Input Protection in Low-Distortion Opamp Circuits

Author: Merlin Blencowe. First published on 19 May 2020.
https://www.electronicdesign.com/technologies/analog/article/21131760/input-protection-for-lowdistortion-opamp-circuits

Introduction

A subject rarely found in circuit design textbooks is the matter of non-linear junction capacitance, especially reverse-bias junction capacitance, also called depletion capacitance. When it *is* discussed it is usually with regard to fast switching circuits, and even then it is normally treated as linear. Yet in many situations, junction capacitance will be found to be the principle cause of distortion in otherwise highly-linear analog circuits. This article is principally concerned with two instances of this phenomenon, often encountered simultaneously: input protection circuits, and common-mode distortion in opamps.

Junction Capacitance

We need not go very deeply into the physics of PN junctions since we are concerned here with how to make use of existing devices rather than how to design silicon wafers. Suffice it to say that the interface between P and N contains no net charge and is called the depletion region. It therefore behaves like an insulator, sandwiched between the remaining conductive regions. We therefore have a diode, but also a capacitance. The greater the reverse voltage applied across the diode junction, the wider the depletion region grows, effectively separating the capacitor's 'plates' further. Increasing reverse-bias therefore causes shrinking junction capacitance, but the relationship is not a linear one. The junction capacitance can be estimated using:

$C = C_o/(1+V/V_b)^x$

Where:
C_o = Capacitance with zero bias
V = Applied reverse-bias voltage
V_b = Built-in voltage, about 0.6V to 0.7V
x = Empirical constant <1

A value for C_o is often quoted on datasheets, allowing for relative comparison between different devices. As a real-world example (Fig. 1) shows the above formula fitted to actual measured data for a 1N4148 –a commonly-used signal diode– demonstrating useful agreement. The curve is, of course, unpleasantly non linear.

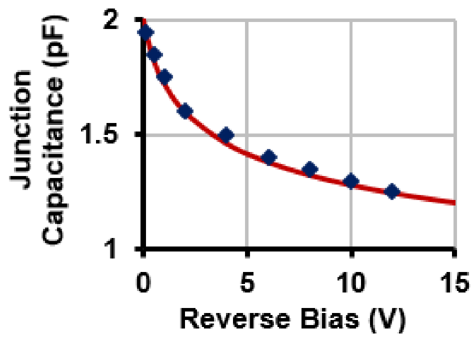

Fig. 1: Measured junction capacitance of a 1N4148 diode. Best fit line: $C = 2\text{pF}/(1+V/0.65)^{0.16}$

Junction Capacitance and Distortion

To appreciate the effect this can have on a linear circuit, consider (Fig. 2) which shows a simple network consisting of a series resistance and a pair of diodes, each reverse biased to the bipolar power rails. Such a network often forms part of an over-voltage protection circuit. Any incoming voltage greater than the supply rails (plus one diode drop) will be clamped, protecting any downstream device. In practice the resistor may be explicitly included to limit the fault current through the diodes, or it may be implicit within the source impedance of whatever the signal source is, or a mixture of both.

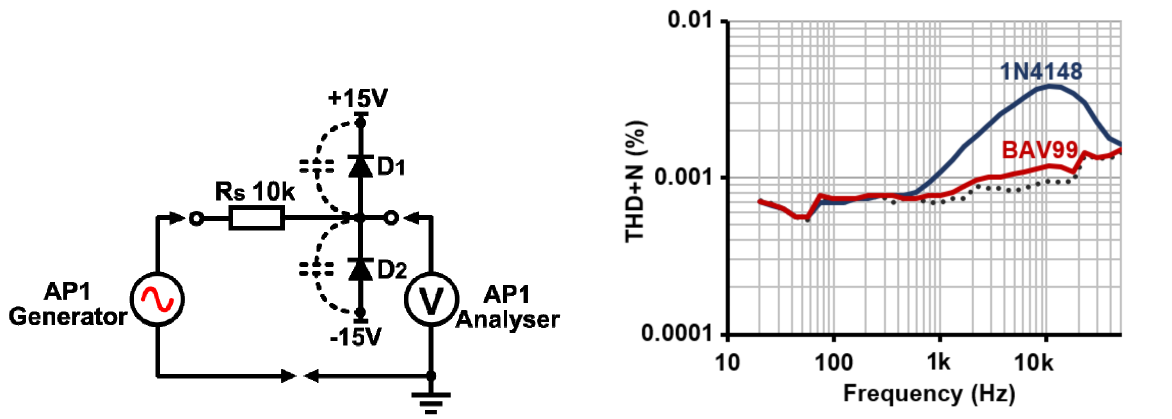

Fig. 2: Typical voltage-clamping protection circuit showing distortion due to diode junction capacitance (20dBu output). Dotted trace is the measurement floor, i.e. diodes removed.

(Fig. 2) also shows the total harmonic distortion plus noise (THD+N) measured for this circuit using an Audio Precision System 1 (80kHz measurement bandwidth) adjusted to achieve 20dBu at the analyser input. This is large enough to maintain a good signal-to-noise ratio for the analyser but still well below the conduction threshold of the diodes. The dotted trace is with the diodes removed, which is the measurement floor of the analyser. Adding a pair of 1N4148s reveals the scale of the problem: they introduce significantly more distortion, mainly odd harmonics (Fig. 3). The distortion declines above 10kHz due to harmonics being filtered out by the input capacitance and bandwidth limit of the analyser.

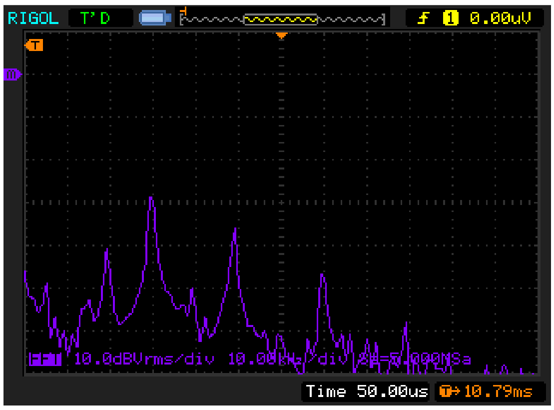

Fig. 3: Dominant odd harmonics produced by the circuit in fig. 2 with 1N4148 diodes at 10kHz (the fundamental has been nulled out by the analyser).

As a reality check that this is indeed caused by junction capacitance, consider (Fig. 4) which shows the 1N4148 capacitance curve from (Fig. 1) earlier, mirrored to represent the two diodes in the test circuit. They are effectively in anti-parallel from the point of view of the signal, so the total will be the instantaneous sum of the two. When applying the 20dBu signal, the total capacitance varies as shown in (Fig. 5), from about 2.4pF to almost 2.6pF, twice every cycle (in reality the diodes are unlikely to be perfectly matched but that is not important to the point being made). The variation itself is a distorted cosine with an RMS value of 56fF.

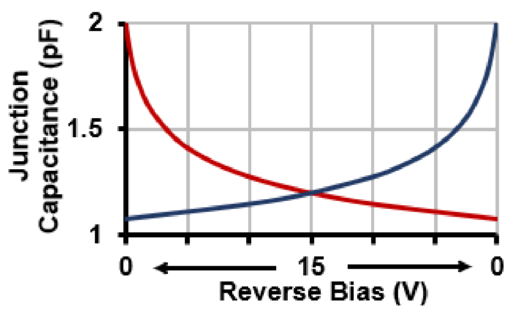

Fig. 4: Representing the variation in capacitance of the two diodes in fig. 2 during signal swing.

That a couple of picofarads varying by mere *femto*farads should have any detectable effect in the audio band might at first seem unlikely. After all, 56fF has a reactance of 284MΩ at 10kHz, which is surely of no consequence? But a 20dBu signal imposed across this reactance draws 27nA of non-linear current through the source impedance, causing a non-linear error voltage to appear across it which is effectively added to the audio signal. In this case the source impedance is 10kΩ, so the error voltage should amount to about 270μV, which is −89dB or 0.0035% THD. The actual measured value was 0.0038%. In other words, the ratio of the capacitive reactance to source impedance gives the level of distortion to be expected.

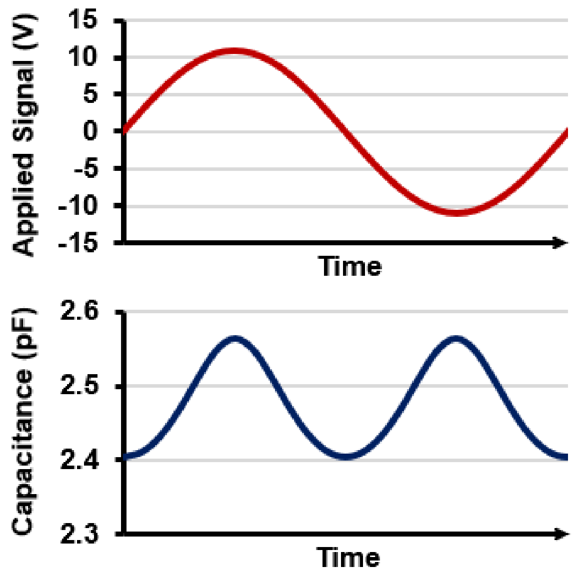

Fig. 5: Estimated variation in total capacitance for the circuit in fig. 2, derived from fig. 4 with a 20dBu applied signal voltage.

That we have identified the problem at least makes it easier to identify a solution: either make the capacitance more linear or make it irrelevant. The former is nontrivial, but we can certainly substitute a different pair of diodes having much lower capacitance. The BAV99 is such a device, containing two diodes with similar specifications to the 1N4148 but less than half the advertised capacitance. As shown in (Fig. 2), they gives a considerably better result.

Common-Mode Distortion

Having witnessed the influence of junction capacitance using discrete diodes, it becomes easier to appreciate the same effect taking place in opamps. Here it is called *common-mode distortion* because it occurs when an opamp is configured in non-inverting mode, meaning there is a common-mode voltage on each input when amplifying a signal. The distortion is caused by exactly the same mechanism of non-linear junction capacitance considered previously, but this time within the opamp itself. It will be due principally to the base-collector capacitance of the internal input transistors, and to any parasitic diodes between the inputs and the substrate.

In inverting mode there is no voltage variation at the inputs and no additional distortion, but in non-inverting mode both inputs follow the signal voltage, leading to nonlinear modulation of the input capacitance. This has led to a general engineering maxim 'always invert', but that is not always convenient, and if we also need overvoltage protection then we may be compounding the problem, as we shall see later. Before doing that, let us first explore common-mode distortion in isolation.

The effect is well demonstrated with a TL07x FET-input opamp as it has relatively large junction capacitances between the inputs and substrate.[1,2] It is also the sort of opamp used when a very large input impedance is needed cheaply, which implies a large source impedance –all the ingredients for common-mode distortion. [Fig. 6a] shows a test circuit using one half of a TL072 arranged for a non-

---

[1] Caldwell, J. (2014) Distortion and source impedance in JFET-input op amps. *Texas Instruments Analog Applications Journal*, Q4, pp4-6

[2] Gross, W. H. (1994) Source Resistance Induced Distortion in Op Amps. *Linear Technology Design Notes*, No 84.

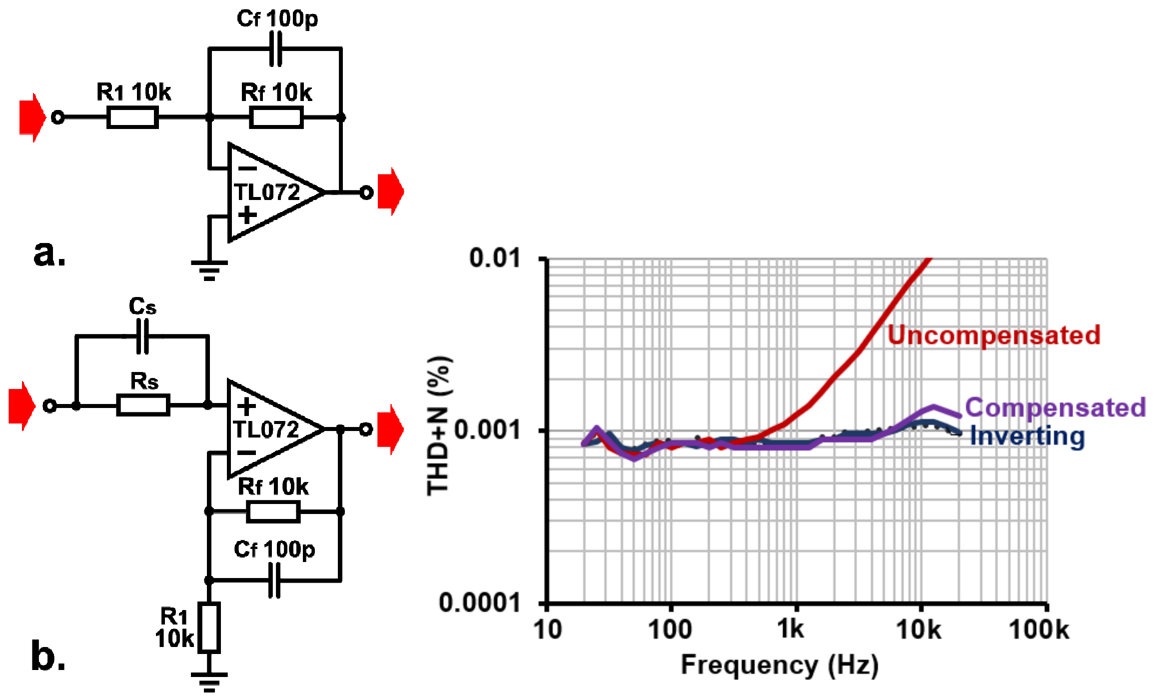

Fig. 6: Inverting opamp stage shows no measurable distortion. Uncompensated non-inverting opamp stage shows common-mode distortion, but this is almost completely compensated when $R_s = R_f||R_1$ and $C_s = C_f$. Dotted trace is the measurement floor; 14dBu input in all cases.

inverting gain of unity (but a *noise gain* of ×2). The same figure shows the measured distortion at 14dBu in/out, which is indistinguishable from the measurement floor.

[Fig. 6b] shows the circuit reconfigured for a non-inverting gain of ×2 (same noise gain). A bias-current path from the input node to ground is not shown but is assumed hereafter. The trace labelled 'uncompensated' is with the same 14dBu input level, and is considerably worsened. This is caused by the non-linear capacitance of the inverting-input pin being modulated by the feedback signal, causing a nonlinear current to be drawn through the feedback path and therefore an error voltage across it. It was strongly dominated by the second harmonic since there is only one junction this time, making the loading more asymmetric. The non-inverting input pin is similarly modulated, but since the source impedance is extremely small (50Ω for the Audio Precision) there is negligible error voltage developed there. A fine example of common-mode distortion.

The foregoing description also hides within it the solution to the problem. Since the non-inverting input is also modulated, putting a suitably matched impedance in series with it will cause an identical error voltage to be developed there, too. These distortion errors, being common mode, will be rejected by the opamp, cancelling out the (unfortunately named?) common-mode distortion. The Thévenin source resistance seen by the inverting-input is R_f and R_1 in parallel, so the required compensating resistance is 5kΩ, plus a parallel capacitor to match C_f. Forgetting to include the extra capacitor will achieve only partial cancellation, although the capacitors *can* be low quality types without spoiling performance. The result is shown by the trace labelled 'compensated', and it is barely worse than the measurement floor.

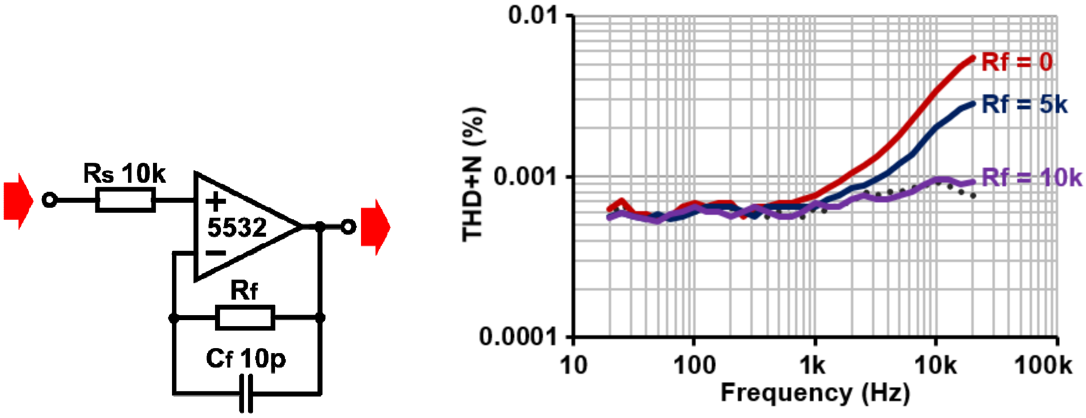

Fig. 7: NE5532 buffer with 10kΩ source resistance result in common-mode distortion. This is compensated when $R_f = R_s$. Dotted trace is the measurement floor; 20dBu input in all cases.

The price we pay for distortion compensation (apart from a few pennies in components) is Johnson noise; in this case the audio-band EIN increased from −102.6dBu with no compensation to −99.7dBu with it. Modern alternative devices like the OPA164x offer an isolated substrate and negligible common-mode distortion, but they are also a lot more expensive. It is left to the designer to decide what is more important.

A unity-gain buffer is the worst culprit for common-mode distortion since it endures the largest common-mode signals at its input. (Fig. 7) shows the result using a bipolar opamp: the NE5532. With a 10kΩ source resistance the distortion is severely degrade, but adding a matching 10kΩ feedback resistor eliminates this completely. A simple cure, but be aware that adding a resistance in the feedback loop will also introduce a pole which may reduce the phase margin and affect stability. Some opamps may therefore require a small capacitance in parallel with R_f. However, it should be sufficient to use a very small value (e.g. 10pF in this case) to avoid the need for a matching capacitance across R_s.

Input Protection plus Common-Mode Distortion

Now, suppose we need an input buffer with overvoltage protection, and suppose we are not free to use premium, low-capacitance devices, perhaps for reasons of cost. (Fig. 8) shows the previous circuit now with protection diodes to each rail –a textbook arrangement. The series resistance R_s is required in some form to limit the current through the diodes during overload. In reality this might be an explicit series resistor, or it may be the implicit source impedance of an input attenuator, or whatever.

The distortion results are also shown in the figure. Here we can see that even with a minimalist 1kΩ

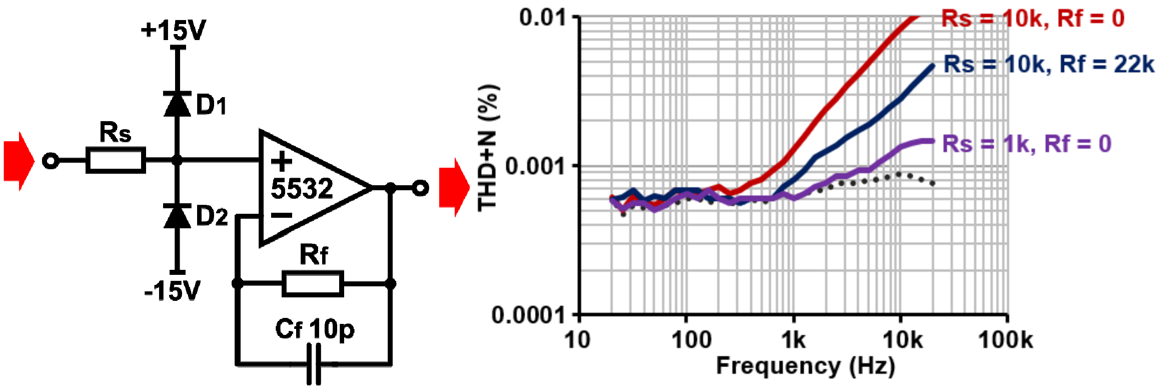

Fig. 8: NE5532 buffer with 1N4148 protection diodes. Complete distortion cancellation is not possible with this circuit. Dotted trace is the measurement floor; 20dBu input in all cases.

source resistance, distortion is already noticeably worse than the measurement floor, because the diodes are now contributing additional junction capacitance distortion. If we need the fault-limiting resistance to be larger, say 10kΩ, distortion becomes woeful, exceeding 0.01% at 20kHz. Adding a matching resistance in the feedback loop can now only partially compensate for all the distortion taking place at the non-inverting input node. Manual adjustment proved that 22kΩ produced the best compensation, but even this is disappointing.

How can we improve on this design? One option is to move the protection diodes to the inverting input. A pair of antiparallel diodes must then also be added between the two inputs to complete the fault current path from input to either rail. Since there is normally no voltage difference between the opamp inputs, i.e. across these diodes, their junction capacitance remains constant. In other words they are bootstrapped. In fact, the 5532 already has these diodes internally as shown in (fig. 9) (when relying on internal diodes the fault current should be limited to <5mA to avoid fusing the internal bond wires[3]). With this arrangement we have a similar situation to previously: one opamp input sees only common-mode distortion, but the other sees common-mode distortion *plus* protection diode distortion, so full distortion cancellation is again not possible. But it does bring down the overall distortion levels and leads to a smaller compensation resistance, meaning less noise. In this case a 3.3kΩ feedback resistor gave optimum cancellation with a 10kΩ source resistance. The reason for showing this circuit is for interest, because it is the approach used in the Audio Precision S1 analyser itself.

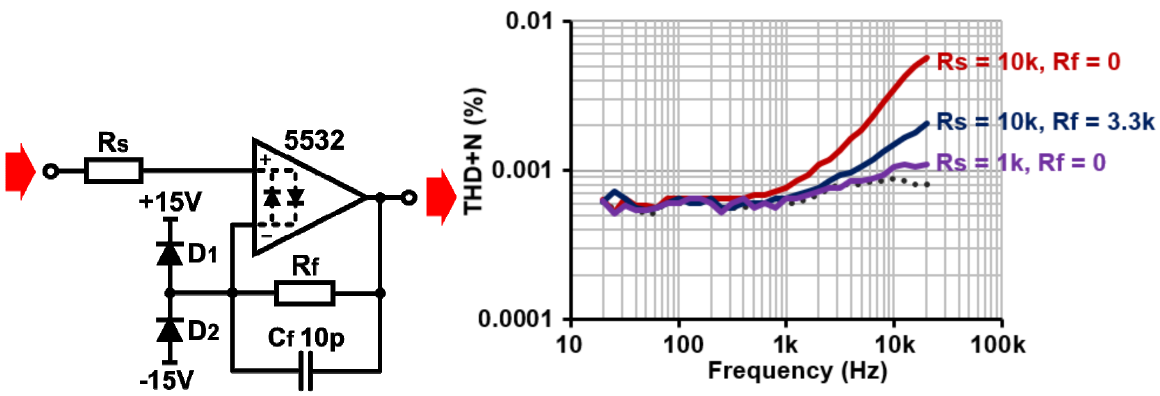

Fig. 9: NE5532 buffer with alternative protection architecture, again using 1N4148 diodes. Complete distortion cancellation is not possible with this circuit. Dotted trace is the measurement floor; 20dBu input in all cases.

[3] Buxton, J. (1994) Simple Techniques Protect Amplifiers from Input Overvoltage. *Analog Dialogue*, Vol 28, No3, pp13-16.

A Better Solution

Can we do better still? And can we do it without needing a distortion meter to find the optimum cancellation resistance? The answer is yes. Keen-eyed readers will already have spotted the clues laid in the previous paragraphs. The best approach is to maintain the *same* junction capacitances at both input pins and combine them with matching source impedances. This pretty much guarantees optimum distortion cancellation with no special tools required, and is independent of the opamp type being used. Fig. 10 shows the circuit. Putting $R_f = R_s = 10\text{k}\Omega$ now produces the same result as a mere $1\text{k}\Omega$ source resistance alone. The remaining rise at high frequencies is mainly due to residual mismatching between the diode pairs. Even atrocious power diodes still yield quite good results this way.

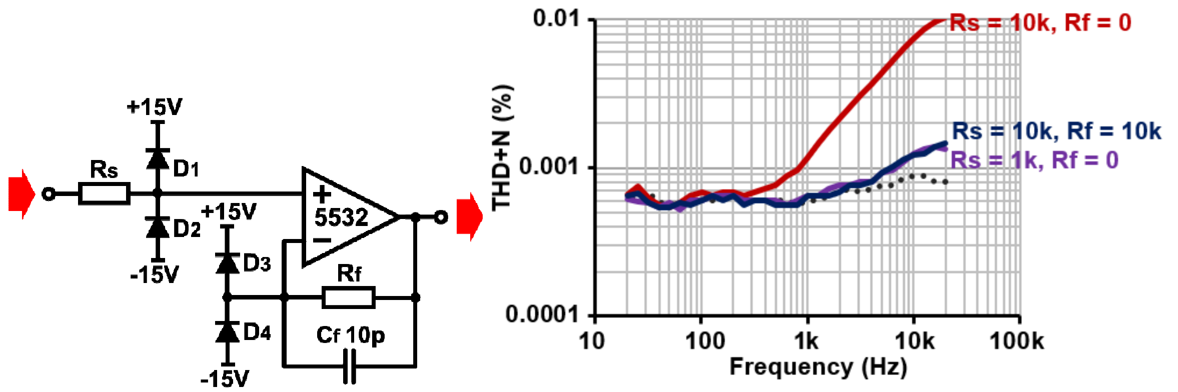

Fig. 10: Improved NE5532 buffer with 1N4148 protection diodes. Complete distortion cancellation is now possible. Dotted trace is the measurement floor; 20dBu input in all cases.

Nevertheless, a final criticism that can be levelled at the previous circuit(s) is that the fault current is pumped into the rail(s), which may not be able to sink it. This can be corrected by returning the protection diodes to dedicated shunt references, e.g. a pair of Zener diodes. Fault current will then be directed safely to ground, and the Zeners can of course be chosen to suit the clamping requirements. It is essential to bias the Zeners with some standing current, otherwise gross distortion will result. Fortunately, the standing current can be very small, less than 1mA if necessary. (Fig. 11) shows the improved circuit. When combined with low-capacitance diodes, ideally in a single package such as a BGX50A (a single package gives some hope of good matching between diode pairs) exemplary performance obtains. As shown in the figure, with proper compensation there is no significant distortion within the audio band.

The circuits shown so far have used a $10\text{k}\Omega$ input resistor which is representative of many real-world interfacing situations. If the Zener clamping option is used then peak input overloads of several hundred volts can be handled this way, provided the Zeners and limiting resistance have sufficient power rating. However, both the NE5532 and TL072 exhibit further HF distortion with source impedances much above $10\text{k}\Omega$, even after compensation, so for very large source impedances other opamps must be tried. For example, the OPA1662 and OPA1678 perform well with source impedances up to at least $100\text{k}\Omega$.

Low Noise, High Linearity

A further option for limiting current without using a simple resistance is to use a current-clamping circuit like the one in (Fig. 12), built from depletion MOSFETs. Under signal conditions the MOSFETs short out their own body diodes and behave like a total resistance of only about $3\text{k}\Omega$, which is distortion compensated by R_f. If the voltage across the MOSFETs exceeds a couple of volts they enter the saturation region and current is limited to the I_{DSS} of about 2mA.[4] The reduced resistances minimise noise contribution while still allowing overloads up to 500Vdc to be tolerated. Of course, if the source impedance is variable, perhaps because it is a switched attenuator or potentiometer, then either we must vary the compensation impedance in sympathy (as was done in the Audio Precision analyser), or else use a compromise value and live with it.

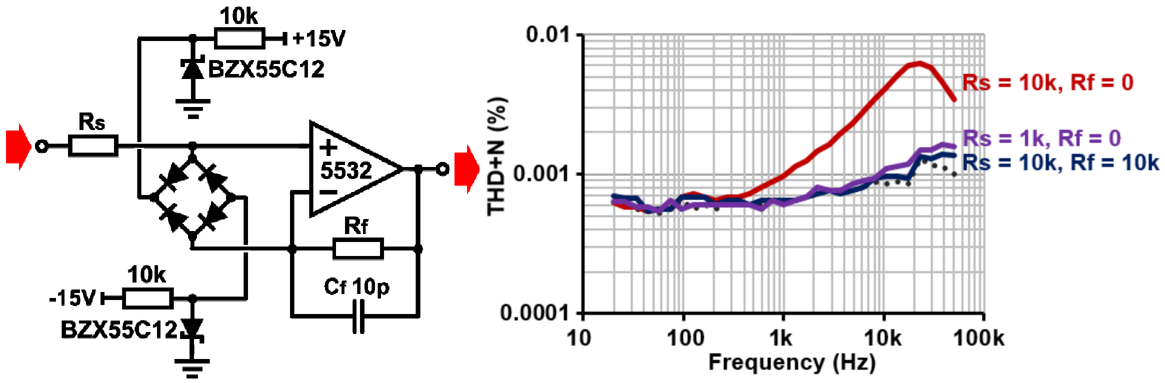

Fig. 11: Superior NE5532 buffer with shunt overvoltage protection. Outstanding results obtain with a BGX50A low-capacitance bridge rectifier package. Dotted trace is the measurement floor; 20dBu input in all cases.

[4] Horowitz, P. & Hill, W. (2015). *The Art of Electronics, 3rd ed.* Iliffe & Sons Ltd., London, p. 362.