Chapter 13

Digital and analogue circuit design

Circuit design is an essential contributor to EMC. In this area we can distinguish the techniques used to control radio frequency emissions from an operating circuit, and those used to control radio frequency and transient immunity of an operating circuit. There are some common points between these two, but there are also some differences which mean that to completely address the compatibility of a product you need to deal with both. Low frequency (mains supply) EMC is also a function of circuit design, principally of the power supply, but does not suffer from the mystery surrounding RF effects and is not considered in this chapter.

13.1 Design for emissions control

Digital circuits are prolific generators of electromagnetic interference. High frequency square waves, rich in harmonics, are distributed throughout the system. The harmonic frequency components reach into the part of the spectrum where cable and enclosure resonance effects are important. Analogue circuits are in general much quieter because high frequency square waves are not normally a feature. A major exception is wide bandwidth video circuits, which transmit broadband signals up to several tens of MHz, or over a hundred MHz for high resolution video. Any analogue design which includes a high frequency oscillator or other high di/dt circuits must follow HF design principles, especially with regard to ground layout.

Some low frequency amplifier circuits can oscillate in the MHz range, especially when driving a capacitive load, and this can cause unexpected emissions. The switching power supply is a serious source of interference at low to medium frequencies since it is essentially a high-power trapezoidal wave oscillator.

13.1.1 The Fourier spectrum

13.1.1.1 The time domain and the frequency domain

Basic to an understanding of why switching circuits cause interference is the concept of the time domain/frequency domain transform. Most circuit designers are used to working with waveforms in the time domain, as viewed on an oscilloscope, but any repeating waveform can also be represented in the frequency domain, for which the basic measuring and display instrument is the spectrum analyser (section 7.1.2). Whereas the oscilloscope shows a segment of the waveform displayed against time, the spectrum analyser will show the same signal displayed against frequency. Thus the relative amplitudes of different frequency components of the signal are instantly seen.

The mathematical tool which allows you to analyse a known time domain waveform in the frequency domain is called the Fourier transform. The necessary equations for the Fourier transform are covered in Appendix D, section D.8. Figure 13.2

shows the spectral amplitude compositions of various types of waveform (phase relationships are rarely of any interest for EMC purposes). The sinewave has only a single component at its fundamental frequency. A square wave with infinitesimal rise and fall times has a series of odd harmonics (3, 5, 7, etc. multiples of the fundamental frequency) extending to infinity. A sawtooth contains both even and odd harmonics.

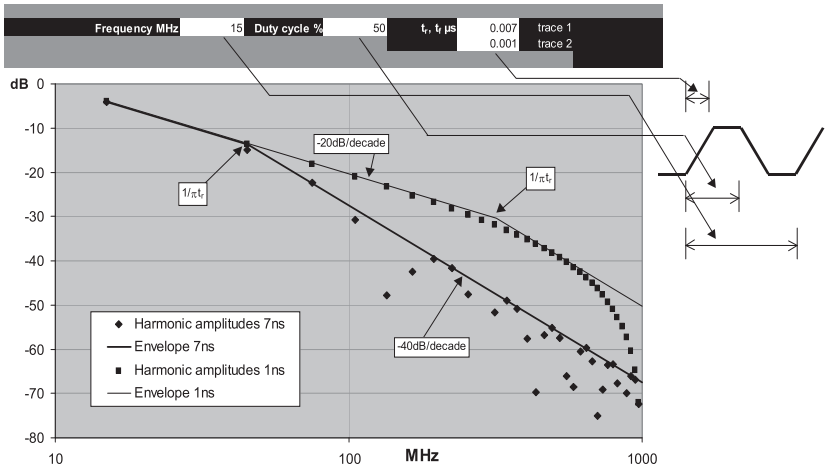


Figure 13.1 Harmonic envelope of a 50% duty cycle trapezoid

Switching waveforms can be represented as trapezoidal; a digital clock waveform is normally a square wave with finite rise and fall times. The harmonic amplitude content of a trapezoid decreases from the fundamental at a rate of 20dB per decade until a breakpoint is reached at $1/\pi t_r$, after which it decreases at 40dB/decade (Figure 13.1). Of related interest is the differentiated trapezoid, which is an impulse with finite rise and fall times. This has the same spectrum as the trapezoid at higher frequencies, but the amplitude of the fundamental and lower order harmonics is reduced and flat with frequency. (This property is intuitively obvious as a differentiator has a rising frequency response of +20dB/decade.) Reducing the trapezoid's duty cycle from 50% has the same effect of decreasing the fundamental and low frequency harmonic content.

Asymmetrical slew rates and duty cycles other than 50% generate even (multiples of 2, 4, 6, etc.) as well as odd harmonics. This feature is important, since differences between high- and low-level output drive and load currents mean that most logic and power switching circuits exhibit different rise and fall times, and it explains the presence and often preponderance of even harmonics at the higher frequencies. Changing the duty cycle even slightly away from 50% will dramatically increase the amplitude of even harmonics with little change to the odd components; conversely a duty cycle of 33.3% causes a null in the odd harmonics.

13.1.2 Radiation coupling

The following section discusses differential and common mode radiation mechanisms from circuits on a PCB, assuming a single source frequency. The equipment is taken to have a plastic enclosure, so that no screening effect is offered and the PCB radiates as if it is a bare board. Conductive enclosures complicate the issue.

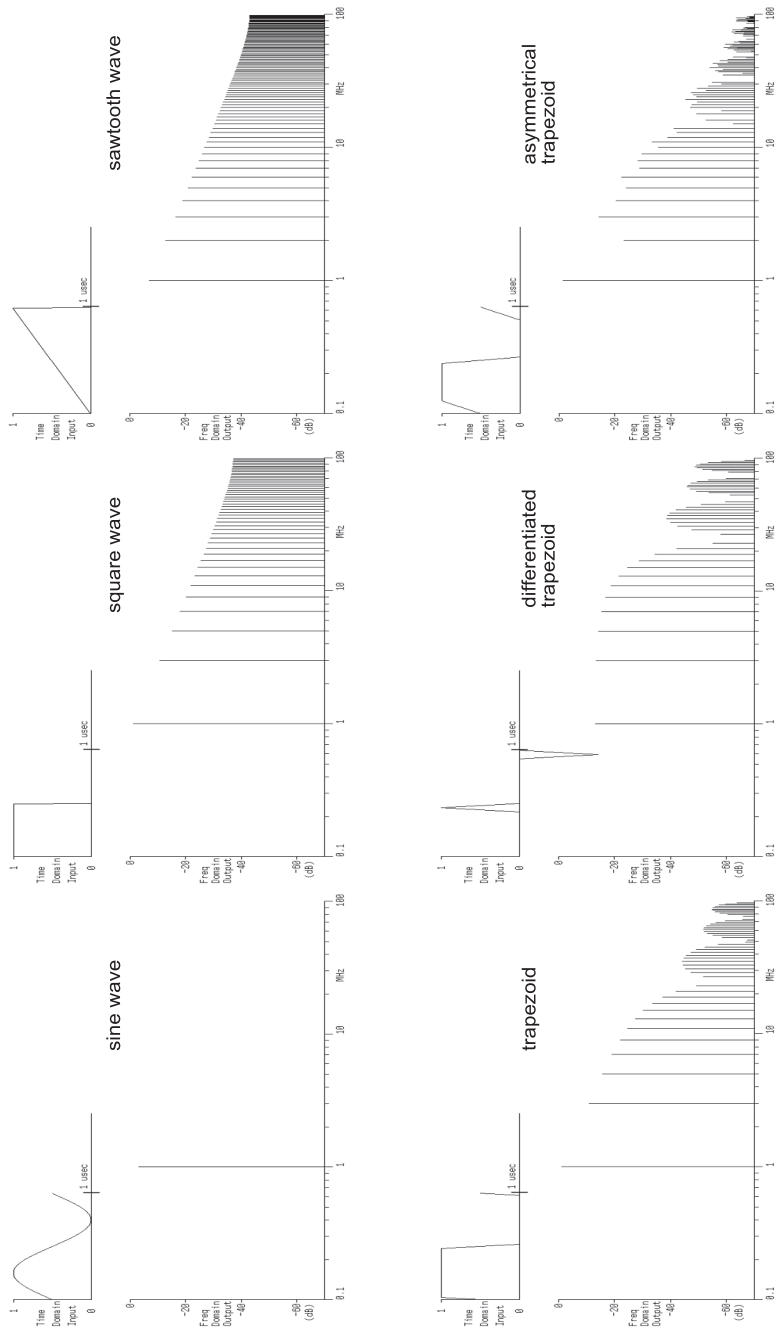


Figure 13.2 Frequency spectra for various waveforms

13.1.2.1 Differential mode radiation

As is shown in section 11.2.1.1, the radiation efficiency of a small loop is proportional to the square of the frequency (+40dB/decade). This relationship holds good until the periphery of the loop approaches a quarter wavelength, for instance about 15cm in epoxy-glass pcb† at 250MHz, at which point the efficiency peaks. Superimposing this characteristic onto the harmonic envelope of a trapezoidal waveform shows that differential mode emissions (primarily due to current loops) will be roughly constant with frequency (Figure 13.3(a)) above a breakpoint determined by the risetime [10].

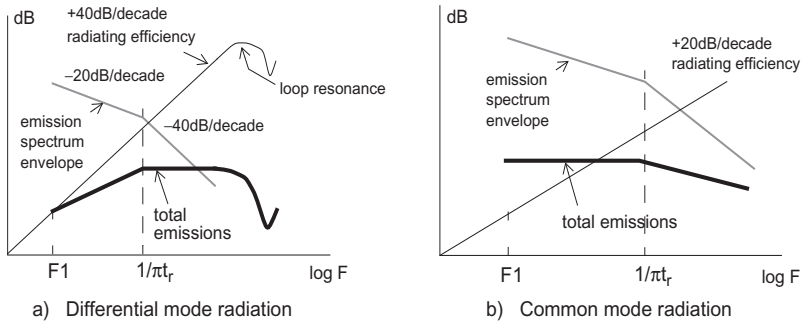


Figure 13.3 Emissions from digital trapezoid waves via different paths

The actual radiated emission envelope at 10m can be derived from equation (11.13) on page 272 provided that the peak-to-peak current, risetime and fundamental frequency are known. The Fourier coefficient at the fundamental frequency F1 is 0.64, therefore the emission at F1 will be:

$$E = 20\log_{10} . [119 . 10^{-6} (f^2 . A . I_{pk})] \text{ dB}\mu\text{V/m} \quad (13.1)$$

from which the +20dB/decade line to the breakpoint at $1/\pi t_r$ is drawn. In fact, by combining the known rise and fall times and transient output current capability for a given logic family with the trapezoid's Fourier spectrum at various fundamental frequencies, the maximum radiated emission can be calculated for different loop areas. If these figures are compared with the EN Class B radiated emissions limit (30dBμV/m at 10m up to 230MHz), a table of maximum allowable loop area for various logic families and clock frequencies can be derived. Table 13.1 shows such a table. The working for an example from this table (74AC at 30MHz) is shown underneath.

The ΔI figure in Table 13.1 is the dynamic switching current that can be supplied by a device to charge or discharge its node capacitance. It is not the same as the steady-state output high or low load current capability. Some manufacturers' data will give these figures as part of the overall family characteristics, as for t_r and t_f, which are typical figures, and are not the same as maximum propagation delay specifications.

† The effect of the epoxy-glass dielectric is to slow the wave propagation in the PCB and hence reduce the effective wavelength, by a factor proportional to $\sqrt{\varepsilon_r}$ if the track is on an inner layer; see Appendix D section D.5.3

Table 13.1 Differential mode emission: allowable loop area

Logic family	t_r/t_f ns	ΔI mA	Loop area cm^2 at clock frequency 4MHz	10MHz	30MHz	100MHz
4000B CMOS @ 5V	40	6	1000	400	–	–
74HC	6	20	45	18	6	–
74AC	3	80	5.5	2.2	0.75	0.25
74F	3	80	5.5	2.2	0.75	0.25
FPGA (example)	2.1	16	20	8	2.6	1.25

Loop area for 30dBmV/m 30MHz-230MHz, 37dBmV/m 230-1000MHz at 10m

Working: take 74AC family with F_{clk} = 30MHz as example. Worst case is at 150MHz (5th harmonic)

Fourier analysis of the source current using section D.8 on page 527 with $(t + t_f)/T = 0.5$, T = 33.3ns, t_r = 3ns and I = 80mA gives $I_{(5)}$, the current at the 5th harmonic, as 7.1mA.

From equation (11.13), for a field strength E of 30dBμV/m and $I_{(5)}$ at 150MHz as above, the allowable loop area A is **0.75cm^2**

Layout and construction implications

The implication of these figures is that for clock frequencies above 30MHz, or for fast logic families (AC or F), a ground plane layout is essential as the loop area restrictions cannot be met in any other way. Even this is insufficient if you are using fast logic at clock frequencies above 30MHz. The loop area introduced by the device package dimensions exceeds the allowed limit and extra measures (some form of shielding) are unavoidable. This information can be useful in the system definition stages of a design, when the extra costs inherent in choosing a higher clock speed (versus, for example, multiple processors running at lower speeds) can be estimated.

Table 13.1 applies to a single radiating loop. Often, only a few circuits dominate the radiated emissions profile. Radiation from these loops can be added on a root mean square basis, so that for *n* similar loops the emission is proportional to $\sqrt{n}$. If the loops carry signals at different frequencies then their emissions will not add.

Do not make the mistake of thinking that if your circuit layout satisfies the conditions in Table 13.1 then your radiated emissions will be below the limit. Total radiation is frequently dominated by common mode emissions, as we are about to discover, and Table 13.1 only relates to differential mode emissions. But if the circuit does *not* satisfy these criteria, then extra precautions will definitely be needed.

13.1.2.2 Common mode radiation

Common mode radiation which is due mainly to cables and large metallic structures increases at a rate linearly proportional to frequency (ignoring resonances), as shown in equation (11.14) on page 274. There are two factors which make common mode coupling the major source of radiated emissions:

- cable radiation is much more effective than from a small loop, and so a smaller common mode current (of the order of microamps) is needed for the same field strength;
- because of typical lengths, cable resonance tends to fall within the range 30–100MHz, in which range radiation is enhanced over that of the short cable model.

A similar calculation to that done for differential mode can be done for cable radiation on the basis of the model shown in Figure 13.4. This assumes that the cable is driven by a common mode voltage developed across a ground track which forms part of a logic circuit. The ground track carries the current ΔI which is separated into its frequency components by Fourier analysis, and this current then generates a noise voltage differential V_N of $\Delta I \cdot j\omega \cdot L$ between the ground reference (assumed to be at one end of the track) and the cable connection (assumed to be at the other). A factor of –20dB is allowed for lossy coupling to the ground reference; such coupling is often due to stray capacitance and structural resonances and is very difficult to quantify. At some frequencies the loss will be much greater, but at others (for instance if a series resonance is involved) it could be much less, approaching 0dB. The cable impedance is assumed to be a resistive 150Ω and constant with frequency – this is a value which represents the average of typical cables between their resonant extremes. PCB dimensions are assumed to be negligible compared to the cable dimensions.

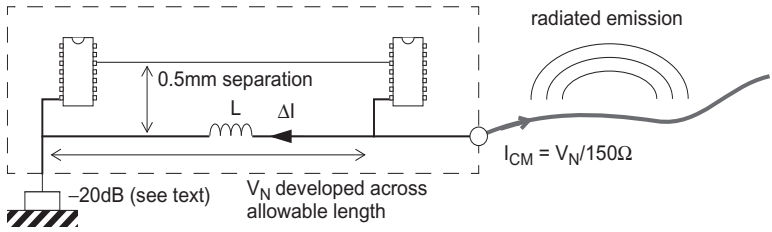


Figure 13.4 Common mode emission model

Track length implications

The inductance L is crucial to the level of noise that is emitted. In the model, it is calculated from the length of a 0.5mm wide track separated from its signal track by 0.5mm, so that mutual inductance cancellation reduces the overall inductance. Table 13.2 tabulates the resulting allowable track lengths versus clock frequency and logic family as before, for a radiated field strength corresponding to the EN Class B limits.

This model should not be taken too seriously for prediction purposes. Too many factors have been simplified: cable resonance and impedance variations with frequency and layout, track and circuit resonance and self-capacitance, and resonance and variability of the coupling path to ground have all been omitted. The purpose of the model is to demonstrate that logic circuit emissions are normally dominated by common mode factors. Common mode currents can be combated by:

- ensuring that logic currents do not flow between the ground reference point and the point of connection to external cables;
- filtering all cable interfaces to a “clean” ground as discussed in section 12.2.4.1;
- screening cables with the screen connection made to a “clean” ground;
- minimizing ground noise voltages by using low inductance ground layout, preferably involving a ground plane.

Table 13.2 shows that the maximum allowable track length for the higher frequencies and faster logic families is impracticable (fractions of a millimetre!). Therefore one or

Table 13.2 Common mode emission: allowable track length

Logic family	t_r/t_f ns	ΔI mA	Track length cm at clock frequency 4MHz	10MHz	30MHz	100MHz
4000B CMOS @ 5V	40	6	180	75	–	–
74HC	6	20	8.5	3.2	1	–
74AC	3	80	1.0	0.4	0.14	0.05
74F	3	80	1.0	0.4	0.14	0.05
FPGA (example)	2.1	16	3.6	1.4	0.47	0.22

Allowable track length for 30dBμV/m 30MHz–230MHz, 37dBμV/m 230–1000MHz at 10m; cable length = 1m; layout: parallel 0.5mm tracks 0.5mm apart (2.8nH/cm)

Working: take 74HC family with F_{clk} = 10MHz as an example. Worst case is at 90MHz (9th harmonic).

From equation (11.14), for a field strength E of 30dBμV/m and 1m cable length, I_{CM} must be 2.8μA.

From $V_N = I_{CM} \cdot 150$ and including 20dB coupling attenuation, V_N = 4.18mV.

Fourier analysis of the source current using section D.8 on page 527 with $(t + t_r)/T = 0.5$, T = 100ns, t_r = 6ns and I = 20mA gives $I_{(9)}$, the current at the 9th harmonic, as 0.826mA.

Now from $L = V_N/2 \cdot \pi \cdot f \cdot I_{(9)}$, the allowable inductance across which V_N will be developed at $I_{(9)}$ and 90MHz is 8.95nH which at 2.8nH/cm gives **3.19cm** allowed.

a combination of the above techniques will be essential to bring such circuits into compliance. Clearly, if the ground reference point in Figure 13.4 is moved to be adjacent to the cable interface, no driving voltage is developed and the cable becomes benign. This is the purpose of the interface clean ground structure discussed in section 12.2.4.1. With this approach the common mode emissions are due only to common mode currents flowing directly in the PCB tracks. This is not to say that the effects of such currents are negligible – with high clock frequencies and unshielded products PCB common mode radiation can certainly present a real problem. The model can be modified to remove the cable and instead treat the PCB itself as the radiating structure. This has been explored in section 11.2.1.2.

13.1.3 Emissions from digital circuits

13.1.3.1 Choice of logic family

The damage as far as emissions are concerned is done by switching edges which have a fast rise or fall time (note that this is not the same as propagation delay and is rarely specified in data sheets; where it is, it is usually a maximum figure). Using the slowest risetime compatible with reliable operation will minimize the amplitude of the higher-order harmonics where radiation is more efficient. Figure 13.1 (earlier) shows the calculated harmonic amplitudes for a 15MHz clock with risetimes of 1 and 7ns. An improvement approaching 20dB is possible at frequencies around 400MHz by slowing the risetime.

The advice based on this premise must be, use the slowest logic family that will do the job; don't use fast logic when it is unnecessary. FPGAs with selectable slew-rate and drive strength interfaces should be programmed for the lowest and slowest values that work. Treat with caution any proposal to substitute devices from a faster logic family, such as replacing 74HC parts with 74AC. Where parts of the circuit must

operate at high speed, use fast logic only for those parts and limit the areas over which fast clock signals are distributed. This doesn't just apply to clock circuits; data transmission is also subject to the same rules, and so slew-rate limiting on data bus drivers is always a good practice for EMC purposes.

The graph in Figure 13.5 shows the measured harmonic envelope of a 10MHz 50% duty cycle waveform for three devices of different logic families in the same circuit. Note the emphasis in the harmonics above 200MHz for the 74AC and 74F types.

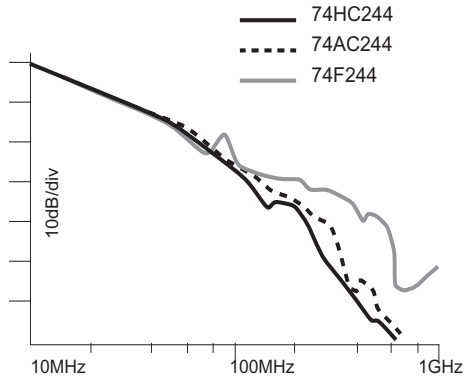


Figure 13.5 Comparison of harmonic envelopes of different logic families

13.1.3.2 Dealing with clock emissions

All the energy in clock signals is concentrated at a few specific frequencies, with the result that clock signal levels are 10–20dB higher than the rest of the digital circuit radiation (see for example Figure 13.7, which shows emissions dominated by a 40MHz clock oscillator). Since the commercial radiated emissions standards do not distinguish between narrowband and broadband, these narrowband emissions should be minimized first, by proper layout, grounding and buffering of clock lines. Then pay attention to other broadband sources, especially data/address buses and backplanes, and video or high-speed data links.

As good practice, don't drive circuits with clock signals whose transition times are substantially faster than is necessary. Certainly there are many high-performance circuits where clock timing is critical and transitions must be as fast as possible, but this is not universal. Where circuit constraints allow it, you should deliberately slow clock edges to minimize harmonic generation. This can be done in three ways: series

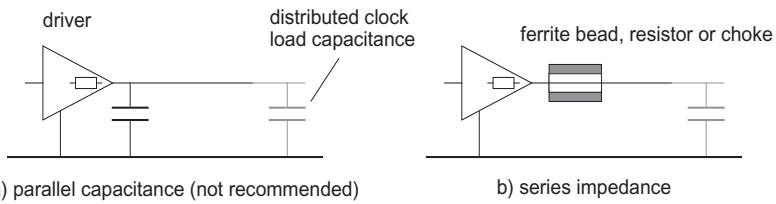


Figure 13.6 Controlling the clock edges

impedance, parallel capacitance or by using a low-performance buffer. Figure 13.6 shows the first two of these. Generally, slugging the clock output with a parallel capacitor is undesirable because although it has the desired effect of reducing the dv/dt feeding into the clock line, it increases the capacitive loading on the driver and hence increases the di/dt drawn from its supply pins; the overall effect may be to worsen the emissions rather than improve them.

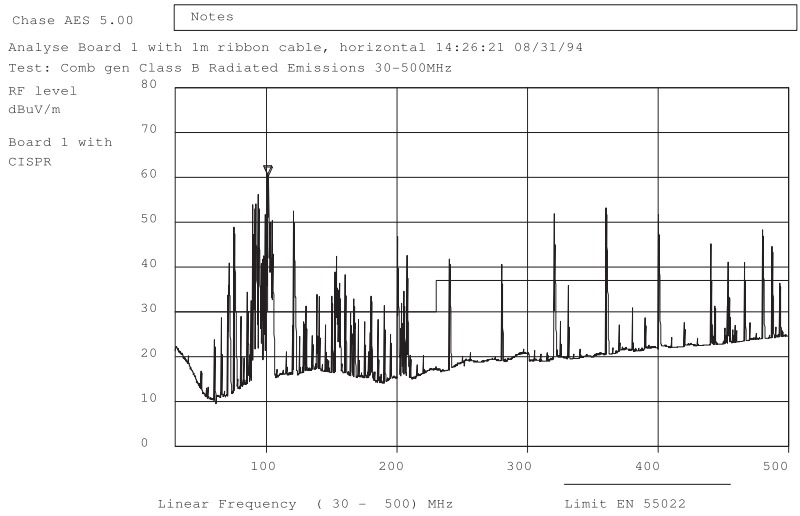


Figure 13.7 Typical emissions plot showing clock harmonics at multiples of 40MHz

Stopper resistors

It is preferable to increase the series impedance of the driver output at the harmonic frequencies, and this can best be done with a low-value resistor or a small ferrite impeder in series with the output (Figure 13.6(b)). Low-loss inductors are less helpful as they tend to introduce ringing.

In fact, series resistors can be used all over a digital board, not just in clock lines; they are small enough (0603 or 0402 size, or in multiple arrays) to be used in every line of an entire bus. The effect of a series resistor of 22–47Ω at the output of each pin of a bus driver or parallel interface on an ASIC or FPGA, in concert with the distributed capacitance of the bus line being driven, is to damp the bus edge transitions as well as to reduce the V_{CC} spikes by isolating the driver output from the capacitance. All that is necessary is to ensure that the extra resistance does not cause a violation of timing constraints. You needn't restrict use of these resistors just to the expected high frequency lines; the effect of internal bond wire inductance is to infect the chip's own power rails with clock noise (see the section on ground bounce, 13.3.3) and this means that even low frequency output port drivers can be carriers of this noise away from the device. Series stopper resistors near the device output pins will prevent it from propagating far across the board.

Silicon-level solutions

Some IC manufacturers are addressing the problem of RF emissions at the chip level – Ramdani [124] gives a review of efforts over the last few decades. By careful attention to the internal switching regime of VLSI devices, noise currents appearing at the pins can be minimized. The transition times can be optimized rather than minimized for a given application [75]. Asynchronous design can be combined with spectrum spreading using multi-frequency clocks [113]. Alternatively, the shape of the waveform can be "softened" at the transition from the rising/falling edge to the steady-state high or low value, to reduce higher order harmonic content without substantially reducing the logic switching speed. Revised package design and smaller packages can allow the decoupling capacitor (section 13.1.4) to be placed as close as possible to the chip, without the internal leadframe's inductance negating its effect; also, the reduction in operating silicon area gained from shrinking silicon design rules can be used to put a respectable-sized decoupling capacitor (say 1nF) actually on the silicon with a small series impedance to the package power pin [121].

Spread-spectrum clock generation

One possible alternative to damping the clock edges is a technique more often associated with radio transmission, known as spread-spectrum clock generation. In this technique the clock oscillator itself is frequency modulated by 0.5–2% by a waveform selected for the most even spectral spreading, which was patented in 1996 by Lexmark [82]. This results in a wider distribution of the spectral energy associated with each clock harmonic (Figure 13.8) so that the level measured in a constant bandwidth falls, it is claimed by between 10 and 20dB [102]. This is achieved without any extra effort in layout and without slowing the clock risetimes. The technique can be applied not only to digital clocks but to any periodic signal, and has been used in switchmode power supplies to the same end.

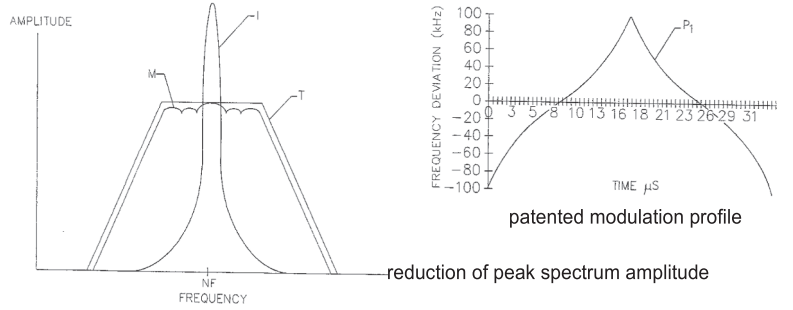


Figure 13.8 The effect of spectrum spreading [82]

The clock frequency will exhibit some jitter and therefore the technique may be restricted in applications which need accurate clock timing such as telecoms or video products, but there will undoubtedly be many applications which can use it, and it may present itself as an attractive "quick fix" for remedial work. One effect of spreading both up and down in frequency is potentially to violate clock set-up and hold constraints – a higher instantaneous frequency giving tighter and perhaps unacceptable timing margins. This can be avoided by "down-spreading", that is, by modulating only downwards in frequency [57]. Another problem is caused by PLL (phase-lock-loop) frequency multiplication, which may not track the desired modulation waveform accurately, so that the expected best-case reduction in emissions does not materialize.

The technique does not reduce the total amount of radiated energy, and it has been compared to "getting rid of a cow-pat by stamping on it". To an EMC purist, of course, it should not be regarded as a substitute for proper design and layout in the first place.

Backplanes and daughter board connectors

Buses that drive several devices or backplanes which themselves drive several boards, carry much higher switching currents (because of the extra load capacitance) than circuits which are compact and/or lightly loaded. Products which incorporate a backplane are more prone to high radiated emissions. A high-speed backplane should always use a multilayer board with a ground plane, and daughter board connectors should include a ground pin adjacent to every high-speed clock, data or address pin (Figure 13.9). If this is impractical or too expensive, multiple distributed ground returns

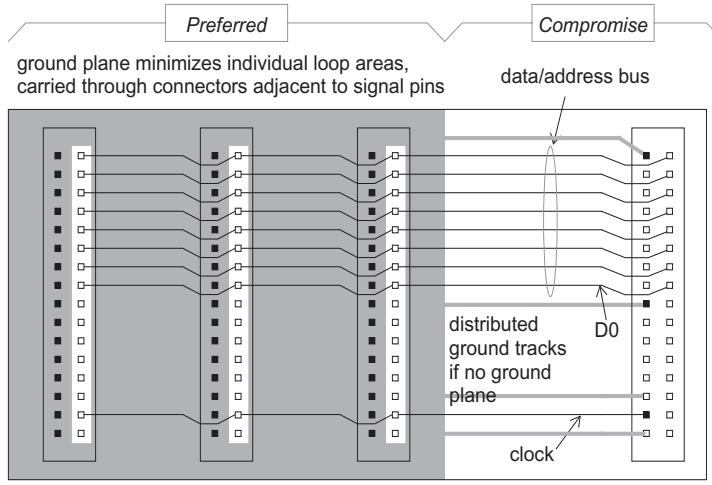


Figure 13.9 Backplane layout

can be used to minimize loop areas, although this is at best a second-rate compromise and would be unacceptable if transmission line distribution were involved. The least significant data/address bit usually has the highest frequency component of a bus and should be run closest to its ground return. Clock distribution tracks must *always* have an adjacent ground return. The capacitive load on the clock signal at each daughter board should be minimized, by having a buffer on the board for local clock distribution.

Low-voltage logic families

The move from 5V V_{CC} to 3.3V and below has been driven by the need to reduce the speed-power product of digital circuits, so that more performance can be squeezed out of a lower power consumption. The consequences for EMC performance are equivocal. On the one hand, reducing the voltage swing of external clocks and data might seem to make these signals less aggressive for RF emissions. But actually, RF emissions are driven much more by circuit currents than voltages.

In fact, the trend to lower supply voltages has merely meant that circuits have become faster and more complex for the same power consumption. Current drain from the supply rails has not reduced and can sometimes be greater, and this is associated

with increased supply ΔI noise which is a serious RF emissions hazard itself. In addition, the need for multiple low-voltage supplies (often 3V3, 2V5 and 1V8 are all needed for a single VLSI device) causes headaches for PCB layout when several power planes need to be referenced to the same 0V plane. All in all, moving to low-voltage logic has not made the EMC designer's life any easier.

13.1.3.3 I_{CC} at switching

The Fourier components of the output waveform are not the only reason for emissions of logic circuits. The current taken from the V_{CC} pin or pins related to driving the outputs has two components (Figure 13.10):

- the current fed through the output pin(s) and used for charging or discharging the capacitance of each output node;
- the current taken through the totem-pole output stage transistors as they switch, which is not passed through the output pin(s).

Both of these components appear at the V_{CC} pin and have to be decoupled, as discussed shortly in section 13.1.4. The second component does not involve the signal tracks and is not affected by the quality of their routing or layout, but it is significantly affected by the nature of the decoupling regime and the use of ground and power planes. Logic families which control this unwanted current (sometimes known as "delta-I" noise) are widely touted as helping to reduce emissions overall.

As well as the current associated with output drivers, there are similar supply currents associated with switching each of the internal nodes of the device. Individually, these are orders of magnitude less than the output drive currents. But in a typical VLSI device there are many more of them: thousands at least, and if as is usual the device is synchronously clocked then all of the current transitions occur simultaneously, at the clock edge. Therefore the supply current of any VLSI device tends to be dominated by short but potentially high amplitude current spikes ("simultaneous switching noise") at each transition of the internal clock.

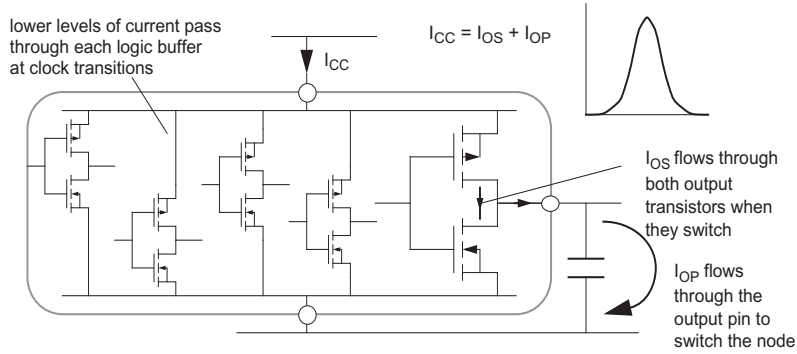


Figure 13.10 Current through the supply pin

Since these current transients are almost entirely a function of the internal design of the chip, it is for this reason as much as any other that different manufacturers' implementations of a particular device may be noisier or quieter in EMC terms. In any case, their existence ensures that high quality supply decoupling (section 13.1.4) in conjunction with carefully laid out power and 0V planes is essential for VLSI devices.

13.1.3.4 Ringing on transmission lines

If you transmit data or clocks down long lines, these must be terminated to prevent ringing. Ringing is generated on the transitions of digital signals when a portion of the signal is reflected back down the line due to a mismatch between the line impedance and the terminating impedance. A similar mismatch at the driving end will re-reflect a further portion towards the receiver, which gets reflected back again; and so on. Severe ringing will affect the data transfer, by causing spurious transitions, if it exceeds the device's input noise margin.

Aside from its effect on noise margins, ringing may also be a source of radiated interference in its own right. The amplitude of the ringing depends on the degree of mismatch at either end of the line while its frequency depends on the electrical length of the line (Figure 13.11). The question is, at what length does a PCB track or cable need to be treated as a transmission line? This can be determined by whether the logic transition has been completed before the reflected part has made the round trip. A digital driver-receiver combination should be analysed in terms of its transmission line behaviour if:

$$2 \times t_{PD} \times \text{line length} > \text{transition time} \qquad (13.2)$$

where t_{PD} is the line propagation delay in ns per unit length [106], which depends on the dielectric constant of the board material and can be calculated from section D.5.3 (Appendix D)

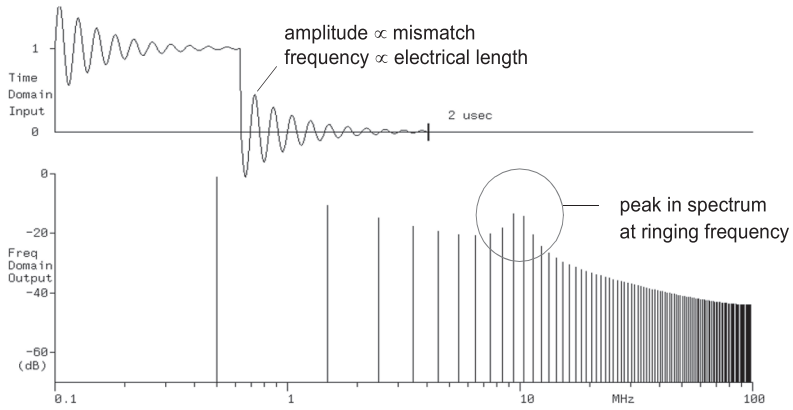


Figure 13.11 Ringing due to a mismatched transmission line

This means matching the track's characteristic impedance to the source and load impedances, and may require extra components to terminate the line at either end. Most digital circuit data and application handbooks include advice and formulae for designing transmission line systems in fast logic. Table 13.3 is included as an aid to deciding whether the particular circuit you are concerned with should incorporate transmission line principles.

13.1.4 Digital circuit decoupling

No matter how good the V_{CC} and ground connections are, they will introduce an impedance which will create switching noise from the transient switching currents

Table 13.3 Critical transmission line length

Logic family	t_r/t_f ns	Critical line length
4000B CMOS @ 5V	40	3 m
74HC	6	45 cm
74AC	3	22.5 cm
74F	3	22.5 cm
FPGA (example)	1.4	10.5 cm

Line length calculated for dielectric constant = 4.5 (FR4 epoxy glass), t_{PD} = 0.067ns/cm

taken from the V_{CC} pins, as shown in Figure 13.10. The purpose of a decoupling capacitor is to maintain a low dynamic impedance from the individual IC supply voltage to ground. This minimizes the local supply voltage droop when a fast current pulse is taken from it, and more importantly it minimizes the lengths of track which carry high di/dt currents.

13.1.4.1 Component placement

Placement is critical; unless you are using closely spaced power and ground planes, the capacitor must be tracked close to the circuit it is decoupling (see Figure 13.12). "Close" in this context means less than half an inch for fast logic such as 74AC, especially when high-current devices such as bus drivers are involved; if all outputs on an octal bus buffer are heavily loaded and the output changes from $\#FF_H$ to $\#00_H$ or vice-versa, a current pulse which may exceed an amp passes through the supply pins. For low-current, slow devices such as 4000B-series CMOS the requirement is more relaxed; but a typical ASIC or FPGA clocked at hundreds of MHz needs a capacitor adjacent to every pin, most conveniently mounted on the opposite side of the board under the package.

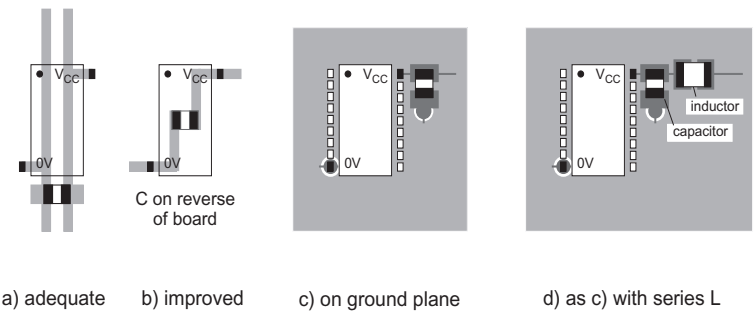


Figure 13.12 Decoupling capacitor positioning

When decoupling capacitors are used with closely spaced (<0.5mm) power and ground planes, as is usual for high-performance multilayer designs, the positioning of

individual capacitors is less important [87], although considerations of current distribution still have an effect as discussed later (page 348). It is still normal to use at least one capacitor next to each power pin, but the effect of the planes is to act as a transmission line that interconnects all such decouplers in parallel, with the minimum of inductance. Therefore rather than primarily decoupling individual devices, capacitors on a power/ground plane network are effectively reducing the impedance between the planes over as wide a bandwidth as possible. Thus the injected high frequency currents shown in Figure 13.10 create the lowest possible noise voltages.

To achieve this each decoupler should present a low impedance especially at high frequency where the planes' dimensions become a significant fraction of a wavelength. Since this is normally in the region of a few hundred MHz for typical sized boards, the capacitor is operating mainly above self resonance, and its actual contribution is to place its small package inductance in parallel with all the others on the plane. So, as we are about to see, the component's inductance is its most important characteristic. It will have lowest impedance at its self resonant frequency, which will be determined by its value and its mounted self-inductance: $F_{res} = 1/(2\pi\cdot\sqrt{L_m\cdot C})$, where L_m is the sum of the package inductance and the PCB mounting inductance due to pads and vias to the planes. For instance, a 2200pF 0402 capacitor with about 1.5nH of total mounted inductance will have a SRF of 88MHz. At this frequency, it should be closer to the device it is decoupling than roughly one tenth of a quarter wavelength [22] in order that the propagation delay across the planes in the x-y direction should not effectively detach the capacitor from the device pin: in FR4 glass fibre PCB material, this example corresponds to $(300/88)/(40\cdot\sqrt{4.5}) = 4$cm. Generally, then, capacitors can be placed around the periphery of a device rather than directly underneath it with little compromise. Larger value capacitors with lower SRFs are even less critical.

13.1.4.2 Component selection

This leads to the view that the crucial factor when selecting capacitor type for high-speed logic decoupling is package or lead inductance rather than absolute value. Minimum inductance offers a low impedance to fast pulses. Multilayer chip ceramics (MLCCs) are preferred. The overall inductance of each connection is the sum of package, via and track inductances.

Capacitance value for low frequencies

The minimum required value for the total parallel capacitance on a particular power rail depends on the specification for tolerable supply voltage deviation, and can be calculated from equation (13.3):

$$C_{min} = \Delta I \cdot \Delta t / \Delta V \qquad (13.3)$$

ΔI and Δt can to a first order be taken from the figures in Table 13.1 and Table 13.2 while ΔV depends on your judgement of permissible supply voltage droop at the capacitor. Typically on a 5V supply a power rail droop of 0.25V is reasonable; lower voltage supplies can tolerate less droop. For an octal buffer taking 50mA per output and switching in 6ns, the required capacitance for $\Delta V = 0.25$V is 9.6nF. For smaller devices and faster switching times, less capacitance is required, and often the optimum capacitance value is as low as 1000pF. Then, the question becomes how many discrete capacitors of what value do you need to provide this total with the least overall inductance; invariably, more small packages are better than a few large ones.

In low voltage (1.2–2.5V) circuits you can get high value (tens or even hundreds of μF) multilayer ceramic components in small packages, which have the advantage of low inductance and low series resistance, and are very useful as bulk decouplers as long as you take note of the issue of voltage coefficient of capacitance (see page 408).

PCB layout for high frequencies

Assuming you will be using small chip multilayer ceramics, with a package size of 0402, 0603 or at a pinch 0805, the determining factor for the inductance of the component is its PCB layout. This is illustrated in Figure 13.13(a). The two pads of the capacitor are taken through vias to the relevant power and ground planes; the length of the pad-to-via track, separation of the vias, the dimensions of the capacitor and the height of each via between the surface pad and the buried plane determine the overall loop area, and hence the inductance contributed by the layout. If the board costs are to stay low then the vias should not be placed in the pads, so layout 'X' is the best option. If you double the vias (b), a plus and minus pair close to each other on each side of the component pads, this nearly halves the overall inductance and is largely cost-free. Archambeault [24] has pointed out that an even better approach is to create a doublet of a pair of capacitors with alternating opposing vias (c): here the proximity of the plus and minus pair of vias is used to best effect.

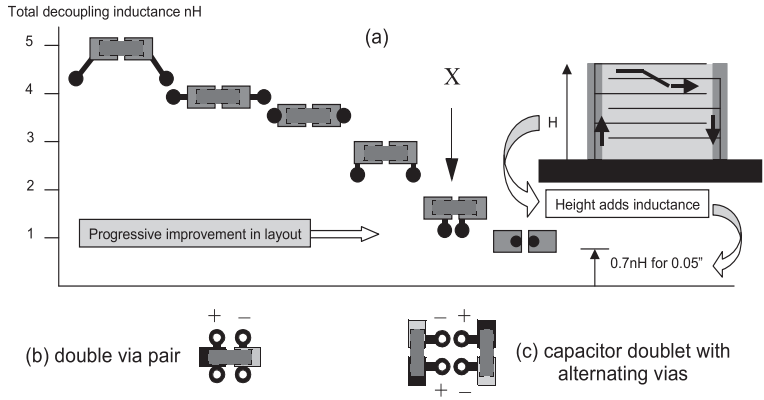


Figure 13.13 Decoupling capacitor track layout for minimum inductance

Multiple capacitors

The practice of applying decoupling capacitors of different values in parallel is controversial. Historically, you would use a high-value component for low frequencies and a low value for high frequencies. Unfortunately this creates a situation where the inductance of the larger component causes a parallel resonance, and hence a high impedance, with the capacitance of the smaller component. So in the mid-frequencies there is an unwanted peak in the decoupling impedance. One approach with modern MLCCs is to use a substantial number of fairly high-value components of the same value (say 1μF) across the board in the smallest package available. On the other hand, a high equivalent series resistance (ESR) at resonance is desirable to damp the related impedance peaks; having many capacitors with the same self-resonance reduces the overall ESR and leads to higher impedance at the peaks [23]. Choosing multiple values

maintains a higher ESR over a wider frequency range, improving the damping. But this means a larger spread of components in the bill of materials. In the end, simulating the decoupling effects across frequency with all parasitics included is the most appropriate design method.

13.1.4.3 Use of series inductors

Liberally sprinkling a board with decoupling capacitors of itself is not necessarily the optimum decoupling regime. This is particularly the case where one or a few ICs contribute the majority of the offending di/dt through its or their power supply pins. A common example of this situation is the use of a single-chip microprocessor with no other digital components; the main emissions source is the clock harmonic current drawn through the V_{CC} pin. When one decoupling capacitor is positioned close to the processor package and others are positioned elsewhere on the board, intuitively it would seem that the noise currents would flow through and be decoupled by the adjacent capacitor. But in practice this is not so at all frequencies; the inductance of the interconnecting tracks forms a series tuned circuit with the remote decoupling capacitors, and at the resonant frequencies the noise currents flowing towards the remote capacitors are actually greater than if these capacitors were not present [58]. This in turn causes worse emissions at these frequencies when the capacitors are added. (Note that this is an issue for situations with a ground plane but no power plane; it doesn't arise when a power plane is implemented between the various decoupling points.)

This effect can be seen through analysis of the equivalent circuit (with arbitrary but representative circuit values, see Figure 13.14). The interference source is modelled as a frequency-dependent voltage generator feeding the decoupling and power rail network through a source impedance of 100Ω. The "far end" of the power supply is modelled as a 10μF capacitor in parallel with 1Ω. C1 is the local decoupling capacitor, C2 is the remote one placed across another IC, represented in the model by 100Ω. (This value of resistance may or may not represent the RF impedance of an IC; it has little effect on the analysis.) Track and parasitic inductances are included as shown. The two results graphs show the decoupling effect represented as a ratio of I_{IN} to I_1 or I_2 in the appropriate length of track, l_1 or l_2. The left-hand graph shows the decoupling effect in the track between C1 and C2; the right-hand graph shows the effect downstream of C2. The solid line shows the decoupling due to the equivalent circuit as drawn, with the dotted lines representing various changes.

Analysing the results of the model

Firstly, the resonant peaks in both graphs' solid lines giving an *increase* in track current (dB > 0) at the lower frequencies are evident. This is due to resonances of L_{+1} and L_{+2} with C1 and C2. Above 100MHz all the curves are flat, the actual decoupling being determined by the ratio of inductances since all capacitors are operating above self-resonance. The action happens mostly in the 2–50MHz frequency range.

Increasing C1 from 4n7 to 47n improves the decoupling in the mid-frequencies at the expense of shifting the resonances lower – which is not necessarily undesirable. Because it is a larger capacitor, the self-inductance L_{LD1} has been increased to 10nH and this worsens the decoupling above 20MHz.

Removing C2, although worsening the decoupling of track l_2, in fact improves the decoupling of l_1 above 7MHz, because the impedance looking into l_1 has increased and therefore less noise current flows into it. The impact of this modification in real life will

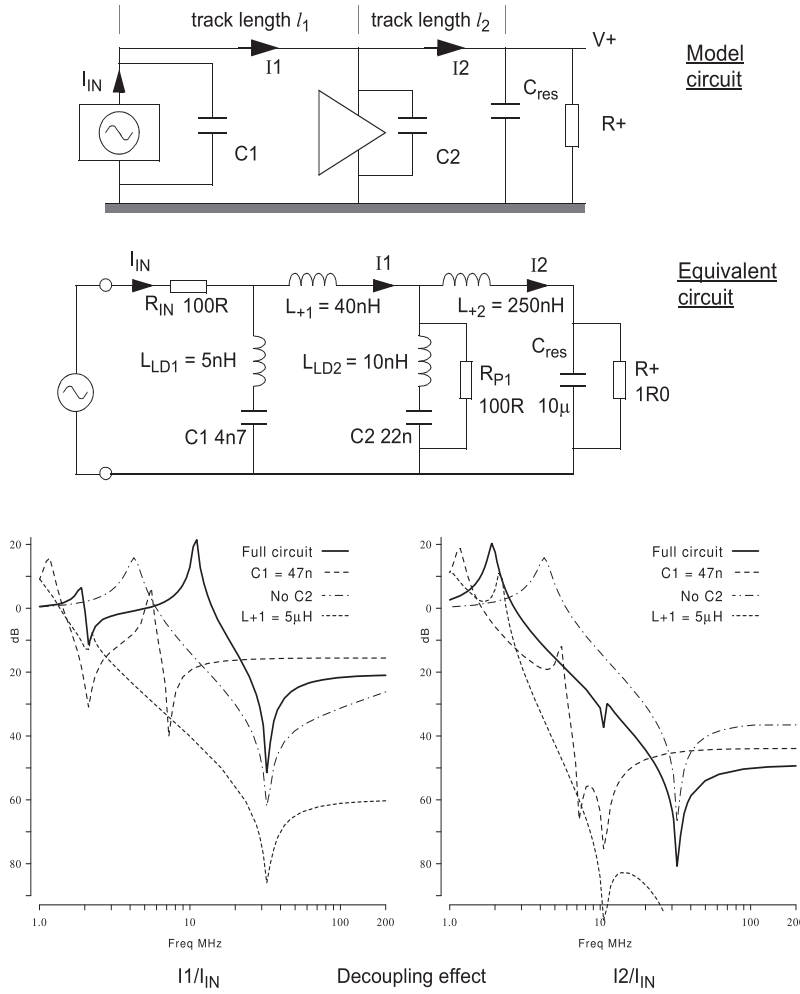


Figure 13.14 Analysing the decoupling equivalent circuit

depend on the relative significance of l_1 and l_2 in actually coupling interference out of the PCB.

The most substantial improvement occurs when the impedance looking into l_1 is significantly increased. This can only be achieved by inserting a discrete inductor; in the analysis the inductor value was chosen to be 5µH, but this is fairly arbitrary. You can see that the improvement over most of the frequency range is in the order of 40dB. This technique is shown schematically in Figure 13.12(d). As a general design rule, you should plan to include such series inductors (available as chip ferrite components, taking up negligible extra space at low cost) in the $+V_{CC}$ feed to every IC which is expected to contribute substantially to the noise pollution of the supply lines.

Power planes

The addition of series inductors is still helpful even where both power and ground planes are used. In fact segmentation of power planes, as discussed in section 12.2.3.3, requires their use to partition one plane segment from another.

If the board has four or more layers and includes a power plane, the inductance of the coupled planes can be so low that resonances are not analysed according to the simple model above. However, the placement of decoupling capacitors on a ground-and-power plane board is still significant. Imagine the situation shown in Figure 13.15: an IC is mounted somewhere on a four-layer board and it injects I_{CC} noise into the power and ground planes (A). Elsewhere on the board is a decoupling capacitor.

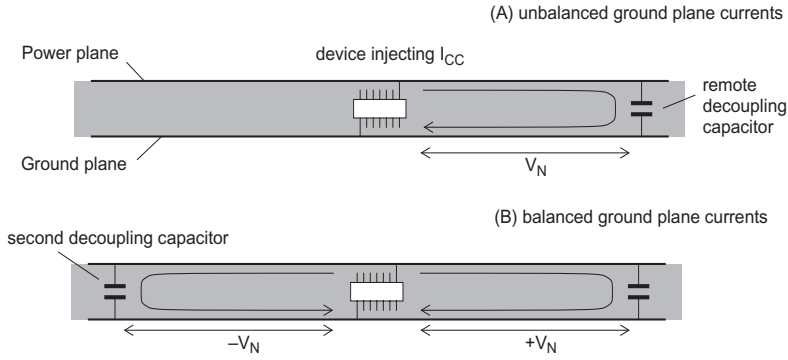


Figure 13.15 Unbalanced power/ground plane currents

The current that flows between the two will create a noise voltage across the impedance of the plane and this noise voltage will create common mode emissions, as discussed in section 11.2.1, either by direct radiation from the board or by coupling to connected cables. If, now, another identical capacitor is placed in a position on the board exactly opposite from the first (B) with respect to the device, then a counterbalancing noise voltage $-V_N$ is produced. The effect of this will be to cancel the first $+V_N$ and to reduce the board's emissions.

This requires careful layout of decoupling capacitors to balance the board, which is impractical for all but the simplest of circuits, but distributing many capacitors across the board, particularly close to high-noise devices, will control the resulting currents as far as possible. As discussed earlier, keeping the smallest capacitors within one tenth of a quarter wavelength (at their self-resonant frequency) from the pin they are decoupling will minimize the distance that the decoupling currents must travel through the power planes [22]. But a more practical solution is found by separating the relevant power and ground planes as little as possible; a layer separation of 0.1mm or less will create a very small "spreading inductance" offered by the plane structure between the decoupling capacitors and the IC power pins, so that the unavoidable inductance of the capacitors and vias dominates, and the positioning of the capacitors becomes less important.

The I_{CC} noise from individual devices can be reduced by applying ferrite impeders in each supply connection as shown in Figure 13.12(d). It is also helpful to group together devices that generate high I_{CC} noise so that their interconnections are short, they can all use a small segmented power plane, and to keep them away from the edge of the board.

13.1.5 Analogue circuits: emissions

In general analogue circuits do not exhibit the high di/dt and fast risetimes that characterize digital circuits, and are therefore less responsible for excessive emissions. Analogue circuits which deliberately generate high frequency signals (remembering that the emissions regulatory regime currently begins at 150kHz, and in some cases lower than this) need to follow the same layout, decoupling and grounding rules as already outlined. It is also possible for low frequency analogue circuits to operate unintentionally outside their design bandwidth.

13.1.5.1 Instability

Analogue amplifier circuits may oscillate in the MHz region and thereby cause interference for a number of reasons:

- feedback-loop instability;
- poor decoupling;
- output stage instability.

Capacitive coupling due to poor layout and common-impedance coupling are also sources of oscillation. Any prototype amplifier circuit should be checked for HF instability, whatever its nominal bandwidth, in its final configuration and layout. Feedback instability is due to too much feedback near the unity-gain frequency, where the amplifier's phase margin is approaching a critical value. It may be linked with incorrect compensation of an uncompensated op-amp.

Even if the circuit doesn't actually oscillate, near-instability is a potential problem for RF immunity as well: if the amplifier circuit is configured so that its stability margin is reduced at frequencies of a few MHz, then RF coupled into the circuit at around these frequencies will be much more likely to cause susceptibility effects (see section 13.3.6.4).

13.1.5.2 Decoupling

Power supply rejection ratio falls with increasing frequency, and power supply coupling to the input at high frequencies can be significant in wideband circuits. The higher the impedance of the power rails, the more unwanted coupling exists.

This is cured by good decoupling at the power supply pins, but typical 0.01–0.1μF decoupling capacitors may resonate with the parasitic inductance of long power leads in the MHz region, so decoupling-related instability problems may show up in this range. Paralleling a low-value capacitor with a 1–10μF electrolytic capacitor will drop the resonant frequency and stray circuit Q to a manageable level. The electrolytic's series inductance could resonate with the ceramic capacitor and actually worsen the situation, although the losses of a few ohms in series with the capacitor usually prevent this, which is a good argument for not using low-loss 1–10μF ceramic capacitors in this case. The input stages of multi-stage high gain amplifiers may need additional resistance or a ferrite bead in series with each stage's supply to improve decoupling from the power rails.

13.1.5.3 Output stage instability

Capacitive loads cause a phase lag in the output voltage by acting in combination with the op-amp's open-loop output resistance (Figure 13.16). This increased phase shift reduces the phase margin of a feedback circuit, possibly by enough to cause oscillation. A typical capacitive load, often invisible to the designer because it is not treated as a

component, is a length of unterminated coaxial cable. Until the length starts to approach a quarter-wavelength at the frequency of interest, coax looks like a capacitor: for instance, 10m of the popular RG58C/U 50Ω type will be about 1000pF. Other capacitive loads may be found within a system, such as a sample-and-hold circuit, but the cable load is of special interest for EMC since it couples any RF oscillations both into and out of the circuit. To cure output instability, decouple the capacitance from the output with a low-value series resistor, and add high frequency feedback with a small direct feedback capacitor C_F which compensates for the phase lag caused by C_L [30]. When the critical frequency is high a ferrite chip is an acceptable substitute for R_S.

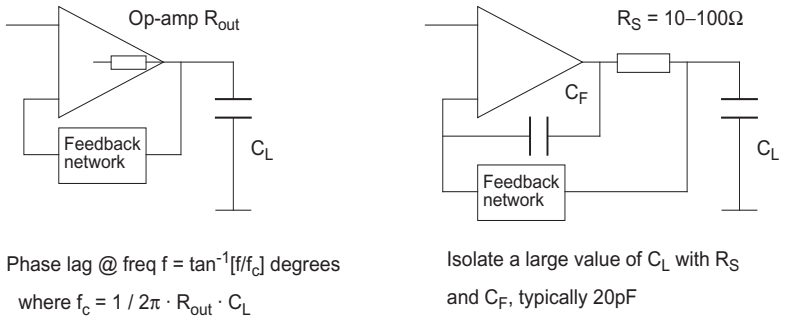


Figure 13.16 Instability due to capacitive loads

13.2 The power switching converter

Power switching circuits of all sorts present extreme difficulties in containing generated interference [156]. Typical switching frequencies of 50–500kHz and their harmonics can be emitted by both differential and common mode conducted and radiated mechanisms. Switching waveform asymmetry normally ensures that both odd and even harmonics are present. A further cause of broadband noise may be due to reverse recovery transitions of rectifier diodes.

If the fundamental frequency is stable and well controlled then a spectrum of narrowband emissions is produced which can extend well beyond 30MHz when the waveform transition times are fast. A measurement bandwidth of 9kHz means that individual harmonics can be distinguished if the fundamental frequency is greater than about 20kHz. Self-oscillating converters will normally show drift or modulation due to input or output ripple, which has the effect of broadening individual harmonic lines so that an emission "envelope", rather than narrowband harmonics, is measured. This is also observed with variable on-time converters, which don't have a fixed frequency, such as many PFC stages.

13.2.1 Switchmode off-line mains supplies

Figure 13.17 shows a typical flyback mains switching supply with the major emission paths marked; topologies may differ, or the transformer may be replaced by an inductor but the fundamental interference mechanisms are common to all designs.

13.2.1.1 Magnetic field radiation from loops and components

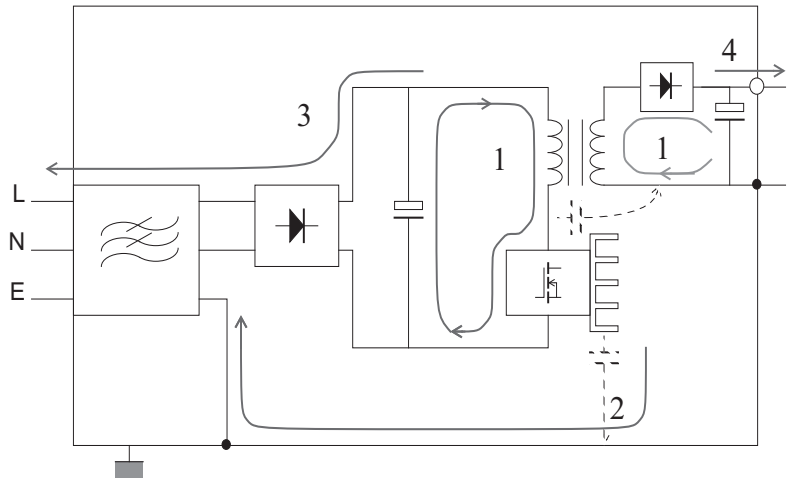
Near field magnetic radiation isn't usually a problem for commercial products, since most standards (there are exceptions) don't have a test for this, but it can be a significant issue for those that do, and for military applications (see section 7.2.1.2). If the field is from a loop which is carrying a high di/dt it can be controlled by reducing the loop area or by reducing di/dt. With low output voltages, the output rectifier and smoothing circuit may be a greater culprit in this respect than the input circuit. Loop area is a function of layout and physical component dimensions. di/dt is a trade-off against switching frequency and power losses in the switch. It can to some extent be controlled by slowing the rate-of-rise of the drive waveform to the switch.

Unfortunately, the trend towards minimizing power losses and increasing frequencies goes directly against the requirements for low EMI. The lowest harmonic content for a given frequency is given by a sinewave: sinusoidal converters (such as the series resonant converter, see [115]) or other variants with reduced risetimes have reduced EM emissions.

Magnetic component construction and layout

As explained in section 15.1.2, screening will have little effect on the magnetic field radiation due to this current loop, although it will reduce the associated electric field. The transformer (or inductor) core should be in the form of a closed magnetic circuit in order to restrict magnetic radiation from this source. A toroid is the optimum from this point of view, but may not be practical because of winding difficulties or power losses; if you use a gapped core such as the popular E-core type, the gap should be directly underneath the windings since the greatest magnetic leakage flux is to be found here.

To reduce the external flux, another method is to surround the outside of the component with a single turn of copper tape; sometimes known as a "flux band", the



1: H-field radiation from high di/dt loop
2: Capacitive coupling from high dv/dt node to E
3: Differential mode voltage passed via DC link to L/N
4: Conducted and/or radiated on output

Figure 13.17 Switching supply emission paths

effect of this is for the induced current in the "shorted turn" to partially cancel the flux outside the assembly. The copper tape doesn't have to be connected to anything for this to work, but if it is connected to the same DC 0V as the outer winding it will also then act as an electric field screen for that winding as per the next section.

Fast risetimes together with high currents can extend direct radiation both from the switching magnetics and from the circuit structure to above 30MHz, making it a threat for the radiated emissions test. Additionally, such radiation can couple into the output or mains leads and be responsible for conducted interference (over 150kHz to 30MHz) if the overall power supply layout is poor. You should always keep any wiring which leaves the enclosure well away from the transformer or inductor. In the same vein, coupling between the various inductors in the supply needs to be avoided. Stray leakage flux from the transformer could couple with the core of the input filter choke, which would dramatically reduce the attenuation of the filter of which it is a part. Both magnetic and electric field coupling should be considered.

13.2.1.2 Capacitive coupling from high dv/dt nodes

High dv/dt at the switching point (normally the drain of the switching transistor) will couple capacitively to ground and create common mode interference currents. The solution is to minimize dv/dt and coupling capacitance, or provide a preferential route for the capacitive currents through appropriate screening (Figure 13.18).

dv/dt is reduced by a snubber and by keeping a low transformer leakage inductance and di/dt. These objectives are also desirable, if not essential, to minimize stress on the switching device, although they increase power losses. The snubber capacitor is calculated to allow a defined dv/dt with the maximum load as reflected through the transformer; the series resistor must be included to limit the discharge current through the switching device when it switches on. You can if necessary include a diode in parallel with the resistor to allow a higher resistor value and hence lower switching device ratings.

Capacitive screening

Capacitive coupling is reduced by providing appropriate electrostatic screens, particularly in the transformer and on the device heatsink; in mains-to-low-voltage SMPSs these are the two points which carry the highest dv/dt. Note the proper connection of the screen: to either supply rail, which allows circulating currents to return to their source, not to the safety earth. On the transformer, an external foil screen connected to 0V (as above) will also reduce coupling of the high dv/dt on the outside of the winding to other parts of the circuit.

Even if the transformer is not screened, its construction can aid or hinder capacitive coupling C_{IW} from primary to secondary, and from the whole component to others around it (Figure 13.18(b)). Separating the windings onto different bobbins reduces their coupling capacitance but increases leakage inductance. Coupling is greatest between nodes of high dv/dt; so the end of the winding which is connected to V_{CC} or 0V can partially screen the rest of the winding in a multi-layer design, hence the need to ensure that ends A and D in Figure 13.18(b) are adjacent, not ends B and C. Conversely, if you don't take care with this aspect, the transformer can be the major contributor to the circuit's emissions.

If the transformer design is less than optimum but cannot be improved, a capacitor $C_{P\text{-}S}$ across primary and secondary 0V nodes is another solution. This acts as a capacitive potential divider working against the internal interwinding capacitance, so

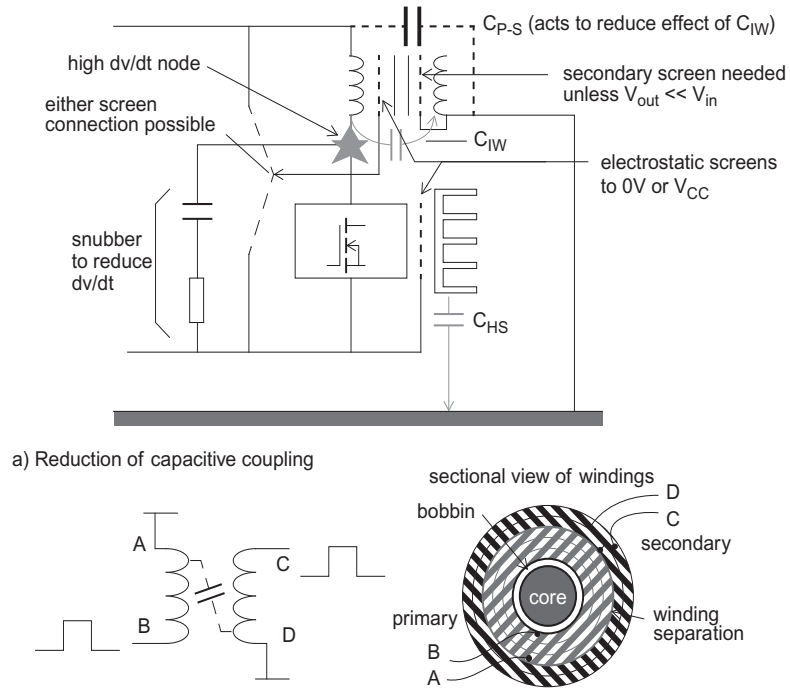


Figure 13.18 Common mode capacitive coupling

that if, for instance, the interwinding capacitance C_{IW} is 100pF and C_{P-S} is 1000pF, a 20dB reduction in emissions through this route should be expected. In a mains supply C_{P-S} is of course across the transformer isolation barrier, so needs to be safety rated and its value is restricted by leakage concerns; and it is critical to place it with the least parasitic lead and track inductance to the transformer pins.

Switching device heatsinks can also cause problems. Particularly with high-voltage transistors, the switching node (drain of a MOSFET or collector of an IGBT) carries the highest dv/dt but is closely coupled to the heatsink tab: indeed with a non-isolated device it *is* the heatsink tab. Even with iso-tab devices there is a significant capacitance within the package between the semiconductor and the tab, which appears in series with the heatsink-to-chassis capacitance C_{HS}. This means that the one structure that you should never bolt the heatsink to is the equipment chassis or case, since this references the switching noise directly to the ground node, shorting C_{HS}. Unfortunately, that is usually the most thermally desirable place for it.

If at all possible, the metalwork of the heatsink should be directly connected to the 0V rail of the circuit that uses it. When this is impossible, for thermal or safety reasons, there should be an electrostatic screen between the device and the heatsink, connected to this 0V; thermal pads with integral foil screens are available for this purpose. Where this is also impossible, aim for the minimum achievable coupling capacitance between

the switching node and the heatsink and expect to have to work harder at (and pay more for) your filter design.

An alternative which is becoming increasingly attractive is to use a thermal polymer [131] for a heatsink. Although these materials are less thermally conductive than aluminium or copper this may be less of a disadvantage if the cooling is more by convection than conduction, and there are potential advantages in manufacturability. From the EMC point of view the capacitances discussed above can be reduced 5–10 times, which if the heatsink is the dominant coupling path can be extremely helpful.

Physical separation of parts carrying high dv/dt is desirable [156] although hard to arrange in compact products. Extra electric field screening to 0V of the offending component(s) with small formed tinplate parts is an alternative.

13.2.1.3 The 20MHz peak

When you compare conducted emissions plots from many different types of product using switchmode mains power supplies, a striking similarity emerges – the existence of a peak in the emissions profile in the region of 5–25MHz, usually consisting of high order harmonics of the switching circuit [153]. If the level of the peak is below the limit then no further effort is needed, but if it is higher there usually follows a protracted period of remedial work, much of which appears to have no useful effect. Why is there this hump, and why is it so hard to get rid of?

The answer is best illustrated by discussing the overall testing equivalent circuit. There are four significant parts to this circuit (Figure 13.19(a)): the equipment under test including a mains filter, its mains lead, the LISN, and the ground plane. The EUT internal circuit is represented by a common mode noise source V_N referred to the chassis and capacitively feeding the common mode equivalent of the mains filter (see Figure 14.27 on page 416). This capacitance C_{CM}, equivalent to the combination of C_{IW}, C_{HS} and any other strays as discussed above, is the main coupling path in most cases. Good circuit design and layout minimizes its value, since it forms a potential divider with C_Y, the combined phase-to-earth capacitance of the mains filter. In most portable apparatus C_Y's maximum value is set by safety earth leakage current limits.

All but the simplest filters will include a series common mode choke L_{CM}. In this equivalent circuit a single winding is shown, since live and neutral are regarded as identical and with the conventional double-wound compensating choke the two windings reinforce each other. An important extra parasitic component is the winding capacitance C_W which appears across the windings and limits the HF performance of the choke.

Provided that it is electrically short (less than a tenth of a wavelength at the highest frequency, i.e. 1m for 30MHz) the transmission line effects of the mains cable can be ignored and the cable is represented by its lumped equivalent. This is dominated by the self-inductance due to the cable length, modified by mutual inductance and capacitance between the conductors – shown by capacitor C_{CBL} and a transformer of self-inductance L_{CBL} and mutual coupling M.

The EUT capacitance to the ground plane C_S is also part of the overall circuit. For EUTs with no safety earth it is a crucial part of the common mode coupling path. Any safety earth connection appears in parallel with this capacitance.

With this description of the equivalent circuit, it is easy to see that there are several possibilities for resonances to occur, and the main problem is determining which components are critical. The easiest way to do this is to model the circuit. The results of the modelling (Figure 13.19(b), solid line) show exactly the sort of hump in the

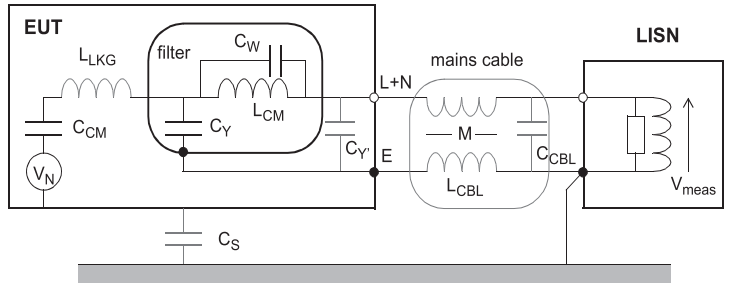


Figure 13.19 Analysis of the high frequency emissions peak

transfer function from V_N to the LISN measurement point that is expected from the measurements. The self-capacitance of the choke is in fact resonating with the inductance of the mains cable, which of course is affected by cable length and the mutual coupling of the wires within it. The effect of this resonance can be substantially reduced by placing a subsidiary Y-capacitor of 100–470pF on the mains cable side of the common mode choke, $C_{Y'}$ in Figure 13.19(a). This lowers the frequency of the resonance and reduces its amplitude.

Another more difficult problem occurs when the common mode coupling path *within* the EUT includes some inductance. This is again quite usual with switched-mode power supplies when the inductance in question is the leakage inductance of the switching transformer. The leakage inductance L_{LKG} appears in series with C_{CM} as shown in Figure 13.19(a) and the resonant frequency of this combination is typically in the 5–10MHz range. Figure 13.19(b), dotted line, shows the effect of this L_{LKG} resonance. Depending on the relative values of the stray capacitances and inductances, the resonant peaks may be separate as shown or may overlap and reinforce each other.

Because C_{CM} is also largely determined by the interwinding capacitance of the transformer, the resonance is impossible to shift without a complete redesign of the transformer – usually not an option by the time EMC compliance testing of an established design is being carried out. Deliberately increasing the leakage inductance – although undesirable for other reasons – may have a beneficial effect on the amplitude of the coupling peak. An alternative is to include a small extra C-M choke before C_Y in the mains filter.

13.2.1.4 Ringing on the switching edge

Resonance in the coupling path is not the only reason for peaks in the emissions profile. Another very common feature is the presence of ringing on the switching waveform – analogous to that described in section 13.1.3.4, but with a different cause. Although you would like to think that the switching waveform would be a nice clean trapezoid, in practice the switching transient (usually the turn-off edge) excites parasitic resonances in the circuit so that the waveform looks more like Figure 13.20. The faster the edge, the more energy is expended in the ringing, which couples out in both common and differential mode and causes a peak in the spectrum. Because the ringing frequency is affected by parasitic capacitance in the switching devices, which can change dynamically over the switching period, the actual peak may be more or less spread out – Figure 13.20 shows a particularly poorly designed circuit (without a supply filter) with a high degree of stray circuit inductance and capacitance and hence a low frequency and sharp peak.

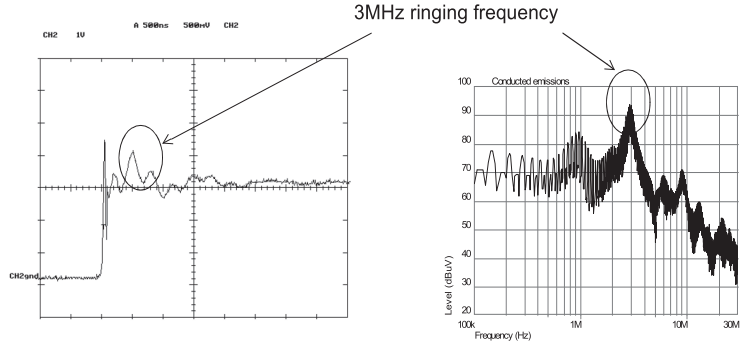


Figure 13.20 A ringing SMPS waveform and its spectrum

The diagnosis of this type of peak is quite simple since you can observe the ring on an oscilloscope and correlate its frequency with that in the spectrum. Fixing it is harder. Parasitics, which include transformer leakage inductance and circuit layout inductance, should be kept to a minimum, and the switching edge speed should be limited as far as possible without compromising efficiency. The circuit should include snubbers at appropriate high dv/dt nodes; on occasion a series ferrite in the switching current path may be effective.

13.2.1.5 Differential mode interference

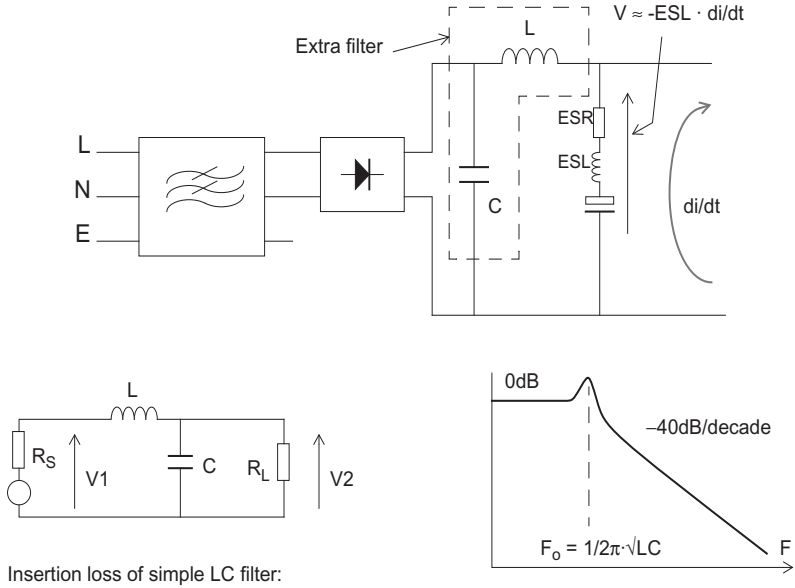
Differential mode interference is caused by the voltage developed across the finite impedance of the input reservoir capacitor with high di/dt due to the switched power. It is often the dominant interference source for the lower switching harmonics. Choosing a capacitor with low equivalent series impedance (ESL and ESR) will improve matters, but it is impossible to obtain a low enough impedance in a practical capacitor to make generated noise negligible.

Extra series inductance and parallel capacitance on the input side will attenuate the voltage passed to the input terminals. A capacitor on its own will be ineffective at low frequencies because of the low source impedance. Series inductors of more than a few tens of microhenries are difficult to realize at high DC currents (remembering that the inductor must not saturate at the peak ripple current, which is much higher than the DC

average current), and multiple sections with smaller inductors may be more effective than a single section. When several parallel reservoir capacitors are used, one of these may be separated from the others by the series inductor; this will have little effect on the overall reservoir but will offer a large attenuation to the higher frequency harmonics at little extra cost.

Figure 13.21 demonstrates filtering arrangements. The LC network may also be placed on the input side of the rectifier. This will have the advantage of attenuating broadband noise caused by the rectifier diodes switching at the line frequency. The mains input filter itself (see section 14.2.3) will offer some differential mode rejection. It is also possible to choose switching converter topologies with input inductors such as the Cúk circuit [66] which reduce, but do not eliminate (because of stray capacitance across the inductor), fast di/dt transitions in the input and/or output waveforms.

When you are testing the performance of a differential mode filter, be sure always to check it at the maximum operating input power. Not only do the higher switching currents generate more noise, but the peak mains input current may drive the filter inductor(s) into saturation and make it ineffective.



Insertion loss of simple LC filter:

$$|V1/V2| = |1 - \omega^2 LC \cdot (R_L/(R_S + R_L)) + j\omega \cdot ((CR_LR_S + L)/(R_S + R_L))|$$

For low R_S and high R_L this reduces to

$$|V1/V2| = |1 - \omega^2 LC|$$

Standard mains impedance R_L is approximated by 100Ω//100μH differentially, which approaches 100Ω above 300kHz. Note resonance of LC at low frequency gives insertion gain

Figure 13.21 Differential mode filtering

Diagnosing differential versus common mode

A handy feature when you are investigating SMPS noise is to observe the effect of load current on the noise profile. Generally – and of course there are exceptions – differential

mode noise increases with load, whereas common mode noise doesn't. This is due to the different sources: principally, differential mode noise is due to ripple across the reservoir capacitor, which increases as more current is drawn through it. Common mode noise is dominated by the switching voltage edge dv/dt, which to a first order doesn't change from light load to full load. You may of course use the technique described in Figure 7.12 on page 158 for more precision, but varying the load is usually a quick and easy first call.

13.2.1.6 Output noise

Switching spikes are a feature of the DC output of all switching supplies, partly because of the finite impedance of the output reservoir (differential mode) and partly due to coupling across the isolating transformer interwinding capacitance (common mode). Such spikes are conducted out of the unit on the output lines, and may re-radiate onto other leads or be coupled to the ground connection. A low-ESL reservoir capacitor is preferable, but good differential mode suppression can be obtained, as with the input, with a high frequency L-section filter. 20–40dB is obtainable with a ferrite bead and 0.1µF capacitor above 1MHz.

Common mode spikes will be unaffected by adding a differential output filter, and this is one way to diagnose which mode of interference is being generated. To remove common mode interference, place a capacitor of 1–10nF between the primary and secondary 0V nodes ($C_{P\text{-}S}$ in Figure 13.18): this acts as a voltage divider against the transformer's interwinding capacitance, but its applicability is limited by the isolation voltage and leakage capacitance requirements you face from input to output. If this is inadequate or impossible, a common mode choke on the output lines before the power supply feeds any load circuit is another option.

The abrupt reverse recovery characteristic of the output rectifier diode(s) can create extra high frequency ringing and transients. When the current is switched from conduction to blocking through a conventional diode there is a short period during which the forward current continues to flow, which causes undesirable extra power dissipation in the diode. Fast recovery diodes are especially designed to minimize the period during which this happens, to allow higher switching speeds. Unfortunately, although these devices don't take long to recover, when they do, there is a very fast switch-off of the remanent current – in other words, a high di/dt. This excites the resonant circuits created by the transformer secondary leakage inductance along with stray capacitance, leading to high levels of emission at the resonant frequencies which may, for smaller supplies, be in the hundreds of MHz, modulated by the switching frequency. These can be attenuated by choosing soft recovery diodes or by paralleling the diodes with an RC snubber.

13.2.2 DC-DC converters

Onboard small point-of-load DC-DC converters are rarely a threat for conducted emissions since they are well decoupled from the interfaces. Nevertheless, you need to reduce the radiating effect of the main switching current as much as possible. If you could reduce the PCB area of this current loop to zero, and use an ideal DC link capacitor with zero RF impedance, there would be no problem. In reality, optimizing the layout for minimum radiating efficiency is necessary. This means that you need to identify the path of the most significant switching current.

The "hot loop", i.e. the one creating the highest emissions, is that which carries the highest di/dt, that is the loop in which the current is switched instantaneously (or at least, as fast as possible) between zero and its peak. Figure 13.22(a) and (b) identify this

loop for the common buck and boost configurations; you can similarly identify one or two high di/dt loops for any other configuration. Note that generally this loop does not include the switching inductor, whose purpose (as long as the converter is operating in continuous mode) is to maintain continuity of current flow during the switching cycle. Its current waveform is usually a triangle shape rather than an abrupt switch on-off.

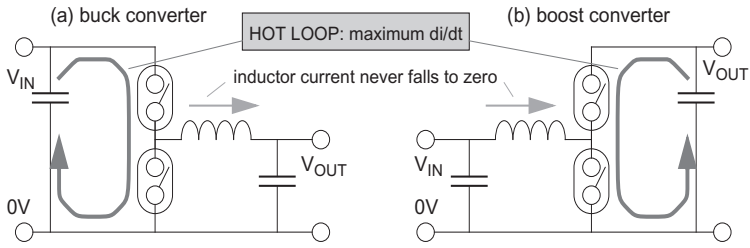


Figure 13.22 DC-DC converter current loops

One or other capacitor generally does carry the hot loop current, and this must be chosen both for small size (to reduce loop area) and low parasitic impedance. The total loop area includes both this capacitor and the relevant switching circuit. Well-designed ICs with integrated switches help here; for emissions, a power supply with integrated synchronous switches, optimized pinout and careful internal switch control will outperform a non-synchronous power supply IC with an external Schottky diode, and both will outperform a design with external MOSFETs [95].

A simple and effective way to reduce both the inductance of the "hot loop" and its radiating efficiency, is to use an image plane (see section 15.1.5) directly underneath the loop traces. This plane is typically connected to a GND (0V) plane but it doesn't have to be. The magnetic flux caused by the source currents in the loop induces an opposing current in the adjacent surface of the plane. The net effect is to largely cancel the external radiated magnetic field and nearly eliminate the overall inductance of the loop; the closer this image layer is to the source layer, the better the cancellation. On a four-layer board, if the source traces are on the top side then layer 2 can carry the image plane and may be no more than 0.2mm below it, with a thin layer separation.

Because the image plane carries induced currents it will have a noise voltage developed across it, and therefore it should not be part of the main 0V plane; but in the x-y direction it must be continuous across the whole area of the loop that it is covering. It can be connected at one point to the 0V plane (on another layer) of the rest of the board, or it can be cut out of the 0V plane on the same layer – but you should make sure that no other traces on other layers cross the break in the plane.

The driver stage

After the main switching current, the next source of trouble for emissions is often the MOSFET gate drive signal(s), for circuits that use switches external to the controller. Power MOSFETs and IGBTs will have several nanofarads gate-to-source capacitance – the higher the power level, the larger the FET area and the higher its gate capacitance; and the gate-to-drain capacitance adds to this (the "Miller" capacitance) especially when the device switches a high voltage on its drain.

To drive this capacitance to switch the device on and off in a short time, requires a gate drive current in the order of amps:

$$I = C \cdot dV/dt \qquad (13.4)$$

That is, to switch a gate capacitance of 5nF through 5V in 10ns requires 2.5 A. The di/dt in this drive path can be as significant as that in the main switching circuit. Therefore, you need to minimize the current loop which provides this signal. Figure 13.23 gives an example for the case of a half-bridge switching circuit where two MOSFETs are switched from two drivers in one package. All these loops will benefit from a small PCB area, coupled with an image plane beneath the traces.

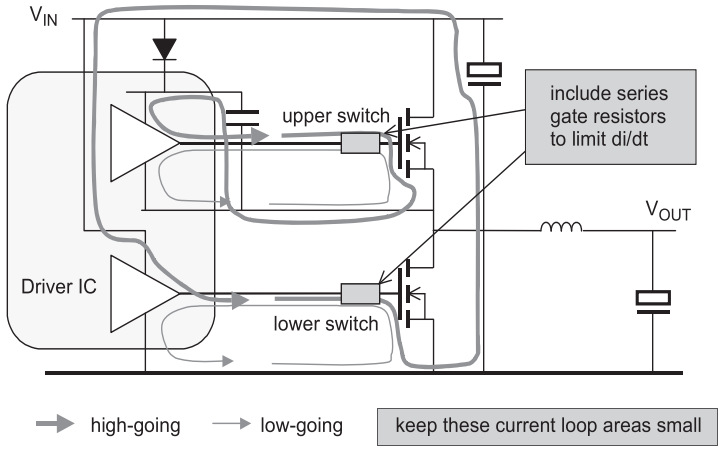


Figure 13.23 External MOSFET (or IGBT) drive circuit

A further technique to control the impact of fast drive signals, is to incorporate a small impedance in series with each gate. This acts in concert with the gate capacitance to limit the peak drive current into the MOSFET. Clearly there is a compromise to be made here, since too much impedance will slow the switching time unacceptably and create too much dissipation in the converter. Even so, including a series component in the prototype PCB layout is good practice; it can be populated initially with a low value resistor, for instance 2.2 ohms, and this can be adjusted during the initial evaluation to find the optimum balance between low emissions and low losses.

13.2.3 Other power switching circuits

The SMPS is not the only source of RF emissions, although because of its ubiquity it is the one most frequently encountered in the mains conducted emissions test. Any circuit that switches significant power at frequencies from tens to several hundreds of kHz is also capable of serious RF emissions problems, such as:

- variable speed motor drives;
- Class D audio power amplifiers;
- high-efficiency lighting converters;
- grid-connected power inverters for renewable energy;
- phase angle or PWM lighting and heating controllers.

The underlying purpose of all these circuits is to switch power as fast as possible so that minimum power is lost in the switching device, usually a power MOSFET, IGBT or triac. As we have seen, the faster the switching the greater the content of high frequency harmonics in the switched current and voltage waveforms.

Invariably in the design of these products you must take similar precautions to those outlined above for power supplies. Both the differential and common mode coupling paths must be addressed; nodes of high dv/dt must be identified and designed to minimize capacitive coupling, and loops of high di/dt must be identified and designed to minimize inductive coupling. PCB track layout must take into account the threat of common impedance coupling and you will often need to implement single point grounding schemes to deal with this. Both differential and common mode filtering will always be needed and such filters are often a significant cost, size and weight overhead. A generalized equivalent circuit for any of these products is as shown in Figure 13.24, from which you can see that although the main switching circuit may be coupled to the output of the device through filters, there is still potential for a substantial interference signal to be radiated from the output leads or conducted back through the mains.

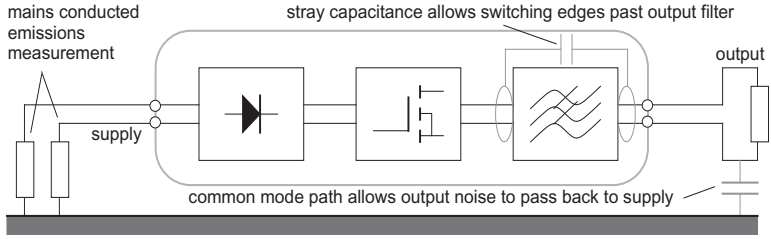


Figure 13.24 General power switching circuit

The installation issues raised by these types of systems are briefly covered in section 16.5.

13.3 Design for immunity

Because the microprocessor is a state machine, processor-based circuits are prone to corruption by fast transients which can cause the execution of false states. Great care is necessary to prevent any clocked circuit (not just microprocessor-based) from being susceptible to incoming transient interference. Analogue signals are more affected by continuous interference, which is rectified by non-linear circuit elements and causes bias or signal level shifts. The immunity of analogue circuits is improved by minimizing amplifier bandwidth, maximizing the signal level, using balanced configurations and electrically isolating I/O that will be connected to "dirty" external circuits.

13.3.1 Digital circuits: interference paths

Most of the critical interference in microprocessor circuits is ground-borne, whether it is common mode RF or transients, and the damage is done by conversion of common mode ground noise to differential mode noise at susceptible signal nodes. This occurs because of a high transfer impedance from common to differential mode due to poor PCB layout. Differential mode interference will not propagate far into the circuit from the external interfaces. Therefore, firstly you should concentrate on the best PCB tracking to minimize mode conversion – that is, appropriate use of a ground plane; and secondly, lay out the physical design of the product to keep ground interference

currents away from the logic circuits. Filter the I/O leads or isolate them, to define a preferential safe current path for interference. Radiated RF fields that generate differential mode voltages internally are dealt with in the same way as differential RF emissions, by minimizing circuit loop area with a ground plane, and by restricting the bandwidth of susceptible circuits where this is feasible.

13.3.1.1 Interference paths – transients

A typical microprocessor-based product, including power supply, operator interface, processor board, enclosure and external connections can be represented at high frequency [28] by the layout shown in Figure 13.25. The 0V rail will appear as a network of inductances with associated stray capacitances to the enclosure. If it's a 0V plane, then the inductances will be lower but the total capacitance will be higher: the principle is the same. An incoming common mode transient current on the mains can travel through the circuit's 0V rail, generating ground differential spikes as it goes, through any or all of several paths as shown (observe the influential effect that stray capacitance has on these paths):

1: primary-to-secondary capacitance through the power supply to 0V, through the equipment and then to case;

2: as above, but then out via an external connection;

3: direct to case, then via stray capacitance to 0V and out via an external connection.

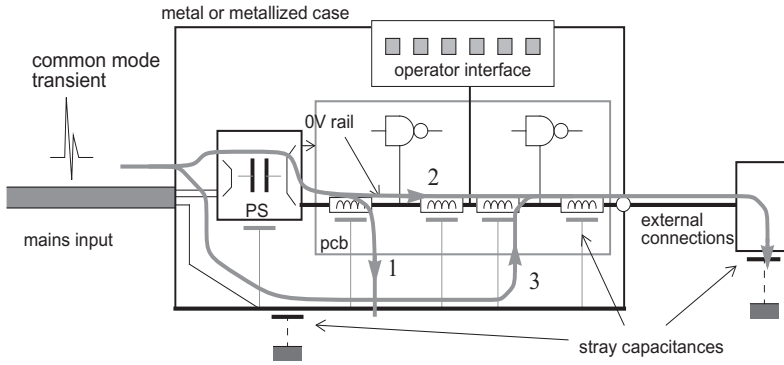


Figure 13.25 Representative high frequency equivalent circuit: transients

If there are no external connections, (1) is the only problem and can be cured by a mains filter and/or by an electrostatic screen to the case in the mains transformer. (2) arises because an external connection can provide a lower impedance route to ground than case capacitance. You cannot control the impedance to ground of external connections, so you have to accept that this route will exist and ensure that the transient current has a preferential path via the case to the interface which does not take in the circuit. This is achieved by ensuring that the case structure is well bonded together, i.e. it presents a low impedance path to the transient, and by filtering interfaces to the case at the point of entry/exit (see section 14.2.4 and section 12.2.4).

If the enclosure is non-conductive then transient currents have no choice but to flow through the circuit. Local grouping of interfaces is essential, as well as liberal use of

common mode chokes to increase the impedance seen looking from the circuit into the interface cable; but you must expect some interference current to flow through the PCB itself, so a low transfer impedance for minimum mode conversion, as provided by a good 0V plane, is also needed. Effectively, removing the protection offered by a metal case places much greater stress on the PCB design.

With external connections, route (3) can actually be *worsened* by a mains filter, since at high frequencies parts of the enclosure can float with respect to true ground; and in the transient immunity test the safety earth line is tested as much as the live and neutral. The case, which is connected directly to the safety earth and/or indirectly at high frequencies to live and neutral through the mains filter, will have transients injected straight into it. Even large conducting structures may exhibit high impedances, and hence voltage differentials when subjected to fast transients. The safety earth – the green and yellow wire – is *not* a reference point at HF (refer back to section 12.1.3.4 for ground wire impedances) and if this is the only "earth" connection, the voltage on the enclosure is defined by a complex network of inductances (connected cables) and stray capacitances which are impossible to predict. To deal with this problem when the case is inadequately grounded, the external interfaces must be filtered to the case, just as is the power supply input. This provides a preferential path through the metalwork, and not through the circuit, for the interference currents. This will also deal with transient currents injected via the interface.

In all cases, grouping all I/O leads together with the supply lead (see Figure 12.26 on page 327) will offer low-inductance paths that bypass the circuit and prevent transient currents from flowing through the PCB tracks.

13.3.1.2 Interference paths – ESD

An electrostatic discharge can occur to any exposed part of the equipment. Common trouble spots as shown in Figure 13.26 are keyboards and controls (1), external cables (2) and accessible metalwork (3). A discharge to a nearby conductive object (which could be an ungrounded metal panel on the equipment itself) causes high local transient currents which will then also induce currents within the equipment by inductive or common impedance coupling.

Because there are many potential points of discharge, the possible routes that the discharge current can take are widespread. Many of them will include part of the PCB ground layout, via stray capacitance, external equipment or exposed circuitry, and the induced transient ground differentials will cause maloperation in the same way as described above for injected common mode transient bursts. The discharge current will take the route (or routes) of least inductance. If the enclosure is well bonded to ground then this will be the natural sink point. If it is not, or if it is non-conductive, then the routes of least inductance will be via the connecting cables.

In any case, it will be hard to prevent the PCB from being subject to some of the ESD currents, and the di/dt of an ESD stress is the highest of all EMC phenomena (10^9 or 10^{10} A/s), so this again puts the emphasis on good layout practice to minimize mode conversion transfer impedance. If the edge of the PCB may be exposed, as in card frames, then a useful trick is to run a "guard trace" around it, unconnected to any circuitry, and separately bond this to ground to act as a sacrificial path.

When the enclosure consists of several conductive panels then these must all be well bonded together, following the discussion of section 15.1 for shielded enclosures. All metallic covers and panels must be bonded together with a low impedance connection (metal-to-metal, <2.5mΩ at DC) in at least two places. If this is not done

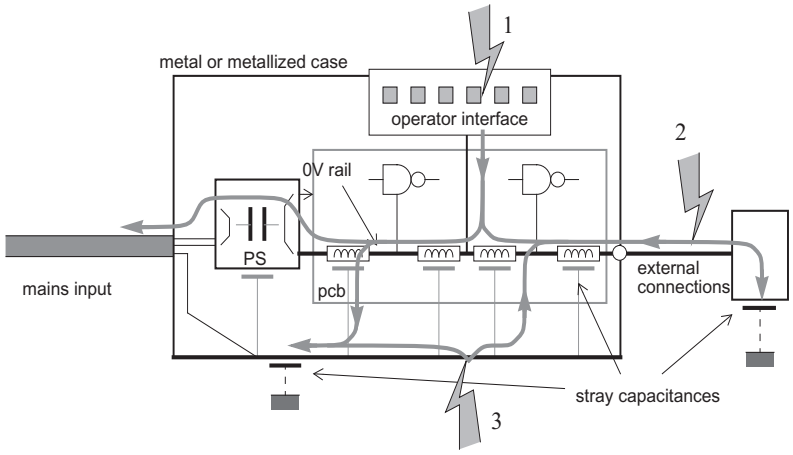


Figure 13.26 Representative high frequency equivalent circuit: ESD

then the edges of the panels will create very high transient fields as the discharge current attempts to cross them. If they are interconnected by lengths of wire, the current through the wire will cause a high magnetic field around it which will couple effectively with nearby PCB tracks.

The discharge edge has an extremely fast risetime (sub-nanosecond, see section 11.3.3.1) and so stray capacitive coupling is essentially transparent to it, whilst even short ground connections of a few nH will present a high impedance. For this reason the presence or absence of a safety ground wire (which has a high inductance) will often make little difference to the system response to ESD.

13.3.1.3 Transient and ESD protection

Techniques to guard against corruption by transients and ESD are generally similar to those used to prevent RF emissions, and the same components will serve both purposes. Specific strategies aim to prevent incoming transient and ESD currents from flowing through the circuit, and instead to absorb or divert them harmlessly past the circuit (Figure 13.27).

Design for no discharge

To achieve this for ESD the simplest approach is not to allow an ESD event to happen at all, which means that all around the EUT enclosure there are no accessible conductive surfaces. “Accessible” in this context means accessible to an air discharge from a finger charged to typically 8kV (15kV for certain more rigorous standards), so that creepage and clearance paths from the outside to circuits on the inside must also be considered. As a general rule, you should aim to maintain at least a 2–4mm clearance gap between tracks on the edges of PCBs and the inside surfaces of enclosures near apertures, and/or 8mm of creepage path from accessible outer surfaces to inner conducting parts, including conductive coatings on the plastic itself. For small products this can be quite a challenge. Particular issues arise with

- enclosure joints and ventilation slots: no metal chassis, or worse, internal plastic metallization, should be near them

- displays and LEDs: should not pass unprotected through holes in a plastic panel, but should be placed behind clear windows with adequate creepage distance around them
- keyboards: see below
- connectors: metal shells and pins should be recessed enough to be inaccessible, plug-in cards should preferably be protected by access doors or flaps

Designing to control the discharge path

If on the other hand you have a (grounded) metal case:

- keep all external interfaces physically near each other;
- filter all interfaces to ground at their point of entry;
- if this is not possible, isolate susceptible interfaces with a common mode ferrite choke or opto-couplers;
- use screened cable with the screen connected directly to the interface ground;
- screen PCBs from gaps in the enclosure or external discharge points with extra internally grounded plates.

For reduced ESD susceptibility, the circuit 0V needs to remain stable during the ESD event. A low inductance 0V network is essential, but this must also be coupled (by capacitors or directly) to a master reference ground structure to prevent significant transient potential differences between the structure and the circuit during an ESD event.

I/O cables and internal wiring may provide coupling paths for the current, just as they are routes into and out of the equipment for common mode RF interference. The

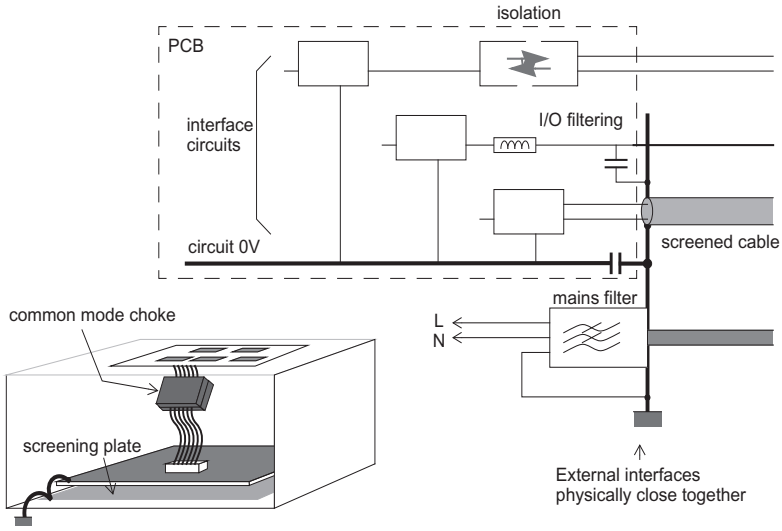


Figure 13.27 Transient and ESD protection

best way to eliminate susceptibility of internal harnesses and cables is not to have any, through economical design of the board interconnections. If you must have internal wiring, make sure that its routing is controlled to avoid close coupling with any particularly susceptible (or emissive) circuits. External cables must have their shields well decoupled to the ground structure, following the rules in section 14.1.7, that is 360° bonding of cable screens to connector backshells and no pigtails [135].

Insulated enclosures make the control of ESD currents harder to achieve, and a well-designed and low inductance circuit ground is essential. But, if the enclosure can be designed to have no apertures which provide air gap or creepage paths to the interior then no direct discharge will be able to occur, provided the material's dielectric strength is high enough. You will still need to protect against the field of an indirect discharge, though.

The operator interface

Keyboards present an operator interface which is frequently exposed to ESD. Keyboard cables should be foil-and-braid shielded which is 360° grounded at both ends to the low-inductance chassis metalwork. Plastic key caps will call for internal metal or foil shielding between the keys and the base PCB which is connected directly to the cable shield, to divert transients away from the circuitry; or the whole keyboard should be protected with an overall rubber sheet. The shield ground should be coupled to the circuit ground at the cable entry point via a 10–100nF capacitor to prevent ground potential separation during an ESD event. A membrane keypad with a polyester surface material has an inherently high dielectric strength and is therefore resistant to ESD, but for optimum protection it should incorporate a ground plane to provide a bleed path for the accumulated charge and to improve RF immunity.

This ground plane must be "hidden" from possible discharges by sealing it behind the membrane surface. The edges of ground plane layers are a potential problem with air discharge if there is only a short distance between the conductive plane and the edge of the membrane assembly. Such ground layers should have an insulating border around their edge of preferably 1cm or more to increase the creepage distance. The ground layer should be taken to chassis directly via a separate short, wide tab, not carried with the keypad signal connections (Figure 13.28). Membranes without such a ground plane must be regarded as unprotected interfaces, and all lines that connect to the driver on the PCB must be heavily RC filtered and preferably provided with ESD surge protection.

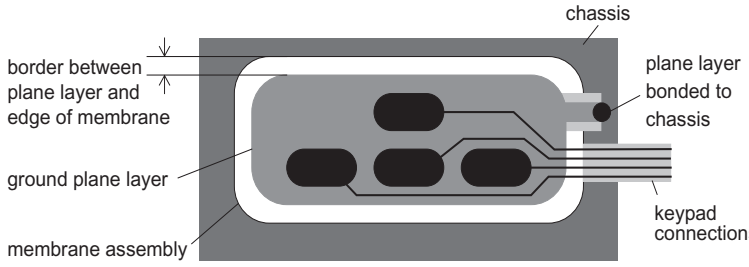


Figure 13.28 Ground plane on a membrane keypad

Transient protection devices for ESD

Interfaces may be particularly susceptible to an ESD event, if they are easily accessible to personnel, or can be approached by the connector of another device carrying a different charge level. In these circumstances transient protection of the interface pins is necessary. Section 14.2.5 discusses the general approach of transient suppressors; for ESD purposes the devices used must be very fast but do not have to carry or dissipate high energy levels. The peak current of an 8kV discharge event is only 30A for a nanosecond or so. The purpose of the suppressor is to divert current from a susceptible device to a suitable grounding structure, which will normally be the chassis but could be the interface power planes if these can carry the current without disrupting the circuit operation. This can be achieved either with a transzorb-type (zener or varistor) device, surface mounted and in a small (0805 or 1206) package for minimum lead inductance, or with steering diodes which dump the applied current into the 0V or supply rails. Figure 13.29 shows these methods.

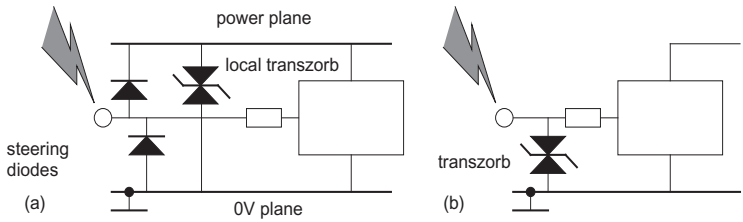


Figure 13.29 ESD interface protection

The problem facing designers who have to deal with high-speed digital interfaces such as the USB or DVI specifications, is that although the individual pins at these interfaces may need ESD protection there is precious little leeway for adding extra parallel components. Any such component will add capacitance, and for interfaces which employ matched transmission lines such capacitance is undesirable since it degrades the matching. This has led to the development of some very low-capacitance diodes specifically for the purpose; below 1pF is possible, which should satisfy most specifications. Transzorbs cannot be made with such low capacitances, but steering diodes can, since these are inherently operated with reverse bias under normal conditions. Packages with several diode pairs are available for protecting several lines at once.

In contrast to a transzorb, a steering diode does not absorb the ESD energy itself but merely diverts it to a local sink which can cope with the energy without harm. This means that such a sink must be explicitly provided: if it is not within the diode package itself then a separate transzorb and/or decoupling capacitor is necessary across the power and 0V planes near to the interface. This safely dissipates the current passed through the steering diodes without allowing a significant power voltage rise across the protected IC. Bearing in mind the fast rise time of the current pulse, there should be minimum inductance in the path created by the diodes, planes and transzorb.

13.3.1.4 Vulnerable circuit points

Power supply

CMOS devices are fairly tolerant of slow V_{CC} changes, but this is less true of other technologies. You should consider the possible impact of voltage dips, interruptions

and variations, and check the circuit operation under a wide variety of conditions. Surges can cause both short-term overvoltages and undervoltages on the supply rail. Voltages exceeding the devices' absolute maximum ratings (typically 7V for 5V supplies, or 4.6V for 3V3) can cause CMOS latchup: if the voltage on any pin with respect to 0V is taken outside its absolute maximum rating, even briefly, this can evoke a mechanism within the silicon whereby a low impedance is presented across the power supply. The mechanism is similar to the SCR or thyristor effect of a p-n-p-n junction. If this happens, DC current is taken from the supply, limited primarily by the supply's own source impedance, and this current maintains the fault state after the transient. The device is said to be "latched up" and can easily draw enough current to overheat and burn out.

The phenomenon requires a certain minimum amount of energy to create the initial low impedance, but this energy may well be provided by an ESD strike, especially to unprotected I/O pins, or to supply pins that have inadequate decoupling. Devices embedded within a circuit are usually safe provided their supply decoupling is good, but interface ICs should be protected by external transient limiting unless their datasheet specifically mentions ESD withstand capabilities; good practice is also to implement current limiting of the supply, perhaps as simple as a resistor of a few ohms in series – suitably decoupled, of course.

Oscillator, reset and interrupt inputs

The clock oscillator input pin is high impedance and particularly likely to be affected. A transient coupled onto this pin will generate a shortened processor clock period, which is liable to cause a spurious instruction or data access, thus corrupting the program counter or memory. The reset and edge-triggered interrupt inputs are also sensitive and glitches here have widespread effects on the processor operation. On many microcontrollers, the reset pin is specified for a minimum active duration of several clock cycles. If it is pulsed for a shorter period, an incomplete reset occurs, with unpredictable results.

Always treat these pins on a microcontroller with the greatest care (Figure 13.30): never run long lines to them, put in buffering resistors where possible, keep the tracks closely coupled to the ground plane.

I/O ports

Input/output ports connected to long tracks or external connections are vulnerable on three counts. Small amounts of noise can give spurious data. Higher levels can corrupt the contents of control or data registers, while direct ESD can cause CMOS latch-up.

13.3.2 Logic noise immunity

The ability of a logic element to operate correctly in a noisy environment involves more than the commonly quoted static noise margins. To create a problem an externally generated transient must cause a change of state in an element which then propagates through the system. Systems with clocked storage elements or those operating fast enough for the transient to appear as a signal are more susceptible than slow systems or those without storage elements (combinational logic only).

13.3.2.1 Dynamic noise margin

The effect of a fast transient will depend on the peak voltage coupled into the logic input, and also on the speed of response of the element. Any pulse positive-going from 0V but below the logic switching threshold (typically 1.4V for 74HCT, 50% V_{CC} for

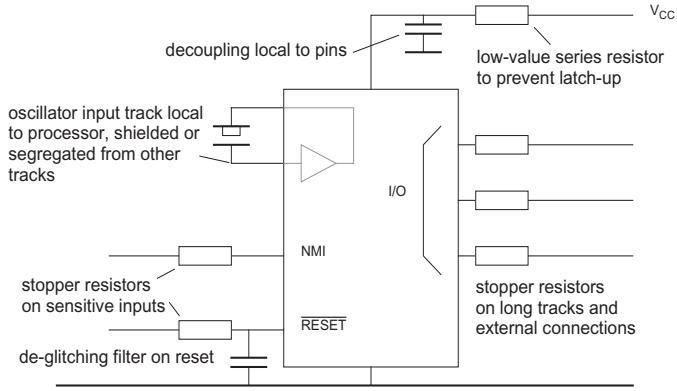


Figure 13.30 Immunity precautions around microcontrollers

74HC circuits) will not cause the element input to switch from 0 to 1 and will not be propagated into the system. Conversely a pulse above the threshold will cause the element to switch. But a pulse which is shorter than the element's response time will need a higher voltage to cause switchover, and therefore the graph shown in Figure 13.31 can be constructed, which illustrates the susceptibility of different logic families versus pulse width and amplitude. Bear in mind that switching and ESD transients may lie within the 1–5ns range. Here is another argument for slow logic!

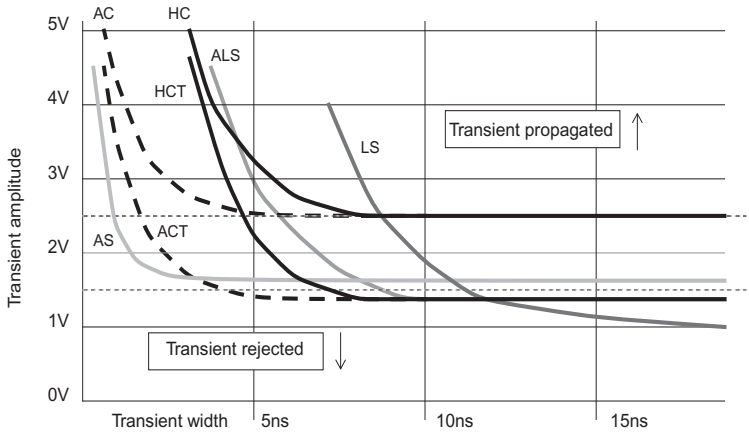


Figure 13.31 Dynamic noise margins

With synchronous logic, the time of arrival of the transient with respect to the system clock (assuming it corrupts the data line rather than the clock line, due to the former's usually greater area) is important. If the transient does not coincide with the active clock edge then an incorrect value on the data line will not propagate through the system. Thus you can expand the graphs of Figure 13.31 to incorporate another

dimension of elapsed time from the clock edge. Tront [144] has simulated a combinational logic circuit with a flip-flop in 3-micron CMOS technology, and generated a series of "upset windows" in this way to describe the susceptibility of that particular circuit to interference. Such a simulation process, using the simulation package SPICE3, can pinpoint those parts of a circuit which have a high degree of susceptibility.

13.3.2.2 Transient coupling

The amplitude of any pulse coupled differentially into a logic input will depend on the loop area of the differential coupling path which is subjected to the transient field $H_{transient}$ due to transient ground currents $I_{transient}$, and also on the impedance of the driving circuit – less voltage is capacitively coupled into a lower impedance. If sensitive signal tracks are run close to their ground returns as recommended for emissions, then the resulting loop area is small and little interference is coupled differentially into the sensitive input (Figure 13.32). This is another way of stating that the transfer impedance is lowest and best for tightly coupled signal and return tracks.

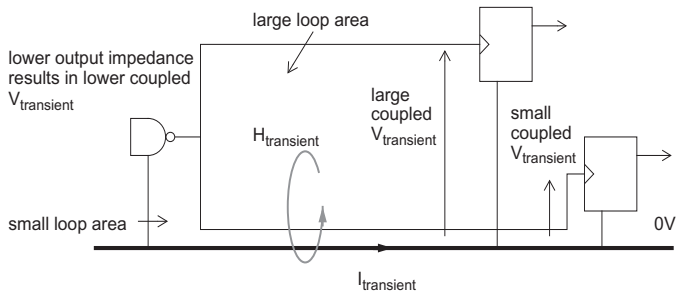


Figure 13.32 Transient coupling via signal/return current loops

13.3.2.3 Susceptibility to RF

Susceptibility of digital circuits to RF, as opposed to transient interference, tends to be most marked in the 20–200MHz region. Susceptibility at the component level is broadband, although there are normally peaks at various frequencies due to resonances in the coupling path. As the frequency increases into the microwave region, component response drops as parasitic capacitances offer alternative shunt paths for the RF energy, and the coupling becomes less efficient. Prediction of the level of RF susceptibility of digital circuits using simulation is possible for small-scale integrated circuits [150] but the modelling of the RF circuit parameters of VLSI devices requires considerable effort, and the resources needed to develop such models for microprocessors and their associated peripherals is overwhelmed by the rate of introduction of new devices.

The first effect of RFI to be noticed on dynamically active logic circuits is timing jitter at the transitions (Figure 13.33). The RF signal is added to the input voltage as it switches through the non-linear active transition region of the device, which causes either an advance or a delay on the transition edge as it is seen by the input. If the circuit timing is critical then this itself can cause maloperation. Thus, to increase RF immunity the first design approach is to keep the system timing as relaxed as possible throughout.

As the RF level increases then actual spurious logic level transitions occur and these will propagate through the circuit if the logic response speed is high enough. Different thresholds can be found for logic 0 and logic 1 effects because of the different driver output impedances in the two states. Active high transition clock inputs are far more sensitive when held at logic 0 than when at logic 1.

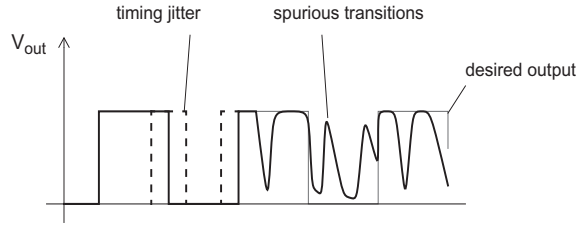


Figure 13.33 RF effects on logic circuits

13.3.3 Signal integrity and ground bounce

Many of the aspects of PCB layout for digital circuits already discussed fall easily under another heading, that of signal integrity. Simply put, this refers to design of the interconnections within digital circuits such that they don't compromise the integrity of signal transmission between them. This is not principally a matter of external EMC, more it is a case of ensuring that the circuit actually works: that is, it doesn't interfere with itself. But good signal integrity has the beneficial side effect of reducing the circuit's exposure to external disturbances, particularly transients and RF.

Design aspects that are relevant for signal integrity in high-speed systems include:

- transmission line matching of interconnecting traces (see section 13.1.3.4);
- control of ground bounce at internal IC nodes and external pins.

We have already looked at transmission lines. Ground bounce is a threat in several ways. Consider the equivalent circuit of an IC including its internal bond wires as shown in Figure 13.34. This representation assumes perfect (zero inductance) ground and power planes, but takes into account the inductance of the bond wires between the chip and its connecting pins, and the inductance of the connecting pins to the planes.

At each output transition, a current spike is drawn through the GND pin (see section 13.1.3.3) partly to charge/discharge the output node capacitance and partly because of ΔI currents. This current passes through the ground inductance (the same thing happens with the opposite polarity to the supply pins) and creates a voltage pulse at the internal 0V node of the IC. Thus, compared to the (assumed perfect) 0V rail of the ground plane, the IC's internal 0V experiences a "bounce" at each current transition. To put it in context, if the total ground inductance is 10nH and di/dt is 50mA/ns the amplitude of the bounce is 0.5V.

This voltage disturbance is shared with all the circuits on the chip. In a complex device this can mean that one part of the IC is capable of affecting other parts, since the rogue voltage appears in series with other inputs and therefore affects the dynamic logic zero threshold, possibly leading to mis-clocking or mis-reading of logic levels. The chip manufacturer can only mitigate the problem by offering several 0V pins so that the bond wire inductances are paralleled, and/or by providing several distinct 0V nodes within the chip to remove the common impedance coupling. The cures available to the

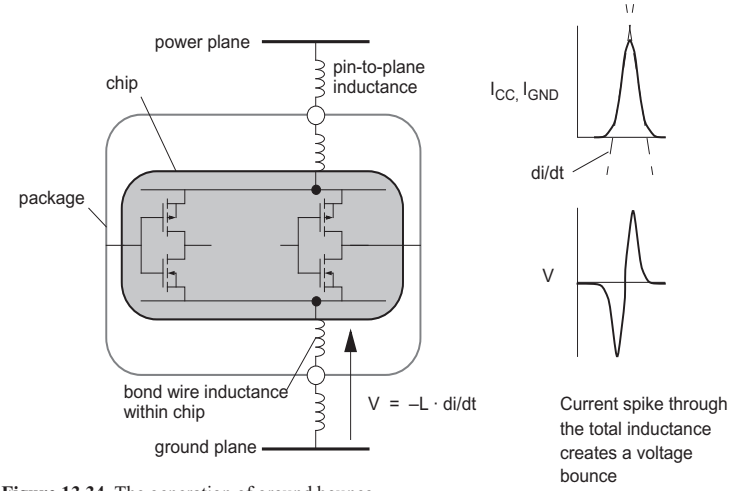


Figure 13.34 The generation of ground bounce

designer are not to overload the outputs with capacitance, and to make sure that every 0V pin is solidly bonded to the ground plane by local, minimum inductance vias.

Figure 13.34 assumes that the ground and power planes do not themselves suffer from voltage bounce. But this is only true if they have zero impedance, and practical power distribution networks will not meet this ideal. If there is too much inductance in the supply traces between ICs, then the ground bounce problem is exported to create voltage differentials between the 0V pins of these ICs with a comparable result, that of different ICs in the same system interfering with each other (Figure 13.35).

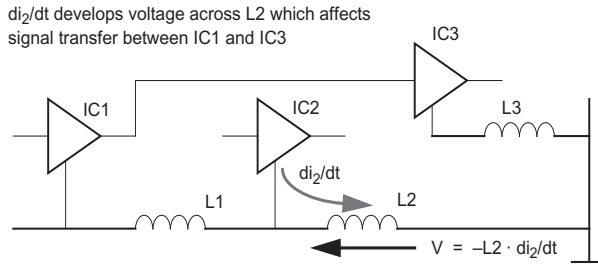


Figure 13.35 System-developed ground bounce

Excessive ground bounce has a number of consequences:

- in the extreme, it will prevent the circuit working;
- more usually, it will create errors and strange effects which may be attributed to other causes such as marginal timing;
- the degradation in thresholds may not show effects under normal conditions but will make the circuit more susceptible to external interference;
- the voltage developed along the supply rails makes the circuit inherently more noisy and contributes to higher RF emissions (*cf* section 13.1.2.2).

13.3.4 The microprocessor watchdog

Circuit techniques to minimize the amplitude and control the path of disruptive interference go a long way towards "hardening" a microprocessor circuit against corruption. But they cannot *eliminate* the risk. The coincidence of a sufficiently high-amplitude transient with a vulnerable point in the data transfer is an entirely statistical affair. The most cost-effective way to ensure the reliability of a microprocessor-based product is to accept that the program *will* occasionally be corrupted, and to provide a means whereby the program flow can be automatically recovered, preferably transparently to the user. This is the function of the microprocessor watchdog [18].

Some micros on the market include built-in watchdog devices, which may take the form of an illegal-opcode trap, or a timer which is repetitively reset by accessing a specific register address. If such a watchdog is available, it should be used, as long as it is well matched to the processor's operation; otherwise, one must be designed-in to the circuit.

13.3.4.1 Basic operation

The most serious result of a transient corruption is that the processor program counter or address register is upset, so that it starts interpreting data or empty memory as valid instructions. This causes the processor to enter an endless loop, either doing nothing or performing a few meaningless or, in the worst case, dangerous instructions. A similar effect can happen if the stack register or memory is corrupted. Either way, the processor will appear to be catatonic, in a state of "dynamic halt".

A watchdog guards against this eventuality by requiring the processor to execute a specific simple operation regularly, regardless of what else it is doing, on pain of consequent reset. The watchdog is actually a timer whose output is linked to the RESET input, and which itself is being constantly re-triggered by the operation the processor performs, normally writing to a spare output port. This operation is shown schematically in Figure 13.36.

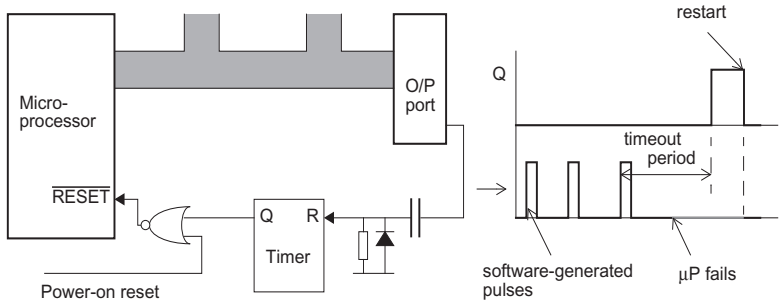


Figure 13.36 Watchdog operation

13.3.4.2 Timeout period

If the timer does not receive a "kick" from the output port for more than its timeout period, its output goes low ("barks") and forces the microprocessor into reset. The timeout period must be long enough so that the processor does not have to interrupt time-critical tasks to service the watchdog, and so that there is time for the processor to

start the servicing routine when it comes out of reset (otherwise it would be continually barking and the system would never restart properly). On the other hand, it must not be so long that the operation of the equipment could be corrupted for a dangerous period. There is no one timeout period which is right for all applications, but usually it is somewhere between 10ms and 1s.

13.3.4.3 Timer hardware

The watchdog circuit has to exceed the reliability of the rest of the circuit and so the simpler it is, the better. A standard timer IC is quite adequate, but the timeout period may have an unacceptably wide variation in tolerance, besides needing extra discrete components. A digital divider such as the 4060B fed from a high frequency clock and periodically reset by the report pulses is a more attractive option, since no other components are needed. The divider logic could instead be incorporated into an ASIC if this is present for other purposes. The clock has to have an assured reliability in the presence of transient interference, but such a clock may well already be present or could be derived from the unsmoothed AC input at 50/60Hz.

An extra advantage of the digital divider approach is that its output in the absence of re-triggering is a stream of pulses rather than a one-shot. Thus if the micro fails to be reset after the first pulse, or more probably is derailed by another burst of interference before it can re-trigger the watchdog, the watchdog will continue to bark until it achieves success (Figure 13.37). This is far more reliable than a monostable watchdog that only barks once and then shuts up.

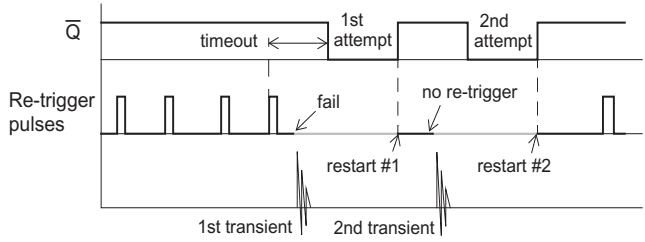


Figure 13.37 The advantage of an astable watchdog

A programmable timer must not be used to fulfil the watchdog function, however attractive it may be in terms of component count. It is quite possible that the transient corruption could result in the timer being programmed off, thereby completely silencing the watchdog. Similarly, it is unsafe to disable the watchdog from the program while performing long operations; corruption during this period will not be recoverable. It is better to insert extra watchdog "kicks" during such long sequences.

13.3.4.4 Connection to the microprocessor

Figure 13.36 shows the watchdog's $\overline{Q}$ output being fed directly to the $\overline{\text{RESET}}$ input along with the power-on reset (POR) signal. In many cases it will be possible and preferable to trigger the timer's output from the POR signal, in order to assure a defined reset pulse width at the micro on power-up.

It is essential to use the RESET input and not some other signal to the micro such as an interrupt, even a non-maskable one. The processor may be in any conceivable state when the watchdog barks, and it must be returned to a fully characterized state.

The only state which can guarantee a proper restart is reset. If the software must know that it was the watchdog that was responsible for the reset, this should be achieved by reading a separate latched input port during initialization.

13.3.4.5 Source of the re-trigger pulse

Equally important is that the micro should not be able to carry on kicking the watchdog when it is catatonic. This demands AC coupling to the timer's re-trigger input, as shown by the R–C–D network in Figure 13.36. This ensures that only an edge will re-trigger the watchdog, and prevents an output which is stuck high or low from holding the timer off. The same effect is achieved with a timer whose re-trigger input is edge- rather than level-sensitive.

Using a programmable port output in conjunction with AC coupling is attractive for two reasons. It needs two separate instructions to set and clear it, making it very much less likely to be toggled by the processor executing an endless loop; this is in contrast to designs which use an address decoder to produce a pulse whenever a given address is accessed, which practice is susceptible to the processor rampaging uncontrolled through the address space. Secondly, if the programmable port device is itself corrupted but processor operation otherwise continues properly, then the re-trigger pulses may cease even though the processor is attempting to write to the port. The ensuing reset will ensure that the port is fully re-initialized. As a matter of software policy, programmable peripheral devices should be periodically re-initialized anyway.

13.3.4.6 Generation of the re-trigger pulses in software

If possible, two independent software modules should be used to generate the two edges of the report pulse (Figure 13.38). With a port output as described above, both edges are necessary to keep the watchdog held off. This minimizes the chance of a rogue software loop generating a valid re-trigger pulse. At least one edge should only be generated at one place in the code; if a real-time "tick" interrupt is used, this edge could be conveniently placed at the entry to the interrupt service routine, whilst the other is placed in the background service module. This has the added advantage of guarding against the interrupt being accidentally masked off.

Placing the watchdog re-trigger pulse(s) in software is the most critical part of watchdog design and repays careful analysis. On the one hand, too many calls in

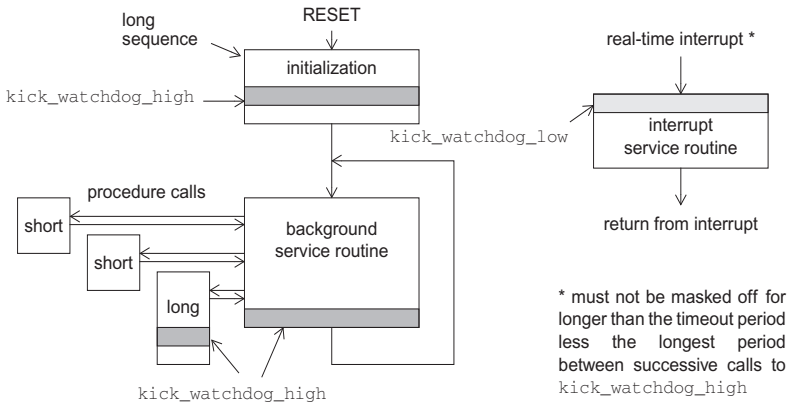


Figure 13.38 Software routine for watchdog re-trigger

different modules to the pulse generating routine will degrade the security and efficiency of the watchdog; but on the other hand, any non-trivial application software will have execution times that vary and will use different modules at different times, so that pulses will have to be generated from several different places. Two frequent critical points are on initialization, and when writing to non-volatile (EEPROM) memory. These processes may take several tens of milliseconds. Analysing the optimum placement of re-trigger pulses, and ensuring that under all correct operating conditions they are generated within the timeout period, is not a small task.

Ideally, the microprocessor should only refresh the watchdog if it is certain that *all* its software routines are running normally. You may be able to identify the most critical routines and define a byte of corresponding task flags. When each task runs it sets its own flag; a watchdog refresh task running in the main loop keeps monitoring the flags, and when all have been set it toggles the watchdog refresh and clears the byte to zero.

13.3.4.7 Testing the watchdog

This is not at all simple, since the whole of the rest of the circuit design is bent towards making sure the watchdog never barks. Creating artificial conditions in the software is unsatisfactory because the tested system is then unrepresentative. An adequate procedure for most purposes is to subject the equipment to repeated transient pulses which are of a sufficient level to corrupt the processor's operation predictably, if necessary using specially "weakened" hardware. For safety critical systems you may have to perform a statistical analysis to determine the pulse repetition rate and duration of test that will establish acceptable performance. An LED on the watchdog output is useful to detect its barks. A particularly vulnerable condition is the application of a burst of spikes, so that the processor is hit again just as it is recovering from the last one. This is unhappily a common occurrence in practice.

As well as testing the reliability of the watchdog, a link to disable it must be included in order to test new versions of software.

13.3.5 Defensive programming

Some precautions against interference can be taken in software. Standard techniques of data validation and error correction should be widely used. Hardware performance can also be improved by well-thought-out software. Some means of disabling software error-checking is useful when optimizing the equipment hardware against interference, as otherwise weak points in the hardware will be masked by the software's recovery capabilities. For example, software which does not recognize digital inputs until three polls have given the same result will be impervious to transients which are shorter than this. If your testing uses only short bursts or single transients the equipment will appear to be immune, but longer bursts will cause maloperation which might have been prevented by improving the hardware immunity.

By no means all microprocessor faults are due to interference. Other sources are intermittent connections, marginal hardware design, software bugs, meta-stability of asynchronous circuits, etc. Defensive programming is also an important technique in safety-critical design, as discussed in Chapter 6. Nevertheless it is necessary to prepare for interference-induced errors, and typical software techniques to enhance immunity are to:

- type-check and range-check all input data;
- sample input data several times and either subject it to rate-of-change checks, for analogue data, or validate it, for digital data;

- incorporate parity checking and data checksums in all data transmission;
- avoid recursion and use token passing;
- protect data blocks in volatile memory with error-detecting and correcting algorithms;
- wherever possible rely on level- rather than edge-triggered interrupts;
- periodically re-initialize programmable interface chips (PIAs, ACIAs, etc.).

When a failure occurs on testing, it is often difficult to find out what actually happened. Connecting an in-circuit emulator while undergoing ESD or EFT testing is not recommended since it drastically worsens the interference coupling (and can be devastating for the emulator). Usually, you have to resort to deducing the internal state after failure from the state of the I/O and bus lines. If you can program spare I/O pins to give a diagnostic status indication with LEDs this is made easier. Alternatively or in addition, unused non-volatile memory can be used to store a diagnostic trace, which is recovered after failure [53].

13.3.5.1 Input data validation and averaging

If you can set known limits on the figures that enter as digital input to the software then you can reject data which are outside those limits. When, as in most control or monitoring applications, each sensor inputs a continuous stream of data, this is simply a question of taking no action on false data. Since the most likely reason for false data is corruption by a noise burst or transient, subsequent data in the stream will probably be correct and nothing is lost by ignoring the bad item. Data-logging applications might require a flag on the bad data rather than merely to ignore it.

This technique can be extended if there is a known limit to the maximum rate-of-change of the data. An input which exceeds this limit can be ignored even though it may be still within the range limits. Software averaging on a stream of data to smooth out process noise fluctuations can also help remove or mitigate the effect of invalid data.

You should take care when using sophisticated software for error detection not to lock out genuine errors which need flagging or corrective action, such as a sensor failure. The more complex the software algorithm is, the more it needs to be tested to ensure that these abnormal conditions are properly handled.

13.3.5.2 Digital inputs

A similar checking process should be applied to digital inputs. In this case, there are only two states to check so range testing is inappropriate. Instead, given that the input ports are being polled at a sufficiently high rate, compare successive input values with each other and take no action until two or three consecutive values agree. This way, the processor will be “blind” to occasional transients which may coincide with the polling time slot. (This method is, of course, nothing more than a variant of the standard technique of “de-bouncing” switch inputs.) It does mean that the polling rate must be two or three times faster than the minimum required for the specified response time, which in turn may require a faster microprocessor than originally envisaged.

13.3.5.3 Interrupts

For similar reasons to those outlined above, it is preferable not to rely on edge-sensitive interrupt inputs. Such an interrupt can be set by a noise spike as readily as by its proper signal. Undoubtedly edge-sensitive interrupts are necessary in some applications, but in these cases you should treat them in the same way as clock inputs

to latches or flip-flops, and take extra precautions in layout and drive impedance to minimize their noise susceptibility – treat their circuits as absolutely critical. If there is a choice in the design implementation, then favour a level-sensitive interrupt input.

If an interrupt is non-maskable, excessive noise on its input pin can cause a burst of consecutive interrupts which will quickly result in a stack overflow. This should be dealt with by monitoring the value of the stack pointer at the beginning of the interrupt routine, and forcing a hardware reset if it gets too large.

13.3.5.4 Token passing and recursion

Software structures tend to mimic business management structures: the big decisions are made at the top, but enacted by those at the bottom [53]. The danger comes when those at the bottom execute tasks without authorization from the top. Interference could force a microcontroller into a low-level sub-routine, perhaps by corrupting its program counter, accidentally initiating an action which has serious consequences.

Recursion is the act of a program or subroutine calling or referencing a part of itself. The technique is susceptible to transient corruption as nesting results in a chain of calls held as a list on the program stack, which could be vulnerable to overflow as above. If recursion is vital for your software, be very clear as to the maximum depth allowed and the consequent memory size needed to support this.

Token passing (Figure 13.39) provides a means whereby each sub-routine can check the authority of the calling routine. The top-level (decision maker) routine sends a specific byte value (token) to the first sub-routine, which checks it against a list of valid tokens and if necessary passes it on. Normally, the tokens match and the action is executed successfully, but if not, a reset is forced. There is of course a price to pay in terms of code size and execution time, for which reason you may want to use token passing selectively rather than globally.

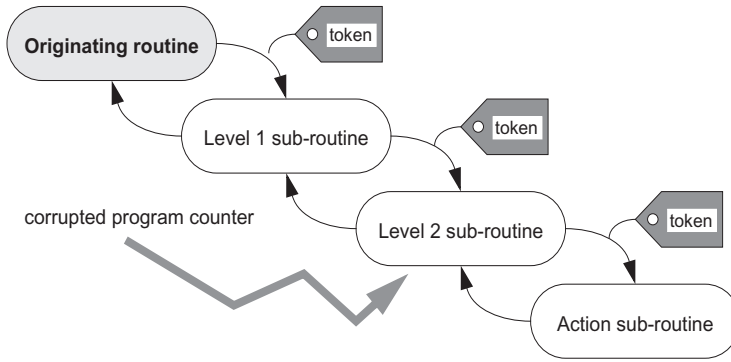


Figure 13.39 Token passing through sub-routines

13.3.5.5 Data and memory protection

Volatile memory (RAM, as distinct from ROM; EEPROM and flash memory could be categorized as volatile since its data may be corrupted while being written) is susceptible to various forms of data corruption. These can be detected by placing critical data in tables in RAM. Each table is then protected by a checksum, which is stored with the table. Checksum-checking diagnostics can be run by the background

routine automatically at whatever interval is deemed necessary to catch RAM corruption, and an error can be flagged or a software reset can be generated as required. The absolute values of RAM data do not need to be known provided that the checksum is recalculated every time a table is modified. Beware that the diagnostic routine is not interrupted by a genuine table modification or vice versa, or errors will start appearing from nowhere! Of course, the actual partitioning of data into tables is a critical system design decision, as it will affect the overall robustness of the system.

Some non-volatile memory devices include a software locking mechanism, which disables the write facility. Full use should be made of this feature, unlocking only for those infrequent times when fresh data has to be written. If the power fails or the processor crashes in the middle of writing to non-volatile memory, the stored data will be unavoidably corrupted. To manage the problem, use a COPY–MODIFY–STORE–ERASE sequence to prevent critical data from being part overwritten.

13.3.5.6 Unused program memory

One of the threats discussed in the section on watchdogs above was the possibility of the microprocessor accessing unused memory space due to corruption of its program counter. If it does this, it will interpret whatever data it finds as a program instruction. In such circumstances it would be useful if this action had a predictable outcome.

An 8-bit bus access to a non-existent address returns the data $\#FF_H$, provided there is a passive pull-up on the bus. Nothing can be done about this. However, un-programmed ROM also returns $\#FF_H$ and this can be changed. A good approach is to convert all unused $\#FF_H$ locations to the processor's one-byte NOP (no operation) instruction (Figure 13.40). The last few locations in ROM can be programmed with a JMP RESET instruction, normally three bytes, which will have the effect of resetting the processor. Then, if the processor is corrupted and accesses anywhere in unused memory, it finds a string of NOP instructions and executes these (safely) until it reaches the JMP RESET, at which point it restarts. An alternative is to block-fill the unused locations with the STOP instruction rather than the NOP, which then relies on the hardware watchdog restarting the system.

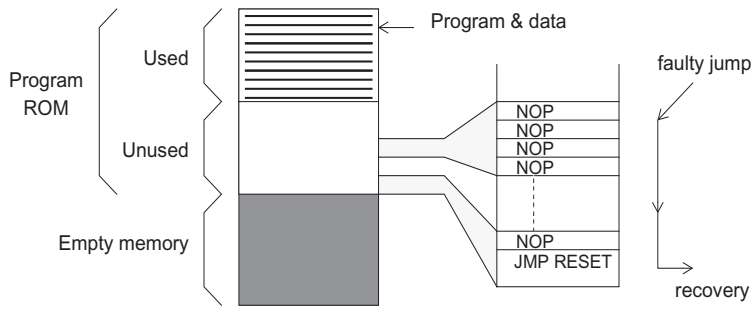


Figure 13.40 Protecting unused program memory with NOPs

The effectiveness of this technique depends on how much of the total possible memory space is filled with NOPs or STOPs, since the processor can be corrupted to a random address. (In practice, the corruption is often less than random, due perhaps to particularly susceptible states or a particular data/address line being the weakest point.)

If the processor accesses an empty bus, its action will depend on the meaning of the $\#FF_H$ instruction. The relative cheapness of large ROMs and EPROMs means that you could consider using these, and filling the entire memory map with ROM, even if your program requirements are small.

13.3.5.7 Re-initialization

As well as RAM data, you must remember to guard against corruption of the set-up conditions of programmable devices such as I/O ports or UARTs. Many programmers assume erroneously that once an internal device control register has been set up (usually in the initialization routine) it will stay that way forever. Experience shows that control registers can change their contents, even though they are not directly connected to an external bus, as a result of interference. This may have consequences that are not obvious to the processor: for instance if an output port is re-programmed as an input, the processor will happily continue writing data to it oblivious of its ineffectiveness.

The safest course is to periodically re-initialize all critical registers, perhaps in the main idling routine if one exists. Timers, of course, cannot be protected in this way. The period between successive re-initializations depends on how long the software can tolerate a corrupt register, versus the software overhead associated with the re-initialization.

13.3.6 Transient and RF immunity – analogue circuits

Analogue circuits in general are not as susceptible to transient upset as digital, but may be more susceptible to demodulation of RF energy. This can show itself as a DC bias shift which results in measurement non-linearities or non-operation, or as detection of modulation, which is particularly noticeable in audio and video circuits. Such bias shift does not affect digital circuit operation until the bias is enough to corrupt logic levels, at which point operation ceases completely. Improvements in immunity result from attention to the four areas as set out below. The greatest RF signal levels are those coupled in via external interface cables and so interface circuits should receive the first attention.

Analogue immunity principles

- minimize circuit bandwidth
- maximize signal levels
- ensure a good circuit stability margin
- use balanced signal configurations
- isolate particularly susceptible paths

13.3.6.1 Audio rectification

This is a term used rather loosely to describe the detection of RF signals by low frequency circuits. It is responsible for most of the ill effects of RF susceptibility of both analogue and digital products.

When a circuit is fed an RF signal that is well outside its normal bandwidth, the circuit can respond either linearly or non-linearly (Figure 13.41). If the signal level is

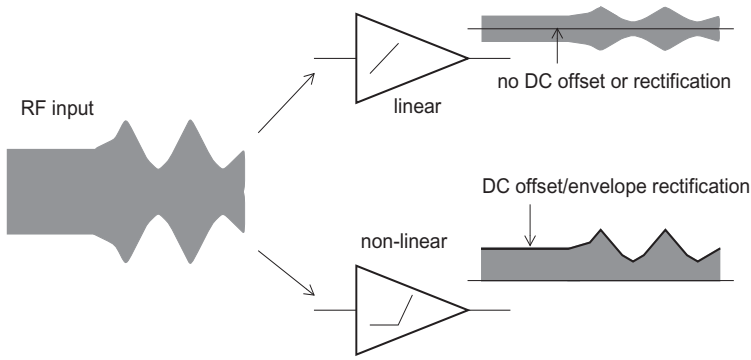


Figure 13.41 RF demodulation by non-linear circuits

low enough for it to stay linear, it will pass from input to output without affecting the wanted signals or the circuit's operation. If the level drives the circuit into non-linearity, then the envelope of the signal (perhaps severely distorted) will appear on the circuit's output. At this point it will be inseparable from the wanted signal, and indeed the wanted signal will itself be affected by the circuit's forced non-linearity.

The response of the circuit depends on its linear dynamic range and on the level of the interfering signal at the point in the circuit which has the most non-linearity. Usually this is its input. All other factors being equal, a circuit which has a wide dynamic range will be more immune to RF than one which has not.

13.3.6.2 Bandwidth, level and balance

The level of the interfering signal can be reduced by restricting the operating bandwidth to the minimum acceptable. You can achieve this (referring to Figure 13.42) by input RC or LC filtering (1), feedback RC filtering (2), and low-value (10–33pF) capacitors (3) or resistors (4) directly at the input terminals. RC filters may degrade stability (see section 13.3.6.4) or worsen the circuit's common mode rejection (CMR) properties, and the value of C must be kept low to avoid this, but an improvement in RF rejection of between 10 to 35dB over the range 0.15 to 150MHz has been reported [140] by including a 27pF feedback capacitor on an ordinary inverting op-amp circuit. High frequency CMR is determined by the imbalance between capacitances on balanced inputs, and the typical low-value tolerance of 10% may make capacitive filtering unacceptable for this reason alone. If an increased input resistance would be too high and might affect circuit DC conditions, a lossy ferrite-cored choke or bead is an alternative series element.

Localized low-value series resistance (4) is a useful approach at the inputs of op-amps and comparators – their high input impedance allows a series resistor of up to a few hundred ohms to be applied at each input, literally at the pins. In conjunction with the input capacitance, this attenuates any induced RF voltages that may be induced along the input tracks. Since surface mount resistors are extremely cheap, it is worth implementing this as a standard technique for all analogue circuits.

You should design for signal level to be as high as possible throughout, consistent with other circuit constraints, but at the same time impedances should also be maintained as low as possible to minimize capacitive coupling and these requirements may conflict. The decision will be influenced by whether inductive coupling is

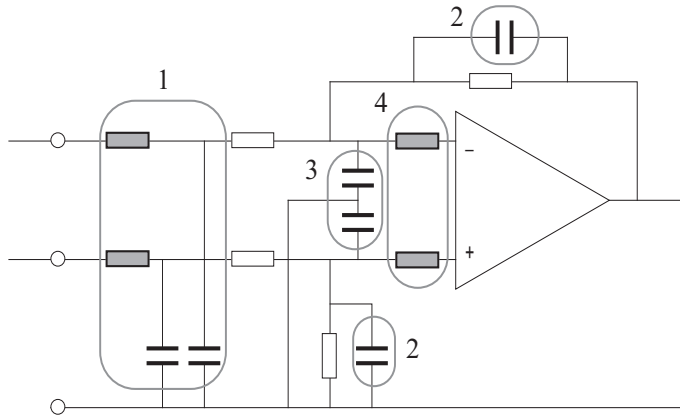


Figure 13.42 Bandwidth limitation in a differential amplifier

expected to be a major interference contributor. If it is (because circuit loop areas cannot be made acceptably small), then higher impedances will result in lower coupled interference levels. Refer to the discussion on inductive and capacitive coupling (section 11.1.1).

Balanced circuit configurations allow maximum advantage to be taken of the inherent common mode rejection (CMR) of op-amp circuits. But note that CMR is poorer at high frequencies and is affected by capacitive and layout imbalances, so it is unwise to trust too much in balanced circuits for good RF and transient immunity.

A common fix for improving the immunity of a discrete transistor circuit is to incorporate either or both of a resistor or ferrite chip in series with the base and a low-value (10–33pF) capacitor directly across the base-emitter junction (Figure 13.43). The effect of this is to reduce the RF applied across the junction, where the non-linearity occurs, but the components need to be mounted right next to the transistor connections. As a matter of standard design practice, you should put such components in place wherever there is a circuit input that could be exposed to incoming RF.

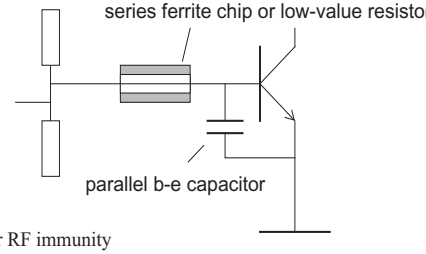


Figure 13.43 Increasing discrete transistor RF immunity

13.3.6.3 The effect of op-amp type and operating conditions

It has been observed [150] that in the frequency range 1–20MHz, mean values of demodulated RFI are 10 to 20dB lower for BiFET op-amps than for bipolar op-amps. This may well have to do with the inherent non-linearity of the bipolar input base-emitter junction. A research project carried out by the author [154] investigated the

contributions to uncertainty of the immunity tests, and one aspect was the non-linearity of response of op-amp circuits with injected interference. It was clear that:

- different op-amp technologies have different characteristic transfer functions of injected RF to DC offset, with CMOS types showing less effect than bipolar;
- for any given type, different responses can be observed at different frequencies;
- the DC operating conditions often have a marked effect on the response, and this is particularly noticeable when single-supply devices are operated close to their lower common mode voltage limit;
- only one of the devices that were investigated, the LM301A which has the oldest design, showed a response that was approaching linearity over the range tested. All others showed non-linearity to a greater or lesser extent.

A sample of the results illustrating these points is shown in Figure 13.44. Notice particularly that for the bipolar LM358 (the National part's response is shown, but Philips and Harris parts performed in the same way) in single supply mode with the input near the 0V rail, as soon as the RF voltage is enough to force the input outside its declared common mode range the output swings immediately to the positive supply rail and stays there. This doesn't happen when the same part is used in dual supply mode with the input at mid-rail, and it doesn't happen with the CMOS TLC272, although it does happen with the BiCMOS CA3240. A general observation is that you should avoid operating op-amp inputs very close to one end or the other of their common mode range, as this makes them much more susceptible to RF.

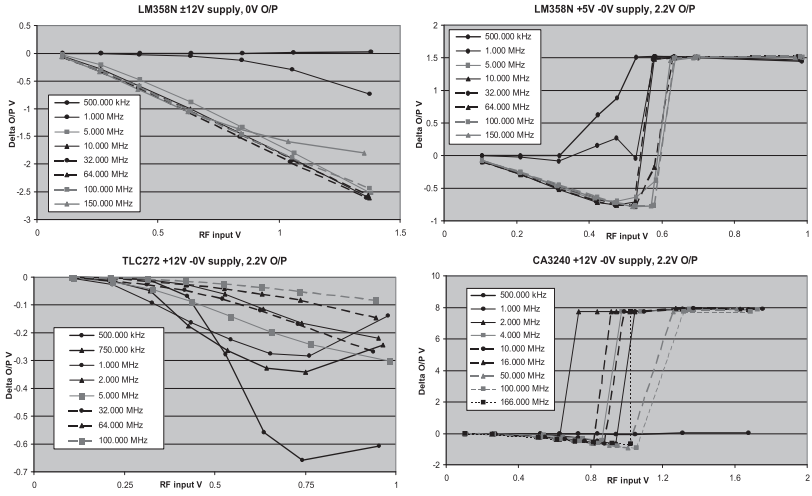


Figure 13.44 Op-amp DC response versus RF input voltage (from [154] Annex F)
NB the Y axis in each case is the deviation from the undisturbed DC output level

13.3.6.4 Stability

Op-amps with gain bandwidths of hundreds of MHz or even GHz are routinely used in many applications. The stability of wideband amplifiers was briefly mentioned in section 13.1.5.1 in the context of emissions, but a quasi-stable amplifier is also a threat to immunity. If the circuit is close to oscillation but not actually unstable at a particular radio frequency, this is equivalent to saying it has a peak in its frequency response at that frequency (Figure 13.45). If an interfering RF signal is applied at this peak the amplifier will happily respond to it, most probably saturating and corrupting its desired signal. This is not an uncommon occurrence particularly on the conducted immunity test in the range 5–30MHz, where peaks in susceptibility cannot be explained by the usual structural resonances of higher frequencies.

Apart from inadequate design of feedback circuitry, amplifier instability is usually a result of bad circuit layout or poor supply decoupling. You need to pay particular attention to these aspects with wideband devices. Also, low-value capacitors at the input as recommended above in Figure 13.42 can affect the feedback to worsen the stability of the circuit, so they should be used with caution; the input pin series resistor method is preferable. Stability problems can be diagnosed without access to RF immunity test facilities by feeding a square wave with a fast rising edge into the circuit, and looking for ringing on the output.

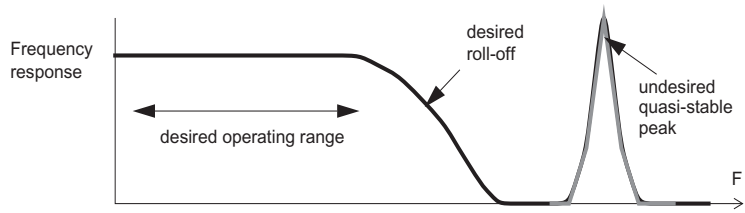


Figure 13.45 Amplifier quasi-stability

13.3.6.5 Isolation

Signals may be isolated at input or output with either an opto-coupler or a transformer (Figure 13.46). The ultimate expression of the former is fibre optic data transmission; given that the major interference coupling route is via the connected cables, using optical fibre instead of wire completely removes this route. This leaves only direct coupling to the enclosure, and coupling via the power cable, each of which is easier to deal with than electrical signal interfaces. Signal processing techniques will be needed to ensure accurate transmission of precision AC or DC signals, which increases the overall cost and board area.

Coupling capacitance

Isolation breaks the electrical ground connection and therefore substantially removes common mode noise injection, as well as allowing a DC or low frequency AC potential difference to exist. However there is still a residual coupling capacitance between primary and secondary which will compromise the isolation at high frequencies or high rates of common mode dv/dt. This capacitance is typically 2–3pF per device for an opto-coupler; where several channels are isolated the overall coupling capacitance (from one ground to the other) rises to several tens of pF. This common mode impedance is a few tens of ohms at 100MHz, which is not much of a barrier!

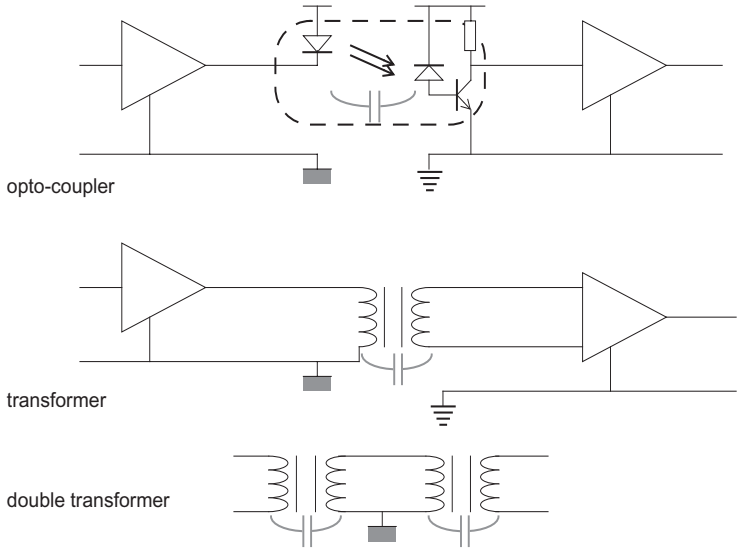


Figure 13.46 Signal isolation

Electrostatically screened transformers and opto-couplers are available where the screen reduces the coupling of common mode signals into the receiving circuit, and hence improves the common mode transient immunity *of that (local) circuit*. This improvement is gained at the expense of increasing the overall capacitance across the isolation barrier, and hence reducing the impedance of the transient or RF coupling path to the rest of the unit. A somewhat expensive solution to this problem is to use two unscreened transformers in series, with the intervening coupled circuit separately grounded.

A particular problem with opto-couplers, not suffered by transformers, is that high-level but short transients can couple into the base of the receiving photo-transistor and cause its saturation. This then effectively "blinds" the coupler for however long it takes for the device to come out of saturation, which may be several microseconds; the overall impact on the circuit is to apparently stretch the duration of the transient to a potentially unacceptable extent. The same effect happens with RF interference, and is due to capacitive coupling directly to the base (Figure 13.47), especially when this is brought out to a separate termination on the package. The problem is not as bad with photodiode couplers, and can be mitigated with a phototransistor by connecting its base to emitter with a low-value resistor, at the expense of sensitivity.

Brøndegaard Nielsen [108] investigated a number of widely available opto-couplers in the presence of fast transient and RF interference. He concluded that

> 2kV transients always caused some short wrong signals on the optocoupler output ... you have to accept a wrong signal of up to 1μs on each single transient of the burst, if you want the optocoupler to insulate fast transient voltages ... if you cannot accept a short time wrong signal from an optocoupler, you have to limit the dV/dt value on the galvanic barrier by common mode filters, or you have to limit the signal bandwidth of the transferred signal by a differential mode

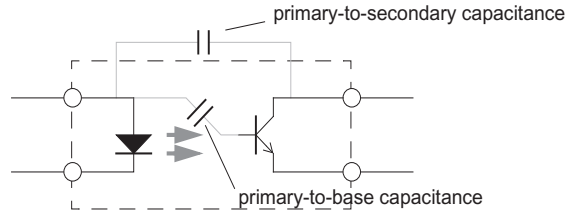


Figure 13.47 Parasitic coupling in the opto-coupler

> filter after the output of the optocoupler ... the most transient susceptible optocouplers were found among the types with an external base pin.

It is best to minimize the number of channels by using serial rather than parallel data transmission. Do not compromise the isolation further by running tracks from one circuit near to tracks from the other. Figure 11.4 on page 263 will give some idea of the degree of mutual coupling versus track spacing. Isolated sections must be laid out to prevent any coupling to other parts of the circuit, including ground planes: the main (non-isolated) 0V plane should stop halfway across the isolator, and no part of the isolated circuit should overlap it.

Digital couplers

A relatively recent device to appear is neither an opto- nor a standalone transformer coupler. Exemplified by Analog Devices' *i*Coupler® range, these components use a micro-transformer at chip scale which passes data and/or power at very high frequency across a magnetically coupled isolation barrier. To system designers these are very attractive components, particularly those which pass power across the barrier as well as data. Unfortunately, in order to do this, the power switching circuit has to operate between 180 and 300MHz [44] and the signal interface uses encoded 1ns pulses at data transitions [93]. As such, these devices are easily capable of exceeding radiated emissions limits due to the common mode signal across the barrier.

To deal with this you need to implement a significant amount of capacitance between the isolated and non-isolated sides of the system, typically by overlapping the two sets of planes on adjacent layers in a multilayer board, exactly as proscribed above. This in turn reduces the value of the device in isolating the system from high frequency incoming interference. Although they appear to be a useful single-chip solution for many high speed data isolation purposes, you would be well advised to study the above references carefully before deciding to design them in.