paX – a Power Amplif
with Error Correction

Part 1: EC and the output stage

Jan Didden

Most solid state power amplifiers employ some form of global negative feedback to reduce non-linearities and output impedance. In some cases, designers exploit alternatives like feedforward to circumvent perceived disadvantages of global negative feedback. The present design uses error correction as (re)defined by Malcolm Hawksford around 1984 [1].

In part 1 of this article the author discusses error correction for audio amplifiers and presents an audio output stage based on error correction. Part 2 will extend the principle to the voltage amplifier stage and present a complete error correction power amplifier. In a separate article next month the protection circuitry is discussed in more detail.

Negative feedback is not negative

This is not an article against negative feedback (nfb). Negative feedback, as a general principle, is one of the most powerful tools available to the designer to build amplifiers that are transparent to the signal they amplify. With that I mean that they do not add anything to, or subtract from, the input signal. In reality, circuits are never ideal, but the changes to the signal can be made so small that inaudibility of the change is pretty much guaranteed.

Negative feedback is also fully understood, and although bad-sounding feedback amplifiers are still being sold, it is totally unnecessary. So, you may ask, why bother with error correction

(ec)? For one thing, it is different; and different roads to the same goal are often enjoyable to travel and explore. Secondly, although we will see that ec is in many respects another face of nfb, with similar advantages and disadvantages, there are some interesting differences and different challenges leading to better understanding of the processes inside any feedback amplifier.

What the H.ec?

Let's start by narrowing down what we mean by ec in the context of this article. I refer specifically to Malcolm Hawksford's paper referenced above. I will use the term 'H.ec' to refer to Mr. Hawksford topology, and 'ec' to refer in general to error correction.

Figure 1 is the basic topology from that paper. We see an amplifier block N, on which the error correction will be ap-
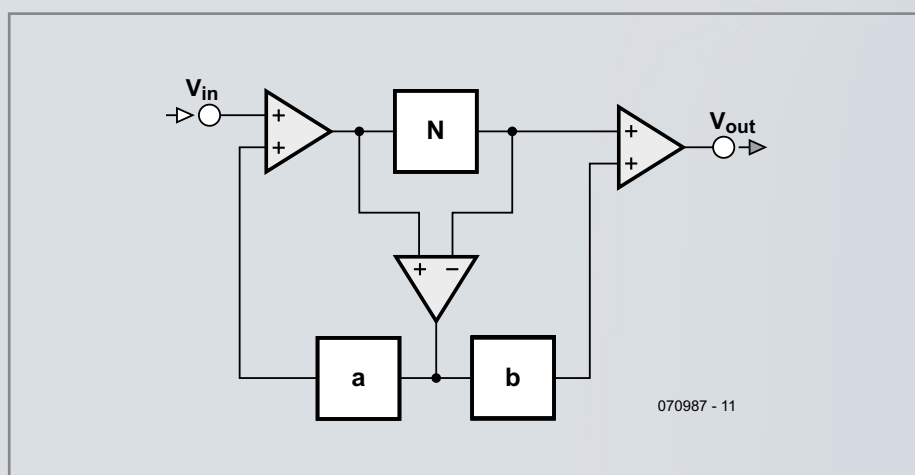

070987 - 11

**Figure 1.** Hawksford's general feedback/feedforward/error correction structure.

plied. The difference between the input and output of that amplifier is either added to the input via block 'a' (error feedback), or to the output via block 'b' (error feedforward). All the summing and differencing blocks have unity gain (1×). Of course it is important that the error correction signal is added in just the right proportion to cancel the original error. In this design I use the correction via block 'a', added to the input.

We will further assume that the amplifier block 'N' is an output stage from an audio power amplifier with a gain only just below 1. The object here is to make this gain exactly 1, independent of output load, frequency and signal level. If we succeed, we will have an ideal output stage. Of course we won't get that far, but as we will see, we can get pretty close.

We now get to our conceptual circuit in **Figure 2**. The output stage we want to linearise is 'A', with a gain of 'about 1'. In practice, most output stage gains are anywhere between 0.92 and 0.98 depending on the topology, loading, signal frequency etcetera. We see immediately that V_c, the signal that is fed back to the input, is $V_{out} - V_e$, where V_e is the effective input signal to the output stage. We can now calculate V_e as

$$V_e = V_{in} - (V_{out} - V_e).$$

We also know that $V_e = V_{out}/A$, and if we plug that into the former equation and rearrange terms, we get:
$V_{out} = V_{in}$.
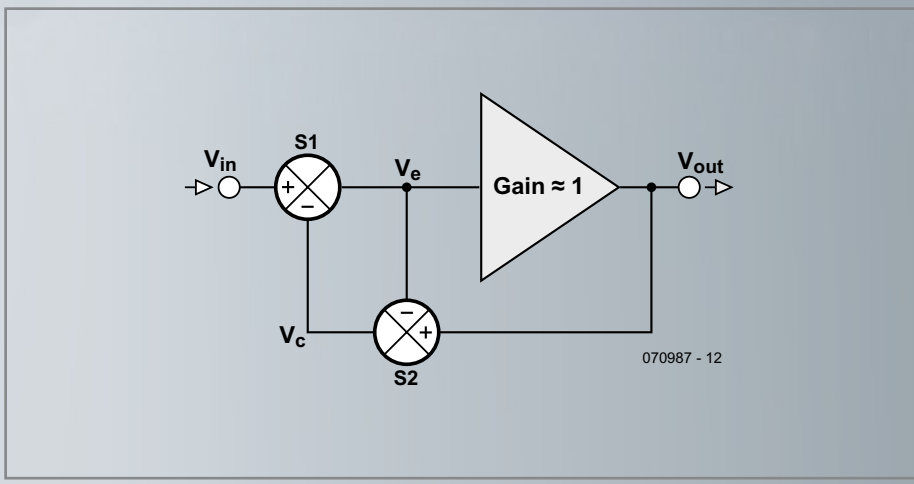
Figure 2. Error correction concept.

070987 - 12

What is significant here is that the actual amplifier gain *A* is no longer part of the equation. Whatever the shortcomings, non-linearities or errors of *A*, we got rid of all of those. This is our ideal output stage!

If you are familiar with practical audio amplifiers, you probably are feeling a bit uneasy now. An ideal power stage? That would be the first one ever! And you are right — in reality, we can't make that ideal.

The reason is that we assumed that the summers in Figure 2 are ideal summers, without flaws. That cannot be. Those summers consist of passive and (most probably) active devices that have their own non-linearities, so the basic accuracy will not be ideal. Their characteristics will also vary with frequency, so the correction accuracy will vary with frequency. As the output stage gain will also vary with frequency and load, the amount of correction required will vary with frequency and load, meaning that signal levels in the summers vary with frequency and load as well. This further leads to accuracy limitations.

Nevertheless, it is possible to greatly improve the performance of the output stage with relative simple means, as we will see.

Thermal memories

There was one other goal I had with this amplifier. In most power amplifiers, the thermal bias compensation is obtained by mounting a transistor in the bias circuit on the same heatsink as the output devices. In that way, if the output devices heat up and start to draw more current, the bias transistor also heats up. That causes the bias voltage to decrease, and the object is to dimension this thermal feedback loop such that the bias current in the output devices remains stable with temperature. But because it
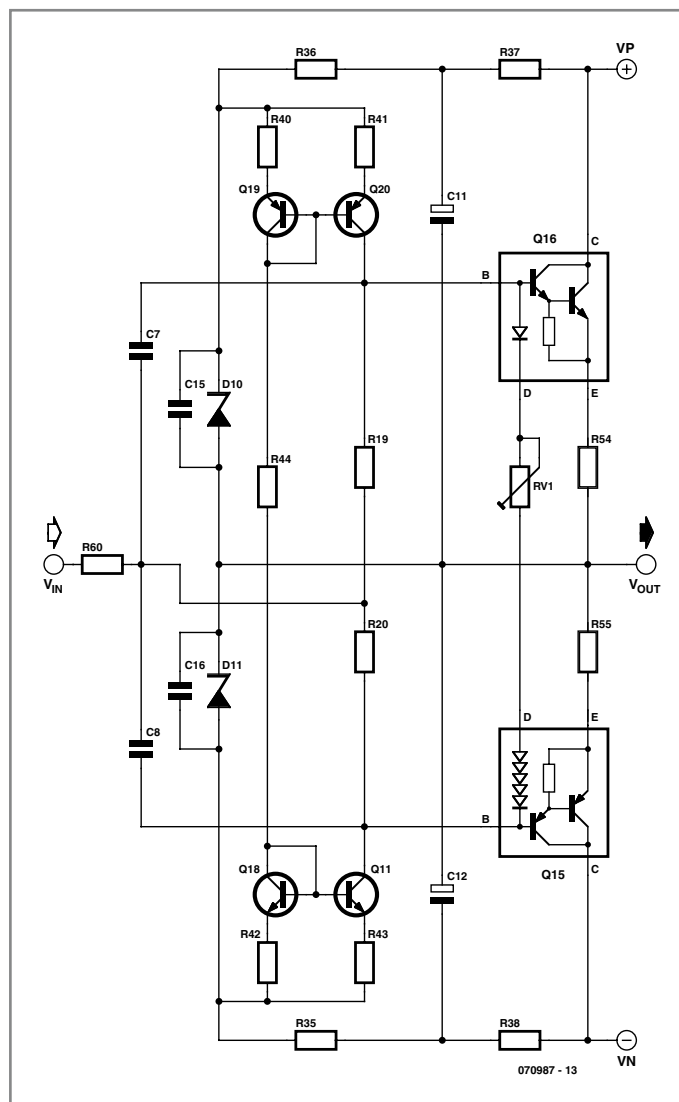

Figure 3. Basic paX output stage.

takes time for the heatsink to heat and cool and transfer this changing temperature to the bias transistor, this loop reacts relatively slowly at several seconds or more.

The first time I read about this was in an article by a French audio (not fashion) designer authoring under the pseudonym Hephaïstos [2]. He realized

that the dissipation levels in the output and driver devices vary with signal output levels, related to music level variations, and that they are much faster than the reaction time of the thermal bias compensation circuit. When a high level signal burst appears, the devices' biasing point would shift, and would return to the earlier state only seconds after the high level burst had disappeared. Such a burst is too short for the thermal feedback to adjust the bias. A smaller signal after the burst would thus be processed with different operating conditions than before. He called this 'thermal distortion'.

In my design I wanted to get rid of this problem as well. Therefore I selected Sanken STD03N and STD03P devices for the output stage. These are Darlingtons, with a bias diode integrated on the transistor chip. By using this diode in the bias circuit, it can track thermal cycles in the output devices instantaneously, thus hopefully eliminating thermal distortion. Because it is a Darlington, any (pre) driver dissipation would also be low enough to avoid memory distortion.

There's one catch: the on-chip sense diodes need to be run with a specific current to make their thermal bias changes in millivolts per degree, equal to the V_{be} changes of the driver and output devices in millivolts per degree. This requires a different bias circuit from the usual V_{be} multiplier.

Error correction basics

Figure 3 shows the basic topology of the output stage without error correction. You will notice the integrated bias diodes in the output devices. To get a precise matching between these diodes and the output Darlington's b-e junctions, Sanken's engineers used a string of five Schottky diodes in the P-device and a single silicon diode in the N-device.
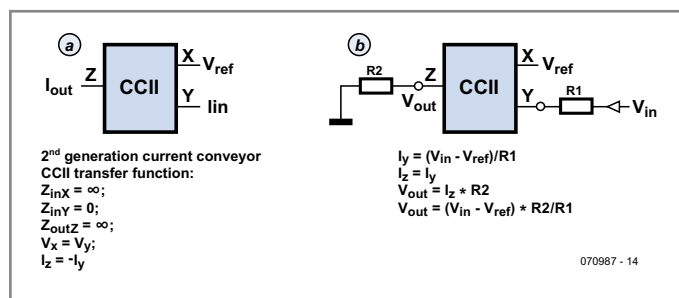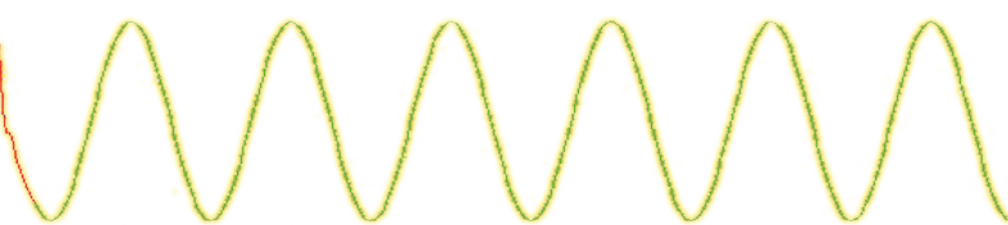

Figure 4. Current Conveyor basics.

The datasheet specifies a bias current of 2.5 mA through the diodes, and this current is set by the current mirror formed by Q19-Q20 and Q18-Q11. The current is set with R44 and mirrored into the diodes. To keep this current stable, the supply voltage for the current mirrors is regulated by zener diodes D10-D11 to 15 V. The bias for the zeners is coming from the main supply via R36-R37 and R35-R38. To keep exactly that supply stable, the junction between those resistors is bootstrapped from the output via C11-C12. The result is that with varying output signal levels, the zener diodes can always provide a constant supply voltage for the current mirrors.

The datasheet for the output devices also specifies a quiescent current of 40 mA through the output Darlingtons for best thermal tracking. Although the diodes track the changes in the Darlington V_{be} quite well, the absolute value of the diode threshold values varies from unit to unit. Thus, it is necessary to be able to adjust it, and that is the purpose of RV1. So, we should set RV1 for 40 mA through the output stage.
OK, so now we have a nice and stable output stage, but we need to drive it with a signal. That is done through R60 to the junction of R19-R20. Suppose that the input signal goes positive: the junction of R19-R20 goes positive so there will be less current through R19. Since the current coming out of the collector of Q20 is fixed, there will be more current into the base of the Darlington and the output signal will also go positive, following the input signal. In this case there will also be more current through R20 so less from the base of the P-device.

This will work quite well for DC and low frequencies and nominal, resistive (8 ohms) loads, where the Darlingtons have a very high gain and the current mirrors are almost perfect. However, with increasing frequency and increasing load currents, the Darlingtons need more input current, so capacitors C7 and C8 are added to bypass R19-R20 so the input source can directly drive the output devices.

This output stage is quite simple, reasonably linear and satisfies the requirement for accurate temperature tracking of the instantaneous output device dissipation.
The next step is to wrap error correc-

tion around it. Because the stage has no excess gain, you cannot use a global nfb loop around it. The nfb loop that normally includes the output stage of a feedback amplifier of course relies on the excess gain in the voltage amplifying stage to work its distortion-reduction trick.

The Current Conveyor

For ec we need to derive the difference between the input and output signal, and add that difference to the input signal, as shown in Figure 2.
Subtracting two signals that are of similar level, and adding two signals that are of similar level is trivial with opamp circuits. We can use one opamp (S2) to subtract V_{out} from V_e, and another (S1) to add the resulting V_c to V_{in}. But in the spirit of ec I didn't want to use an obvious high global feedback element in this circuit.

After a lot of head scratching and many Internet searches, I came up with the idea to use a current conveyor [3]. The basic circuit is shown in **Figure 4a**. This is a very interesting circuit: basically, whatever you send into the Y terminal comes out in opposite direction from the Z terminal, which works as a current source.
Hence the name *current conveyor*: the current at the input is conveyed to the output. This particular type is called a second-generation conveyor, generally shown as a 'CCII'. Terminal X is a reference terminal for the current input, and the voltage at Y will be kept the same as that at X.
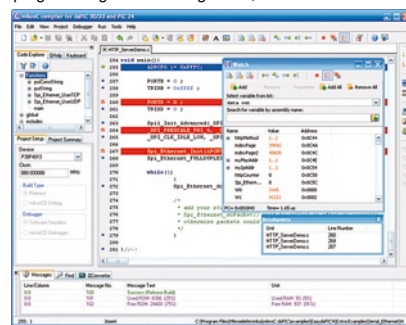
What is this good for? As an example (**Figure 4b**), if you connect a signal to Y via a resistor, you know that the current into that resistor will be the signal minus V_x (remember, $V_Y = V_X$). So the current into Y is this voltage across the resistor divided by the resistance. That same current comes out of Z, so you have now converted a signal level to a current that can be converted back to a signal via another resistor.

If the second resistor on Z is twice the input resistor on Y, you have a gain of 2, without a feedback loop. You can set the gain of this circuit with just two resistor ratios, as shown in Figure 4b. The nice thing here is that we can accurately set a gain (or attenuation, if that would be required) without having to resort to nfb.

How H.ec differs from Global Negative Feedback - or not.

Error correction works by feeding back (part of) the output signal to the input of the amplifier. It has many similarities to a classical negative feedback loop. But how is it different?

Two sides of the same medal. Figure 1a is the basic H.ec topology. You can identify two feedback loops: one positive feedback loop via V_e to V_c and back to V_e; and a negative feedback loop from the amplifier output to V_c. That particular arrangement makes H.ec so interesting. Let's assume that the output stage 'A' is ideal and has a perfect gain of 1. In that case, the correction signal at V_c is 0, because there is no difference between V_e and V_{out}. We could just throw away the whole error correction part. But if the gain A is just below 1, V_c is a negative signal which is subtracted from V_{in} and the effect is that V_e increases. This works as a positive feedback loop. In the case where the gain A would be larger than 1, the signal at V_c is a positive signal, which will be subtracted from V_{in} to decrease V_e. In this case, the system works as a negative feedback loop. So, what you see is a combined feedback loop that can act either as a positive or negative feedback loop, or not at all, depending on the error of the main amplifier block. It looks as if the phase and level of the loop adjusts itself as needed. The validity of treating a combined feedback loop by 'summing' the individual loop contributions has been established by several researchers in the past. Gerald Graeme, at the time analogue IC development manager of Burr-Brown, wrote several articles about it [1].

How does this differ from a 'classical' negative feedback loop? We can take the two feedback loops in H.ec apart as in **Figure 1b**. Error correction can be seen as a classical global nfb loop with a positive feedback loop embedded. For comparison, **Figure 1c** shows the global nfb topology for a unity-gain amplifier.

Less complexity is better. In the classical case, the high loop gain, necessary for negative feedback to reduce distortion, is obtained by means of a high-gain (ideally infinite-gain) forward amplifier block. In the case of H.ec, this high forward gain is obtained by a positive feedback loop, and the forward amplifier block can have any open-loop gain. To keep the task of a practical error correction circuit as small as possible (and therefore optimise the ec circuit linearity), the error signals should be minimal. For this, an open-loop gain equal to the required closed-loop gain is best. That means that with H.ec, similar results as with classical feedback can be reached with a much simpler amplifier. In the case of a unity gain amplifier, just an emitter follower would be sufficient, while in the classical case, even if we wanted a closed-loop gain of 1, we would still need a very high-gain amplifier block. If nothing else, this leads to the practical advantage

that the forward amplifier can be much simpler. Of course, there is no such thing as a free lunch (TANSTAFL): we have diverted some of the complexity to the two summers that must be precise circuits to make it all work. As always in audio engineering, the final advantage depends on how smart you are at implementing the various circuit parts.

Less gain is better. There is another advantage to H.ec. When a negative feedback amplifier clips, the feedback loop tries to make the output following the input by increasing the input signal more and more. But that output cannot increase beyond clipping, and as a result the amplifier is overdriven. This can be quite drastic: consider an amp with a 1 V input signal and an open-loop gain of 60 dB (1000 times) and a closed-loop gain of 30 dB (30 times) that clips at 30 V_{out}.

If the output signal is 30 V, **the effective input signal** to the 60 dB gain amplifier ($V_{in} - V_{feedback}$) is 30 V divided by the open-loop gain, or about 30 mV. If we overdrive this amplifier by doubling the input signal to 2 V, the feedback signal cannot increase due to clipping. So the effective input signal to the 60 dB gain amplifier jumps from 30 mV to 1 V! (2 V input – 1 V feedback). The clipping will become very hard, and the output wave looks as a sine wave with the tops sheared off. There is a danger of heavy internal stage overdrive, and it is possible that it takes some time, after the clipping ends, until all internal stages are returned to their stable operation point. Many amplifiers will have measures to try to avoid these conditions but it does add to the complexity. With an H.ec amplifier, the mechanism is similar. The error correction signal will increase and the input signal to the amplifier block is increased in an attempt to enable the output to follow the input. But now the amplifier block is a low-gain block, typically having a gain of only the closed-loop gain, for example only 30 dB, instead of the very high open-loop gain of 60 dB of the negative feedback amplifier. As a result, clipping and internal overdrive is not as hard. In my opinion, this is an important advantage, which is shared with valve amplifiers. Valved amps of moderate power output can sound quite pleasant, despite relatively high distortion factors, which is to a large extend caused by their soft clipping characteristics. It is also one of the reasons why very high power (300 W or more) solid state amplifiers sound better than 50 W amps with the same distortion: the high power amp clips far less than the low power amp, so the hard clipping, which can make the sound so harsh, is greatly reduced.

[1] "Generalized opamp model simplifies analysis of complex feedback schemes" – Gerald G. Graeme, EDN, 15 April 1993, pp. 175.
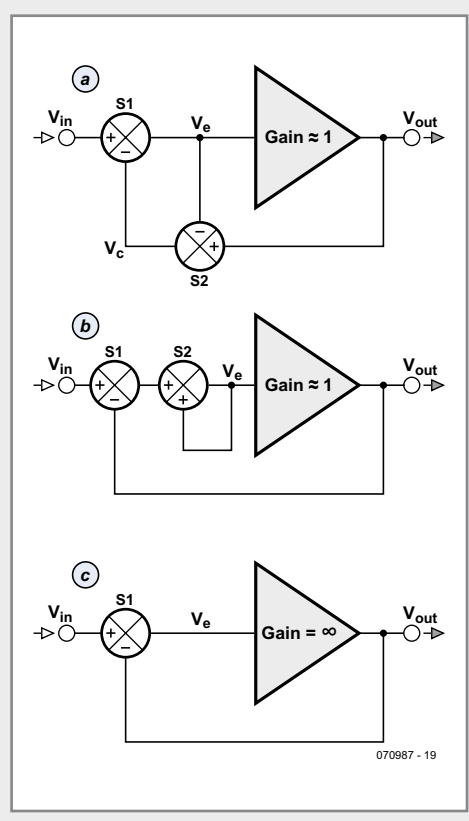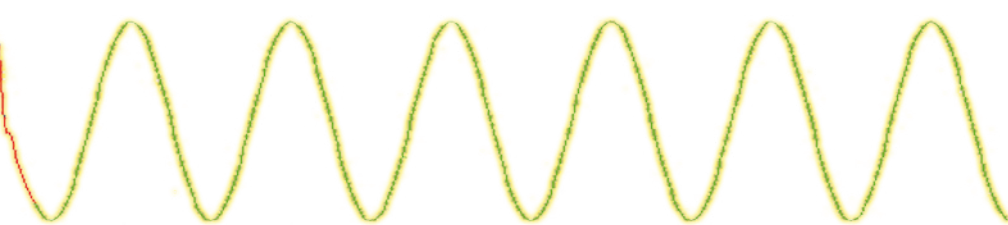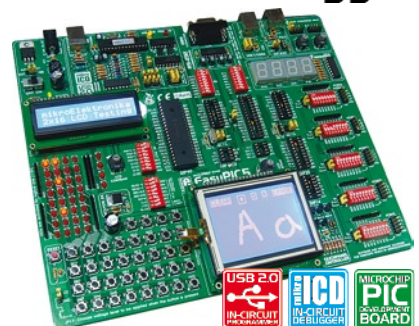

Figure 1a, top, is the basic error correction topology. In Figure 1b, middle, the two feedback loops are separated. The two figures are equivalent. Figure 1c, bottom, is a classical global nfb amplifier.
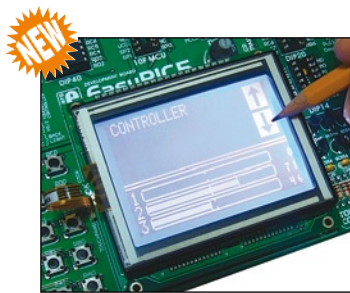
The current conveyer does not use global feedback; it is an open-loop circuit. As an aside, current conveyors consist internally mainly of current mirrors; and it can be argued that current mirrors use 100% local feedback. But we won't get into that discussion here!

There is our summer circuit: we will use equal resistors for unity gain (1).

V_c back to V_e is realized with the connection of Z to X in Figure 5.

So, at this point, we need to get practical: where do we find this CCII thing? There are CCII's around, often designed in low-voltage CMOS IC technology, and mostly used in laboratories in academic surroundings. They might as well be made from 'unobtanium'.
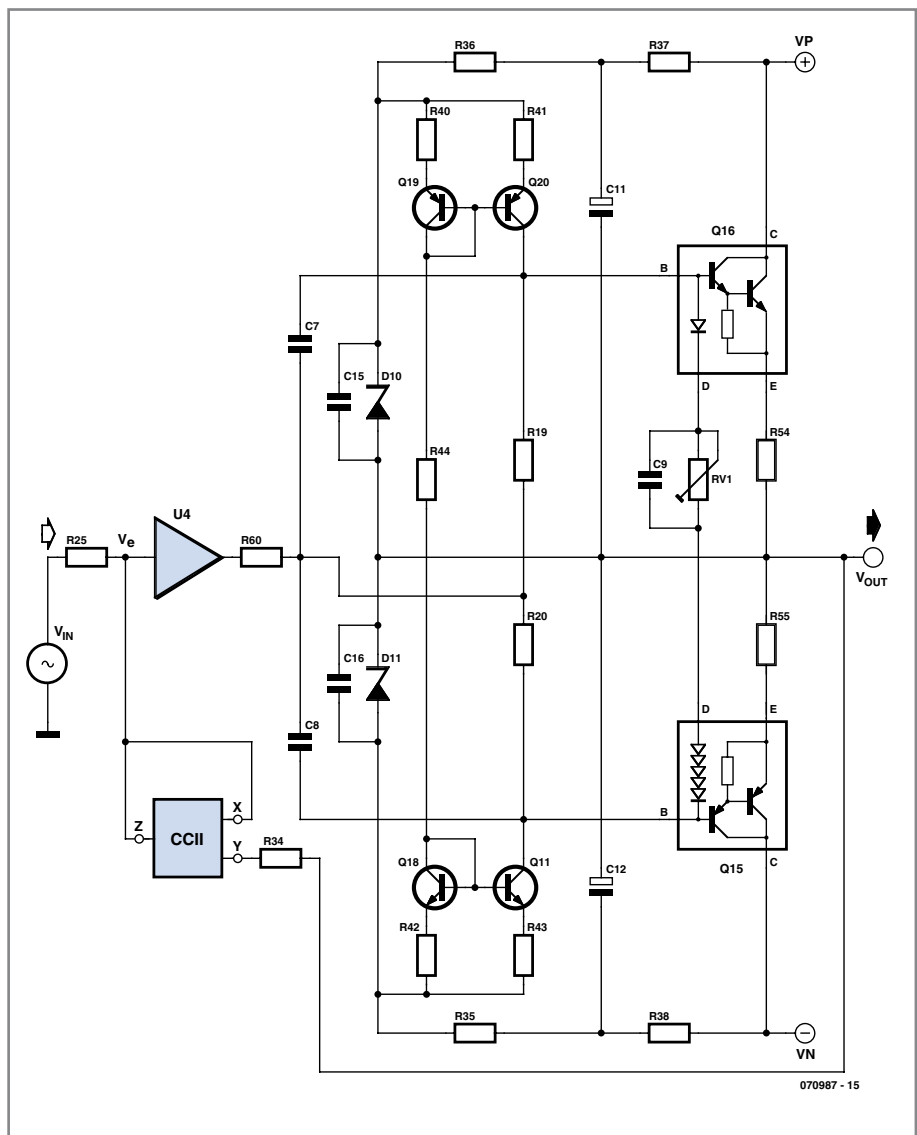
Figure 5. Output stage with error correction.

The error correction loop around my output stage is shown in **Figure 5**. The error between the input and the output appears across R34 and develops a current that is 'conveyed' to R25. This adds the correction signal to the V_{in} to make the H.ec work as we discussed before in Figure 2. The positive feedback loop in Figure 2 from V_e through

But, interestingly, CCII's are part and parcel of current feedback opamps and so-called 'diamond transistors'.

Figure 6 shows the simplified diagram of the internals of an AD844 current feedback opamp. We recognise the CCII terminals: pin 3 is the high-impedance X terminal, pin 2 the low-
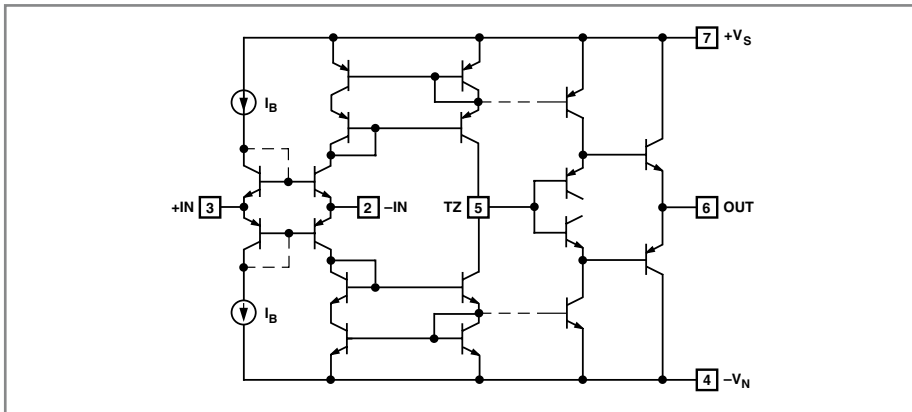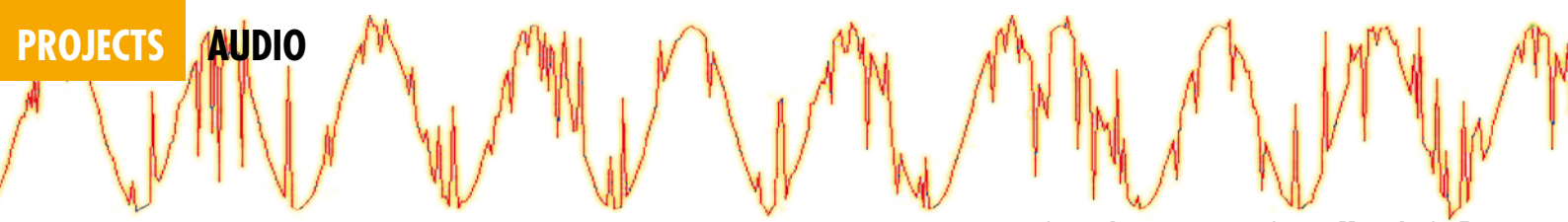
Figure 6. Basic AD844 CCII diagram plus output buffer.

impedance current input Y and pin 5 labelled Tz is the current source output Z. It is readily seen that whatever current is injected into pin 2 (Y) comes out at pin 5 (Z) but in opposite direction. It seems this part has been custom designed for us! Other similar parts are Maxim's MAX435 and MAX436 (now obsolete) and TI's OPA860, which superseded the OPA660 and OPA2660.

We can use the input stage for the CCII function, and the buffer to drive the output stage to isolate the input summer from the output stage's non-linear input impedance. We see that the current output terminal is already internally connected to the buffer input, again, just what we need!

Figure 7 shows the full output stage circuit, with component values. **Note**
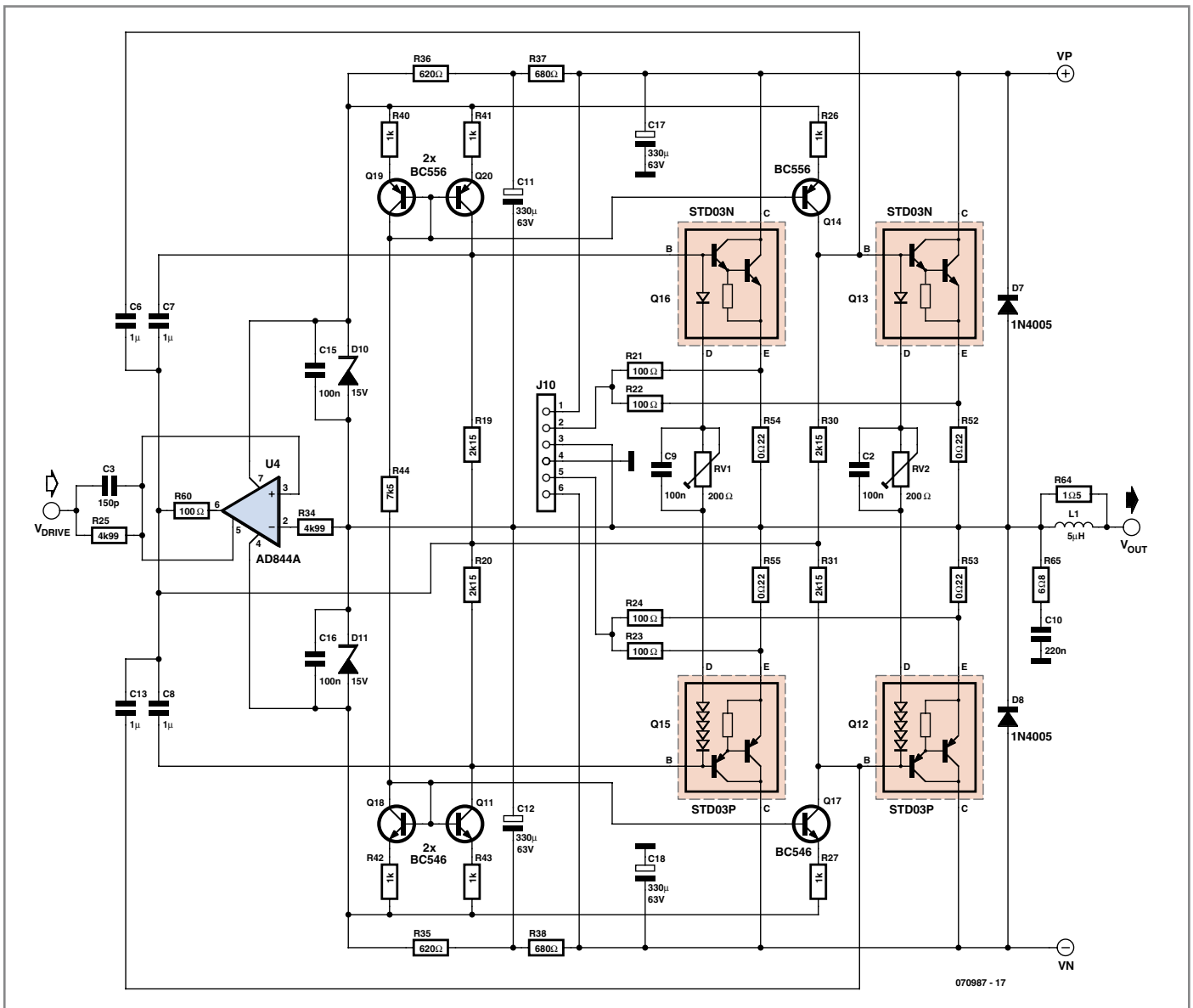

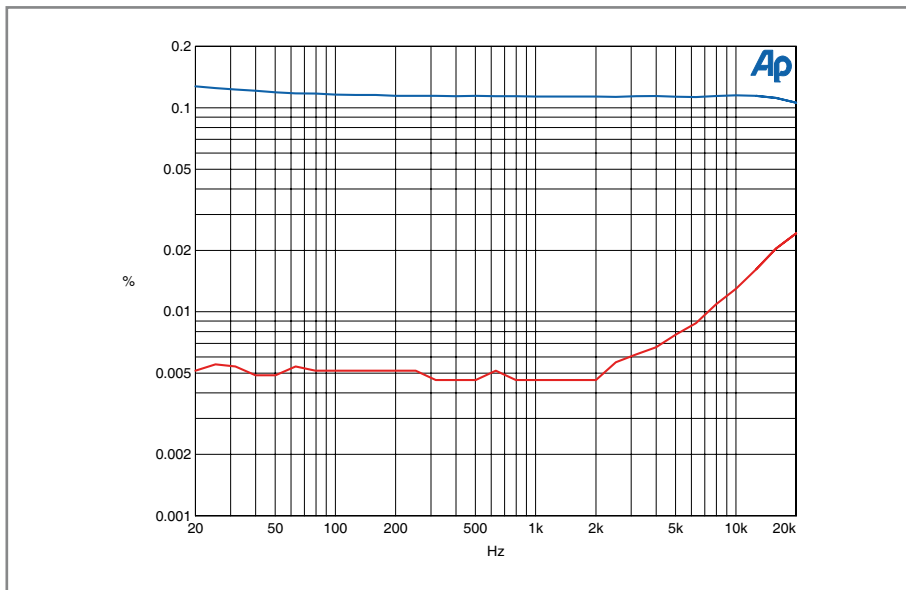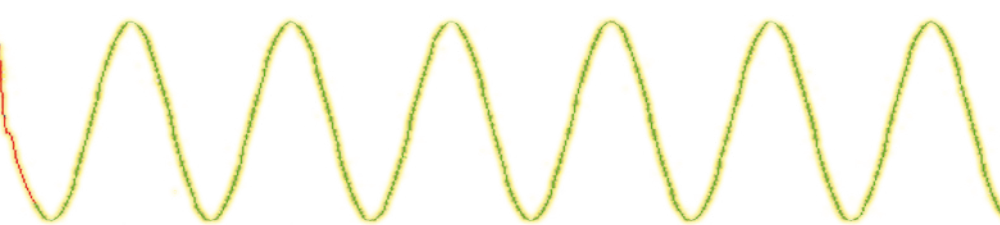Figure 7. Full unity-gain output stage with Hawksford error correction.

Figure 8. Output stage distortion curve without (top) and with (bottom) error correction. P_{out} = 50 W into 8 Ω.

Putting it all together

that the amplifier input is at pin 5 of the AD844, not at pin 2 or pin 3 as we would expect in a classical opamp circuit. This is NOT an opamp circuit, but I'm sure a lot of people will be confused by this...

Putting it all together

This output stage has two pairs of output devices. The design goal was

100 watts in 8 Ω and 200 watts in 4 Ω. For a pure resistive load, one pair of devices would have been sufficient. However, speakers are not purely resistive; depending on the signal frequency they can act capacitive or inductive, especially if there is a complex crossover filter. This leads to phase shifts between the output voltage and output current, so you can get the situation that the output voltage is negative,
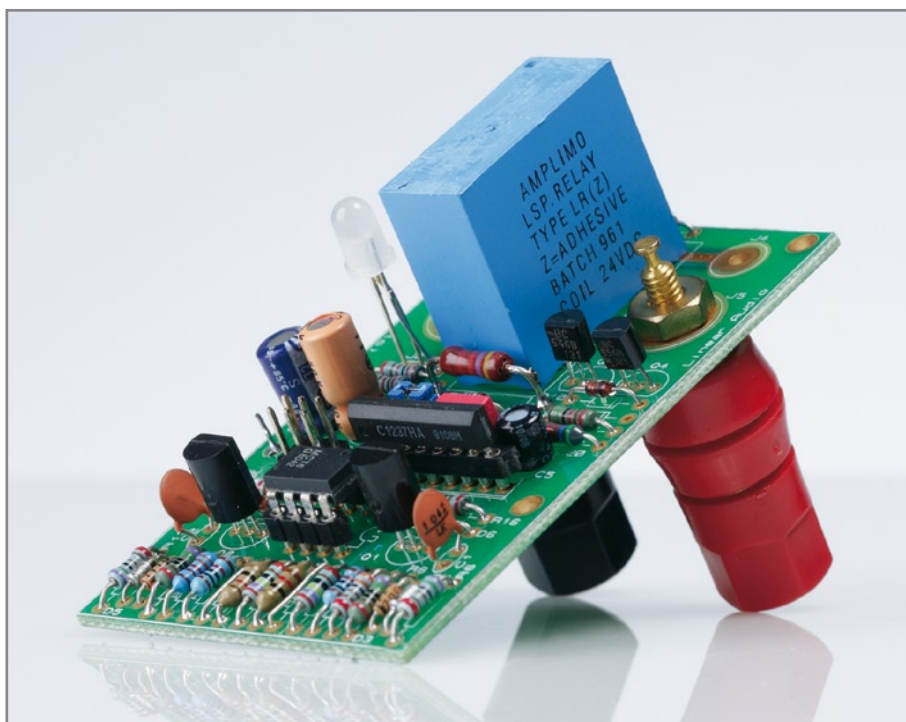

Figure 9. Prototype amplifier module and output/protection board.

but that the current is coming from the positive side (the N-device). The N-device will have a quite large V_{ce}, and the current it is allowed to source is much smaller with a large V_{ce} than what you would think from the allowable dissipation.

Further details about the safe operation area of the output devices will be given in a separate article next month. Anyway, because of the dual output devices, there is also an additional current source to bias the thermal tracking diodes, as well as an extra bias adjust trimmer.

as the current conveyor will exhibit phase shift. In a 'classical' feedback amplifier, if the phase shift becomes too large, it will turn the nfb into pfb which, as we all know, will lead to instability and even oscillations. In H.ec this is also the case of course (it shares many attributes with a feedback amplifier), so we need to roll off the loop gain for higher frequencies, just as in a classical nfb amplifier. C3 does just that by decreasing the effective correction impedance (R25//C3) with increasing frequency.

Also remember that this stage needs

output stage can perfectly stand on its own when driven by a suitable voltage amplifier (Vas) stage. So, let's take a break here; we'll attack that Vas, and the power supply, in the next instalment and develop a full-fledged, high quality audio power amp. Stay tuned.

(070987-I)

Literature and note

[1] Hawksford, M.J., 'Distortion correction in audio power amplifiers', JAES, Vol. 29, No.1/2. pp. 27-30, Jan/Feb 1981.

[2] Hephaïstos, 'Thermal Distortion – it exists,

Figuur 10. De opgebouwde versterker-module op z'n koellichaam.

There are a few extra components in the circuit which we haven't mentioned before.

R60 is a small resistor in series with the AD844 buffer output stage. A large part of the output drive goes via the two capacitors C2 and C4 to bypass the current mirrors. R60 isolates the buffer output from capacitive loads ensuring stability.

Another capacitor, C3, is placed across R25. As the frequency increases, the loop through the output stage as well

to be driven from a low impedance source, because the source output impedance forms part of the ec scaling resistor R25.

Finally, there is the 6-pin connector J10 and some associated resistors. This is the connection to the protection board which we will discuss separately. It provides the V_{ce} and I_c related information of the output devices to the protection circuitry.

This output stage is pretty linear as attested by the curves in **Figure 8**. The

I've seen it', L'Audiophile, no. 32, May 1984. (Ed.: Hephaistos was the Greek god of fire, metalworking, stonemasonry and the art of sculpture)

[3] In the early 90's, an IC designer, Doug Wadsworth, designed a current conveyor for audio on his own money (PA630). The chip, the 'Swift Current' chip, eventually found its way into Wadia DACs. It is no longer available for other parties, but I had bought some from him and knew they were a well kept secret for hi-end audio.