Single ended flyback

The term ‘single-ended’ just refers to a circuit using a single switch. There are versions of flybacks and forward converters that use multiple switches which we’ll cover later.

The basic circuit is shown in **Figure 7.3** where Q1 turns ON allowing current to ramp up in the primary. The secondary winding is phased such that while Q1 in ON, the voltage on D1 anode is negative so it does not conduct. In this situation, T1 just looks like an inductor as in our boost converter designs. Similarly, when Q1 turns OFF, the voltage on its drain ‘flies’ higher than the input rail. The transformer action of T1 now proportions the secondary voltage according to the turns ratio and the flyback voltage forward biasing D1 providing current for the output. Note that the output voltage is the turns ratio multiplied by the voltage that the primary is allowed to go to, the ‘flyback voltage’ not the input voltage. The same equations used in the boost converter generally apply but modified by the insertion of T1. Let’s go through a design to see the detail.

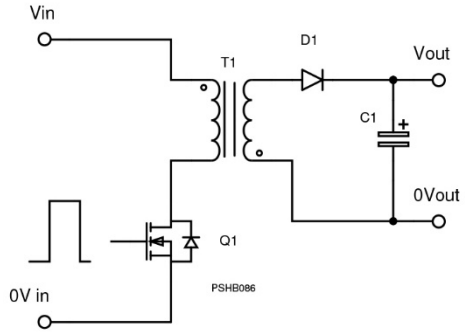


Figure 7.3 Single ended flyback converter outline

Let’s assume a 50W supply producing 13.8V at 3.6A with its input from a mains transformer rectified and filtered giving a voltage ranging from 30 to 50V with a nominal of 40V. We choose this range because higher voltages give lower currents and smaller smoothing capacitors but we don’t want to exceed 60V which is considered unsafe by safety agencies. Although the mains transformer provides safety isolation, we want another level of isolation in the flyback converter because the transformer is also feeding other circuitry and we want to separate ground systems for example.

As ever, we have to think what mode of operation we want, CCM, DCM or CRM. We’re going to use a fixed frequency controller so CRM which needs variable frequency is not appropriate. A converter operating in DCM at 50W will have high peak currents in the transformer, output diode and capacitor so CCM would be the normal choice. The converter will then cross over to DCM at lower loads and/or higher input voltages so we need to decide at which point. This is a question of judgement but a useful approach as described by Billings in his book ‘Switchmode Power Supply Handbook’ is to make the primary current waveform at full load and minimum input proportioned like **Figure 7.4** where the converter will cross to DCM at about 50% load. The average current during

the ON time is the sum of the peak value of the 'square' part of the waveform and the average of the 'peak' part which is half the peak value. With this waveform, the average is twice the 'square part'. This is a suggested starting point but if EMI is too high, you could cross to DCM at lower loads or if component stress is too high you could cross higher. A flyback in CCM is more difficult to stabilise due to an effect called the 'right half plane zero'. We'll look at this later in the section on loop compensation but stabilisation is possible and when the converter crosses into DCM it becomes less prone to instability anyway.

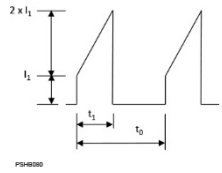


Figure 7.4 Recommended flyback waveform

Another assumption that needs to be made is the maximum ON-time duty cycle. If it is high, you have more adjustment range but then the OFF time is very short producing very high peak currents and voltages. As our input range is not so wide, let's fix the maximum to be just under 50%. Actually, over 50% in some control schemes produces an effect called sub-harmonic oscillation which can be controlled but is best avoided. A controller IC can be chosen that limits at 50% duty cycle, such as the UCC38C44 from TI.

We now have starting point, we want SW1 drain voltage waveform to look like **Figure 7.5** at maximum load and minimum input voltage. The duty cycle is 50% so by the rule of volt-second equality, the 'flyback' voltage must also be 30V. (I'm ignoring Q1 and resistive volt-drops). The output voltage is 13.8V, say 14.5V before D2 and as we've mentioned, the output voltage is the transform of the flyback voltage so the turns ratio of T1 must be 30/14.5 = 2.07

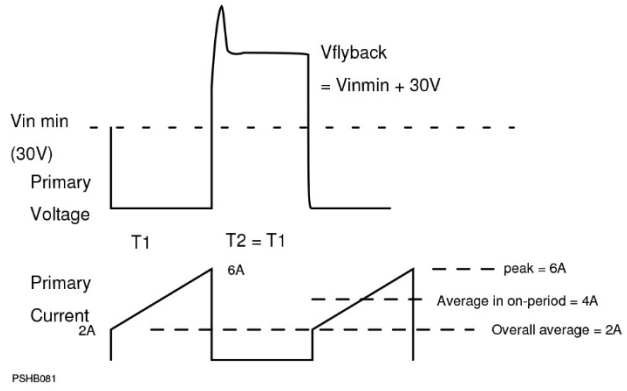


Figure 7.5 Flyback primary voltage and current waveforms - 50% duty cycle and Vin minimum

To determine the primary turns *n*, we need to look at the practicalities. We'll be using a ferrite core for low loss so from our magnetics section Equation 7:

$$n = \frac{ET_{on}}{A_e B}$$

Where *E* is applied voltage, T_{on} = maximum ON-time in microseconds, A_e is the effective cross section of the core in square millimetres and *B* is our maximum working flux density in teslas, say 0.1.

We shall see that when we select CCM there will also be flux density produced by the DC component of the transformer primary current. We'll put a figure on this later.

We can choose our switching frequency defining maximum ON-time at 50% duty cycle and a common figure is 100kHz. Much higher in a simple circuit gives high switching losses and much lower gives big magnetics so it's a good compromise. Some commercial power supply units for the amateur market allow external trimming of the frequency to avoid 'birdies' on a particular RF channel. You would limit this to a very small adjustment value of say one or two hundred Hz as the birdies would be high multiples of the switching frequency and would move away rapidly. Another reason to keep around 100kHz is that statutory limits for conducted emissions (FCC-0871, EN55022) generally start at 150kHz so keeping the fundamental emission below this figure is beneficial. So, 50% duty cycle at 100kHz is 5µs ON-time. We can try a common E30 core with A_e of 60mm^2 giving us:

n = 30 x 5/(60 x 0.1) = 25 turns which seems reasonable.

If we look at the primary current waveform at maximum load and V_{inmin} in Figure 7.5 we need to know the rms value to see what gauge of wire we need. The input power is about 60W so the average current in the input which all passes through the transformer must be 60W/30V = 2A. At 50% duty cycle, the average during the ON-time must therefore be 4A. Inspecting the waveform you can see that the average must be half way up the ramp current. We decided arbitrarily that we wanted the ramp height to be twice the 'pedestal' so the start of the pedestal must be at 2A and the finish at 6A, From Appendix 3, this waveform has an rms value of 2.93A.

For 5A per square millimeter rule of thumb we need about 0.6mm^2 wire which is about 0.8mm diameter wire or 20 AWG. The EE30 core bobbin has a winding area of about 88 mm^2 so allowing half for the primary, 25T of 20 AWG fits nicely.

We know we want a turns ratio of 2.07 which would give 12.08 turns for the secondary. Obviously, we want a round number so if it were 12 what would happen? This would transform to a flyback voltage of 14.5 x 25/12 = 30.2V which doesn't work as volt second equality would force the ON-time to be higher than 50% which we don't want and our controller IC can't reach.

Volt second equality 30 x T_{on} = 30.2 x T_{off}

so T_{on}/T_{off} = 1.006 = 50.2% duty cycle.

We therefore have to round up the turns to 13 so now the flyback voltage is 27.9V and from volt second equality:

30 x T_{on} = 27.9 x T_{off}

So T_{on}/T_{off} = 0.92 = 48.4% duty cycle maximum which is fine.

Turns ratio n = 25/13 = 1.92

Our 12-turn secondary is providing an average of 3.6A, the DC output, but the current waveform through the winding and the rectifier diode is far from DC. If we take the worst-case condition of maximum load at minimum input voltage, we know the transformer primary current waveform from Figure 7.5. As the current stops, peaking at 6A, the secondary current starts at a value which 6A multiplied by the transformer turns ratio which is 1.92 giving a peak current of 11.5A, reducing to 3.84A. Again, from Appendix 3, you can see that this gives an rms value of about 5.6A. For 5 A/mm^2 we would need about 1mm^2 wire which is around 1.12mm diameter or 18 AWG. This thickness of wire would be difficult to wind so multiple strands of thinner wire might be easier.

If we had let the converter be in DCM at 50% duty cycle, the primary peak current would be 8A, the secondary peak current over 15A and secondary rms over 6A. This illustrates the stress that can be caused in DCM. The output capacitor 'sinks' the 6A AC component of the transformer secondary current showing the typical case that capacitor ripple current in a DCM flyback is around 1.5 – 2 times the DC output current. At higher input voltages and lower duty cycles, the situation is even worse as the peak currents remain the same (power x cycle time = $½ Li^2$, which all haven't changed) but pulse width is now shorter giving higher rms values for the same average values.

We haven't talked about primary inductance yet. An easy way to get to it is to look at Figure 7.5 again. The ramp of current is 4A start to finish in about 5µs at 30V input so from $L = ET/i$:

L = 30 x 5µs/4 = 37µH

A typical E30 core in a typical 3C90 material has an A_L with no gap of 1900 so 25 turns would produce over a millihenry according to $A_L = nH/n^2$ so we need a gap. From Equation 9, the gap length l_g is:

$$l_g = \frac{\mu_0 \mu_r n^2 A_e 10^3}{L} \qquad (mm, mm^2, \mu H)$$

= 1.3mm total or 0.64mm in each leg of the E-core which is reasonable. See **Figure 7.6** for gap arrangements.

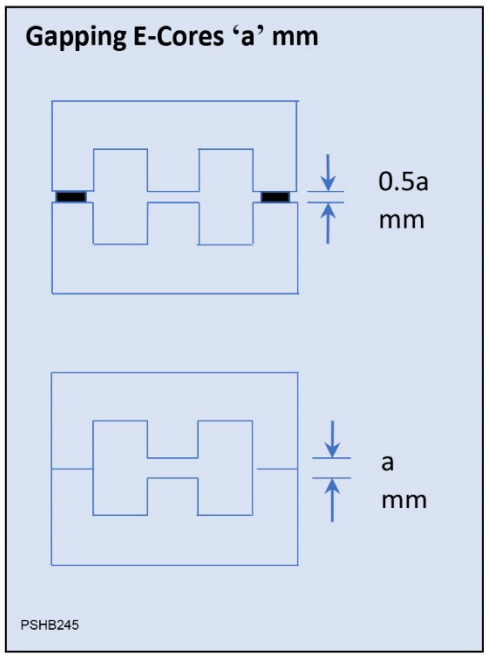


Figure 7.6 Gapping E-cores

We have just looked at the one condition of minimum input voltage and maximum load. At other input voltages while the converter is in CCM the duty cycle reduces according to:

$$D = \frac{V_{out} N}{V_{in} + V_{out} N} \qquad \textbf{Equation 10}$$

N is the turns ratio, primary to secondary, diode voltage drops are ignored. You can see that at 50V input the duty cycle becomes 0.36 giving an ON-time of 3.6µs or 0.4 and 4µs at 40V input.

In the analysis, assumptions were made about maximum duty cycle under particular conditions but what forces this to happen? How does the controller know to stop the ON-time at the right duty

cycle? We have feedback from the output to the controller through the error amplifier IC2 and the optocoupler. In a feedback loop if you force one condition to one of your assumptions then the loop settles according to your analysis. In this case, we monitor the output voltage to be what we want, 13.8V, and if lower we allow a longer ON-time to compensate. If higher, the ON-time is curtailed. The feedback is operating relatively slowly, typically taking a millisecond to react so many cycles of switching occur before the correction takes effect. This is why converters have a transient response time after a change to load or input voltage which affects output voltage. Because of this lag, if the input voltage suddenly changes to its maximum value from minimum say, the duty cycle stays at maximum until the correction occurs. In this case the transformer sees more volt-seconds transiently and hence higher flux density. For example, 50V with a transient maximum ON-time of 5µs would give a transient flux density *B* of:

B = 50 x 5µs/ 60mm^2 x 25 = 0.16 tesla.

This is still well within the limit of the material of about 0.35 tesla but is a reason why you would not run the core at its maximum flux density under steady state conditions.

I mentioned before that there is also a contribution to flux density from the DC component of primary current when in CM. If, as is the case, our air gap is defining the effective permeability of the core, then we can say from Equation 6 that:

$$B_{DC} = \frac{\mu_0 . N . i_{DC} . 1000}{gap} \quad \text{(mm, tesla, amps)}$$

i_{dc} is the constant part of the primary current which doesn't ramp which is the 'pedestal' current = 2A in Figure 7.5. Therefore:

B_{dc} = 4 x π x 10^{-7} x 25 x 2 x 10^3/1.3

= 0.043T.

This is added to our previously calculated 0.16T to give around 0.2T total which is still well below the saturation limit of about 0.35T.

Let's see when the circuit goes into DCM. Looking at our current waveform Figure 7.5 for maximum load we always want the same peak current because energy = 1/2Li^2 and we want the same energy each cycle = Pout x cycle time (not ON-time). DCM happens when the waveform just becomes triangular as in **Figure 7.7** with the same peak current 4A. From *E = iL/t* :

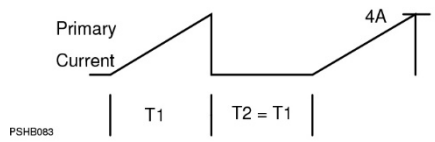


Figure 7.7 Flyback primary current on the edge of CCM/DCM at 50% duty cycle

$V_{in} = 6 \times 37 \times 10^{-6}/t_{on}$

And: $t_{on} = 10 \times 10^{-6} \times DC$

So: $V_{in} = (6 \times 37 \times 10^{-6})/(10 \times 10^{-6} \times DC) = 22.2/DC$ or $DC = 22.2/V_{in}$

Equation 10. still holds as we are still just in CM so:

$D = (14.5 \times 1.92)/(V_{in} + (14.5 \times 1.92)) = 27.84/(V_{in} + 27.84)$

So, combining and rearranging:

V_{in} = 109V meaning that the converter will not go into DCM in its normal input range at maximum load. What about load? At nominal input of say 40V, we know the duty cycle from before is 0.41 so t_{on} is 4.1µs.

At 40V input, ramp current $i = ET/L = V_{in} \times T_{on}/L$ = 4.43A

DCM starts when the 'pedestal' on the waveform just disappears with the peak current still 4.43A

From Energy = $\frac{1}{2}Li^2$ = Power (P) x cycle time (T) or:

$$P_{in} = \frac{i_{out}V_{out}}{efficiency} = \frac{Li_{pk}^{\ 2}}{2T} \qquad \textbf{Equation 11}$$

$i_{out} = 0.9 \times 37 \times 10^{-6} \times 4.3^2 /(14.5 \times 2 \times 10 \times 10^{-6}) = 2.12A$

So, the converter goes into DCM at just over 50% load at nominal input as intended assuming 90% efficiency.

The analysis can be done from different starting points and one way is to set a critical primary inductance L_{pcrit} for the onset of CM at a particular load. This inductance is given by:

$$L_{pcrit} = \frac{R_{out}N^2}{2.f_{sw}(\frac{V_{in}}{V_{in}+V_{out}N})^2}$$

N is the primary to secondary turns ratio.

Above L_{crit}, the converter is guaranteed to work in CM.

Power output

In Equation 11. It looks like L, the primary inductance is proportional to power available. This is not the case as if L increases, i_{peak} decreases (remember $i=Et/L$) and i_{peak} is squared in Equation 11 so for more power we actually want less inductance, everything else being equal.

Diode ratings

The output diode sees an average current equal to the DC output of 3.6A. As the diode has a nominally constant forward voltage drop, the power dissipated could be expected to be 0.6 x 3.6A for a silicon type. Because the peak value is 11.5A however, the diode is probably dropping much more voltage than you expect and therefore dissipating significantly more. For example, a BYW80-150 diode rated for an average of 8A would actually drop around 0.9V at the peak of 11.5A.

The diode also sees quite a high reverse voltage. When the primary is storing energy, the input voltage is transformed to the transformer secondary in the turns ratio. So, at the highest input of 50V, 1.92 x 50V = -96V appears on the anode of the diode. The diode cathode however is already at +13.8V so the total reverse voltage is 110V. Inevitably there are ‘spikes’ (more on this later) on the waveforms so a diode rated for at least 150V is necessary, this for just a 13.8V output!

Apart from current and voltage rating, the diode needs to be a ‘fast’ type. Standard diodes store charge internally when they are forward biased. When a reverse bias is then applied, this stored

charge is swept out and because the reverse voltage can be high, this can result in a high peak current, the so-called ‘reverse recovery’ current. This causes high dissipation in the diode and reflects to current in the transformer primary, causing more dissipation, EMI and control circuit problems. It’s therefore necessary to use a ‘fast recovery’ diode with minimal stored charge. A Schottky diode has a low forward voltage drop and does not store charge in the same way but does have self-capacitance which produces a similar but lesser effect. High voltage Schottky diodes are more expensive but there are other choices. SBR (Super Barrier rectifier) diodes have a nice combination of reverse recovery and voltage rating while silicon carbide diodes are better still though currently expensive. We’ll choose an SBR type SBR10U150C which is 10A , 150V rated. It is actually a dual diode but if we put the two elements in parallel, the overall effect is lower stress and lower voltage drop. It has a forward voltage drop of about 0.63V at 5A.

Schottky diodes also have a problem whereby their reverse leakage current can be high, especially at high temperature. It can be measured in milliamps and if the reverse voltage is high, the dissipation can be fractions of a watt. As the effect is worse with temperature, thermal runaway can occur unless the diode has good heatsinking.

We’ve mentioned synchronous rectification before, using MOSFETs. These can be used in simple flyback circuits but controlling them under all conditions gets difficult. A signal has to be generated with the right timing to switch the MOSFET. In CCM this is not too difficult and can be derived from a transformer winding or drive to the main MOSFET switch. In DCM, necessarily the transformer voltage has collapsed before the end of each cycle and there is no clean signal to use to turn the MOSFET OFF. All practical flyback circuits at fixed frequency will enter DCM at some light load. Specialised ICs can provide the function by literally detecting reverse current through the MOSFET and turning it off but this adds cost and complexity. If the timing of the MOSFET control is not optimum under all conditions, then the efficiency advantages are quickly lost.

Capacitor ratings

Output capacitor considerations are the same as discussed in our section on boost converters. Remember that the AC portion of the transformer output current passes through the ESR of the output capacitor to give output voltage ripple. Several capacitors in parallel may be necessary to give a sufficiently low value of ripple and are probably more economical than one high performance capacitor. The ripple current is at least the same as the output DC current so for example, six 470µF, 25V capacitors in the FR series from Panasonic will have a combined rating of 6.2A and 8 milliohms giving around 50mV of ripple.

Transistor ratings

The switching transistor is almost exclusively a MOSFET these days although there is still use for bipolar types at very low power where they can be very low cost. Assuming a MOSFET, the device, like the output diode, has to be rated for the average and peak currents seen. Unlike a diode or bipolar transistor, it has an ON-resistance rather than a saturation voltage. This means that you need to use the square of the rms value of the primary current multiplied by the ON-resistance to get power dissipated. Note that the ON-resistance quoted for a typical device will be specified at 25°C and increases significantly at high temperatures. The IRF630 for example has a value of 0.4 ohms maximum at 25°C but nearly doubles to 0.7 ohms at a junction temperature of 100°C. See **Figure 7.8**.

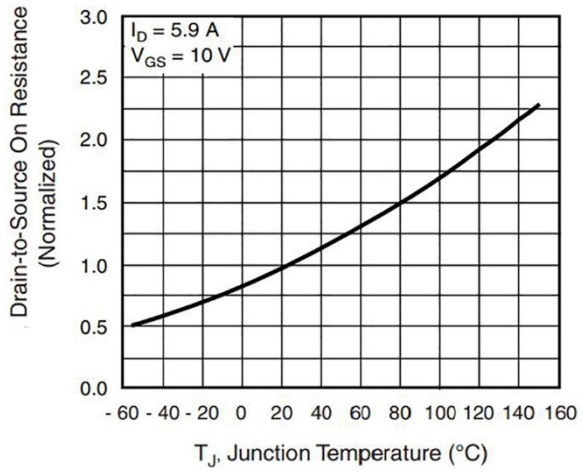


Figure 7.8 Variation of ON-resistance with temperature IRF630

Our example has a primary rms current of 2.93A so the dissipation in the channel resistance of the MOSFET would be 2.93^2 x 0.7= 6W at high junction temperatures.

There is further dissipation in the MOSFET during the switching transitions when there is momentarily high voltage and current. This can be very significant if not controlled and can easily match the conduction losses. A 'snubber' circuit helps minimise these losses.

The voltage rating of the MOSFET is important. The drain sees the highest input voltage plus the 'flyback' voltage on the primary plus any spikes. In our example, this would be 50V + 28V + spikes. A 100V device might be OK but 150V would be safer. An IRF630 would be OK with a V_{DS} max of 200V.

Snubber circuits

I've hinted that there are voltage and current 'spikes' on waveforms coming from different sources. Reverse recovery currents from diodes are a major source but also stray/parasitic inductances contribute. With any rapidly changing current you get a voltage induced according to $E = -L\ di/dt$

where *L* could be track or wiring inductance or transformer leakage inductance. We want rapidly changing currents to maximise efficiency.

Transformer leakage inductance can be seen as a separate inductor in series with each winding which plays no part in voltage or current transformation across windings. The value is typically 1-5% of the winding inductance and therefore in our example could be up to around 2µH. This dwarfs track inductances which are around 1nH per millimetre. When our MOSFET switches OFF, we might want the peak current of 6A to fall to zero within say 50ns representing a *di/dt* of 6/50 x 10^{-9} producing 240V across 2µH! This adds to the drain and flyback voltage. The spike produced tends to be a fast, resonant ring with stray capacitance such as in the MOSFET drain-source capacitance.

Clearly, we don't want to have to use an expensive 500V MOSFET with consequent high ON-resistance to cope with this so a way needs to be found to reduce or clamp the voltage to a reasonable level. The transformer can be wound carefully to minimise leakage by tightly coupling and interleaving windings. If we slow the *di/dt* in the MOSFET however, we risk losing efficiency so the common solution is to add one of the snubber or clamp circuits as shown in **Figure 7.9**.

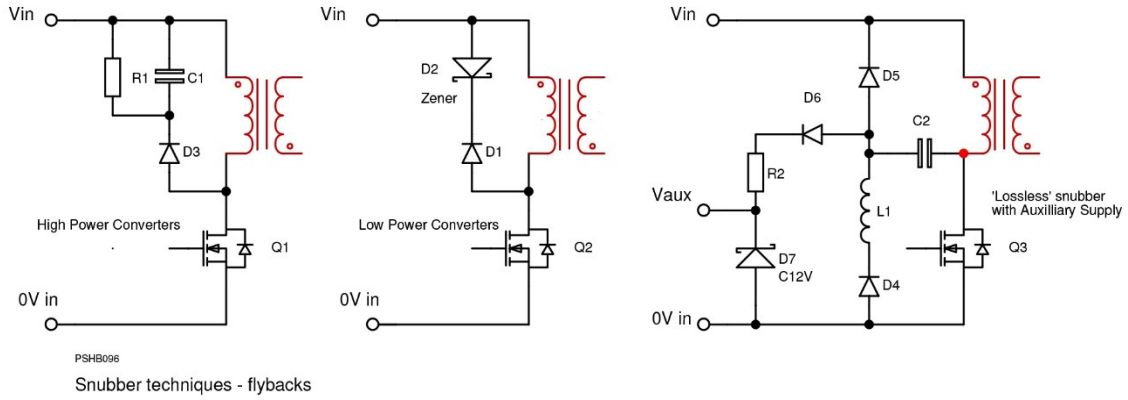


Figure 7.9 Snubber techniques for flyback circuits

For very low power, a simple Zener can clamp the voltage as shown in in Figure 7.9 (middle) for low power converters. This is a lossy solution but enables a lower voltage MOSFET to be used. The Zener is chosen to clamp the voltage at a safe value for the MOSFET but not so low that it dissipates too much. A more common circuit is shown in Figure 7.9 (left) for higher power converters where capacitor C1 charging through D3 absorbs the transient, discharging relatively slowly through R1 when Q1 turns ON. When 'tuned' for an acceptable overshoot, this can be effective. One technique for selecting components is to first measure the leakage inductance L_{lk} of the transformer (see Figure 7.42). Use this to get the power P_l that needs to be dissipated to remove all the energy from the leakage inductance from:

$$P_l = \frac{1}{2} L_{lk} {I_p}^2 F_s$$

Where I_p is the peak primary current and F_s is the switching frequency. If the voltage overshoot is allowed to be high by light snubbing, the snubber network dissipates less but the MOSFET voltage rating may be exceeded.

In practice, you can choose a MOSFET to allow some voltage spike V_x above the reflected output voltage or 'flyback' voltage V_f in which case the power to be dissipated P_{max} is:

$$P_{max} = P_i(1 + \frac{V_f}{V_x})$$

The voltage Vx is set by the resistor value given by:

$$R = \frac{2V_x(V_f + V_x^{max})}{F_s {I_p}^2}$$

The capacitor value is not critical and might be around 10nF.

There are many other snubber networks which can be used including types which re-circulate the energy absorbed during the transient to increase efficiency. Figure 7.9 (right) shows a 'lossless' snubber circuit which does this at the expense of some complexity and cost. Because the energy is not dissipated in this circuit, it can be used to provide an auxiliary power supply for the control IC, a nice bonus.

We now have the main components specified for our flyback converter so now let's complete the schematic with the remaining detail as in **Figure 7.10**.

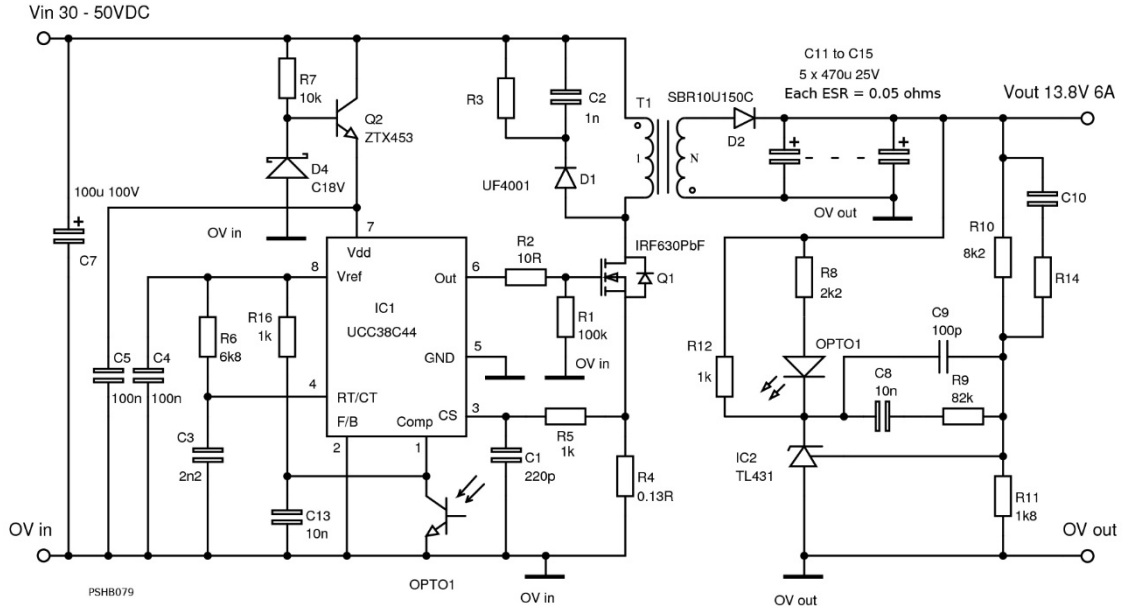


Figure 7.10 Example isolated flyback circuit

The controller chosen is from the tried and tested 38xx series which have been around for decades in different guises and is very low cost. We choose the UCC38C44 which has a 50% duty cycle limit.

Q2 is a general purpose NPN transistor with a Vce rating of at least 60V. A ZTX453 would do. Zener D4 on Q2 base is 18V, above the minimum start-up voltage of the IC. R7 provides bias current to D4, 10K is fine as at 30V input there would be 1mA flowing which is plenty of base current for Q2 as IC1 only needs about 20mA in operation and has a gain of 40 minimum.

C4 and C5 are for decoupling and 100nF ceramics close to IC1 are fine. Sometimes a bigger electrolytic is fitted for C5 as well as but is unnecessary when IC1 is fed from a 'hard' source such as Q2. A rule of thumb is to make C5 at least ten times the gate capacitance C_{iss} of the MOSFET. If we use an IRF630 it has C_{iss} of 800 pF then 100nF is plenty. R6 and C3 set the operating frequency of IC1 and from the data sheet, 6k8 and 2n2 give us 100kHz.

R2 provides a little slowing of the drive to Q1 which can help with EMI and stability, 10R is typical. R1 pulls down the gate of Q2 on start-up so that there is no transient condition where it might turn on. 100k is typical.

R4 passes Q1 current to give a scaled voltage to IC1. We know our peak current will be 6A and the operating maximum voltage for the current sense is about 0.8V so R = 0.8/6 = 0.13 ohms. The signal on C1 is actually used to terminate the ON-pulse to the MOSFET gate so it should be as clean as possible and represent the transformer magnetising current accurately. We'll talk about what this current sense is doing in the circuit later. To clean up the signal, R5 and C1 form a low pass filter. It should be just enough to attenuate edge spikes with a RC time constant of no more than about 200ns for a 100kHz switch frequency. 1k and 220pF are about right.

Snubber components C2 and R3 are chosen as described earlier to attenuate any ringing on the MOSFET drain to acceptable levels and D1 is a small fast diode rated at about 100V such as UF4001.

OPTO1 transfers a signal to IC1 representing the error in the sensed output voltage from what it should be. As the opto transistor turns more or less on, it varies the voltage on its emitter. This voltage is the input to the IC on the 'comp' pin which is internally compared with a ramp voltage to set the pulse width. The comp pin needs 0V to 2V for full control of the output pulse width from 0 to 50% duty cycle. The opto transistor which drives the comp pin is fed by R16 from V_{ref} which is 5V. The V_{ref} pin can only source 5mA maximum so we make R16 1k so that when the opto transistor is fully on it can only draw about 4mA from V_{ref}. The current will change from 0 to 2mA through the opto transistor and R16 to produce 0-2V as required. The CTR of the opto, the ratio of collector current to diode current is nominally 100% so the opto diode current to produce 0-2mA transistor current is also nominally 0 – 2mA

We'll start with C13 at 10nF and justify this later in the section on control loops. C13 looks increasingly like a low impedance at high frequency so across OPTO1 transistor, it reduces the loop gain as the frequency increases, necessary for stability.

R8 sets the effective voltage gain of the optocoupler network from the cathode of IC2 a TL-431, to the opto transistor collector. IC2 minimum voltage is 2.5V and the opto diode drops about 1.5V so IC2 cathode can only swing 2.5 to about 12.3V or alternatively R8 drops 0 to 9.8V maximum with 0 – 2mA current swing. This sets a maximum value of R8 to 4k9. In practice, we want a standing current through IC2 of a few milliamps to get it into a 'linear' region so R12 at 1k ensures that at least 1.5mA must flow before 1.5V is dropped across it allowing the opto diode to start conducting. If R8 is say 2.2k, this easily allows the fixed 1.5mA and our desired opto diode adjustment range of 0-2mA which causes a voltage drop across R8 of 3.3V to 7.7V. Because the voltage drop across the opto diode is nominally constant at 1.5V, the voltage at IC2 cathode varies between (13.8V-1.5V –(3.3 to 7.7V)) = 9V to 4.6V.

Error amplifier IC2 acts like an op-amp and must have enough gain at DC so that any small error detected on its input produces an output voltage within the range 3.3 to 7.7V to correct the error. The voltage gain of a TL-431 with no feedback is around 55dB or 560 so its input would vary (7.7V-3.3V)/560=7.85mV. Because of the voltage setting pot down effect of R10 and R11 this represents about 43.6mV at the output. Put another way, a 43.9mV change in the output swings the duty cycle from minimum to maximum.

Unfortunately, this high sensitivity can lead to instability as delays around the control loop can add enough phase shift at high frequencies, along with significant IC2 gain to produce an oscillator. Reducing the gain with just a feedback resistor R9 across IC2 spoils the DC output accuracy and can still leave instability. Just C8 would not affect DC but limits the high frequency performance quickly so that the control loop is slow to react to load changes. Adding C8 as 10nF with R9 in series and 'tuning' R9 to 82k can leave the DC performance unchanged with gain-reducing feedback 'switched in' at higher frequencies as C8 becomes a low impedance. C9 is a small value around 100pF to help with any noise pick up at the sensitive gate input of IC2.

Optocoupler CTR values are very variable from part to part and with temperature, operating conditions and time so you might like to redo calculations with the extreme values of CTR that might occur.

We'll look more closely at the theory of loop stabilisation later with some examples and justification for the selection of the components around IC2.

R10 and R11 'pot down' the output voltage to the 2.5V reference pin of programmable Zener diode IC2. R10 and R11 can also function as a minimum load for the converter but typically only sink a few milliamps for minimal heating and subsequent drift so for a 13.8V output and 2.5V reference, R11 could be 1k8 and R10 8k2, nice standard values.

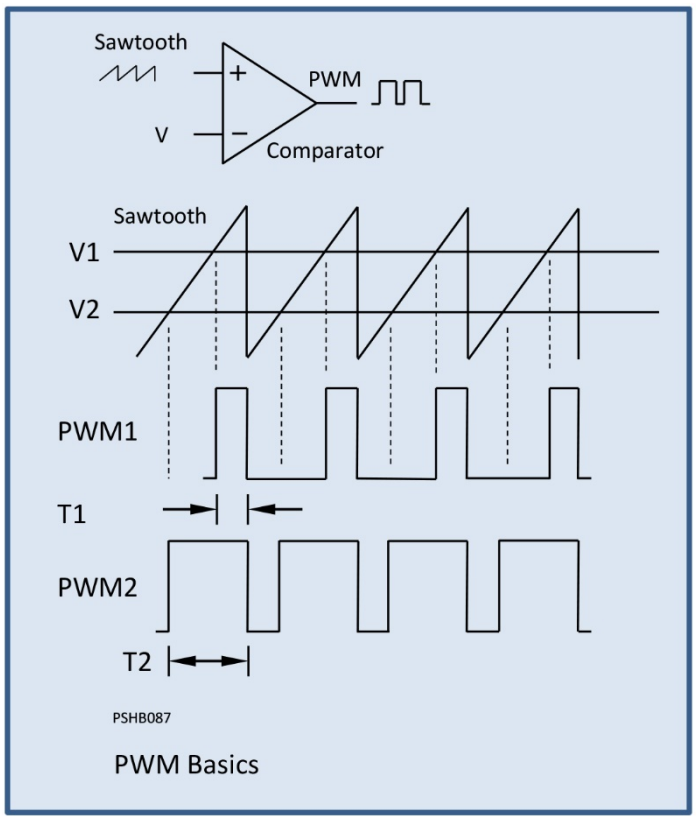
R14 and C10 are included as a provision which can give a 'phase advance' effect to help with stability especially when there is a 'right half plane zero' present. The later section on loop stability explains

how these values can be chosen. One thing you might notice about the feedback circuit is that there are actually two loops! One is through IC2 controlled by the pot-down from the output but the other comes from changes in the output directly injecting current into the opto through R8. These are called the 'slow' loop and 'fast' loop respectively. It makes for challenging analysis but actually can give good performance. Transient load changes affecting the output voltage are quickly compensated for via the fast loop while slower load changes come through the slow loop. To make analysis easier, you can feed R8 from a stabilised voltage, perhaps a resistor-Zener off the output. This can make the circuit more predictable. You can see the arrangement in the feedback section of **Figure 9.4**.

We've now got a working converter that reacts to changes in load voltage with compensating changes in MOSFET pulse width. In the IC chosen, the current sense signal is used to terminate the pulse in what is called 'Current Mode' control. An alternative is 'Voltage Mode' control as well as different versions of current mode control - this needs a little explanation.

Voltage mode & current mode control

The method of generating a variable pulse width from an error signal is quite simple. A saw-tooth waveform is generated and fed into a comparator along with the error signal as shown in **Figure 7.11**.



PWM Basics

Figure 7.11 PWM basics

In the classical scheme, an error signal is compared with an internally generated saw-tooth waveform. If the output voltage is correct, the error amplifier is designed to have a positive output voltage value V2 say, giving a MOSFET gate drive like PWM2. If the output is too high, the error amplifier output goes to voltage V1 for example and the comparator outputs narrower pulses like PWM1 to bring the output back down. If the output is too low, the opposite happens and wider pulses are produced. This scheme is called 'Voltage mode' control as the output voltage error directly controls the pulse width. The scheme has some shortcomings though. If the input voltage changes, the correction to the pulse width has to propagate all the way round the control loop from input, through the transformer, through the error amplifier, through the optocoupler if fitted, back to the IC, through its internal circuitry and to the MOSFET gate drive. All this adds delay meaning that the output doesn't respond quickly to input voltage changes and over- and under-shoots. Also, any external capacitor fitted to the output affects the delay around the loop. To keep this loop fast but stable requires careful and complex shaping of the loop gain and phase shift changes with frequency, achieved with RC networks around the error amplifier.

A better scheme is to use the inherent saw-tooth waveform generated by the current in the MOSFET primary. This current should linearly rise as energy is stored in the transformer primary in a flyback converter. The current is sensed and a similar shaped voltage produced to use in our pulse width modulation comparator. See **Figure 7.12**.

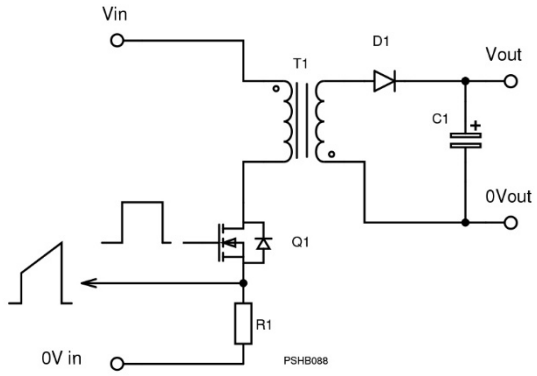


Figure 7.12 Current mode control uses ramp from R1 for PWM

By using this saw tooth, and turning off the MOSFET when it reaches a set level, we are effectively controlling the peak current in the MOSFET. For fixed frequency and transformer inductance this controls output power. (*Power = f Li²/2*). The control loop is now working to keep the output voltage correct through the error amplifier and the demanded power correct ie the output load so we have two control loops in operation. How does this help? Well, if the input voltage changes, this naturally tries to produce a higher current from *i = ET/L* but *i* is now controlled so *T* must reduce that is the ON-time of the MOSFET. This means that correction for input voltage changes is immediate and doesn't have to propagate around the loop making for faster more accurate response. Another advantage is that in current mode control, the effect of the transformer inductance is controlled in

the current loop and doesn't appear then in the outer voltage control loop, making it easier to stabilise.

Another major advantage of current mode control is that because the primary current is controlled it can be limited. This gives us automatic protection of the MOSFET from current stress. In a voltage mode converter, you would want this protection anyway so normally the current sense signal is present so it's easy to implement current mode. The over-current protection doesn't apply to the output however as it's a power limit (peak current *i* is limited which limits power from $Power = f Li^2/2$ if *f* and *L* are fixed. *So,* for a fixed inductance and frequency, if a short circuit is applied to the output for example, the output voltage is zero so the converter will try to supply very high current to maintain the set power. In practice, the shorting or dropping of output voltage can be detected in the control loop and the converter can be put into a 'hiccup' mode or similar to limit average output current to a safe value.

Another advantage claimed is that you can parallel converters easily by connecting their current sense signals directly together forcing each to give the same peak current and power. The converter outputs are paralleled so if each is giving the same power for the same output voltage, they must be sharing the output current accurately. In practice, I've never seen this done as it would mean bringing out the sensitive current sense signal and a common ground to make the parallel connection. The risk of noise pickup and instability is extreme.

If there is any down-side to current mode control, it is susceptibility to noise on the current sense signal. The waveform needs to be clean and any edge spikes suppressed. In CCM operation when the slope of the current waveform is shallow, this is a particular risk. If noise is picked up, the effect can be chaotic operation or even component failure.

Self-oscillating flyback – variable frequency

As promised, we'll now look at an even simpler flyback circuit that doesn't use a control IC. It is often used at low power, up to about 10W and has the lowest component count for an isolated, regulated converter with good efficiency and low noise. A circuit is shown in Figure 6.46 and can be regarded as a power oscillator. It inherently works in CRM, the critical conduction mode on the boundary between continuous and discontinuous modes so the waveforms are as in **Figure 7.13**

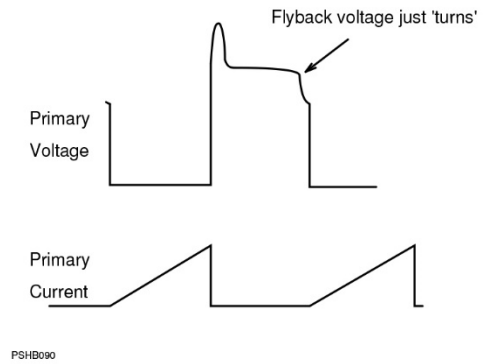


Figure 7.13 Flyback critical conduction mode waveforms

Referring back to Figure 6.46, initially, R5 provides a bias to the MOSFET gate to turn it on. Once oscillating R5 plays little further part if high enough in value. With Q1 ON and the full input supply connected across the transformer winding T1b, the current starts to linearly rise according to $i = Et/L$ and a voltage appears across T1a phased so that the top end goes positive, reinforcing the gate drive voltage through R1 and C3. The voltage on the output winding is negative so D1 is reverse biased. The rising current through R4 stores increasing energy in T1 and the current drops an increasing voltage across R4 so that when it exceeds about 0.6V on the base of U1, the opto transistor starts to turn on, pulling the MOSFET gate down, switching it OFF. The MOSFET drain voltage starts to rise and 'flies-back' to a voltage higher than the drain producing a positive voltage for the output on the anode of D1 and a reinforcing negative voltage on the gate of Q1 via winding T1a. Output capacitor C1 is now charged and load current produced from the energy in T1 also replacing energy in C1 which was passed to the load when D1 was reverse biased. The flyback voltage is maintained until all of the energy in the transformer is used up and the drain voltage begins to fall back towards the supply. The fall in voltage produces a positive going rise in voltage on winding T1a, turning the MOSFET ON and repeating the whole cycle. Output voltage regulation is achieved through the optocoupler. If the output voltage is too high, the opto diode conducts through D2, and switches the opto transistor ON slightly. This means that less voltage is needed from the drop across R4 to switch the MOSFET OFF so the ON pulse is cut short at a lower peak MOSFET current, reducing the energy transferred and consequently reducing the output voltage, correcting the measured error.

The net result is that the MOSFET ON-time is fixed by the winding inductance, input voltage and output voltage feedback. The OFF-time is variable with load as it stays off until all energy is transferred. The converter therefore varies its effective switching frequency with input voltage and load between a chosen minimum and a high value. Typical circuits can vary between tens and hundreds of kHz between extremes. A nice feature of the circuit is that the switch-on point for the transistor is always at zero current as all energy has been transferred in the previous cycle. Also, at this point the forward current in D1 has fallen to zero so there is no reverse recovery current transient. This makes a 'quiet' switching edge. The transistor does switch OFF at the peak primary current potentially producing noise but a snubber can be added to reduce this.

What sets the nominal flyback voltage on the MOSFET? When the MOSFET is OFF and D1 conducting, the secondary winding is clamped to $Vout + 0.6V$ by the feedback and clamping effect of

the output capacitor which can only change its voltage slowly over many switching cycles. The primary winding is therefore also clamped to a voltage which is the secondary voltage multiplied by the turns ratio. In practice, there will be a spike and ringing which may not matter for low power. If it is too high for the drain voltage limit, a simple clamp or snubber is fitted. Let's run through the calculations to get a practical circuit. We shall design for 5V out at 1W with an input range of 6 to 18V. This might be used to provide a constant supply to some logic from a 12V car battery that might rise on charging and fall to a low value when cranking.

It can be puzzling where to start when the switching frequency, ON and OFF times vary so much but there are fixed points that can be combined with desired conditions that can be set. One of the fixed points is that in CRM, during the OFF time for Q1, the flyback voltage is constant until the next switching cycle. This is because the fall of the flyback voltage *initiates* the next cycle. So, because of volt-second equality for the transformer, T_{on} *x* V_{in} must always equal T_{off} x *Flyback voltage*. We can choose our flyback voltage to be what we want within practical limits. A typical choice is to make the flyback voltage the same as the minimum input voltage so that the duty cycle is 50% at this point. Remember that volt-seconds must be the same so if the voltages are the same, the 'seconds' for ON and OFF times must be the same ie 50% duty cycle. Now we can choose a minimum frequency of operation. Too low and it might be audible, too high and switching losses increase, especially when the load is light and input voltage high when the frequency goes to its maximum. Let's say 50kHz for the minimum frequency. So, at V_{inmin} and maximum load, we have 50% duty cycle at 50kHz so the ON time is 10µs.

Each cycle we need to provide our 1W power output. Say efficiency is 75% so the primary must store then transfer about 1.3W each cycle. Let's say 1.5W for some margin.

This means that the average input current at 6V must be 1.5/6 = 250mA. We know that the wave shape for the primary is a saw tooth with 50% duty cycle so from Appendix 3 and from common sense we can see that the peak value of the waveform must be 4 times the average, that is 1A.

Our primary inductance now comes easily from $L = Et/i$ with E = say 5.5V to allow for the MOSFET drop.

$L_{pri} = 5.5 \times 10 \times 10^{-6}/1 = 55\mu H$

The number of turns *n*, limits our maximum core flux density according to $B = E.\ t/\ (A_e.n)$

We now may have to iterate a little to what is practical. At 1W you would expect a small core so let's try an EF20 with A_e of 32.1mm^2

Assuming a maximum flux density of say 150 mT, with time in microseconds and cross-sectional area in square millimetres, this would give us:

n= E . t/(A_e . B) = 5.5 x 10 x /(32 x 0.15) = 11.46 turns which should fit easily in the winding area.

We said we wanted the flyback voltage to equal the minimum input voltage which is 5.5V on the transformer and we need to round up or down. If we choose 11T, for a 5.5V flyback voltage we have 0.5 volts per turn. For 5V output we need about 5.5V before the rectifier diode so 11 Turns on the secondary would give us exactly that. We now have n_{pri} *and* n_{sec} both 11 turns.

This now requires a core with $A_L = nH/T^2$ = 27.5 x 1000/6^2 = 454

The core gap comes from our Equation 9:

L_g = 4 x π x 10^{-7} x n^2 x A_e x 10^3/L

L_g = 0.09mm. This is quite small and may be a little difficult to control but is workable.

With our assumptions, we have 'forced' the circuit to 50kHz, 50% duty cycle at V_{inmin} and maximum load. Let's see what happens under other conditions.

At lower powers, say 10% load, the circuit could have a shorter ON time to store less energy for a given frequency or it can have a lower frequency for the same ON time with the same peak current. To know how the circuit settles we need another relationship between power and ON and OFF times. The peak primary current at the end of the MOSFET ON-time transforms to the starting peak secondary current during the MOSFET OFF-time in the proportion of the turns ratio. Allowing for the circuit efficiency, we get:

$$i_{pksec} = \frac{Efficiency\,.\;i_{pkpri\,.}\;n_p}{n_s}$$

The transformer secondary current average value is the DC output current so from the wave shape in **Figure 7.14** and Appendix 3, we can say that:

$$i_{DC} = \frac{i_{pksec}\,.\,t_{off}}{2\;.\;t_{cycle}} = \frac{P_{out}}{V_{out}}$$

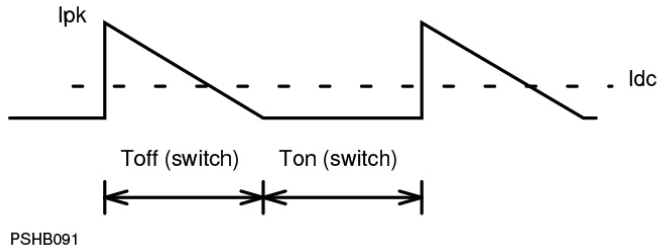


Figure 7.14 Flyback output diode current- discontinuous mode

Combining equations and without going through all the workings, we get:

$$t_{on} = \frac{\frac{2 \,.\, P_{out} \,.\, L}{V_{in} \,.\, efficiency}}{(\frac{1}{V_{in}} + (\frac{n_s}{n_p V_{out}})}$$

$$f = \frac{2 \,.\, P_{out} \,.\, L}{{V_{in}}^2 \,.\, {t_{on}}^2 \,.\, Efficiency}$$

From this we can see that at constant input voltage the ON-time scales with power so is about 1µs in our example at 0.1W load ie a tenth of the 1W value. The frequency however is affected by load and t_{on} and is about 485kHz ie 5 times the 1W value. At high Vin and light load say 18V and 0.1W, t_{on} is 0.3µs and frequency is about 516kHz.

You can see that many approximations have been made such as diode/MOSFET voltage drops and some effects have been ignored such as transformer leakage inductance. The calculations give an approximate result and should therefore only be used as a starting point. This is very typical of power design. Simulation would help but leakage inductance for example will not be known until a practical transformer is built.

To fill in the other components, we said that under worst case conditions our peak primary current is 1A. To be sure that OPTO 1 turns OFF at this level, we need 0.6V across R4 so its value is 0.6/1 = 0.6 ohms.

R3 effectively proportions the effect of R1 voltage on the base of OPTO U1 transistor with the feedback current from the OPTO diode acting on the base of OPTO U1 transistor. It's found by experiment to be about 3k3 ohms.

R5 should be high enough not to dissipate significant power but not so high that start-up takes a long time. 1M ohm with the gate capacitance of Q1 of about 1nF gives a start-up of about 1ms.

R1 along with C3 define the pulses of current needed to switch Q1 sharply ON and OFF. The MOSFET gate needs quite high current pulses to switch so R1 at 22 ohms allows about 0.5 amps to flow transiently going positive. The negative-going voltage pulse from the transformer winding is defined by the turns ratio to the flyback voltage but the positive going voltage varies with the input voltage. So, values for R1 and C3 are a compromise. Their time constant has to be faster than the shortest expected pulse width of around 1 µs, 22R and 470pF are suitable starting points.

D1 is a low power diode and does not have to be fast recovery because in CRM, diode forward current always drops to zero each cycle and there is no charge to 'recover'. A power diode such as 1N4001 would *not* be suitable however as they have what is called forward recovery characteristics where the diode momentarily drops a high voltage as it turns ON before dropping to around 0.6V. This can dramatically drop the average output voltage and increase diode dissipation so a 'fast' diode should still be used such as UF4002 or a Schottky type would be ideal.

C1 will pass a ripple current of about 1.5 times the DC current. The DC is only about 200mA so a capacitor with an ESR of say 100 milliohms would give about 20mV ripple which is fine. Its capacitance could be as low as 47µF. Because the switching frequency varies so much, it's worth checking the capacitor data sheet to see how the ESR varies over the expected frequency range.

D2 is a Zener diode that conducts at the desired output voltage. OPTO U1 diode however is also in series and drops about 1V so for a 5V output, D2 should be about 4V, a standard 3.9V type may be OK. This will not give very accurate regulation as the Zener and OPTO diode voltage will change with temperature and tolerance so the TL-431 circuit used in Figure 7.10 could be used if preferred.

R6 ensures that a minimum current flows before the OPTO diode conducts. This get D2 Zener into its specified operating area for its output to be most accurate. R6 could be used to 'fine-tune' the output voltage a little as it effectively varies output standing current and D2 voltage. 1k is suitable as a start.

R2 sets the gain of the feedback loop: if its high, the output voltage has to move a long way for a given change in OPTO diode current. If the value is low, a small change in output voltage is needed to produce the same change in diode current, the effect of varying gain. The value also limits the current through the opto diode under fault conditions. A starting point could be 470R.

With such a simple converter, the performance depends heavily on the transformer and the way it is wound with inherent leakage inductance, coupling and winding capacitance. Inevitably the circuit values will need 'fine tuning' to get best performance over the input and load range. Playing with a

simulation model is a good way to get a feel for how the circuit responds to component adjustments. You can try for example, the range of optocoupler CTR values that might occur.

Adding additional outputs

A nice thing about flyback converters is that is it easy to add extra outputs, positive or negative, with another winding and just an extra diode and capacitor. Once the 'volts per turn' V/t, for the main output is determined, then other outputs will have the same so for example if the main output is five turns for five volts, ten turns will give approximately double the value. It's not exact however because the V/t relates to the winding, not the DC output. Our 5V output for example will be about 5.6V for a silicon diode at the five-turn winding actually giving 1.12 V/t. Ten turns will therefore give us 11.2V on the second winding or 10.6V after the rectifier diode. You can play with the diodes to affectively change the V/t value. If we had used a Schottky diode for the main output, the V/t value would be more like 1.05 and ten turns would give us 10.5V again reduced by the second output diode drop. We could even use two diodes in series on the main output to increase the V/t to fine tune the second output if adding or subtracting a transformer turn is too coarse. Doing this does change the duty cycle of the control loop however so the effect should be checked on other parameters such as voltage and current stresses and input range.

If the two outputs can share a ground, the second can be 'stacked' on the first so it only needs the difference in number of turns, another five to give our ten total. This makes for a simpler transformer and actually gives a better regulation figure for the second output as half of it is directly regulated. **Figure 7.15** show a possible arrangement for 5V and +10V.

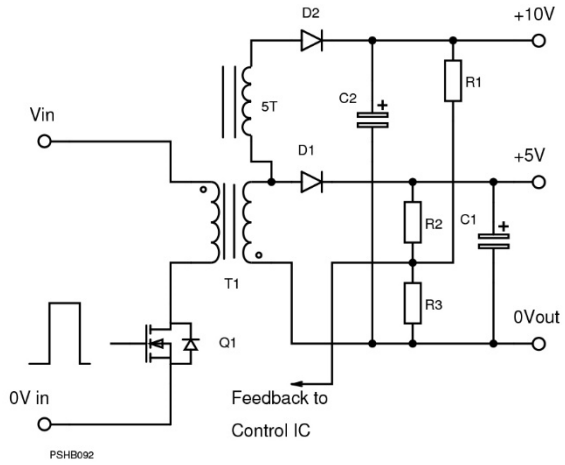
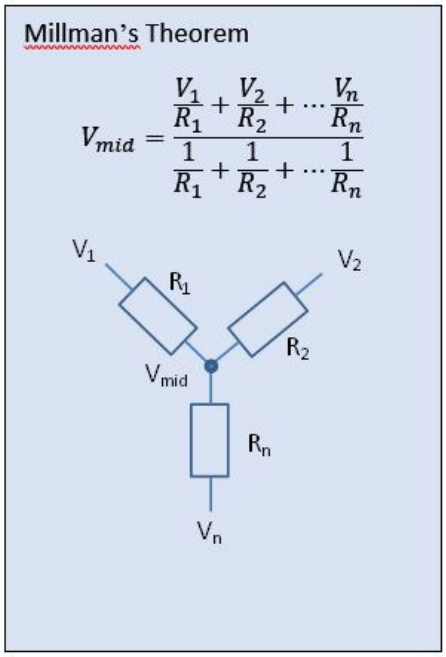


Figure 7.15 Sharing regulation between multiple outputs

A problem occurs if the main output is lightly loaded and particularly if it causes the converter to go into DCM. The duty cycle will decrease and the second output could drop dramatically. If there is a heavy load on the second output while the main output is lightly loaded, the effect is worse still. A

solution is to have a constant preload on the main output but this loses power and efficiency. Another possibility is to share the regulation between outputs. If resistors R1, R2 and R3 are fitted, the control loop acts to keep all outputs approximately correct. Neither output is very accurate and the effect depends on the differential loading of the outputs but it can give a good compromise. There's a little know theorem by Millman that helps set resistor values. It's just a version of Kirchoff's laws but not often used. See **Box 2**.



Box 2 Millman's theorem

Another trick to fine tune voltage is to add a very small inductor in series with the rectifier diode, typically a ferrite bead. The inductor acts as an impedance to the very start of the switching waveform and then saturates at the running current. In saturation, it drops no voltage and dissipates no power because its inductance has collapsed. If the bead is on the main regulated output, the delay it introduces changes the effective volts per turn of the winding and therefore adjusts the other secondary outputs a little, pushing them up, as the duty cycle increases to make up for the short delay introduced on the waveform edge.

The transformer switching waveforms inevitably have spikes on their edges and if the load is light, the spikes significantly add to the output voltage and for a main regulated output, it will then react to 'turn down' the duty cycle, dropping the voltage for secondary outputs. If the main output is loaded sufficiently but the secondary outputs have light load, their voltage can rise substantially because of the spikes on their waveforms and they may need preloads to keep them within bounds. The ferrite bead on a secondary output is a help here because it can 'clip' the edge of the rising waveform that includes the spike.

If secondary outputs are low power, it's common to add cheap linear regulators to effectively provide good regulation at the expense of some drop in overall converter efficiency. A secondary non-isolated switching regulator could even be added to get good regulation at better efficiency.

It could be that the main and secondary outputs from the converter do not need tight regulation. If around 5% load and line regulation can be suffered, feedback from the output is not necessary – it can be taken from a primary referenced winding which 'mirrors' the main output winding see **Figure 7.16**. In fact, this primary winding may already be there as an auxiliary power rail for the primary circuitry. This technique is being increasingly used at low power to save cost and lose the need for an optocoupler with its inherent variability. For this technique to be successful, the primary auxiliary winding is typically sampled for the reflected output voltage at a specific waveform point so any spikes or ringing are ignored. Control ICs are now available that do this automatically.

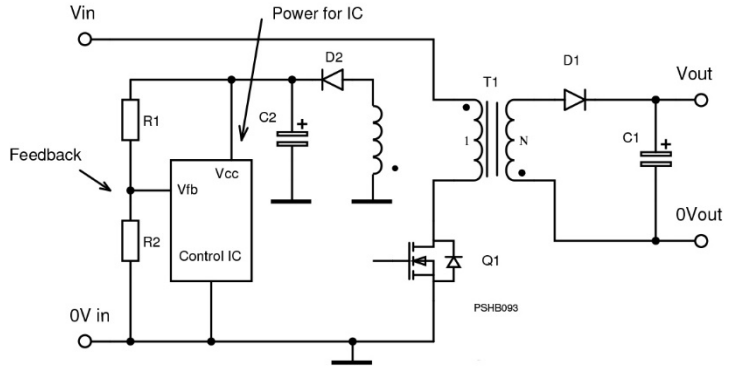


Figure 7.16 Primary referenced winding provides semi-regulation