UNDERSTANDING FEEDBACK

We've all been at a function in a hall or club where someone steps up to the PA system microphone, and suddenly there's an ear-splitting 'howl' from the speakers. That's one example of *feedback*: where some of the energy from the output of an amplifying system is getting back into the input, to change how the system behaves. In this case it's *acoustic feedback* — sound energy from the speakers is getting back to the microphone, so it's being amplified over and over again to produce an oscillation.

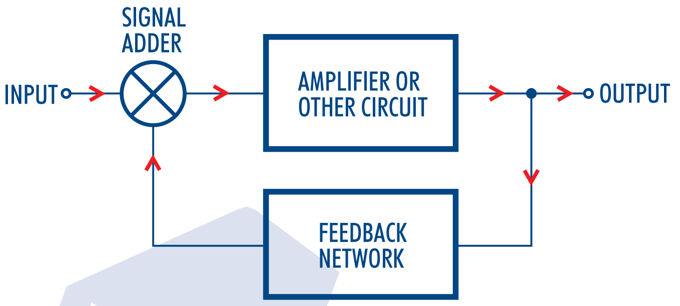


Fig.1: The basic principle of feedback. Some of the output of an amplifier or other circuit is fed back and added or compared with the input, to change or improve the amplifier or circuit's performance.

Feedback is used deliberately in a lot in electronic circuits, to change their performance. The general idea of feedback involves part of the output of an amplifier or other circuit being fed back through a 'feedback network' of some kind and added to (or if you prefer, compared with) the normal input signal, as shown in Fig.1.

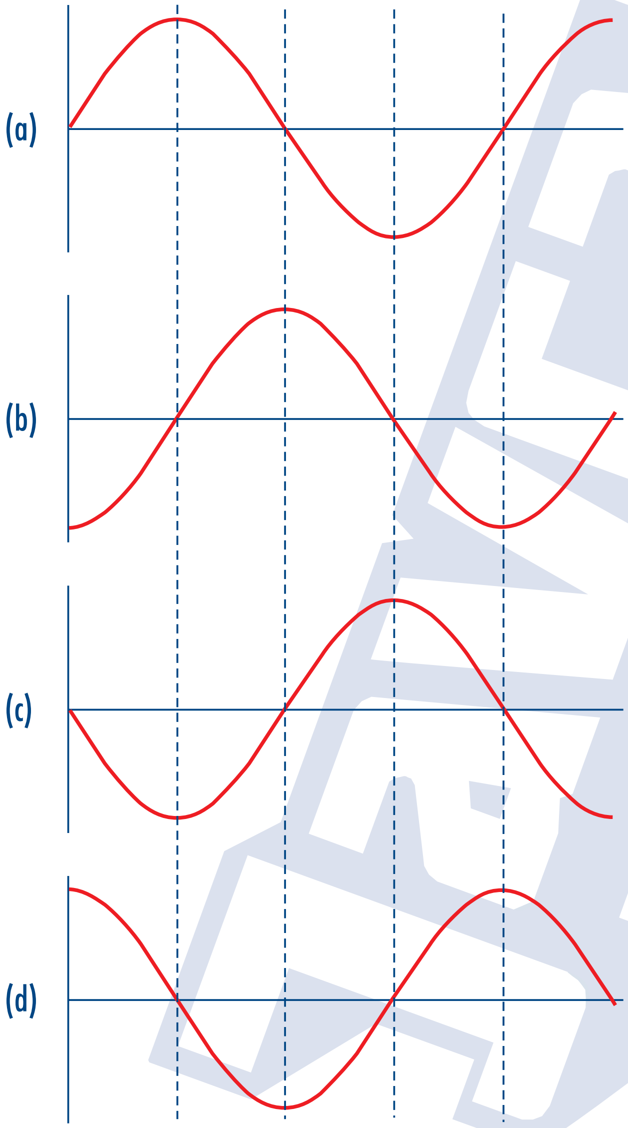


Fig.2: Four different waveforms, all of the same frequency and amplitude but with different timing. In other words, shifted in phase...

Does feedback always cause oscillation, or degrade the performance of a system in some other undesirable way? No, not at all. In fact when it's carefully controlled, feedback can actually improve the performance in various ways. It all depends on the *phase relationship* between the energy that's being fed back to the system's input, and the normal input signal. So you'll be able to understand this properly, let's start by clarifying what we mean by 'phase'.

Phase basics

In Fig.2 we've drawn four sinewave signals. They all have exactly the same frequency and peak-to-peak amplitude, but as you can see they're also quite different in terms of their timing. When signal (a) is zero and about to go positive, signal (b) is at its full negative value while signal (c) is zero but about to go negative, and (d) is at its full positive value.

Similarly when signal (a) has reached its full positive value, (b) has gone to zero but is heading positive, (c) is now at its fully negative value and (d) has gone to zero but heading negative.

Waveforms like this, of the same frequency but with different timing are said to differ in *phase*. So we can describe waveforms (b), (c) and (d) as being 'shifted in phase' with respect to waveform (a).

Since repetitive waveforms like a sinewave repeat themselves every cycle, it's convenient to think of these phase shifts in terms of the corresponding fraction of a cycle. And since a complete cycle is defined as 360°, fractions of a cycle can also be expressed in degrees. For example a quarter of a cycle is 90°, half a cycle 180° and so on.

So in Fig.2 we can say that waveform (b) has a phase shift of 90 with respect to waveform (a). Similarly waveform (c) is shifted 180° with respect to (a), making it completely opposite in polarity. The same applies to waveforms (d) and (b), of course, although they're both shifted by 90° with respect to (a) and (c). Get the idea?

Of course signals can be shifted in phase by *any* amount in passing through an amplifier or other circuit, not just in neat quarters of a cycle. Depending on the behaviour of the circuitry or components they pass through, they can be shifted by anything from a degree or two up to well over a full cycle of 360°.

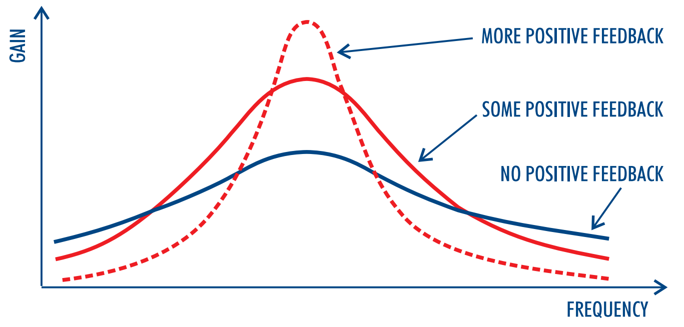


Fig.3: The effect of positive feedback is to exaggerate any unevenness in the gain of an amplifier or tuned circuit. That's why it was used to improve the selectivity and gain of early RF amplifiers.

Combining signals

Glancing back at Fig.2 for a moment, what do you think we'd get if we tried to combine signals (a) and (c)? That's right, nothing — the two waveforms would cancel each other out completely, because they're always of opposite polarity. The same thing would happen if we tried to combine signals (b) and (d).

On the other hand if we tried combining (a) and (b), we'd get a signal with a peak-to-peak value larger than either, and with a phase shift somewhere between the two. The same would apply if we combined (b) and (c), or (c) and (d).

Hopefully you can see from this that when we combine two signals of the same frequency, what we get depends very much on the phase difference between them. Basically they'll either add to, or subtract from each other, to an extent which depends on their relative phase shift.

Broadly speaking, if their relative phase difference is *less than* 90° in either direction ('leading' or lagging'), they'll tend to ADD to each other. On the other hand if it's *more than* 90° in either direction, they'll tend to SUBTRACT from each other.

Positive & negative feedback

You should now be in a good position to understand why engineers talk about two broad kinds of feedback, with quite different effects. Both involve some of the output signal from an amplifier or similar circuit being 'fed back' to its input, but the crucial difference is how the phase of the feedback signal compares with the normal input signal.

If the phase is shifted less than 90° either way of the input signal, so the two tend to *add*, the feedback is said to be POSITIVE. On the other hand if the feedback signal is shifted in phase by more than 90° either way, so the two tend to subtract, the feedback is said to be NEGATIVE. And the effects of positive feedback are very different from those of negative feedback.

Positive feedback

Broadly speaking, positive feedback tends to *increase* the effective gain or amplification of the system, and to exaggerate any unevenness in its behaviour. So if there are any peaks or dips in the system's frequency response, these will be made larger by positive feedback. The peaks will grow higher, and the dips lower.

If the size of the positive feedback signal is increased, a level is soon reached where it's large enough to maintain the system's output by itself — i.e., without any other input signal. The system then *oscillates*, producing a continuous output signal all by itself. (That's what is happening when a PA system 'howls', of course.)

Now positive feedback isn't 'bad', just because it can produce oscillations when we don't want them. It also happens to be the way we produce oscillations when we DO want them; in fact ALL oscillator circuits rely on positive feedback for their operation. So if it wasn't for positive feedback, we wouldn't be able to use oscillator circuits to generate signals.

Early radio engineers also found that carefully controlled positive feedback was very handy for increasing the selectivity or 'Q' of mediocre tuned circuits, and also for getting higher amplification from valves and transistors that weren't particularly good at amplifying high frequencies. In fact positive feedback was so handy for this sort of 'circuit enhancement' that they gave it a special name: *regeneration* (also called 'reaction').

Fig.3 shows how positive feedback can improve the selectivity and gain of an RF amplifying circuit. Of course if the positive feedback is increased too far, the circuit will 'take off' and oscillate — that's why early radios were hard to adjust, and prone to emitting loud 'whistles'.

Negative feedback

Perhaps not surprisingly, *negative* feedback tends to have almost the opposite effect to positive feedback. Instead of increasing the effective gain, it tends to reduce it. And instead of exaggerating any unevenness of circuit behaviour, it tends to smooth out any peaks or dips. It also tends to improve the frequency response (at both ends), reduce distortion by improving linearity, reduce any noise generated by the circuit inside the 'feedback loop', and even reduce its output impedance. Fig.4 shows the way the effective frequency response is smoothed and extended.

So deliberately feeding some of an amplifier's output back to its input — with the correct phasing to produce negative feedback — can improve its performance in a number of ways. It will give it a wider bandwidth, lower distortion, a better signal to noise ratio (SNR), lower output impedance and so on. In short, negative feedback tends to make the amplifier behave more like an 'ideal amplifier', in quite a number of ways.

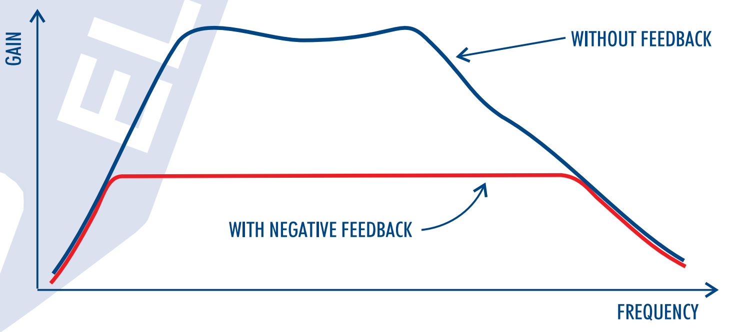


Fig.4: Negative feedback does the opposite of positive feedback – it smoothes out any unevenness in the amplifier or circuit's performance, and extends the frequency response. But there's a price: the gain is reduced as well.

And the only real ‘cost’ of all this improvement is that the amplifier’s gain will also be reduced. So negative feedback makes an excellent tool for improving the performance of amplifiers, as long as they have more gain than you really need.

If negative feedback generally just improves the performance of an amplifier, can you have too much negative feedback? Not in the same way that if you have too much positive feedback, the circuit will burst into oscillation.

There *is* a kind of ‘law of diminishing returns’, though. Once you have more than a certain amount of negative feedback, increasing it further mainly just reduces the overall gain and doesn’t achieve much more in terms of improving performance.

When you apply a lot of negative feedback it also gets harder to ensure that the feedback *remains* negative over all of the working bandwidth of the amplifier circuitry inside the feedback loop. Circuits and components inevitably have phase shifts which vary, and the wider the bandwidth the more they tend to vary. So when you apply lots of feedback it gets harder to ensure that the phase shifts don’t eventually make the feedback swing from negative to positive, especially at high frequencies. That *would* cause oscillation.

With practical amplifiers, then, too much feedback can reduce their *stability margin* and run the risk of them breaking into oscillation. Simply because it’s too hard to stop the feedback turning into *positive* feedback.

Op amps

Operational amplifiers or ‘op amps’ are a good example of the way negative feedback is used to modify and improve the performance of amplifiers. Modern op amps are small high-gain amplifier ICs which are specially made with the intention of having their performance not just improved by negative feedback, but totally controlled using it.

A typical op amp has a very high voltage gain — say between 10,000 and 100,000. But this ‘raw’ gain often varies quite a bit, and the high frequency response generally starts dropping away at quite a modest frequency. See Fig.5, which shows the response of a common TL071 op amp. As you can see without feedback (blue curve), the response drops away quite rapidly above about 5kHz. It’s certainly not hifi!

But this type of amplifier isn’t *meant* to be used without negative feedback, and almost never is. Normally quite a lot of feedback is applied, and as you can see from the red curves this not only brings the amplifier’s gain down to a more practical level (say 1000, 100, or 10), but also gives it a much more useful frequency response at the same time.

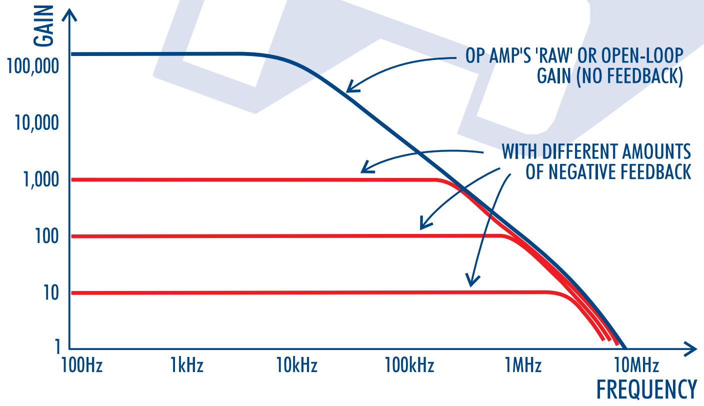
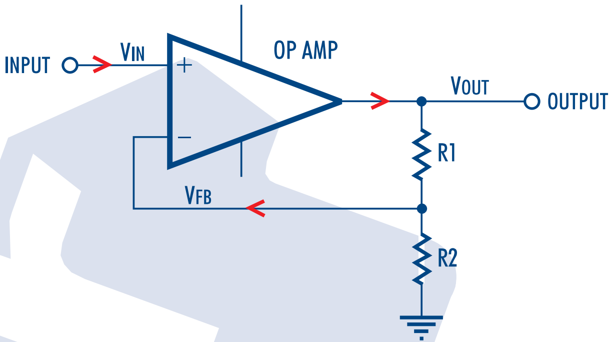


Fig.5: Negative feedback is used to flatten and extend the frequency response of op amps – which are in fact designed to be controlled in this way.



FEEDBACK VOLTAGE $V_{FB} = V_{OUT} \times \left(\frac{R2}{R1+R2}\right)$

IF OP AMP GAIN IS VERY HIGH, THEN $V_{IN} = V_{FB}$

$$\text{SO GAIN} = V_{OUT}/V_{IN} = V_{OUT}/V_{FB} = \frac{R1+R2}{R2}$$

Fig.6: A good example of the way negative feedback is used to control the gain and performance of an op amp. Note that R1 and R2 form the ‘feedback network’ – here just a voltage divider.

As you can see bringing its gain down to even 1000 smooths and extends the frequency response to well over 100kHz, while reducing it right down to 10 makes its gain flat to well beyond 1MHz.

Modern op amps are designed to be used in this way, and are provided with ‘positive’ and ‘negative’ inputs to make it as easy as possible to apply feedback.

Fig.6 shows how easy it is to use an op amp like the TL071 as a wideband voltage amplifier, with its gain controlled by negative feedback. In this case the negative feedback network is simply a voltage divider using resistors R1 and R2, feeding back a fixed proportion of the output to the negative input (for negative feedback).

As you can see, if the ‘raw’ gain of the op amp is very high, it’s actually very easy to work out the ratio of the feedback resistors needed to give a particular final gain with feedback. This is because the difference between the feedback voltage V_{FB} and the input voltage V_{IN} becomes so small that the two have effectively the same value. So the ‘division ratio’ of the feedback voltage divider becomes the main thing controlling the amplifier’s working gain.

To give the amplifier a working gain of 10 you’d simply give R1 a value of say 9kΩ and R2 a value of 1kΩ, for example. Or if you wanted a gain of 50, you’d give R1 a value of 49kΩ and keep R2 at 1kΩ. Get the idea?

This is a very simple example, where we’re only using negative feedback to control the amplifier’s gain and make it as flat as possible. It’s also possible to use a more complex negative feedback network to control and manipulate the frequency response as well, to produce filter and equaliser circuits. But that’s another story.