}_s^s + \frac{d}{dt}\left(\vec{\psi}_s^s\right) \qquad (12.1)$$

The rotor, whether short-circuited or slip ring rotor, can be considered as containing three phase windings. For each rotor winding we can write a differential equation similar to the equations for the stator windings where the index "s" is replaced by "r" (which stands for **r**otor). Written in vector form the equations for all three phases become very compact. We can choose a complex reference frame (*x*,*y*) with its real *x* axis coinciding with the magnetic axis of one of the three phases which becomes the reference phase. In this way we obtain a reference frame rigidly coupled to the rotor which is called the "rotor reference frame" and variables expressed in this reference frame are marked with the upper index "*r*". Expressing the variables in this reference frame the rotor winding is electrically described by:

$$\vec{u}_r^{\,r} = R_r \cdot \vec{i}_r^{\,r} + \frac{d}{dt}\left(\vec{\psi}_r^{\,r}\right) \tag{12.2}$$

The fluxes linking stator and rotor differ from the common air gap flux by the stator and rotor leakage flux, respectively, see Figure 12.3. The stator and rotor leakage flux vectors have the same orientation as the stator and rotor current vectors, respectively.

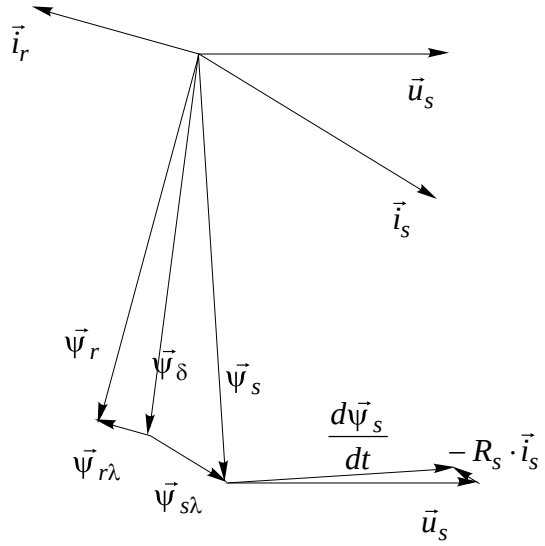


Figure 12.3: Connection between the different magnetic fluxes in the machine.

Since the machine has cylindrical stator and rotor the air gap flux wave coincides with the total mmf and the relationship between the different fluxes can be written:

$$\begin{aligned}\vec{\psi}_s^{\,s} &= L_{s\lambda} \cdot \vec{i}_s^{\,s} + \vec{\psi}_\delta^{\,s} = L_{s\lambda} \cdot \vec{i}_s^{\,s} + L_m \cdot \left(\vec{i}_s^{\,s} + \vec{i}_r^{\,r} \cdot e^{j\theta_r}\right) \\ \vec{\psi}_r^{\,r} &= L_{r\lambda} \cdot \vec{i}_r^{\,r} + \vec{\psi}_\delta^{\,r} = L_{r\lambda} \cdot \vec{i}_r^{\,r} + L_m \cdot \left(\vec{i}_s^{\,s} \cdot e^{-j\theta_r} + \vec{i}_r^{\,r}\right)\end{aligned} \tag{12.3}$$

where

$L_{s\lambda}$ = stator leakage inductance

$L_{r\lambda}$ = rotor leakage inductance

L_m = main inductance of the machine, the same for all directic

For both machines, the AM and the SM, the same expression for the mechanical torque is found:

$$T = \vec{\psi}_s \times \vec{i}_s \tag{12.4}$$

By taking $\vec{\psi}_s$ and $\vec{i}_s$ from equations (12.3) in (12.4) we can express the torque in several different ways:

$$T = \vec{\psi}_s \times \vec{i}_s = \vec{\psi}_r \times \vec{i}_r = -\frac{L_m}{L_r}\vec{\psi}_s \times \vec{i}_r = \frac{L_m}{L_r}\vec{\psi}_r \times \vec{i}_s = \vec{\psi}_\delta \times \vec{i}_s \tag{12.5}$$

The reason why torque may be expressed as in (12.5) is that the leakage fluxes do not contribute to torque generation, as mentioned earlier. Because the stator currents are the variables the torque can be controlled with, the terms in the equation above (framed), which contain the stator current, are interesting to look at.

No matter which one of the expressions above is used for torque control purposes it is necessary to know the magnitude and orientation of the vectors entering the expression. In this respect it is useful to write the electrical equations (12.1) and (12.2) in a reference frame oriented after the rotor flux vector, the air gap flux vector or the stator flux vector. Depending on the reference frame the equations written for the AM differ slightly. The three alternatives will be presented in the following. The first one is a so-called "rotor-oriented" model and relies on equation (12.2) written in rotor flux coordinates. The second one is a simple stator flux model which is useful in most applications except genuine servo applications. The third alternative corresponds to a complete flux observer determining stator and rotor flux as well.

Rotor flux model

The rotor flux reference frame is defined by the rotor flux vector orientation as described by Figure 12.4. It is usually denoted as the (*d*,*q*)-system, where again "*d*" represents the "direct axis" and "*q*" stands for "quadrature axis". This notation is, by the way, generally applied to orthogonal systems oriented after a flux vector.

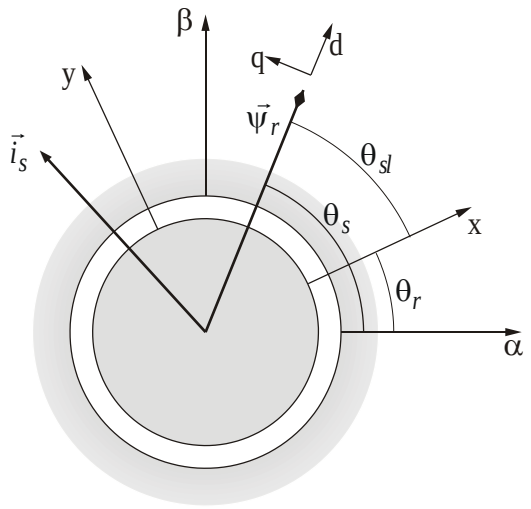


Figure 12.4: Stator current vector and its components in the rotor flux reference frame.

With (12.3) equations (12.1) and (12.2) can be written:

$$\vec{u}_s^{\,s}(t)=R_s\cdot\vec{i}_s^{\,s}(t)+\frac{d\vec{\psi}_s^{\,s}(t)}{dt}=R_s\cdot\vec{i}_s^{\,s}(t)+L_s\frac{d}{dt}\left(\vec{i}_s^{\,s}\right)+L_m\frac{d}{dt}\left(\vec{i}_r^{\,r}\cdot e^{j\theta_r}\right)$$
$$\vec{u}_r^{\,r}(t)=0=R_r\cdot\vec{i}_r^{\,r}(t)+\frac{d\vec{\psi}_r^{\,r}(t)}{dt}=R_r\cdot\vec{i}_r^{\,r}(t)+L_m\frac{d}{dt}\left(\vec{i}_s^{\,s}\cdot e^{-j\theta_r}\right)+L_r\frac{d}{dt}\left(\vec{i}_r^{\,r}\right) \qquad (12.6)$$

In three steps these expressions are transformed into rotor coordinates:

1. in both equations $\vec{i}_r$ is replaced by $\vec{i}_r$ expressed in rotor coordinates with (12.3)
2. all vectors are expressed in rotor coordinates (upper index "*dq*").
3. the derivatives are written explicitly. After that the factor $e^{j\theta_s}$ is cancelled in the stator equation and $e^{j\theta_{sl}}$ in the rotor equation

These steps finally lead to the stator and rotor voltage equation in rotor coordinates:

$$\vec{u}_s^{\,dq}(t)=R_s\cdot\vec{i}_s^{\,dq}(t)+\left(L_s-\frac{L_m^2}{L_r}\right)\cdot\frac{d}{dt}\left(\vec{i}_s^{\,dq}\right)+\frac{L_m}{L_r}\cdot\frac{d}{dt}\left(\vec{\psi}_r^{\,dq}\right)+$$
$$+j\omega_s\left[\left(L_s-\frac{L_m^2}{L_r}\right)\cdot\vec{i}_s^{\,dq}(t)+\frac{L_m}{L_r}\cdot\vec{\psi}_r^{\,dq}\right] \qquad (12.7)$$
$$0=\frac{R_r}{L_r}\cdot\vec{\psi}_r^{\,dq}-\frac{L_m}{L_r}\cdot R_r\cdot\vec{i}_s^{\,dq}(t)+j\omega_{sl}\cdot\vec{\psi}_r^{\,dq}+\frac{d}{dt}\left(\vec{\psi}_r^{\,dq}\right)$$

Equations (12.7) can now be separated into real and imaginary parts. The following aggregated parameters are used to simplify the expressions:

$$\sigma = \left(1 - \frac{L_m^2}{L_r \cdot L_s}\right) \approx \frac{L_{s\lambda} + L_{r\lambda}}{L_s} \approx \frac{L_{s\lambda} + L_{r\lambda}}{L_r}$$
$$\tau_r = \frac{L_r}{R_r} \tag{12.8}$$

and lead to the following four system state equations:

$$\frac{d\psi_{rd}}{dt} = \frac{L_m}{\tau_r} \cdot i_{sd} - \frac{1}{\tau_r}\psi_{rd}$$
$$\frac{d\theta_{sl}}{dt} = \frac{L_m \cdot i_{sq}}{\tau_r \cdot \psi_{rd}}$$
$$\frac{di_{sd}}{dt} = \left[u_{sd} - R_s \cdot i_{sd} - \frac{L_m}{L_r} \cdot \frac{d\psi_{rd}}{dt} + \omega_s \cdot \sigma \cdot L_s \cdot i_{sq}\right] \cdot \frac{1}{\sigma \cdot L_s} \tag{12.9}$$
$$\frac{di_{sq}}{dt} = \left[u_{sq} - R_s \cdot i_{sq} - \omega_s \cdot \left(\sigma \cdot L_s \cdot i_{sd} + \frac{L_m}{L_r} \cdot \psi_{rd}\right)\right] \cdot \frac{1}{\sigma \cdot L_s}$$

The equation for the torque is added to the above expressions:

$$T = \frac{L_m}{L_r}\vec{\psi}_r \times \vec{i}_s = \frac{L_m}{L_r}\psi_{rd} \times i_{sq} \tag{12.10}$$

Equation (12.9) first, describes the direct rotor flux that is the resulting rotor flux, as a function of the stator currents. It is a first order system state equation. If required, the magnitude of the rotor flux can be calculated with this equation. Equation (12.9) first and second, shows that the slip frequency is direct proportional to the quadrature stator current when the rotor flux is constant. Thus the slip angle can be calculated as the time integral of the slip angle velocity. By adding the rotor angle (obtained for example from a resolver) to the slip angle the rotor flux position may be calculated:

$$\theta_s = \theta_r + \theta_{sl} \tag{12.11}$$

A control system which, starting from measurements of the machine currents, keeps track of the rotor flux orientation and magnitude can look like the system in figure Figure 12.5. The block diagram is based on the equations (12.9) first and second one. It contains the blocks for control of the torque and resulting rotor flux magnitude.

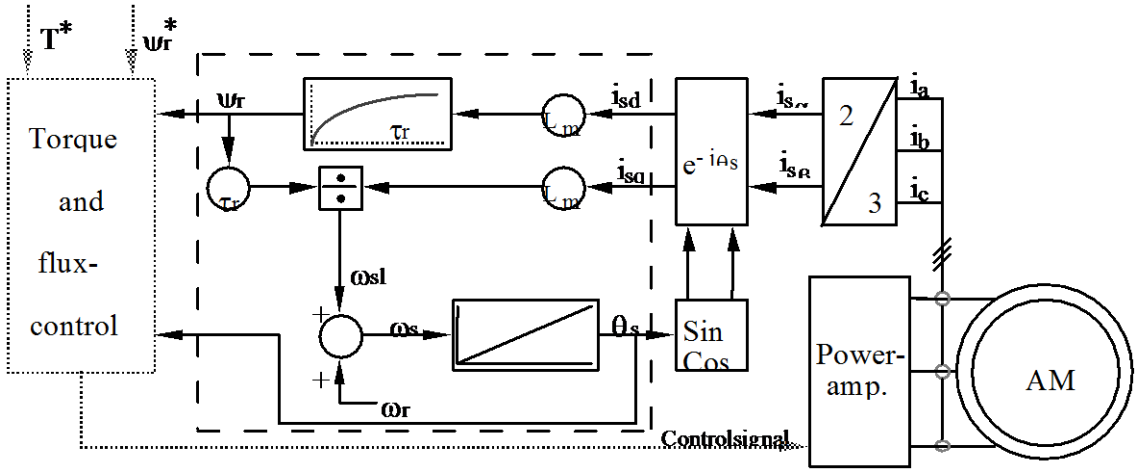


Figure 12.5: Rotor flux model, incorporated in a torque and flux control system.

Equations (12.9) third and fourth one indicate the stator voltages u_{sd} and u_{sq} required to modify the stator current and are therefore used for the design of current controllers for i_{sd} and i_{sq}.

The simple determination of rotor flux amplitude and position has determined a great deal the popularity of the above machine model. Its main disadvantage is that it contains a time constant which may vary considerably due to L_r (which in turn varies with magnetic saturation level) and due to R_r (which itself varies with temperature and frequency - skin effect). Control systems which regularly update the value of the time constant (so-called "on-line parameter estimation") have therefore been applied. An option to reduce the influence of the varying time constant τ_r is to turn over to the stator flux oriented machine model:

Simplified stator flux model

The stator flux reference frame is chosen to have its real axis defined by the stator flux vector in Figure 12.6.

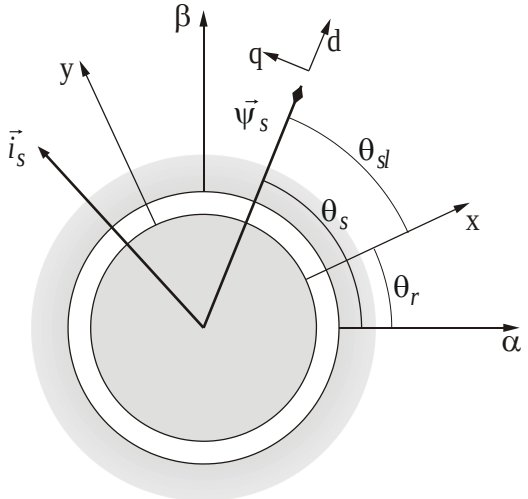


Figure 12.6: Stator flux system of coordinates.

The notations for the axes of the system of coordinates are the same as for the rotor flux reference frame. The stator equation (12.1) with variables expressed in stator flux coordinates will become:

$$\vec{u}_s^{dq} = R_s \cdot \vec{i}_s^{dq} + \frac{d}{dt}\left(\vec{\psi}_s^{dq}\right) + j\omega_s \cdot \vec{\psi}_s^{dq} \qquad (12.12)$$

Note that all vector variables become constant in stator flux coordinates. Therefore the derivative in (12.12) will vanish in the steady state, and the terminal voltage has only two components: the resistive voltage drop in phase with the stator current and the induced emf which leads the stator flux wave at an angle of 90°.

A simplified flux model can be derived that is entirely true for no-load, and a good approximation in many load situations. For this purpose equation (12.12) is rewritten:

$$\vec{u}_s = R \cdot \vec{i}_s + \frac{d\vec{\psi}_s}{dt} = R \cdot \left(\vec{i}_s - \vec{i}_{s0}\right) + R \cdot \vec{i}_{s0} + \frac{d\vec{\psi}_s}{dt}$$

$$\frac{d\vec{\psi}_s}{dt} = -R \cdot \vec{i}_{s0} + \left[\vec{u}_s - R \cdot \left(\vec{i}_s - \vec{i}_{s0}\right)\right] = \left\{ \begin{array}{c} \text{No load running} \\ \\ \vec{i}_s = \vec{i}_{s0} = \dfrac{\vec{\psi}_{s0}}{L_s} \end{array} \right\} = -R \cdot \frac{\vec{\psi}_{s0}}{L_s} + \vec{u}_s = \frac{\vec{\psi}_{s0}}{\tau_s} + \vec{u}_s \qquad (12.13)$$

Note that the lower row contains the expression of a low pass filter with gain and time constant equalling the stator time constant τ_s. The input signal is the stator voltage minus the voltage drop due to the load current $R \cdot \left(\vec{i}_s - \vec{i}_{s0}\right)$. If the machine does not operate under load obviously the load current will be zero so that the entire current is magnetizing current ($\vec{i}_s = \vec{i}_{s0}$). Thus the input signal will be the entire terminal voltage. In conclusion: the filter based on the above model does not consider the load current. It determines the flux for the machine running at actual terminal voltage without load.

The figure below shows a possible realization for the two filters for the components of the stator flux vector. v_a, v_b and v_c are the phase potentials of the AM measured at the machine terminals. Because the input resistance has large values (typically some 100 kΩ) the filters can obtain voltage information measured directly on the phase potentials. A big advantage of this scheme is that non-linarites are automatically accounted for in the converter. The resistance R_p in parallel with the capacitor C should determine a time constant equal to the stator time constant. By coupling this resistance with a switch controlled with relatively high frequency (typically some·10 kHz) the average filter time constant, which lies normally in the range [τ_s = 30 ... 200 ms], can be varied. The switch is a simple and cheap bilateral switch. Controlling all four switches

with the same train of control pulses makes the arrangement to function as a "4 way" potentiometer.

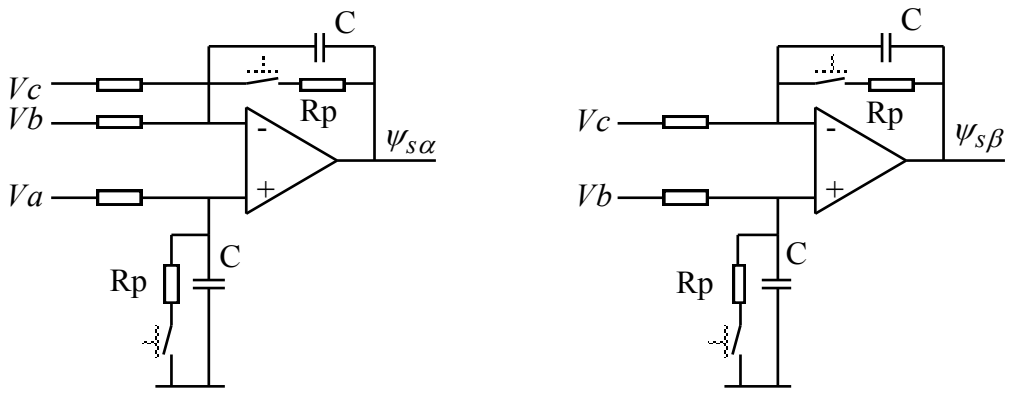


Figure 12.7: Analog filter for AM stator flux for no-load running.

Note that this filter exactly estimates the flux for no-load running if the filter time constant equals the stator time constant. The question concerning error magnitudes for the loaded machine arises and can be answered with the so-called circle diagram of the AM.

Starting with the equations:

$$\begin{aligned}
\frac{d\psi_{sd}}{dt} &= \frac{L_s}{\tau_r}\cdot i_{sd} - \frac{1}{\tau_r}\cdot\psi_{sd} + \left[\sigma\cdot L_s\cdot\frac{di_{sd}}{dt} - \omega_{sl}\cdot\sigma\cdot L_s\cdot i_{sq}\right] = u_{sd} - R_s\cdot i_{sd} \\
\frac{d\theta_{sl}}{dt} &= \frac{L_s\cdot i_{sq} + \sigma\cdot L_s\cdot\tau_r\cdot\dfrac{di_{sq}}{dt}}{\tau_r\cdot\psi_{sd} - \sigma\cdot L_s\cdot\tau_r\cdot i_{sd}} = \frac{u_{sq} - R_s\cdot i_{sq}}{\psi_{sd}} - \omega_r
\end{aligned} \tag{12.14}$$

which are the corresponding equations to (12.9) expressed in stator flux coordinates, it can be shown that in the stator flux system of coordinates (*d*,*q*) the steady state value for the stator current follows a trajectory known as the above mentioned circle diagram for the AM. Figure Figure 12.8 presents a circle diagram for a 5 kW 4-pole AM.

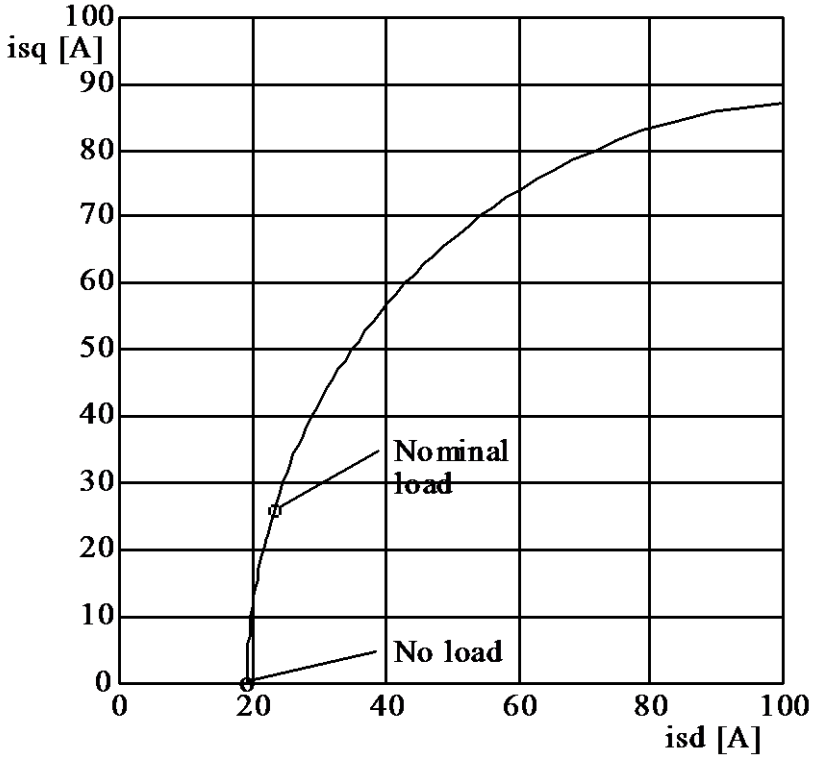


Figure 12.8: Circle diagram for an AM with the parameters Ls = 0.0364 H, Lr = 0.0364 H, Lm = 0.0346 H, Rs = 0.3300 Ω and Rr = 0.3800 Ω. The stator flux is constant □s = 0.7 Vs.

As the diagram reveals, the magnetizing current (i_{sd}) increases at the same time with the torque generating current (i_{sq}). Since resistive voltage drops determined by the load current (load current = difference between the total current and the no-load current) are summed the estimation error grows with increasing load and decreasing frequency. The prior neglected term becomes, relatively seen, a greater part of the terminal voltage. The diagrams in figure Figure 12.9 show the no-load speed versus the estimation error for half and full load.

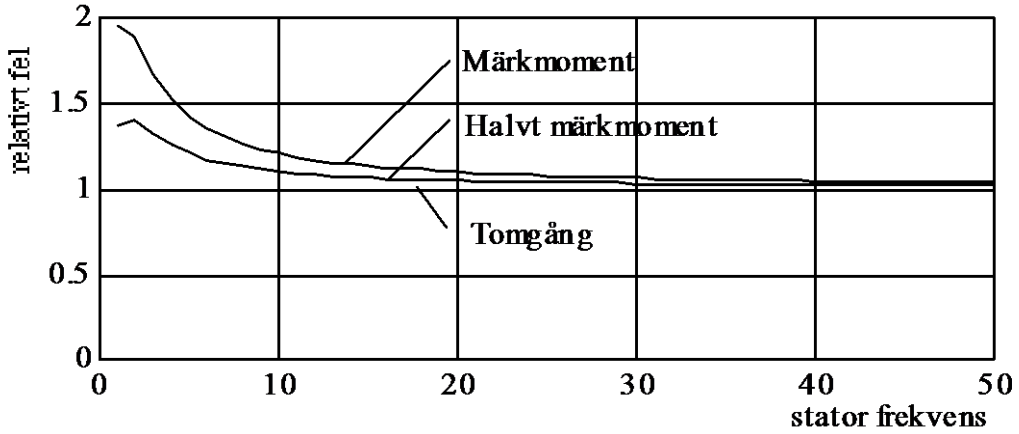


Figure 12.9: Relative flux estimation error for three different mechanical loads.

The graph shows that for full nominal load and low stator frequency the relative error is considerable. Since in many practical cases (pumps, fans etc.) the load torque depends on rotation speed, whereby relatively small torque values occur

at low speed, the described low pass filtering method provides good results. Figure 12.10 shows how the relative error behaves with varying load for a fan.

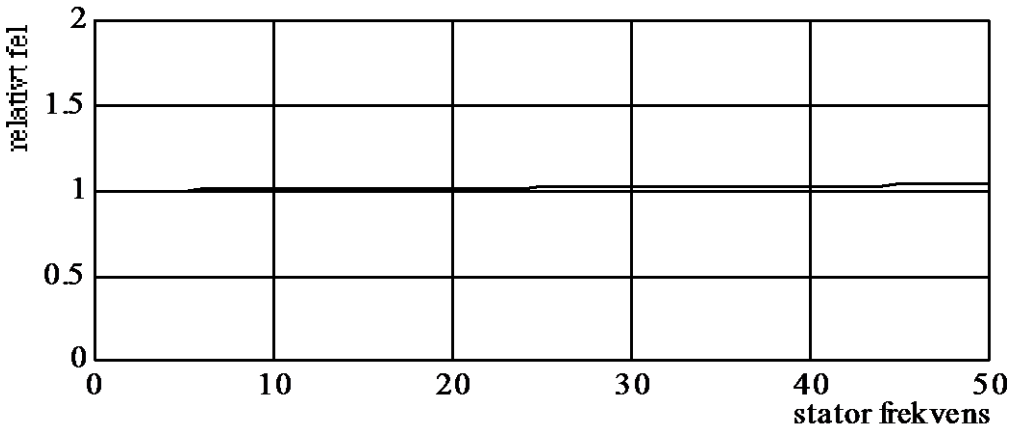


Figure 12.10: Relative flux estimation error for load torque having a quadratic dependency on frequency and reaching its nominal value at 50 Hz.

The types of load described in Figure 12.10 represent approximately 90-95% of the entire market for AM drive applications with speed dependent load. Other applications (for example servo drives) demand higher performance and require therefore more advanced control methods. A good example is the earlier described rotor flux model. The following section will introduce a complete feedback flux model - the flux observer.

Full flux observer

In the following we consider the stator and rotor flux as state variables for the observer model. We start with the vector equations (12.1) and (12.2) in stator coordinates (α,β) and construct the following matrix equation:

$$\begin{aligned} \vec{u}_s &= R_s \cdot \vec{i}_s + \frac{d\vec{\psi}_s}{dt} \\ \underbrace{0 &= R_r \cdot \vec{i}_r + \frac{d\vec{\psi}_r}{dt} - j\omega_r \cdot \vec{\psi}_r}_{\mathbf{u} = \mathbf{R} \cdot \mathbf{i} + \frac{d\Psi}{dt} + \Omega \cdot \Psi} \end{aligned} \qquad (12.15)$$

The observer output signal is the estimated stator current vector that, together with the stator current, is used to construct the corrective signal for the flux estimate.

$$\frac{d\hat{\psi}(t)}{dt} = -\left(\mathbf{R}\cdot\mathbf{L}^{-1}+\Omega\right)\cdot\hat{\psi}(t)+\mathbf{B}\cdot\mathbf{u}(t)+\mathbf{R}\cdot\mathbf{k}\cdot\left(\underline{i}_S(t)-\mathbf{C}\cdot\hat{\psi}(t)\right)=$$

$$=-\left(\mathbf{R}\cdot\mathbf{L}^{-1}+\Omega+\mathbf{R}\cdot\mathbf{k}\cdot\mathbf{C}\right)\cdot\hat{\psi}(t)+\mathbf{B}\cdot\mathbf{u}(t)+\mathbf{R}\cdot\mathbf{k}\cdot\underline{i}_S(t)=$$

$$=\mathbf{A}_{obs}\cdot\hat{\psi}(t)+\mathbf{B}\cdot\mathbf{u}(t)+\mathbf{R}\cdot\mathbf{k}\cdot\underline{i}_S(t) \qquad (12.16)$$

where $\mathbf{R}=\begin{bmatrix} R_S & 0 \\ 0 & R_r \end{bmatrix}$; $\mathbf{L}=\begin{bmatrix} L_S & L_m \\ L_m & L_r \end{bmatrix}$; $\hat{\psi}=\begin{bmatrix} \underline{\psi}_S \\ \underline{\psi}_r \end{bmatrix}$; $\mathbf{B}=\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$\mathbf{u}=\begin{bmatrix} \underline{u}_S \\ 0 \end{bmatrix};\ \mathbf{C}=\mathbf{L}^{-1}(1,:)\ and\ \Omega=\begin{bmatrix} 0 & 0 \\ 0 & -j\cdot\omega_r \end{bmatrix})$$

Note that the correction matrix contains the resistance matrix **R** and a dimensionless matrix **k**. The elements of the **k**-matrix indicate to what extent we can trust the measured and estimated stator current to calculate the flux estimate values. The dynamics of the estimation error is given by the eigenvalues of the matrix $\mathbf{A}_{obs}$. It is theoretically possible to arbitrarily adjust these eigenvalues compared to the eigenvalues of the initial system matrix **A**, by choosing an adequate **k**. matrix. Depending on the actual implementation of the observer very often restrictions to eigenvalue placement are to be adhered to in practice, as we will see later on. If the observer is implemented with a µ-processor the time continuous equations must be transformed into time discrete equations with a sampling interval T_s :

$$\hat{\psi}(k+1)=\mathbf{F}_{obs}\cdot\hat{\psi}(k)+\mathbf{G}_{obs}\cdot\bar{\mathbf{u}}(k,k+1)+\mathbf{K}_{obs}\cdot\bar{\underline{i}}_S(k,k+1) \qquad (12.17)$$

A problem occurring is that the system matrix $\mathbf{F}_{obs}$ depends on the rotor speed. In order to account for the actual rotor speed this matrix must be recalculated for each sampling interval. An exhaustive (but exact) calculation of all elements is time consuming and not desired since in practice signal processing time is limited and might stress the allowable limits even for fast signal processors. The following approximation provides good results in practice without complicating the calculations excessively:

$$\mathbf{F}_{obs}=e^{\mathbf{A}_{obs}\cdot T_s}=e^{-(\mathbf{R}\cdot\mathbf{L}^{-1}+\ +\mathbf{R}\cdot\mathbf{k}\cdot\mathbf{C})\cdot T_s}\approx$$

$$\approx e^{-\ \cdot T_s/2}\cdot e^{-(\mathbf{R}\cdot\mathbf{L}^{-1}+\mathbf{R}\cdot\mathbf{k}\cdot\mathbf{C})\cdot T_s}\cdot e^{-\ \cdot T_s/2}= \qquad (12.18)$$

$$=\begin{bmatrix} 0 & 0 \\ 0 & e^{j\cdot\omega_r\cdot T_s/2} \end{bmatrix}\cdot\mathbf{F}_{obs,\omega_r=0}\cdot\begin{bmatrix} 0 & 0 \\ 0 & e^{j\cdot\omega_r\cdot T_s/2} \end{bmatrix}$$

With this approximation the system matrix needs to be calculated only for ω_r = 0. Then 3 of the 4 matrix elements are multiplied with their corresponding

cosine or sine factors, which in turn may be expressed with a few terms in their series expansion since $\omega_r \cdot T_s$ is a small angle. One can demonstrate that if equation (12.18) is integrated over a sampling interval $[k \cdot T_s, (k+1) \cdot T_s]$ the system matrix may be expressed:

$$\mathbf{F}_{obs} = \left[I - \mathbf{A}_{obs} \cdot \frac{T_s}{2} \right]^{-1} \cdot \left[I + \mathbf{A}_{obs} \cdot \frac{T_s}{2} \right] \approx \left[I + \mathbf{A}_{obs} \cdot \frac{T_s}{2} \right]^2 \qquad (12.19)$$

Equation (12.19) is a very good approximation for the correct values, the errors are typically <0.2% when $\omega_r=0$, and closed algebraic expressions are easy to develop by hand. Thus together with (3-136) very simple algorithms for the time discrete observer can be given. In this way it is possible to fast recalculate $\mathbf{F}_{obs}$ for each new sampling interval and to modify the machine model parameters on-line.

Gopinath style flux observer

A Gopinath style flux observer is a different approach to the problem of estimating the stator and rotor fluxes. It basically consists on combining two flux linkage reduced order observers into a more robust and reliable close loop observer, that provides a good performance over a wider frequency range.

The first part of the Gopinath style observer is a flux estimator based on the stator current. Using equations (12.2) and (12.3) the next expression for the rotor flux is obtained:

$$\frac{\mathrm{d}\Psi_r^x}{dt} = \frac{R_r L_m I_s^x}{L_r} + \Psi_r^x \left(jw_r - jw_x - \frac{R_r}{L_r} \right) \qquad (12.20)$$

In (12.20), *x* represents the reference frame in which the vector quantities are expressed. The dependence on rotor velocity (thus the cross-coupled terms) can be eliminated if the equation is moved into the rotor reference frame, which results in:

$$\frac{\mathrm{d}\Psi_r^r}{dt} = \frac{R_r L_m I_s^r}{L_r} - \frac{R_r}{L_r} \Psi_r^r \qquad (12.21)$$

Notice that in order to implement this model, the rotor position is needed. It may be estimated or measured using a position sensor (e.g. resolver). Generally no position sensor is needed in the control system of an induction machine; therefore implementing this particular flux observer presents a drawback when compared to other flux estimators. Nevertheless, this is compensated by the

good performance obtained on those frequency ranges where other models use to fail. Besides, in practice, when equation (12.20) is implemented in a stationary reference frame, the cross-coupled terms introduce a delay of one sampling time on the feedback signals, which can make the system unstable for sufficiently high values of the rotor speed.

The second part of the combined observer is a flux estimator based on the stator voltage. This part is similar to the one described before, for the Full Flux Observer, but without the corrective term related to the stator current. Starting from equation (12.1) and considering the relation between the fluxes and the stator current obtained from equations (12.3), the rotor flux can be calculated as:

$$\frac{d\Psi_s^s}{dt} = V_s^s - R_s I_s^s \tag{12.22}$$

$$\Psi_s^s = \sigma L_s I_s^s + \frac{L_m}{L_r}\Psi_r^s \tag{12.23}$$

The flux observer based on the stator current provides a good performance at low frequencies and it is mostly sensitive to errors on the mutual inductance. On the other hand the stator voltage model performs better at high frequencies and it is affected by the stator voltage estimation and by mismatches on the rotor leakage inductance. Since the presented models perform better in different frequency ranges it seems a good idea to combine them into a more complex observer.

In order to obtain a soft transition between both models a PI controller is introduced. The input signals for the PI are the rotor fluxes estimated by the models and the output is used as a corrective signal for the stator voltage model as it is shown in the next figure:

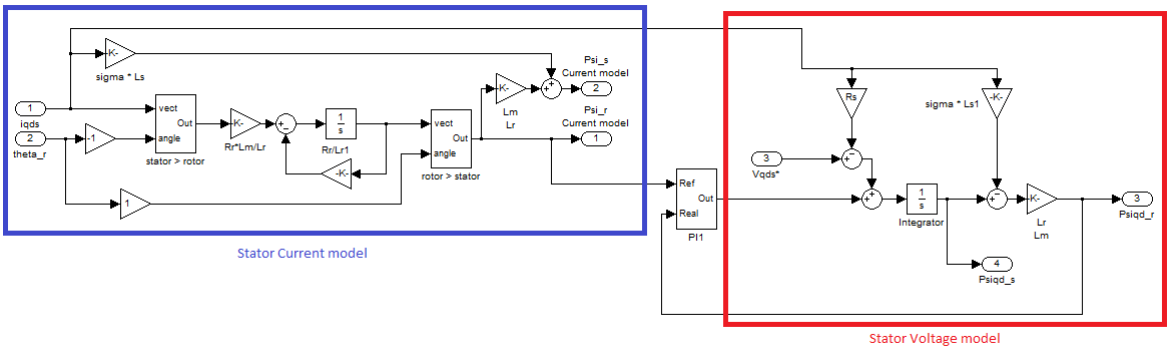


Figure 12.11: Gopinath style stator and rotor flux observer.

The parameters of the PI set the frequency range for the transition. A correct tuning of the PI controller is essential in order to obtain accurate flux estimation over the complete frequency range.

Voltage measurement

Voltage is one of the input signals for the flux observer. It is therefore necessary to measure the voltage delivered by a converter. Since the converter output voltage contains a large amount of harmonics it is not possible to measure this voltage with common sampling methods as for example the A/D converter of a µ-processor. Two different suggestions for practicable solutions have been presented in literature:

4. The reference value for the converter output voltage is taken instead of the real output voltage. Efforts are also made to compensate the time delay but since this is difficult to achieve the performance of the observer will be inherently limited.
5. The converter output voltage is filtered with a higher order low pass filter with the cut-off frequency of some hundreds of Hz, thus limiting the bandwidth of the observer and thereby of the entire drive system.

Another possibility is to use analogue integration of the uppermost relation in (3-134) and to use for the lowermost relation a discrete integration that is using (3-135). In this case the stator voltage cannot be used to calculate the rotor flux since it is not available. This leads the matrix $\mathbf{k}(2)$ to be necessarily chosen to have $\mathbf{G}_{obs}(2)=0$. The analogue integrator contains the signal processor (DSP) in the feedback loop, where the DSP supplies the analogue integrator with the remaining terms of equation (12.17). Thus the method reduces either the bandwidth or the accuracy of the observer as all the methods without direct stator voltage measurement. Figure 12.12 outlines a schematic block diagram for the observer.

An argument which speaks against the practical use of this method is that direct voltage measurement without any galvanic separation is not desired. On the other hand all methods using low pass filtering of the output voltage require direct voltage measurement by means of a galvanically separated transducer. This transducer has to cover the entire bandwidth of the output voltage and is quite expensive. With the above method galvanic separation may be built in the scheme after the integrator. Here much lower potentials occur and the signals will also have lower bandwidths making cheaper galvanic separation possible. Figure 12.13 shows how some variables for a 5 kW 4-pole AM drive evolve under an acceleration with nominal torque starting with -1500 rpm when the above suggested flux observer is employed.

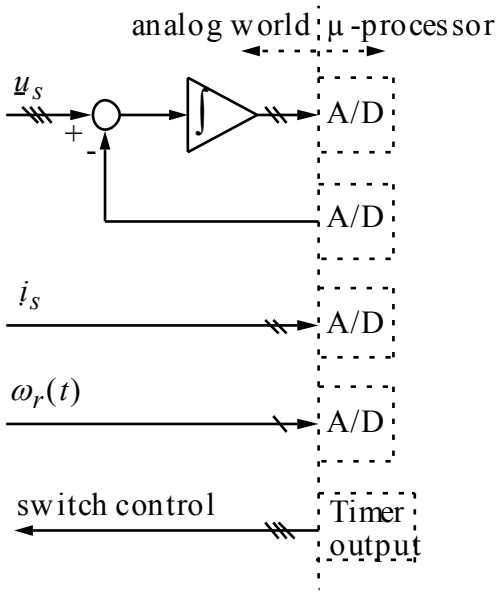


Figure 12.12: Schematic view of the voltage measurement method employed by the above suggested observer.

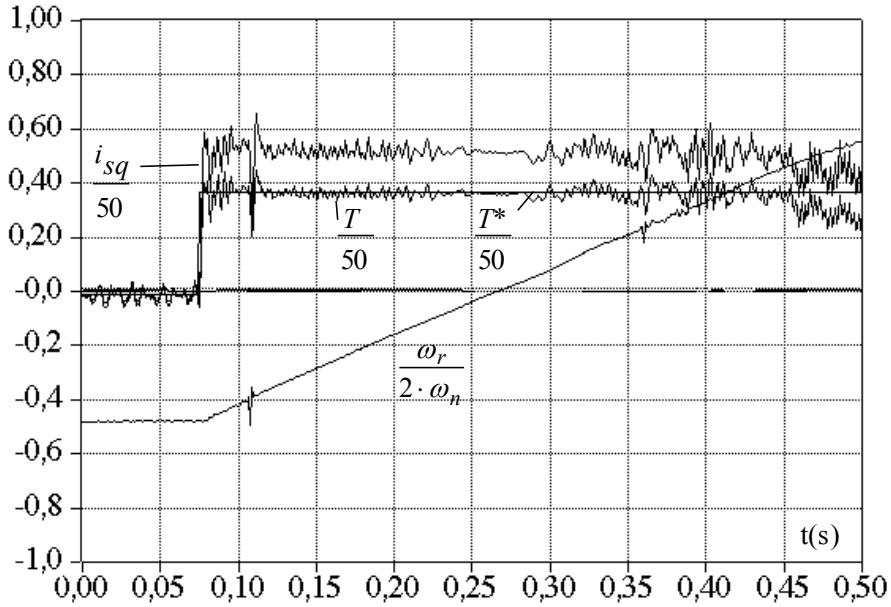


Figure 12.13: AM accelerating from -1500 rpm under the action of nominal torque with flux observer.

12.3 Torque control

The torque control for the AM may, as for the SM, be realized in several quite similar ways. The choice of the implemented method essentially depends on the required torque quality and speed range. Two common methods are presented in the following. One is vector control which is presented in the stator flux oriented reference frame, and the other is DTC (Direct Torque Control), also performed in the stator flux oriented reference frame.

Stator Flux Oriented Vector Control of Torque and Flux

Vector control of an Induction Machine can be performed in any reference frame. The following section describes how it can be done in the stator flux oriented reference frame. There are three advantages with using the stator flux oriented reference frame:

1 In a synchronously rotating reference frame, there are no stationary errors since all reference values are constant in stationarity.

2 The stator flux vector can in most cases be observed with high accuracy and low parameter sensitivity.

3 Direct flux control can be applied.

An elegant way to obtain the control of the stator flux is to use the real part of the stator voltage equation in the stator flux oriented reference frame, and make a first order linearization to obtain a controller for the stator flux:

$$\begin{aligned} u_{sd}(t) &= R_s \cdot i_{sd}(t) + \frac{d\psi_{sd}(t)}{dt} \\ u_{sd}^*(k) &= R_s \cdot \bar{i}_{sd}(k, k+1) + \frac{\psi_{sd}^*(k) - \psi_{sd}(k)}{T_s} = \\ &= \{\psi_{sd} = L_s \cdot i_{sd} \ \textit{in stationarity}\} = \\ &\approx R_s \cdot \frac{\psi_{sd}(k)}{L_s} + \frac{\psi_{sd}^*(k) - \psi_{sd}(k)}{T_s} \end{aligned} \qquad (12.24)$$

The assumption of the *d*-axis current is not true in transient state, but has no significant impact on control performance since the resistive voltage drop is a small part of the applied voltage in the d-axis in transient operation.

Assuming that the stator flux is quasi constant, a suitable reference value for the torque generating component of the stator current can be calculated with from the torque expression:

$$i_{sq}^*(k) = \frac{T^*(k)}{\psi_{sd}(k)} \qquad (12.25)$$

Thereafter, from equations (12.9) an algorithm for the quadrature current can be obtained as

$$u_{sq}=R_s\cdot i_{sq}+\omega_r\cdot\psi_{sd}+\left[\frac{L_s\cdot i_{sq}+\sigma\cdot L_s\cdot\tau_r\cdot\dfrac{di_{sq}}{dt}}{\tau_r\cdot\psi_{sd}-\sigma\cdot L_s\cdot\tau_r\cdot i_{sd}}\right]\cdot\psi_{sd} \tag{12.26}$$

Assuming that $\sigma L_s\cdot i_{sd} << \psi_{sd}$, and that $L_s/\tau_r = L_s/L_r\cdot R_r \approx R_r$, the voltage equation is simplified to:

$$\begin{aligned} u_{sq} &= (R_s+R_r)\cdot i_{sq}+\sigma\cdot L_s\cdot\frac{di_{sq}}{dt}+\omega_r\cdot\psi_{sd}\approx \\ &\approx (R_s+R_r)\cdot i_{sq}+(L_{s\lambda}+L_{r\lambda})\cdot\frac{di_{sq}}{dt}+\omega_r\cdot\psi_{sd} \end{aligned} \tag{12.27}$$

This resembles the equation of a generic single phase load according to Chapter **Error! Reference source not found.**, and the corresponding current controller can be used:

$$u_{sq}{}^*(k)=\left(\frac{L_{s\lambda}+L_{r\lambda}}{T_s}+\frac{R_s+R_r}{2}\right)\cdot\left(\left(i_{sq}^*(k)-\hat{i}_{sq}(k)\right)+\frac{T_s}{\left(\dfrac{L_{s\lambda}+L_{r\lambda}}{R_s+R_r}+\dfrac{T_s}{2}\right)}\cdot\sum_{n=0}^{n=k-1}\left(i_{sq}^*(n)-\hat{i}_{sq}(n)\right)\right)+\omega_r\cdot\psi_{sd}(k) \tag{12.28}$$

The equation in (12.28) describes the PI controller to control torque for the AM by means of adjusting the stator current in a stator flux oriented reference system of coordinates. The integral part contains the stator resistance and a term almost equal to the rotor resistance. The proportional gain is mainly determined by the sum of the quadrature leakage inductances. The feed forward term with the speed dependent voltage is only weighted with the rotor speed. The integral part in the controller allows for a certain flexibility that can be exploited with rotor speed measurement.

If the rotor speed enters the control law for i_{sq} and requires a special device for measurement, the stator frequency which often is already available in the circuits determining θ_s requires only low pass filtering to eliminate fast transients caused by step changes in torque. The signal obtained in this way allows to avoid an additional rotor speed measurement and the steady state error (bias) is eliminated by the integral part of the controller. A simpler procedure is to exclude the speed dependent voltage feed forward term. It also gives satisfactory results since the integral part compensates for the occurring error.

The entire torque and stator flux control concept derived above is presented in the block diagram in Figure 12.14.

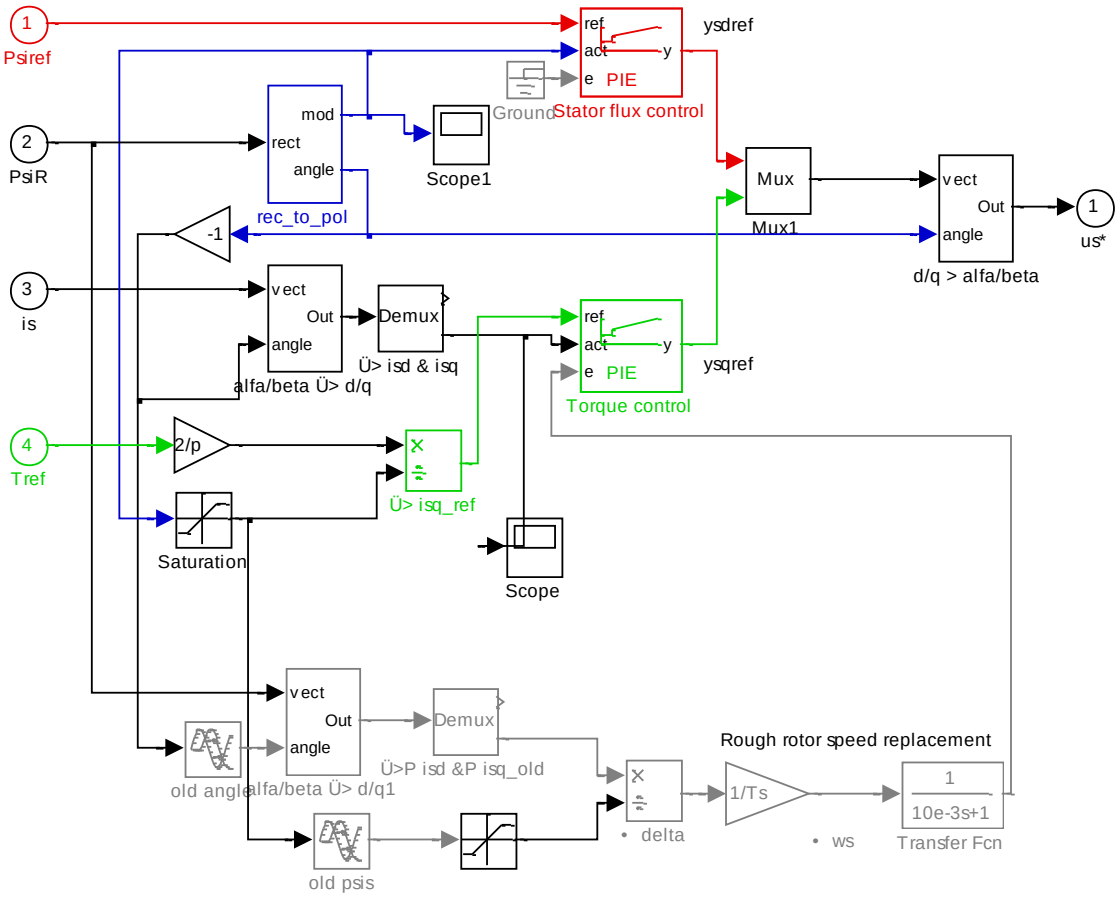


Figure 12.14: Vector controller for induction machine oriented to the stator flux vector.

Direct torque control

In chapter 3.9, a control method called Direct Current Control was applied to a generic three phase load. The same method can be applied to control of an induction machine. Using the stator flux oriented reference frame, the currents along the *d*- and the *q*- axes could be controlled by a DCC controller. A more useful controller is however created of the d-axis current is replaced with the *d*-axis flux, i.e. the stator flux, and if the *q*-axis current is replaced with the torque. This can be done since any control action on the d-axis affects the stator flux, and any control action on the *q*-axis affects the torque, following the definition of the *d*- and the *q*-axes. The DCC method, converted to control the flux and torque directly is used, it is referred to as Direct Torque Control, or DTC. Figure 12.15 shows the contents of a DTC controller for an induction machine.

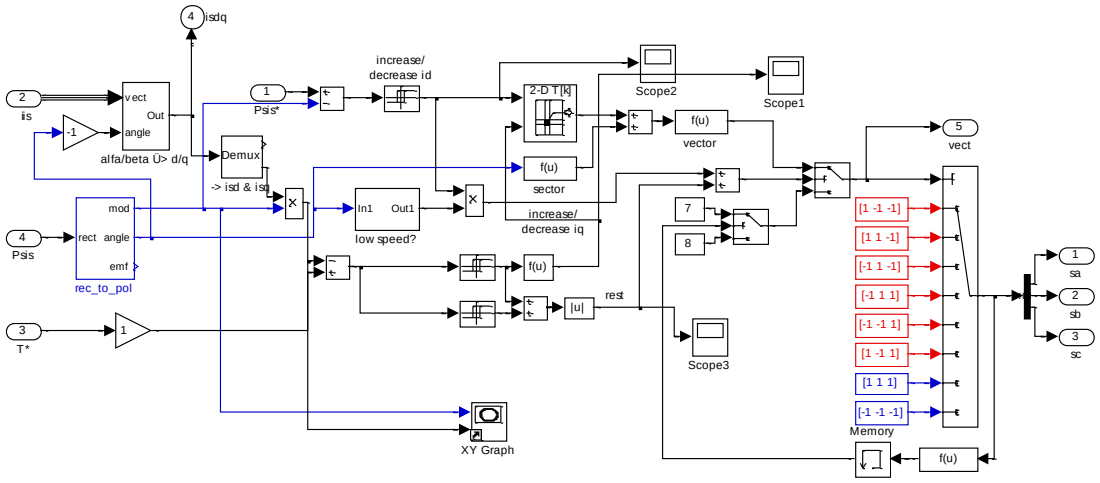


Figure 12.15: Content of a DTC-controller for an induction machine, oriented to the stator reference frame.

Deadbeat Direct Torque and Flux Control (DB-DTFC)

A new alternative control method to Direct Torque Control had been developed recently. This technique shares with DTC that both are aimed to set the torque and the stator flux to the reference levels instead of the currents as in the FOC strategies. However, instead of using a switching table as in DTC, DB-DTFC uses an inverse model of the induction machine. This model is solved in discrete time in order to calculate the voltage time vector needed to generate the desired change on the stator flux and torque over one sampling period.

Starting from the electrical torque equation (12.4) and (12.5) an alternative expression can be obtained:

$$T = \left(\frac{L_m}{\sigma L_s L_r}\right) \Psi_r^s \mathrm{x} \Psi_s^s \tag{12.29}$$

Then the mechanical torque is:

$$Tmech = \frac{P}{2}\left(\frac{L_m}{\sigma L_s L_r}\right) \Psi_r^s \mathrm{x} \Psi_s^s \tag{12.30}$$

When the previous equation is derivate in discrete time, assuming that the rotational speed is constant over one sampling period, the obtained change in torque is presented in the next equation:

$$\dot{Te} = \frac{P}{2}\left(\frac{L_m}{\sigma L_s L_r}\right)\left[\left(\frac{R_s \mathrm{L_r} + \mathrm{R_r L_s}}{\sigma L_s \mathrm{L_r}}\right)\left(\Psi_{\mathrm{rd}}\Psi_{\mathrm{sq}} - \Psi_{\mathrm{sd}}\Psi_{\mathrm{rq}}\right) + \Psi_{\mathrm{rd}}V_{\mathrm{sq}} - \Psi_{\mathrm{rq}}V_{\mathrm{sd}} - w_r(\Psi_{\mathrm{rd}}\Psi_{\mathrm{sd}} + \Psi_{\mathrm{sq}}\Psi_{\mathrm{rq}})\right] \tag{12.31}$$

The previous equation can be reorganized, combined with equation (12.30) and moved into the stator flux reference frame to obtain the following expression

$$V_{\mathrm{sq}}\mathrm{T_s} = \frac{\Psi_{\mathrm{rq}}}{\Psi_{\mathrm{rd}}}V_{\mathrm{sd}}\mathrm{T_s} + w_r\Psi_{\mathrm{sd}}\mathrm{T_s} + \frac{2\sigma L_s \mathrm{L_r}}{\mathrm{PL_m}}\left[\left(\frac{R_s \mathrm{L_r} + \mathrm{R_r L_s}}{\sigma L_s \mathrm{L_r}}\right)\mathrm{Tmech}\mathrm{T_s} + \Delta\mathrm{Tmech}\right] \tag{12.32}$$

In order to simplify further explanations, let us define:

$$m = \frac{\Psi_{\mathrm{rq}}}{\Psi_{\mathrm{rd}}} \tag{12.33}$$

$$b = w_r\Psi_{\mathrm{sd}}\mathrm{T_s} + \frac{2\sigma L_s \mathrm{L_r}}{\mathrm{PL_m}}\left[\left(\frac{R_s \mathrm{L_r} + \mathrm{R_r L_s}}{\sigma L_s \mathrm{L_r}}\right)\mathrm{Tmech}\mathrm{T_s} + \Delta\mathrm{Tmech}\right] \tag{12.34}$$

By doing this, equation becomes:

$$(V_{sq}\mathrm{T_s}) = \mathrm{m}(V_{sd}\mathrm{T_s}) + \mathrm{b} \tag{12.35}$$

This equation represents a straight line on the voltage time space, as it is shown in the next figure:

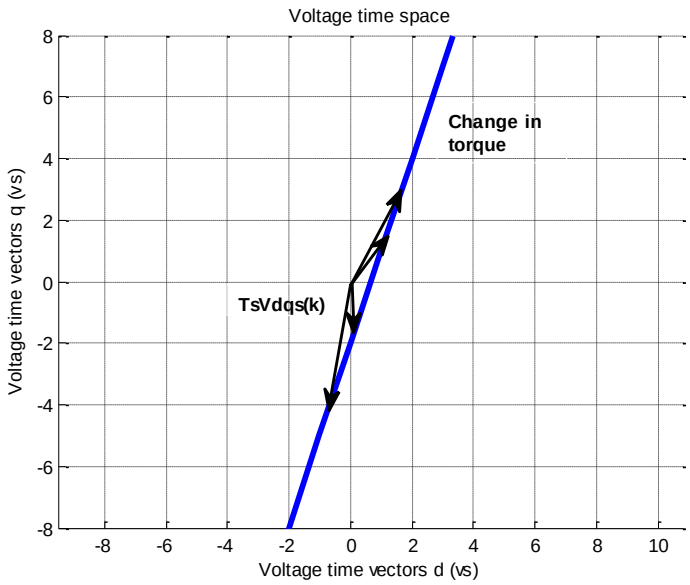


Figure 12.16: Voltage time vectors for the same change in torque.

Notice that there are infinite combinations of voltage time vectors that yield the desired change in torque, each of them resulting in a different stator flux value. Hence, in order to achieve the desired stator flux reference another equation needs to be considered.

If the stator resistance is neglected, the stator flux linkage equals the integral of the stator voltage. In discrete time this can be expressed:

$$V_s \mathrm{T_s} = \Psi_s(k+1) - \Psi_s(k) \tag{12.36}$$

The magnitude of the voltage time vector (VsTs) represents a circle in the voltage time space, as shown in the next figure:

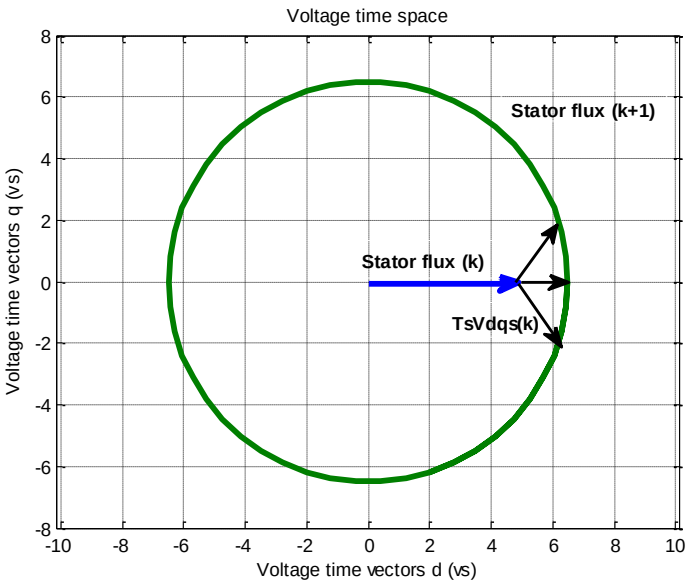


Figure 12.17: Voltage time vectors for the same stator flux magnitude.

As in the previous case, it can be observed that there are several voltage time vectors in the dq plane that take the stator flux to the commanded reference over one sampling period.

Equation can be rewritten to represent the change in the flux magnitude ($\Delta\Psi_{qds}$) as:

$$[\Psi_{sd}(k+1)]^2 = [V_{sd}T_s + \Psi_{sd}(k)]^2 + \left(V_{sq}T_s\right)^2 \qquad (12.37)$$

Where Ψ_{qds}(k+1) is the flux reference (Ψ_{ref}). The graphical representation of this equation is presented next:

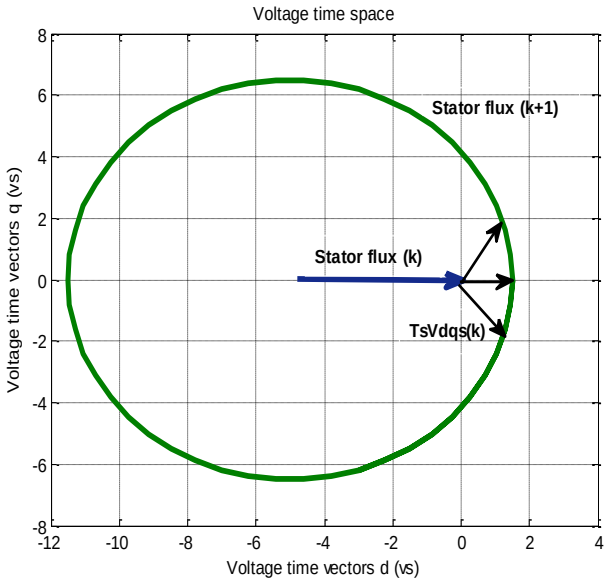


Figure 12.18 Voltage time vectors for the same change in the stator flux.

Using equations (12.32) and (12.37) it is possible to find a voltage vector V_{qds} that solves both equations at the same time, and therefore fulfills the torque and flux requirements. This solution is shown in the following figure:

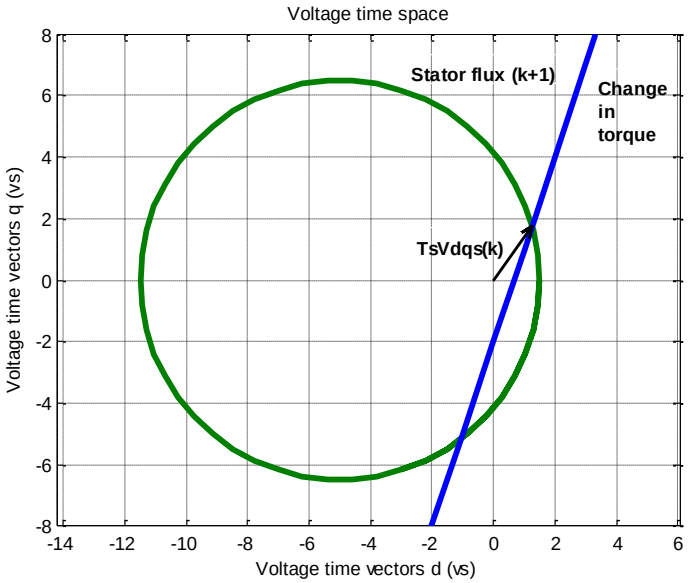


Figure 12.19 Voltage time vectors that achieve both the torque and the flux.

In addition to all the previous calculations, it must be considered that the DC link voltage available in the power converter is limited, meaning that not all the points in the voltage time plane are reachable in one sampling period. The voltage time vectors that the inverter is able to produce are represented in the dq plane as a hexagon with sides equal to $2/3V_{dc}T_s$. When solving equations D and E together, there may be two feasible solutions (as shown in Figure 12.20) that satisfy both torque and flux equations (the intersection between the flux circle and the torque line). However, depending on the DC link voltage, both of them, just one, or none are reachable. In general, the solution with the smallest module for $V_{qds}T_s$ is selected, since it is the one that is most easily reached with a limited DC link voltage. The next figure shows the graphical representation of both equations and the DC link voltage limitation.

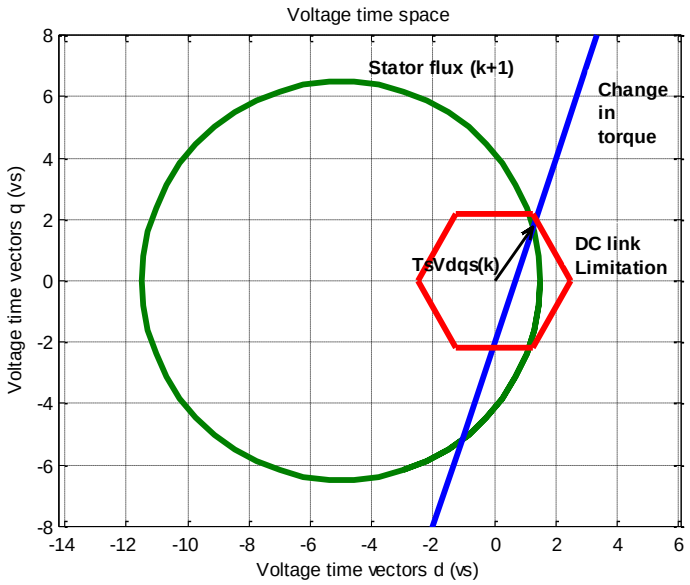


Figure 12.20 Voltage time vectors that achieve both the torque and the flux considering the DC link limitation

In summary, in order to produce a deadbeat response the voltage time vector that lies in the intersection between the torque line and the flux circle and at the same time falls inside the DC link voltage hexagon should be selected.

First the intersection between the torque line and the flux circle is calculated, solving the system of equations composed by (12.32) and (12.35). The possible solutions for (V_{sd} T_s) are:

$$(V_{sd}\mathrm{T_s}) = \frac{(-\Psi_{\mathrm{sd}} - \mathrm{mb}) \pm \sqrt{(1+\mathrm{m}^2)\Psi_{\mathrm{ref}}^2 - (\mathrm{m}\Psi_{\mathrm{sd}} - \mathrm{b})^2}}{(1+\mathrm{m}^2)} \qquad (12.38)$$

From this result V_{qs} can be calculated using equation (12.35).

13 Power Quality Enhancing Systems

Power Quality Enhancing Systems is an application area for power electronic converters that has emerged the last decade. Acronyms like APF (Active Power Filter), UPFC (Unified Power Flow Controller), SVC (Static Var Converter), HVDC (High Voltage Direct Current) and more describe power electronic systems with related functionality used to support the power system in maintaining a high power quality. The HVDC system is furthermore built to transmit active power. This chapter is devoted to a topology that appears in several of those systems, in particular in Active Filters. By understanding the basic functionality described in this chapter, many other variations are more easily understood as well.

13.1 Reactive Power Compensation – Voltage Control

Reactive power compensation is equivalent to voltage control that is by controlling the reactive power flow in a network substantial control of the voltage magnitude is achieved.

Reactive power Q in contrast to the active power P cannot be converted into mechanical work, i.e. kinetic energy or heat etc. Instead, the reactive power represents an energy oscillation between load and generator. Still reactive power is a necessary quantity in electrical engineering, since it is used in electrical motors to build up the magnetic flux used to convert electrical energy into mechanical work. Reactive power is a steady-state property and defined as the imaginary part of the complex valued apparent power S.

$$\bar{S} = \bar{U} \cdot \bar{I}^{*} = P + jQ \tag{13.1}$$

Problem – Losses and large distance transfer

Consider a loss less single distribution line with reactance X_l connecting a stiff (reference) bus to a load bus as in Figure 13.1.

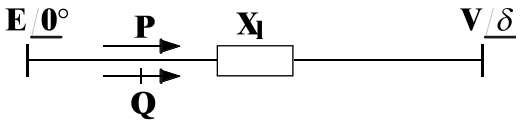


Figure 13.1 Power flow in a single line with reactance Xl.

The transfer of active and reactive power, from the stiff bus, over the line are strongly coupled and given by

$$P = \frac{E \cdot V}{X_l} \cdot \sin\delta$$
$$Q = \frac{E}{X_l}\left(E - V \cdot \cos\delta\right) \tag{13.2}$$

where E is the voltage at the stiff bus, V is the voltage at the load bus and δ is the load angle, i.e. the angle difference between the voltages at the stiff and load bus.

From equation (13.2) it is found that large amount of reactive power transfer results in a significant voltage drop over the line. This is due to the dependency between the load angle δ and the active power transfer. It is also seen that a weak line, i.e. large X_l, reduces the power transfer capability.

The reactive current contributes to the rms-value of the supply current, thus resulting in a higher supply current than needed in order to deliver the active power to the load. This affects the efficiency of the power grid since in reality lines are not loss less; there exist a small resistance in addition to the reactance for every conductor. The losses in the conductors are proportional to the square of the current; therefore it is obvious that also active as well as reactive currents cause losses.

Solution – Local production

Local production of reactive power at a load bus decreases the amount of reactive power transfer over the line, thus reducing the rms-value of the supply current to only correspond to the active power demand at the load bus. Therefore, the losses and the voltage drop across the line are decreased and a more efficient power grid is obtained. As a consequence, the voltage at the load bus may be controlled close to the nominal value by control of the reactive power input.

Shunt capacitor banks provide a simple way of reactive power compensation, see Figure 13.2. In a capacitor excited by a AC voltage the current leads the voltage by 90° in phasor notation, this phase lead corresponds to a injection of reactive current and thus production of reactive power. Shunt capacitor banks are common in distribution networks, where they are connected and disconnected according to the load fluctuation over the day in order to maintain a voltage magnitude close to the nominal.

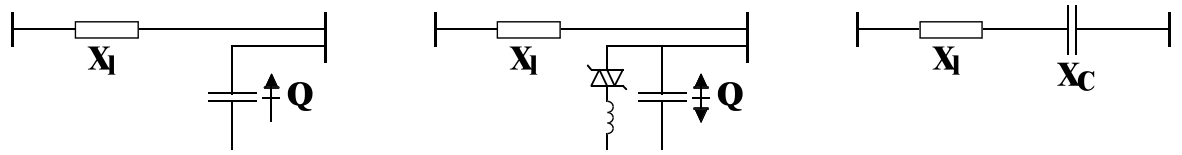


Figure 13.2 Reactive power compensation; shunt (left), TCR (middle) and series (right).

A more sophisticated solution includes a thyristor-controlled reactor (TCR) in parallel with the shunt capacitors. Such a device may even absorb reactive power during light load condition in addition to the production of reactive power. This combination can be treated as a capacitor bank with variable and even negative capacitance. Synchronous condensers, i.e. synchronous motors with no mechanical load, can also control their reactive power output to provide reactive power compensation.

Another form of reactive power compensation involves series capacitors, where capacitor banks are installed in series with the line. Their effect is to reduce the total series impedance of the line seen from the source, thereby reducing line voltage drops and increasing the line load ability.

Lately, even more advanced compensator structures, where power electronics are included, have been proposed. The unified power flow controller (UPFC) consists of two converters connected back to back, one in series with the line and the other one in parallel. The UPFC are intended for use on transmission systems and is very flexible. It can perform shunt reactive compensation as well as voltage magnitude and load angle control and line impedance compensation.

13.2 Load balancing

A balanced power grid is obtained if the three phases are equally loaded, i.e. each phase transfers one third of the total power. This gives the best utilization of the power grid, since during balanced conditions the phases are affected in the same way. Electrical loads can be one, two or three phase loads but since a line feeds several loads, the total load seen from the source is balanced if a large number of loads are distributed to the different phase conductors. Hence for the transmission network and the higher voltage distribution networks the balanced condition holds true. But close to the consumers, the load may be imbalanced due to the low number of loads to be distributed.

Imbalanced loads or systems may be analyzed in steady state by the use of symmetrical components. In accordance with symmetrical components, any balanced or imbalanced three phase quantity are resolved into three sequence components, that is positive-, negative- and zero-sequence components. The positive-sequence component consists of three phasors with equal magnitude and 120° phase displacement in positive sequence, i.e. in *abc* sequence. The negative-sequence component also consists of three phasors with equal

magnitude and 120° phase displacement however in negative sequence, i.e. in *acb* sequence, while the zero-sequence component have zero phase displacement. Figure 13.3 shows an example of an arbitrary unsymmetrical current separated into sequence components.

A balanced system with positive-sequence, *abc* sequence, and rms-value *I* can be represented as a vector in the stationary reference $\alpha\beta$-frame as

$$\vec{i}_p^{\alpha\beta} = \sqrt{3} I e^{j\omega_1 t} \tag{13.3}$$

while for a balanced system with negative-sequence, *acb* sequence, the vector is according to

$$\vec{i}_n^{\alpha\beta} = \sqrt{3} I e^{-j\omega_1 t} \tag{13.4}$$

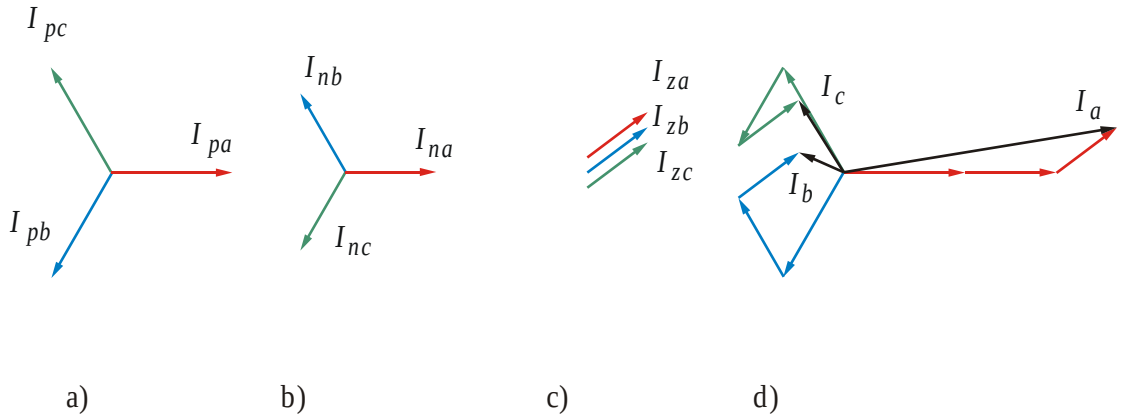


Figure 13.3 Sequences of an arbitrarily unsymmetrical current, a) positive-sequence, b) negative-sequence, c) zero-sequence and d) total current.

Positive- and negative-sequence components can thus be separated by the rotation direction of the vector in the stationary reference $\alpha\beta$-frame, see Figure 13.4. Applying the same transformation on a balanced system with zero-sequence gives

$$\vec{i}_z^{\alpha\beta} = 0 \tag{13.5}$$

This means that zero-sequence components are not represented in the stationary reference $\alpha\beta$-frame.

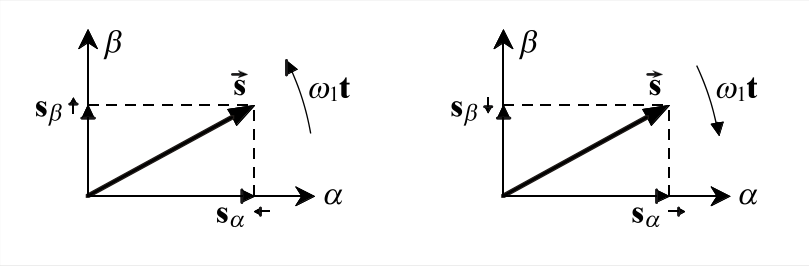


Figure 13.4 Rotation direction of vectors in the stationary αβ-frame, positive-sequence (left) and negative-sequence (right).

In order to obtain a vector transformation that is valid also for zero-sequence components, the transformation can be modified to an $abc \rightarrow \alpha\beta 0$ vector transformation.

Problem – Neutral currents and voltage distortion

Imbalanced loads cause a distorted current distribution among the phase conductors as well as current in the neutral conductor. This leads to higher losses in the phase conductors than for a balanced load and may cause overload of a phase conductor. Neutral currents can result in circulating currents in transformers and additional losses in the windings and in the neutral conductor.

Another aspect of imbalance currents and especially neutral currents is the effect on the voltages at the load due to the self-impedance in the conductors. The voltage drop in the neutral conductor affects the virtual zero potential at the load, resulting in an imbalance of the three phase voltages in terms of magnitude and phase displacement. This may cause over- and/or under-voltage in some of the phases, which in turn may cause malfunction or destruction of the electrical equipment connected to these phases.

Solution – Redirection of load currents

The most obvious solution to load imbalances is to physically connect similar loads to different phase conductors. Load balancing may also be obtained by redirection of load currents by a symmetrization unit, see Figure 13.5. The average power used in the load mainly corresponds to the positive-sequence of the load current due to the balanced supply voltage of positive-sequence. Therefore, the negative- and/or zero-sequence components of the load current may be generated locally and thus eliminating those components from the supply current. The resulting supply current would only contain the positive-sequence component of the load current and the unsymmetrical load is seen as an equivalent symmetrical load from the supply point of view.

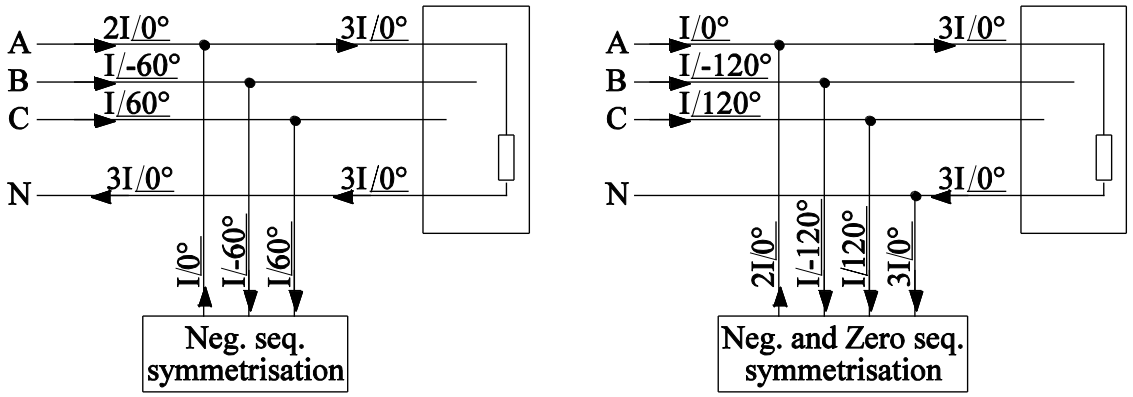


Figure 13.5 Load balancing principle, negative-sequence compensation (left), and negative- and zero-sequence compensation (right).

In Figure 13.5, the basic operation principles of such symmetrisation units are shown. In the left figure, the negative-sequence current of the load is provided by the compensator, thus resulting in only zero- and positive-sequence supply currents. Total symmetrisation is obtained in the right figure by injection of the negative- and zero-sequence currents, resulting in a balanced supply current of positive-sequence and no neutral current.

13.3 Harmonic reduction in distribution networks

Electrical equipment made of power electronics connected to the grid causes, due to their non-linear characteristics, non-sinusoidal grid currents. According to the fourier analysis, every periodic waveform can be treated as the sum of one or several sinusoidal waveforms of different frequencies, i.e. the fundamental frequency and multiples of the fundamental frequency. Since the grid voltage is nearly sinusoidal, the power used by a harmonic generating load corresponds to the fundamental frequency component of the load current. The current harmonics are a non-desired bi-product due to the switched operation of line commutated rectifiers, i.e. diode or thyristor controlled rectifiers.

Current harmonics are in fact a steady-state problem due to their periodic and repetitive behaviour. However, current transients originating from inrush currents at start up or at distinctive shut downs or disconnection of loads can also be treated, in accordance to the fourier analysis, as current harmonics despite their lack of periodicity. Still, transients and their corresponding current harmonics will disappear after a sufficiently long time.

A typical load current and the corresponding harmonic spectrum of a line commutated three phase rectifier can be seen in Figure 13.16. The current contains the fundamental frequency and harmonics of order $6k\pm1$ where k is any positive integer. The current harmonics can further be resolved into sequence components, i.e. positive- and negative-sequences, according to Table 13.2. The absence of zero-sequence harmonics is due to the balanced configuration of

three phase line commutated rectifiers. However, single phase line commutated rectifiers have additional zero-sequence harmonics of order *6k+3* where *k* = [0,1,2…].

Table 13.2 Harmonic sequences of three phase line commutated rectifiers

Sequence	Harmonics
Positive	1, 7, 13, 19, …
Negative	5, 11, 17, 23, …

The load current of a line commutated rectifier can be represented in the stationary reference $\alpha\beta$-frame as

$$\vec{i}_{Load}^{\alpha\beta} = \sqrt{3}I_1 e^{j(\omega_1 t+\phi_1)} + \sum_{h=6k-1}\sqrt{3}I_h e^{-j(h\omega_1 t+\phi_h)} + \sum_{h=6k+1}\sqrt{3}I_h e^{j(h\omega_1 t+\phi_h)} \qquad (13.6)$$

where I_1 and I_h are the rms-values of the fundamental current component and the individual current harmonics, respectively. Equation (13.6) gives a true representation of the current of three phase line commutated rectifiers, whereas for single phase rectifiers the current harmonics of zero-sequence are absent. This is due to the vector transformation used.

Problem – Losses and voltage distortion

Current harmonics result in a distorted voltage, which may cause malfunction of sensitive electrical equipment, i.e. computers. The voltage distortion is due to the voltage drop across the non-zero impedance of cables and transformers. The current harmonics can also cause severe voltage distortion due to resonance oscillations between the self-inductance in the distribution network and the shunt capacitor banks for reactive power compensation.

Current harmonics cause additional losses in the distribution network. Transformers are exposed to thermal stresses due to an increase in iron losses caused by the distorted voltages and currents. This may lead to derating as well as premature ageing of the transformers. For electrical motors, the voltage distorsion causes thermal stresses and torque ripple.

Solution – Harmonic filtering

Current harmonics are often treated as a local problem at least for one feeder in the distribution network. The impedance of the distribution network dampens the harmonic propagation. Therefore, harmonic filtering should be performed nearby the source of the current harmonics for the best result. If this is done, other equipment will be unaffected by the harmonic producing load. Harmonic

filtering or compensation is accomplished by the use of passive filters, shunt active filters (AF) or series active filters.

Passive filters

Passive filters for harmonic reduction provide low impedance paths for the current harmonics. The current harmonics flow into the shunt filters instead of back to the supply. A typical shunt passive filter and the resulting equivalent impedance seen from the load are shown in Figure 13.6. The passive filter consists of series LC filters tuned for specific harmonics in combination with a general high pass filter used to eliminate the rest of the higher order current harmonics.

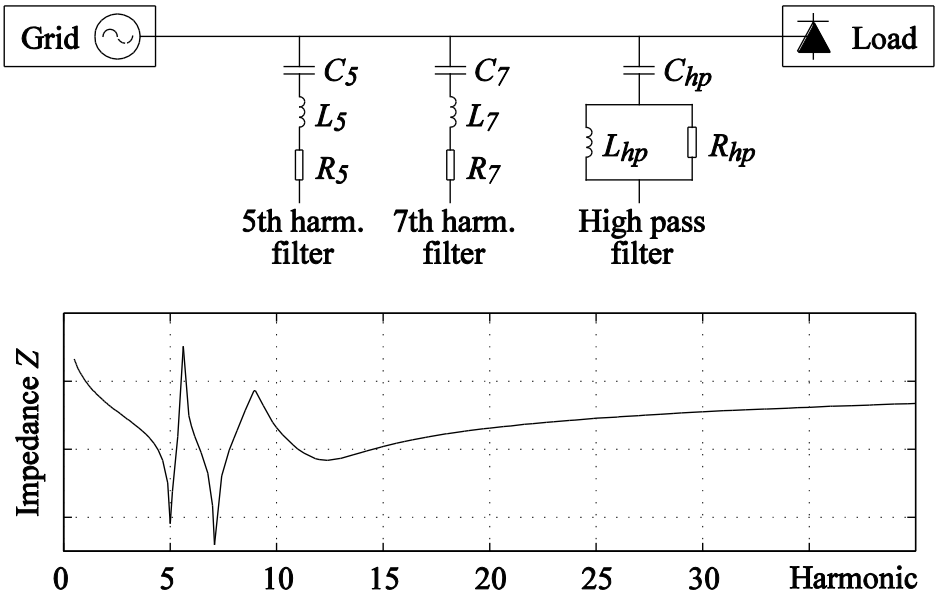


Figure 13.6 Typical passive filter for reduction of current harmonics and the equivalent impedance seen from the load.

The performance of a passive filter is strongly dependent on the system impedance at the harmonic frequencies. The system impedance depends on the distribution network configuration and the loads. Therefore, design of passive filters involves thorough system analysis in order to obtain adequate filtering performance of the filter.

Shunt active filters

A shunt active filter consists of a controllable voltage or current source. The voltage source converter (VSC) based shunt AF is by far the most common type used today, due to its well-known topology and straight forward installation procedure, see Figure 13.7. It consists of a DC-link capacitor C_{dc}, power electronic switches and filter inductors L_f.

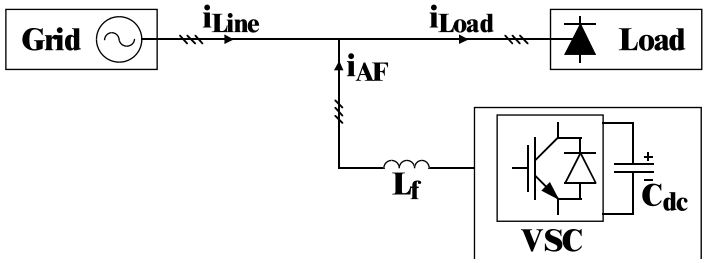


Figure 13.7: Principle configuration of a VSC based shunt active filter.

The operation of shunt AFs is based on injection of current harmonics i_{AF} in phase with the load current i_{Load} harmonics, thus eliminating the harmonic content of the line (supply) current i_{Line}. That is, suppose the load current can be written as the sum of the fundamental current component and the current harmonics according to

$$\vec{i}_{Load} = \vec{i}_{Load,\,fund} + \vec{i}_{Load,harmonics} \tag{13.7}$$

then the injected current by the shunt AF should be

$$\vec{i}_{AF} = \vec{i}_{Load,harmonics} \tag{13.8}$$

The resulting line current is

$$\vec{i}_{Line} = \vec{i}_{Load} - \vec{i}_{AF} = \vec{i}_{Load,\,fund} \tag{13.9}$$

which only contains the fundamental component of the load current and thus free from harmonics. Figure 13.8 shows the ideal phase currents when the shunt AF performs harmonic filtering of a diode rectifier. The injected shunt AF current completely cancels the current harmonics from the load, resulting in a harmonic free line current.

From the load current point of view, the shunt AF can be regarded as a varying shunt impedance. The impedance is zero, or at least small, for the harmonic frequencies and infinite in terms of the fundamental frequency.

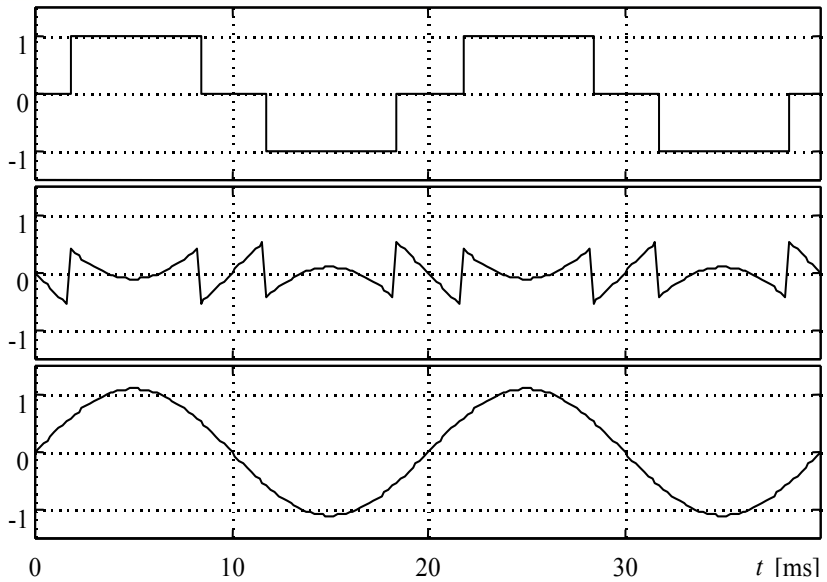


Figure 13.8 Shunt AF operation principle when performing harmonic filtering. Load current (top), filter current (middle) and line current (bottom)

Series active filters

Series active filters are connected in series with the line through a matching transformer. VSCs are appropriate as the controlled source even for series AFs, thus the principle configuration of series AFs are similar to shunt AFs. The filter inductors of shunt AFs are replaced with the series transformer, see Figure 13.9.

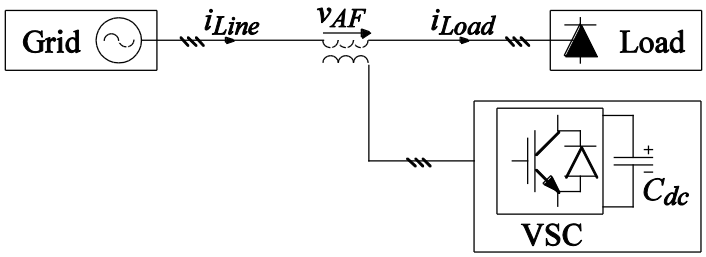


Figure 13.9 Principle configuration of a series AF system.

The operation principle of series AFs is based on isolation of the harmonics in between the load and the supply. This is obtained by the injection of harmonic voltages across the transformer line side. The injected harmonic voltages affect the equivalent line impedance seen from the load. The equivalent line impedance can be regarded as infinite or large for the harmonics whereas it should ideally be zero for the fundamental frequency component. The operation principle of a series AF is illustrated in Figure 13.10.

Harmonic isolation is achieved by means of the infinite impedance for the current harmonics in series with the line. That is, no current harmonics can flow from load to source or from source to load. Current harmonics produced elsewhere on the source side of the series AF cannot interfere with loads on the load side, i.e. passive filters and shunt capacitors.

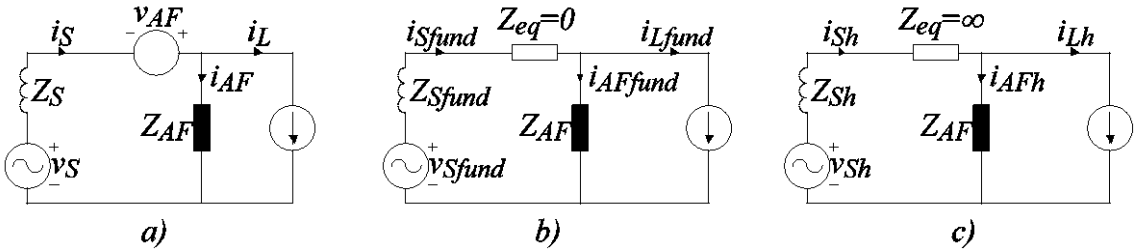


Figure 13.10 Series AF operation principle; a) single phase equivalent of series AF, b) fundamental equivalent circuit, and c) harmonic equivalent circuit.

Hybrid filters

Hybrid filters consists of combinations of shunt/series AFs and shunt passive filters. The main purpose of hybrid filters is to reduce the initial costs and to improve efficiency. The passive filters reduces the bulk harmonic content of the load current, whereas the active filter handles the rest of the harmonic content not filtered by the passive filters. Therefore, the rating of the active filter can be decreased compared to a standalone active filter and thus reducing the initial cost. Figure 13.11 shows two examples of hybrid filter configurations, additional configurations have been reported in the literature.

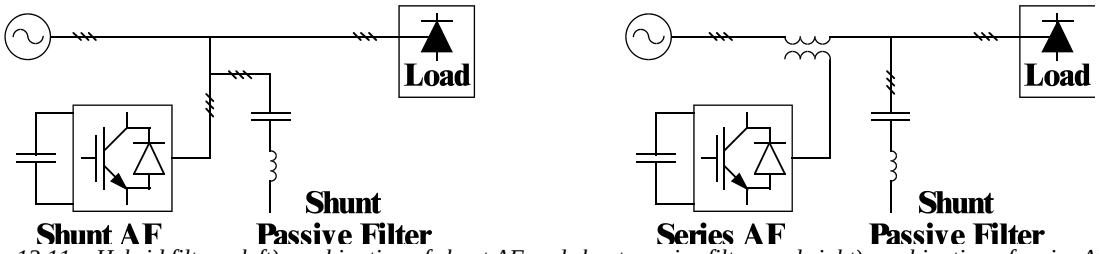


Figure 13.11 Hybrid filters; left) combination of shunt AF and shunt passive filter, and right) combination of series AF and shunt passive filter.

The optimal solution to harmonic reduction in distribution networks in terms of performance characteristics consists of the combination of a shunt and a series AF with a common DC-link, see Figure 13.12. This combination is referred to as the unified power quality conditioner (UPQC) and has the same configuration as the UPFC.

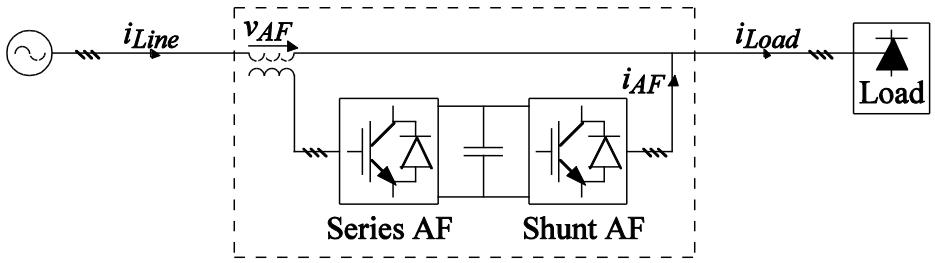


Figure 13.12 Configuration of the unified power quality conditioner (UPQC).

The UPQC for distribution networks differ from the UPFC for transmission networks in terms of operation, purpose and control strategy. The conditioning functions of the UPQC are shared by the series and shunt AFs. The series AF

performs harmonic isolation between supply and load, voltage regulation and voltage flicker/imbalance compensation. The shunt AF performs harmonic current filtering and negative sequence balancing as well as regulation of the DC-link voltage shared by the AFs. The DC-link voltage regulation is obtained by balancing the active power from both the VSCs including all losses in the passive as well as the active components.

Current control in a Shunt Active Filter

The current control system can be made as a "Vector Control" system or a "DCC"-system, and can be directly copied from the control equations for the generic three phase load described in chapter 4.7. In this case the "Vector Control"-algorithm with prediction, to be used with a "slow" computer, is selected - see equation (13.10)

$$\vec{u}_1^{\,*}(k)=\left(\frac{L}{T_s}+\frac{R}{2}\right)\cdot\left(\left(\vec{i}_1^{\,*}(n)-\hat{\vec{i}}_1(n)\right)+\frac{T_s}{\left(\frac{L}{R}+\frac{T_s}{2}\right)}\cdot\sum_{n=0}^{n=k-1}\left(\vec{i}_1^{\,*}(n)-\hat{\vec{i}}_1(n)\right)\right)+\hat{\vec{e}}_{cp}(k) \qquad (13.10)$$

A problem that arises with vector orientation in this particular application is that the voltage $\vec{e}_{cp}$ is not a perfect three-phase system, but contains harmonics that arise from a voltage division over the inductances between the Line Side Converter output voltage $\vec{u}_1$ and the grid voltage $\vec{e}_g$. The vector orientation is aligned with the integral of the voltage in the connection point, $\vec{e}_{cp}$, and the integration usually attenuates the harmonics well. Another problem occurs if the integration of $\vec{e}_{cp}$ is done digitally, since sampling of the voltage is difficult given the high frequency ripple content. Even worse is the situation when the grid voltage is unbalanced, since the vector orientation is supposed to be made with the positive sequence of the voltage $\vec{e}_{cp}$. In those cases more advanced filtering methods must be adopted, to create a good estimate of the positive sequence of the integral of the fundamental voltage in the connection point. It should be noted that this is fully possible. The positive sequence fundamental voltage vector can be detected even in the presence of a 2-phase short circuit of phase to zero in the connection point. Methods for doing this are beyond the scope of this material to describe.

DC link voltage control system

The DC link voltage must be controlled with the help of the active power flow from the grid. When supplying active power to the DC link the converter works like a rectifier. Compared to a pure diode rectifier, the Active Filter Line Side

Converter do switch much faster, usually several kHz, as compared to the three phase diode rectifier that switches at 300 Hz with a 50 Hz grid. The high switching frequency can be regarded as a more frequent update of the charge in the DC link capacitor, and allows a smaller capacitor to be used than with the pure diode rectifier, with maintained DC link voltage ripple.

If this possibility is utilized in extreme, the DC link capacitor can be made very small, typically 10...100 times smaller than the capacitor used with a diode rectifier in the case where a load side converter is present. The draw back with having a very small DC link capacitor is that the DC link ripple becomes more sensitive to undesired variations in the active power flow. As an example; - Assume that an error has caused a loss of active power flow control causing an high active power flow into the DC link. Then the DC link voltage will rise at a rate that is as much faster as the capacitor is smaller. If the rise time is to fast protective control functions may not be able to stop the energy flow and the DC link voltage will be high enough to destroy the power semiconductors on the same DC link.

The DC link voltage control system can be structured as in Figure 13.13. It is here assumed that the DC link current from the converter is updated every sampling interval (T_s), just like the AC side currents are.

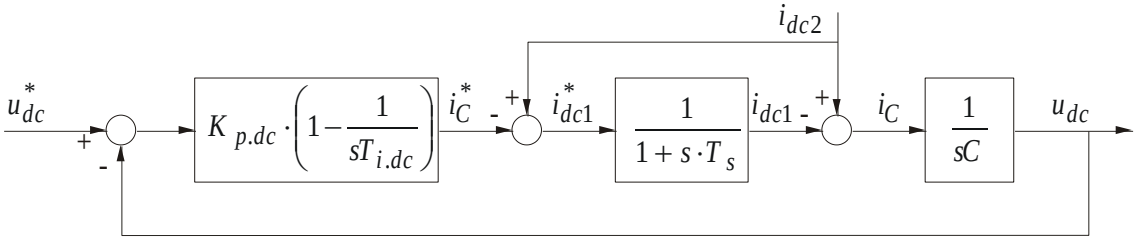


Figure 13.13 DC link voltage control system

Note that the structure of this control system is identical to the speed control system structure where "symmetric optimum" is used for parameter selection. Thus, the same methods for parameter selection can be used. Consequently, the parameters of the DC link voltage controller are selected according to equations (13.11) to (13.13).

$$T_{i.dc} = a^2 \cdot T_s, \text{ where} a > 1 \tag{13.11}$$

$$K_{p.dc} = \frac{a \cdot C}{T_{i.dc}} \tag{13.12}$$

$$\zeta = \frac{a-1}{2} \tag{13.13}$$

The DC link voltage controller depicted in Figure 13.13 gives the reference value of the DC link current entering the Front End Converter. This reference value has to be converted to a reference in the (*d*,*q*) reference frame. This can be derived by using the power equivalence of the two sides of the converter, according to equation (13.14).

$$p(t) = Ri_{1d}^2 + Ri_{1q}^2 + L\frac{di_{1d}}{dt}i_{1d} + L\frac{di_{1q}}{dt}i_{1q} + e_{cp.q}i_{1q} = u_{dc}\cdot i_{dc1} \approx e_{cp.q}i_{1q} \;\Rightarrow$$
$$i_{dc1} = \frac{e_{cp.q}}{u_{dc}}\cdot i_{1q} \qquad (13.14)$$

As long as the DC link voltage control dynamics is relatively slow, then the approximation in equation (13.14) is acceptable. It should be noted that if extreme DC link voltage control performance is desired, then a controller design in the frequency domain is not the way to go, but rather a time domain analysis. Such a design is beyond the scope of this material. The integrating ability of the controller designed with "symmetric optimum" compensates for neglecting the resistive terms in equation (13.14) and the non-optimal dynamics makes the inductive terms less significant. Thus, as will be seen in simulations and experiments, the approximation of equation (13.14) is acceptable.

Figure 13.13 and Figure 13.14 together form the DC link voltage controller according to equation (13.15).

$$i_{1q}^* = \frac{u_{dc}}{e_{cp}}\left(i_{dc2} - K_{p.dc}\cdot\left(1 - \frac{1}{sT_{i.dc}}\right)\cdot\left(u_{dc}^* - u_{dc}\right)\right) \qquad (13.15)$$

In Figure 13.14 an example is given of the DC link voltage control dynamics with a system with the following parameters:

```
>> L=0.01;
>> R=1;
>> Ts=0.0005;
>> Tidc=9*Ts;
>> Kpdc=3*Cdc/Tidc;
>> Cdc=1e-4;
```

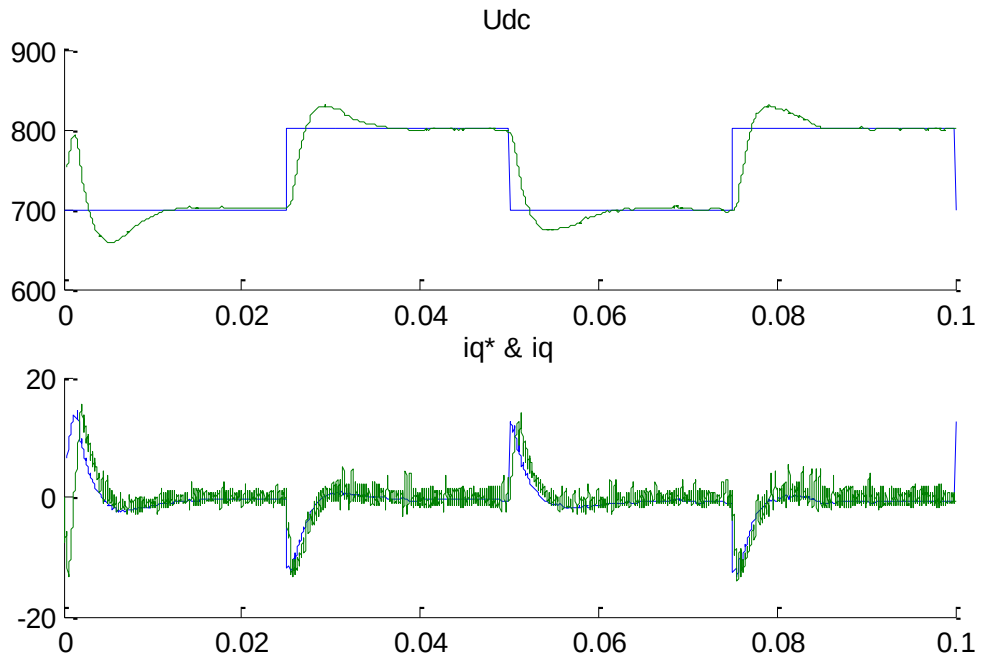


Figure 13.14 DC link voltage with controller according to equation (13.15)

Note the response time of around 10 ms. This DC link voltage controller will now be used in the Active Filtering application described in the following section.

Active filter control

In this section specific properties concerning shunt active filter control are discussed. The operation principle of a shunt AF for grid conditioning purposes, such as reactive power compensation, harmonic filtering and load balancing, is generally discussed in the previous section. It is stated that the basic operation principle of a shunt AF is to inject the undesired components of the load current, in order to cancel their presence in the line current. This is illustrated in Figure 13.15. It is clear that the performance of the shunt AF is determined by the amount of reduction of these undesired components in the line current.

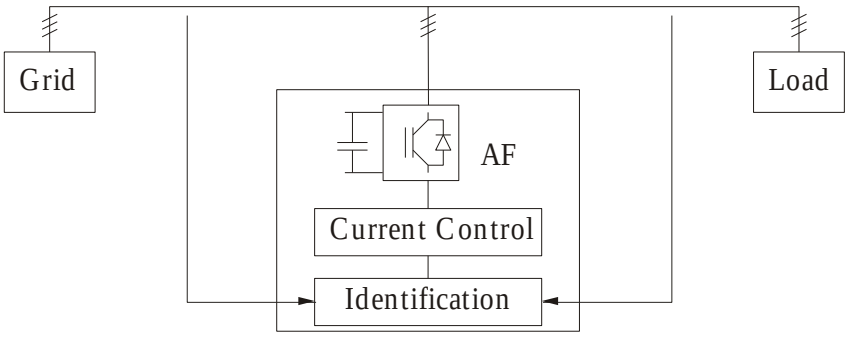


Figure 13.15 Shunt active filter operating principle.

The function of a shunt AF can be divided into two parts. First, the currents that should be compensated for have to be identified or detected. This can be performed either in the frequency domain or in the time domain. Furthermore, the detection can be based either on the load current, the supply (line) current, or the grid voltage. Detection based on the load current, which is used here, is suitable for a shunt AF installed in the vicinity of the load to be conditioned.

The second part consists of the active filter current controller, which should ensure good agreement between the injected currents and the undesired components of the load current. Basically, the filtering performance of a shunt AF is determined by the current control part, at least in steady state operation. However, during transient operation the filtering performance is also affected by the identification part.

13.4 Identification based on the load current

Various methods for identification of the undesired load current components are proposed in the literature. In this context, however, only two of these methods are discussed.

Fast fourier transform

The fast Fourier transform (FFT) method calculates the magnitude and phase of the load current frequency components. Then the component corresponding to the fundamental active current is removed. Finally, the active filter current reference is obtained from the inverse FFT of the remaining frequency components. The advantage of the FFT method is that the magnitudes of the frequency components are known. Hence by manipulation of the magnitudes, overloading of the shunt AF can be prevented. Furthermore, selective conditioning is made possible which is useful in some applications, where the focus is to reduce some specific component of the load current. However, the calculations involved are cumbersome and also the lack of information regarding the sequences, i.e. positive- or negative-sequences, of the conditioning components makes the FFT method less practicable.

Instantaneous currents in synchronous coordinates

This approach for identifying the undesired components of the load current considers the instantaneous currents represented in the synchronously rotating reference *dq*-frame oriented to the integral of the grid voltage vector. The use of synchronous coordinates is convenient, since the line side current controller already utilizes the synchronous reference *dq*-frame.

The load current vector in the synchronous reference *dq*-frame is derived from the stationary reference $\alpha\beta$-frame by subtraction of the corresponding angle difference θ according to

$$\vec{i}_{Load}^{dq} = \vec{i}_{Load}^{\alpha\beta} \cdot e^{-j\theta} \tag{13.16}$$

The angle θ used in the vector transformation is given by the argument of the grid voltage integral, that is

$$\theta = \omega_1 t - \frac{\pi}{2}$$

The representation of the load current vector of in synchronous coordinates is then given by

$$\vec{i}_{Load}^{dq} = \left(I_1 e^{-j\phi_1} + I_{-1} e^{-j2\omega_1 t} + I_{-5} e^{-j6\omega_1 t} + I_7 e^{j6\omega_1 t} \right) \cdot e^{j\frac{\pi}{2}}$$

The DC component of the load current vector in synchronous coordinates is written

$$\vec{i}_{Load,DC}^{dq} = I_1 \cdot \left(\sin\phi_1 + j\cos\phi_1 \right)$$

where the real and imaginary parts correspond to the reactive and active current components, respectively. The AC component of the load current vector corresponds to the negative-sequence current component and the current harmonics, that is

$$\vec{i}_{Load,AC}^{dq} = j \cdot \left(I_{-1} e^{-j2\omega_1 t} + I_{-5} e^{-j6\omega_1 t} + I_7 e^{j6\omega_1 t} \right)$$

The instantaneous currents represented in synchronous coordinates provide complete information of the different load current components. The compensation current should correspond to the fundamental reactive current, i.e. the real part of the DC component, and the AC component

$$\vec{i}_{Comp}^{dq} = \Re\mathrm{e}\left(\vec{i}_{Load,DC}^{dq} \right) + \vec{i}_{Load,AC}^{dq} \tag{13.17}$$

From the discussion in the previous sections, identification by the instantaneous currents represented in the synchronously rotating reference *dq*-frame is selected. The choice is motivated by the convenient relationship with the line current controller. Table 13.3 gives the correspondence between the conditioning components and their representation in the synchronous reference *dq*-frame.

Table 13.3 Representation of the current components in the synchronously rotating reference dq-frame.

Component	Sequence	Frequency	Vector
active	positive	DC	ji_q
reactive	pos.	DC	i_d
asymmetry	negative	2f1	$I_{-1}e^{-j2\omega_1 t}$
5th harmonic	neg.	6f1	$I_{-5}e^{-j6\omega_1 t}$
7th harmonic	pos.	6f1	$I_{7}e^{j6\omega_1 t}$
11th harmonic	neg.	12f1	$I_{-11}e^{-j12\omega_1 t}$
13th harmonic	pos.	12f1	$I_{13}e^{j12\omega_1 t}$
…	…	…	…

13.5 Example with active filtering of a diode rectifier current

In this section, an example is given with active filtering of a load consisting of a six-pulse diode rectifier feeding a resistive/inductive load. The current drawn from the grid to this load contains several harmonics and the purpose of the Active filter is to compensate all components apart from the active power.

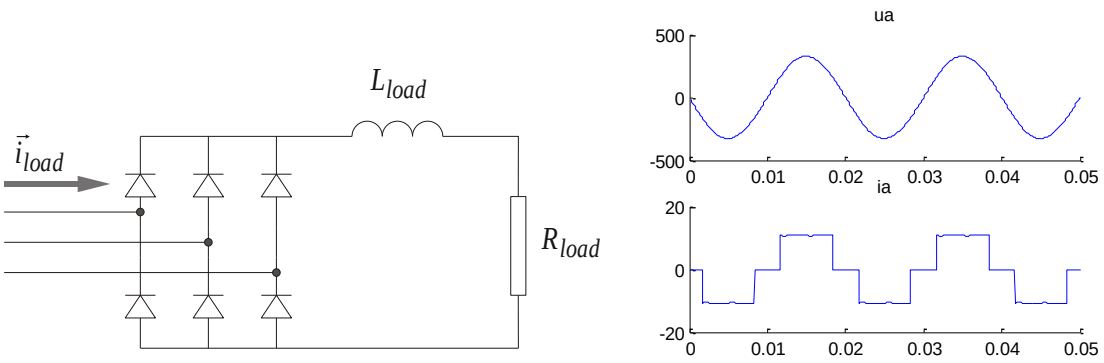


Figure 13.16 Diode rectifier with phase voltage and current

To accomplish the filtering task, the phase currents to the load are measured, converted to vector representation and transformed to the rotating reference frame (*d*,*q*). The *d*-axis component is directly used as a reference for the Active Filter currents, but the *q*-axis component is high pass filtered and then used as a reference for the Active filter *q*-axis current, since the Active filter is not supposed to provide any average active power to the load. A first order low pass filter is used, and the time constant is selected rather long, corresponding to a

low cut-off frequency, typically 10..30 Hz. The same filter time constant is also used for the DC link voltage controller, but as a low pass filter instead. Thus, the low frequency components of the *q*-axis current reference are used for DC link voltage control, and the high frequency components are used for Active Filter function.

$$i^*_{d,ActiveFilter} = i_{d,load}$$
$$i^*_{q,ActiveFilter} = i_{q,load} \cdot \frac{s \cdot T_f}{1 + s \cdot T_f} + \frac{u_{dc}}{e_{cp}} \left(i_{dc2} - K_{p.dc} \cdot \left(1 - \frac{1}{sT_{i.dc}} \right) \cdot \left(u^*_{dc} - u_{dc} \right) \right) \cdot \frac{1}{1 + s \cdot T_f} \quad (13.18)$$

Note that when the low pass filter is introduced in the DC link voltage controller, the filter time constant T_f must be used for selection of the DC link voltage control parameters, just like the filtering time constant for speed sensor filter is used when designing a speed controller in Chapter **Error! Reference source not found.**.

As an example, Active Filtering of the load current is made on a system with the following specifications:

```
>> L=0.01;
>> R=1;
>> Ts=0.00005;
>> Rload=50;
>> Lload=0.1;
>> Tf=10e-3;
>> Tidc=9*Tf;
>> Kpdc=3*Cdc/Tidc;
```

Figure 13.17 shows the load current, Active Filter current and the grid current as functions of time. Figure 13.18 and Figure 13.19 shows the spectral content of the load and grid currents.

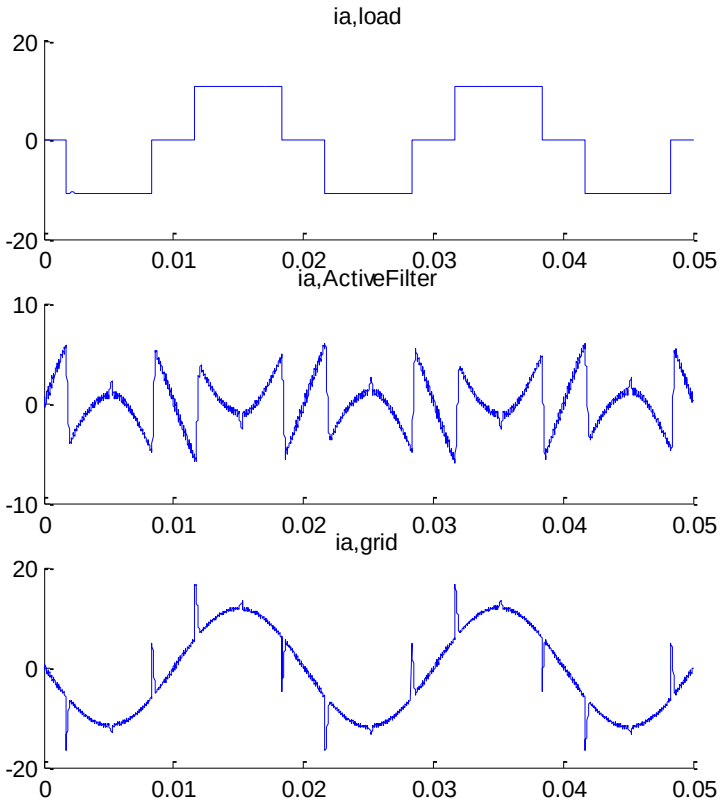


Figure 13.17 Load current, Active Filter current and total current drawn from the grid.

The glitches in the total grid current of Figure 13.17 cannot totally be avoided, but the spectral content of the current is significantly reduced, see Figure 13.18 and Figure 13.19.

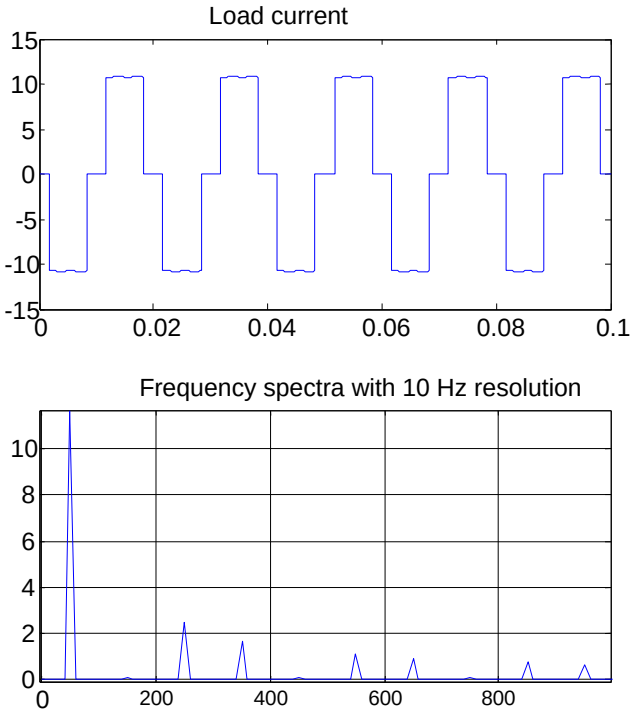


Figure 13.18 Frequency spectra of the load current

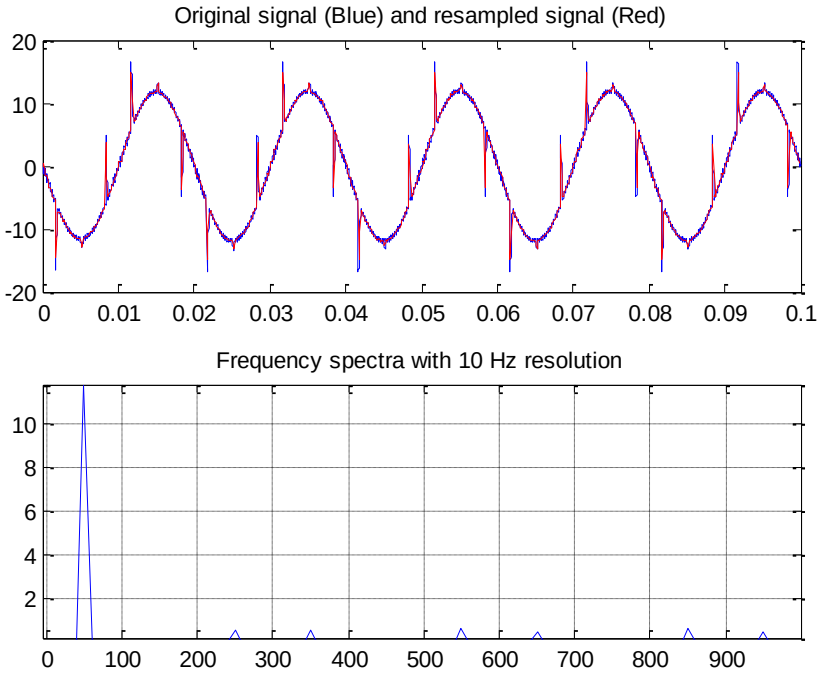


Figure 13.19 Frequency spectra of the grid current.

14 References

[1] B. Jayant Baliga, ”Modern Power Devices”, *Wiley*, New York, USA, 1987.

[2] M. Bojrup, ”Advanced Control of Active Filters in a Battery Charger Application”, *Licentiate’s Thesis, Department of Industrial Electrical Engineering and Automation, Lund Institute of Technology*, Lund, Sweden, Nov. 1999.

[3] Å. Ekström, “High Power Electronics – HVDC and SVC”, *EKC/KTH*, Stockholm, Sweden, June 1990.

[4] Å. Ekström et al, “High Power Electronics in Transmission and Distribution Systems”, Course binders A and B*, ABB T&D University*, Ludvika, Sweden, January/March 2001.

[5] Evox Rifa “Film Capacitors 1997-1998”, 1997.

[6] Evox Rifa “Electrolytic Capacitors 1997-1998”, 1997.

[7] J. Holtz, “Pulsewidth Modulation for Electronic Power Conversion”, *Proc. of the IEEE*, vol. 82, no. 8, pp. 1194-1214, Aug. 1994.

[8] P. Karlsson, "Quasi Resonant DC Link Converters - Analysis and Design for a Battery Charger Application", *Licentiate’s thesis, Department of Industrial Electrical Engineering and Automation, Lund Institute of Technology,* Lund, Sweden, November 1999.

[9] V. Khanna, “Insulated Gate Bipolar Transistors (IGBT)”, *Wiley*, 2003.

[10] A. Kurnia, O.H. Stielau, G. Venkataramanan and D.M. Divan, ”Loss Mechanisms in IGBT’s Under Zero Voltage Switching”, *IEEE-PESC Conf. Rec*., Toledo, Spain, October 1992, pp. 1011-1017.

[11] Colonel W.T. McLyman, ”Transformer and Inductor Design Handbook”, *Marcel-Dekker*, New York, USA, 1978.

[12] A. Mertens, ”Design of a 20 kVA Resonant DC Link IGBT Inverter on the Base of Experimental Device Evaluation”, *EPE ‘91 Conf. Rec*., Firenze, Italy, Sept. 1991, vol. 4, pp. 172-177.

[13] Micrometals “Power Conversion & Line Filter Applications”, 1998.

[14] N. Mohan, T. Undeland and W.P. Robbins, ”Power Electronics: Converters, Applications, and Design”, 2nd edition, *John Wiley and Sons*, Inc., New York, NY, USA, 1995.

[15] National-Arnold Magnetics, “Value Leader in Tape Wound Magnetics”, Product Catalog, 1995.

[16] T. Nilsson, ”Understanding the IGBT properties in practical operation”, Licentiate’s Thesis, Department of Electric Power Engineering division of Electrical Machines and Power Electronics, Chalmers University of Technology, Gothenburg, Sweden, Dec. 1994.

[17] H.W. Ott, ”Noise Reduction Techniques in Electronic Systems”, 2nd ed., *Wiley-Interscience*, New York, USA, 1988.

[18] A. Petterteig and T. Rogne, ”IGBT Turn-Off Losses - In Hard Switching and with a Capacitive Snubber”, *EPE ‘91 Conf. Rec*., Firenze, Italy, Sept. 1991, vol. 0, pp. 203-208.

[19] R.S. Ramshaw, Power Electronics Semiconductor Switches, *Chapman & Hall*, 1993.

[20] H. Skarrie, ”Design of Iron Powder Cores”, Licentiate’s Thesis, Department of Industrial Electrical Engineering and Automation, Lund Institute of Technology, Lund, Sweden, July 2001.

[21] B.G. Streetman, ”Solid State Electronic Devices”, 3rd ed., *Prentice-Hall International*, Englewood Cliffs, USA, 1990.

[22] S.M. Sze, ”Semiconductor Devices; Physics and Technology”, *Wiley*, New York, USA, 1985.

[23] K. Thorborg, ”Power Electronics”, *Prentice-Hall International*, London, UK, 1988.

[24] B.W. Williams, ”Power Electronics; Devices, Drivers, Applications and Passive Components”, *Macmillan*, New York, USA, 1992.

[25] K.J. Åström and B. Wittenmark, ”Computer Controlled Systems; Theory and Design”, 2nd edition, *Prentice Hall, Inc.,*, Englewood Cliffs, NJ, USA, 1990.

15 Vectors in 3 Phase Systems

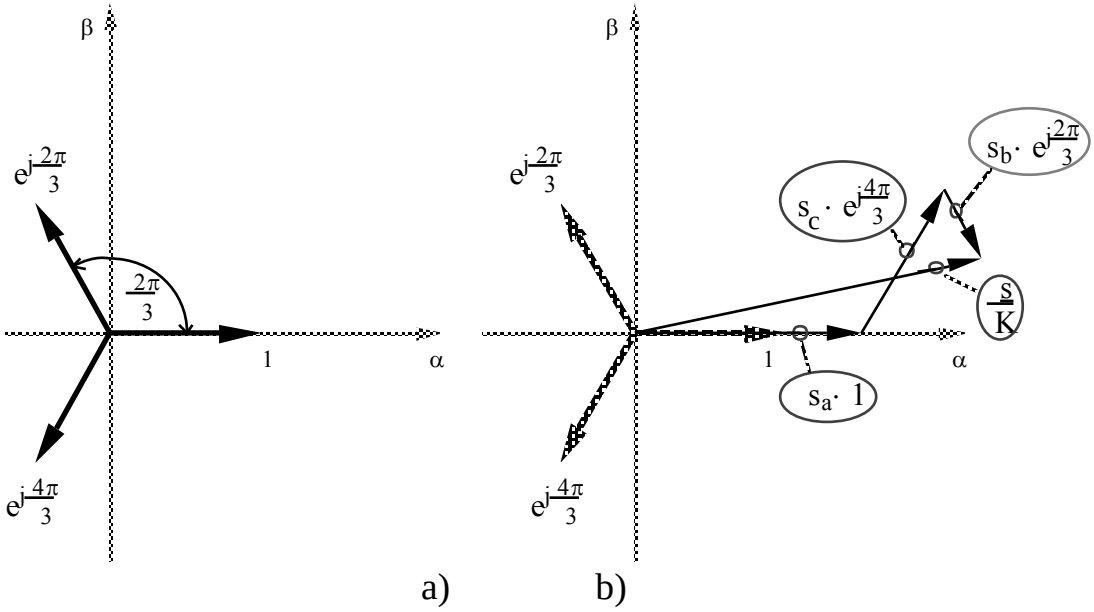


Figure 15.1 a) Reference directions with unity vectors and b) The three phase vector with its components

The impact of three phase voltages, currents and flux linkages can preferably be expressed in vector form. This is accomplished by assigning a reference direction to each phase in a complex plane [α,β]. The three phase quantity reference directions are symmetrically distributed over a turn, see Figure 15.1, and represent e.g. the main magnetic axis of a specific stator winding in a three phase winding.

The reference phase is always assigned along the positive real axis. Assume a three phase system with the phase indexes *a*, *b*, *c*. Let the three phase voltages (u_a, u_b and u_c), currents (i_a, i_b and i_c) and flux linkages (ψ_a, ψ_b and ψ_c) be represented by the general quantities s_a, s_b and s_c. The instantaneous value of the phase quantity s_a is first scaled with the scaling factor *k* (the meaning of which is discussed below) and contributes to the resulting vector *s* in the direction **1**. The other two phase quantities contribute in the same way, but along a different reference direction, e.g s_b in the direction $e^{j_{2\pi/3}}$. The resulting vector $\vec{s}$ can be written as:

$$\vec{s} = s_\alpha + js_\beta = k\left[s_a + s_b e^{j\frac{2\pi}{3}} + s_c e^{j\frac{4\pi}{3}} \right] \qquad (15.1)$$

The design of this vector $\vec{s}$ may seem theoretical and ”taken out of thin air”. It can be given a more physical interpretation if we consider a two-pole three phase winding illustrated with three symmetrically distributed phase windings according to Figure 15.2. The three phase windings are physically (in the real machine) arranged with their main magnetic axis aligned along the same directions as the reference directions of Figure 15.1.

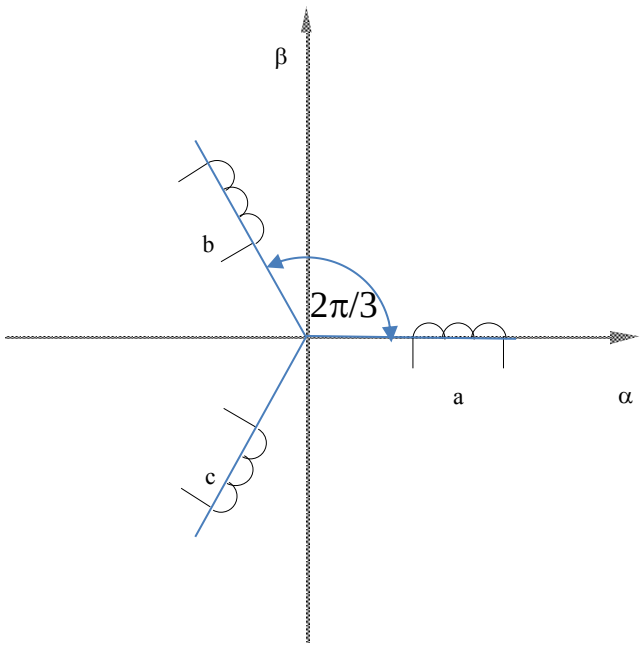


Figure 15.2 Simplified illustration of the winding arrangement of a two-pole three phase ac machine stator.

By replacing $s_a=i_a$, $s_b=i_b$ and $s_c=i_c$ in equation (15.1) and assigning the three phase current specific values it is possible to calculate a vector $\vec{i}$ pointing in any direction in the (α,β)-plane. The three phase currents, each flowing in a physical winding with different main magnetic axes, COULD be replaced with EITHER two phase currents (i_α and i_β) flowing in two alternative phase windings with their main magnetic axes along the α and β direction OR with a single phase current $|\vec{i}|$ in a single phase winding with the main magnetic axis aligned with the instantaneous value of the current vector.

Winding a machine with two orthogonal phase windings (α and β) is not unrealistic and is used both in some special machine but maybe mostly in sensors, like in a resolver (angle sensor).

Winding a machine with just one winding is not meaningful if the ide is that the current vector should rotate. The winding is by nature fixed to the stator. This is the main reason why the three phase winding is the most

appealing – with only three stationary phase windings carrying a three phase ac current it is possible to create a rotating current vector.

The same way of thinking when defining vectors can be applied to the magnetic flux and flux linkage related to each three phase winding. The flux vector is maybe the easiest to comprehend since magnetic flux lines do have a main direction and superposition of three phase contributions into a common resulting flux is imaginable. In the case with currents it is not as easy to comprehend since the phase currents do not have a specific direction (they follow the conductor that is a coil) but is maybe easier to imagine considering that the currents are assigned a direction equal to the main magnetic axis of the winding the current is flowing in. Voltages is maybe the most difficult physical quantity to comprehend that it can be converted into a vector, but a similar thinking as with currents can be used – the voltages do affect the windings and thus contribute along a specific main magnetic direction for each winding.

The selection of a value to the constant k in equation (15.1) is free. The value of k affects the magnitude of the two-phase quantities s_α and s_β in relation to the three phase quantities, and there are mainly two different values of k that are useful each in a different way, called the **power invariant** and **amplitude invariant** transformation.

Power invariance

It is often an advantage if the power expressed in the two phase system can be expressed in the same way as in the three phase system, i.e:

$$p(t) = u_\alpha(t)i_\alpha(t) + u_\beta(t)i_\beta(t) = u_a(t)i_a(t) + u_b(t)i_b(t) + u_c(t)i_c(t) \quad (15.2)$$

Assume that the sum of the three phase quantities is zero i.e:

$$u_a(t)+u_b(t)+u_c(t) = 0 \quad (15.3)$$

$$i_a(t)+i_b(t)+i_c(t) = 0 \quad (15.4)$$

If equation (15.1) is combined with (15.3) and (15.4) it is shown that:

$$u_\alpha = k\frac{3}{2}u_a \qquad i_\alpha = k\frac{3}{2}i_a$$

$$u_\beta = k\frac{\sqrt{3}}{2}(u_b - u_c) \qquad i_\beta = k\frac{\sqrt{3}}{2}(i_b - i_c) \quad (15.5)$$

Inserting (15.5) in (15.2), the resulting power in the three phase system can be expressed as:

$$p(t) = k^2 \frac{3}{2}[u_a(t)i_a(t) + u_b(t)i_b(t) + u_c(t)i_c(t)]$$

The transformation becomes power invariant if:

$$k = \sqrt{\frac{2}{3}}$$

With this value of k inserted in equation (15.1) the power invariant expression for the general vector is:

$$\vec{s} = s_\alpha + js_\beta = \sqrt{\frac{2}{3}}\left[s_a + s_b e^{j\frac{2\pi}{3}} + s_c e^{j\frac{4\pi}{3}} \right] = \sqrt{\frac{3}{2}} s_a + j\frac{1}{\sqrt{2}}(s_b - s_c) \quad (15.6)$$

With equation (15.6) two phase quantities can be calculated out of three phase quantities via power invariance. If equations (15.3) and (15.4) are included, three phase quantities can also be calculated out of two phase quantities. Thereby power invariant three phase to two phase transformation and the opposite can be summarized as:

Power Invariant **Three phase to Two Phase** Transformation:

$$\begin{cases} s_\alpha = \sqrt{\frac{3}{2}} \cdot s_a \\ s_\beta = \frac{1}{\sqrt{2}} \cdot (s_b - s_c) \end{cases}$$

Power Invariant **Two phase to Three Phase** Transformation:

$$\begin{cases} s_a = \sqrt{\frac{2}{3}} s_\alpha \\ s_b = -\frac{1}{\sqrt{6}} \cdot s_\alpha + \frac{1}{\sqrt{2}} \cdot s_\beta \\ s_b = -\frac{1}{\sqrt{6}} \cdot s_\alpha - \frac{1}{\sqrt{2}} \cdot s_\beta \end{cases}$$

Amplitude Invariance

Sometimes when the transfer from three phase to two phase quantities is made in analog electronics it is convenient to maintain the signal amplitude to protect the signal-to-noise ratio. The value of *k* must be:

$$k = \frac{2}{3}$$

With this value of *k* inserted in equation (15.1) the amplitude invariant expression for the general vector is:

$$\vec{s} = s_\alpha + j s_\beta = \frac{2}{3}\left[s_a + s_b e^{j\frac{2\pi}{3}} + s_c e^{j\frac{4\pi}{3}} \right] = s_a + j\frac{1}{\sqrt{3}}(s_b - s_c) \tag{15.7}$$

With equation (15.7) two phase quantities can be calculated out of three phase quantities via power invariance. If equations (15.3) and (15.4) are included, three phase quantities can also be calculated out of two phase

quantities. Thereby amplitude invariant three phase to two phase transformation and the opposite can be summarized as:

Amplitude Invariant **Three phase to Two Phase** Transformation:

$$\begin{cases} s_\alpha = s_a \\ s_\beta = \dfrac{1}{\sqrt{3}} \cdot (s_b - s_c) \end{cases}$$

Amplitude Invariant **Two phase to Three Phase** Transformation:

$$\begin{cases} s_a = s_\alpha \\ s_b = -\dfrac{1}{2} \cdot s_\alpha + \dfrac{\sqrt{3}}{2} \cdot s_\beta \\ s_b = -\dfrac{1}{2} \cdot s_\alpha - \dfrac{\sqrt{3}}{2} \cdot s_\beta \end{cases}$$

With amplitude invariant transformation, the power expression using two phase quantities is altered to:

$$p(t) = u_a(t)i_a(t) + u_b(t)i_b(t) + u_c(t)i_c(t) = \frac{3}{2}\left[u_\alpha(t)i_\alpha(t) + u_\beta(t)i_\beta(t)\right] \quad (15.8)$$

Phasors

In AC theory the notion ”phasors” is used, and has strong similarities with vectors as defined above. There is a relation between these phasors and vectors and it can be shown that phasors are a special case of the more general vector in sinusoidal stationary conditions.

Assume that we have a three phase system, with the voltages and currents according to the following:

$$u_a = \sqrt{2}U\cos\omega t \qquad i_a = \sqrt{2}I\cos(\omega t - \varphi)$$

$$u_b = \sqrt{2}U\cos\left(\omega t - \frac{2\pi}{3}\right) \qquad i_b = \sqrt{2}I\cos\left(\omega t - \frac{2\pi}{3} - \varphi\right)$$

$$u_c = \sqrt{2}U\cos\left(\omega t - \frac{4\pi}{3}\right) \qquad i_c = \sqrt{2}I\cos\left(\omega t - \frac{4\pi}{3} - \varphi\right)$$

If these instantaneous values are inserted into the expression for the general vector $\vec{s}$ then we get the voltage and current vectors:

$$\vec{u}^{\alpha\beta} = \frac{2}{3}\left[u_a + u_b e^{j\frac{2\pi}{3}} + u_c e^{j\frac{4\pi}{3}}\right] =$$

$$= \sqrt{2}U\left\{\cos\omega t + j\frac{1}{\sqrt{3}}\left[\cos\left(\omega t - \frac{2\pi}{3}\right) - \cos\left(\omega t - \frac{4\pi}{3}\right)\right]\right\} = \quad (15.9)$$

$$= \sqrt{2}U\left\{\cos\omega t + j\sin\omega t\right\} =$$

$$= \sqrt{2}Ue^{j\omega t}$$

$$\vec{i}^{\alpha\beta} = \frac{2}{3}\left[i_a + i_b e^{j\frac{2\pi}{3}} + i_c e^{j\frac{4\pi}{3}}\right] =$$

$$= \sqrt{2}I\left\{\cos(\omega t - \varphi) + j\frac{1}{\sqrt{3}}\left[\cos\left(\omega t - \frac{2\pi}{3} - \varphi\right) - \cos\left(\omega t - \frac{4\pi}{3} - \varphi\right)\right]\right\} = \quad (15.10)$$

$$= \sqrt{2}I\left\{\cos(\omega t - \varphi) + j\sin(\omega t - \varphi)\right\} =$$

$$= \sqrt{2}Ie^{j(\omega t - \varphi)}$$

These expressions show that voltage and current vectors, with a sinusoidal symmetrical three phase system with constant amplitude and frequency, rotates with an angular speed ω as seen from the reference frame (α,β)**.**

Coordinate transformation

A vector $\vec{s}$ in e.g. the [α,β]-reference frame can be expressed in any reference frame [x,y] via coordinate transformation.

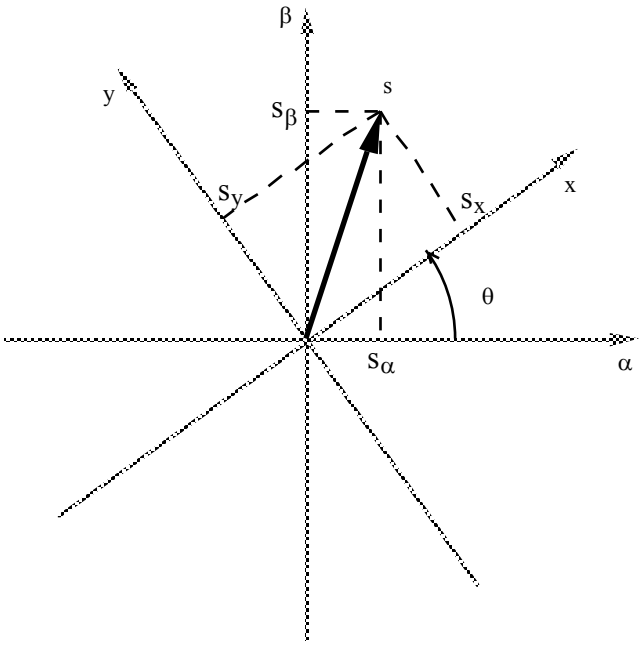


Figure 15.3. The vector $\vec{s}$ *in different reference frames*

Assume and angular difference between the both reference frames θ. The vector $\vec{s}$ can be expressed in both [α,β] or [x,y] - coordinates, note by the superscript according to the following:

$$\begin{aligned}\vec{s}^{\,xy} &= s_x + js_y = \vec{s}^{\,\alpha\beta} e^{-j\theta} = \\ &= (s_\alpha + js_\beta)(\cos\theta - j\sin\theta) = \\ &= (s_\alpha \cos\theta + s_\beta \sin\theta) + j(s_\beta \cos\theta - s_\alpha \sin\theta)\end{aligned}$$

i.e.:

$$\begin{aligned}s_x &= s_\alpha \cos\theta + s_\beta \sin\theta \\ s_y &= s_\beta \cos\theta - s_\alpha \sin\theta\end{aligned} \qquad (15.11)$$

That also can be written in matrix format

$$\begin{bmatrix} s_x \\ s_y \end{bmatrix} = T \begin{bmatrix} s_\alpha \\ s_\beta \end{bmatrix}$$

with

$$T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

This transformation can of course also be inverted by changing the sign of the transformation angle:

$$\begin{aligned}\vec{s}^{\alpha\beta} &= s_\alpha + js_\beta = \vec{s}^{xy} e^{j\theta} = \\ &= (s_x + js_y)(\cos\theta + j\sin\theta) = \\ &= (s_x \cos\theta - s_y \sin\theta) + j(s_y \cos\theta + s_x \sin\theta)\end{aligned}$$

i.e.:

$$\begin{aligned}s_\alpha &= s_x \cos\theta - s_y \sin\theta \\ s_\beta &= s_x \sin\theta + s_y \cos\theta\end{aligned} \tag{15.12}$$

or:

$$\begin{bmatrix} s_\alpha \\ s_\beta \end{bmatrix} = T^{-1} \begin{bmatrix} s_x \\ s_y \end{bmatrix}$$