Linearity Optimization in Negative-Feedback Amplifiers

Martin Lantz
Lund Institute of Technology
martin.lantz@tde.lth.se

ABSTRACT

A method for estimating the linearity in multistage negative-feedback amplifiers has been developed. This analysis method help the designer of high-performance negative-feedback amplifiers to find the appropriate number of amplifying stages needed to fulfill the system specifications. This can be done very early in the design procedure.

The analysis also reveal the constraints on linearity performance, which are dictated by the power supply, and technology. Designers familiar with 'structured electronic design' will find that the nonlinear superposition model fill a gap in the design theory. The design trajectory, thus become more strait forward. Power-series theories are the foundation for this analysis. In order to give better insight memory-less systems are treated in more detail.

INTRODUCTION

This work is the product of a wish to structure electronic design in the same way as have been done and is being done at Delft University of Technology, the Netherlands. One can say that a 'design theory' gradually is developed each time an achievement in this direction is made. Efforts to structure electronic design are rare, however. This is unfortunate since, once the electronic design has been structured it is easy to teach, it gives the designer deeper insight, and it allows close-to-optimum design results to be reached with a minimum of iterations. The alternative, to do a random search of the 'design space', does not provide these benefits and as it appear no others either. The design of negative-feedback amplifiers [1], oscillators, references, integrated continuous-time filters, and frequency demodulation have been structured.

Realizing that the linearity optimization of negative-feedback amplifiers so far has not yet been structured, the author saw an opportunity to contribute to the development of the design theory. When linearity optimization is structured it is possible to give answers to general questions, some of them have of cause already been answered:

- Is the amplifier with the highest loop gain always the most linear?
- Is the most distortion free amplifier the one that has little distortion before the negative feedback is applied?
- Which semiconductor mechanisms influence the linearity?

This work give more detailed answers to questions previously handled without considering multistage realizations:

- How to estimate the linearity in a negative-feedback amplifier?
- Which factors sets the ultimate limit to the linearity performance?
- Which is the highest IMFDR achievable?
- How do MOST perform compared to BJT?
- Does the signal-swing bias ratio (modulation) play a role?
- How many amplifying stages are needed to fulfill the linearity requirement?

NEGATIVE-FEEDBACK AMPLIFIERS

An amplifier performs one of the most basic operations on electronic signals: multiplication with a constant. The task to implement this basic operation is, however, more delicate than one might think. This is even more so when high performance, close to electronic system's essential limitations, is needed.

Amplifiers with high linearity produce signals with very small deviations from those of a linear system. These deviations — known as distortion — have several origins: speed limitation in the amplifier, signals in excess of the signal-handling capability and nonlinearities in the devices that are used to realize the amplifier — that is, transistors.

In the beginning of the twentieth century designers found great difficulty in achieving amplifiers with sufficiently high linearity. In 1927 Black invented a remedy: the negative-feedback amplifier. The linearity performance of such amplifiers are closely related to the amplitude of the loop gain. The operating points of the active devices influence the loop gain. The number of amplifying stages coupled in cascade is, however, the design parameter with the greatest influence on the loop gain.

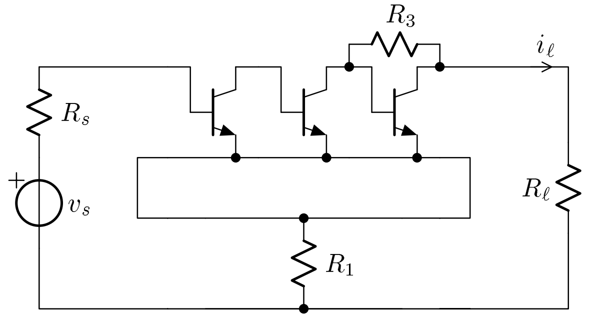

Figure 1: Three-stage transconductance (V→I) amplifier with local shunt feedback R_3 at the output stage.

It does not make sense to use more amplifying stages than needed, since high-frequency compensation becomes more difficult when stages are added. It seems valuable that the designer of multistage negative-feedback amplifiers in an early stage can predict the linearity, allowing the appropriate number of amplifying stages to be chosen. The analysis method developed is designed to provide this information.

MEMORY-LESS NONLINEAR SYSTEMS

In this section a model describing nonlinear memory-less negative-feedback systems is presented. The model allow designers to analytically analyze the nonlinear behavior of amplifiers. Also the circuit-level requirements dictated by system specifications are found. This is a prerequisite for a hierarchical design approach. First a distinction between large- and small-signal distortion mechanisms are made:

Large-signal distortion arise when the signal level is in excess of the bias levels. Such large levels give rise to terribly distorted signals. This type of distortion is also called clipping. Large-signal distortion must be avoided by making the bias currents and voltages sufficiently large. A lower limit to the values of the bias currents and voltages can therefore be found. When reactive elements are present the requirements on the bias levels are frequency dependent, and hence effects like slewing and transient intermodulation distortion need to be considered.

Even when the signal levels are orders of magnitude smaller than the bias levels some distortion will be present. Small-signal distortion has the origin in the nonlinear transfers of active devises. If these transfers are smooth they can be described by power-series representation. Due to their wide spread acceptance and developed theory, power series will be used through out this work. Phenomena, like clipping, which give sudden changes in the transfers are difficult to model if power series are favored, however. Clipping effects is consequently handled separately.

The high-order terms of power series give insignificant contributions if the signal excursion is small. If the signal is sufficiently small the power series can be truncated after the third-order term without introducing too much error. If these truncation errors are acceptable the system is said to be weakly nonlinear. This assumption is accepted here, and hence the analysis developed is limited to second- and third-order transfers.

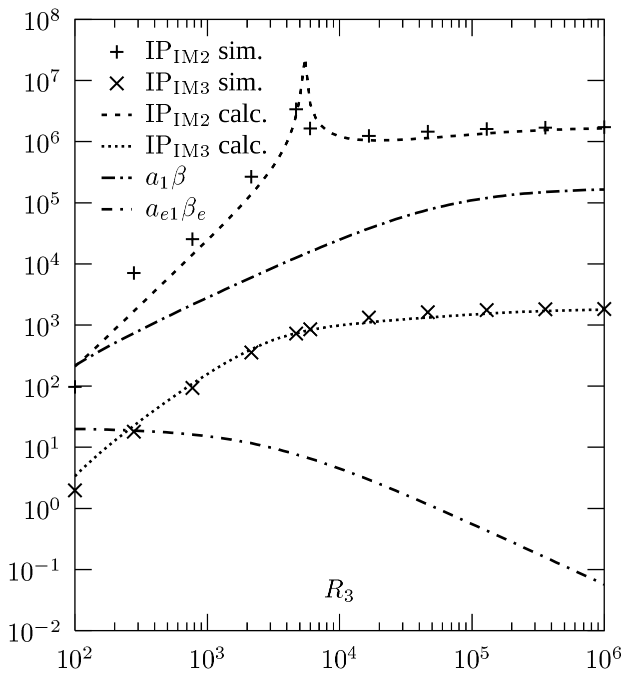

Figure 2: Transconductance amplifier with local shunt feedback R_3 $[\Omega]$ at the output stage. The input-referred intermodulation-distortion intercept points $\mathrm{IP}_{\mathrm{IM2}}$ $[\mathrm{V}]$ and $\mathrm{IP}_{\mathrm{IM3}}$ $[\mathrm{V}]$ are chosen to quantify the linearity — large values signify high linearity. $a_1\beta$ $[\cdot]$ is the global loop gain and $a_{e1}\beta_e$ $[\cdot]$ is the local loop gain.

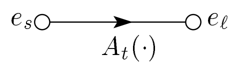

Figure 3: General nonlinear system. $A_t(\cdot) = \sum_n A_{tn}(\cdot)^n$ is a nonlinear transfer function.

When it comes to choosing mathematical tools for analysis of systems described using power series, a distinction between networks including elements with memory and those networks that do not, has to be made. In the general case where elements with memory (frequency-dependent) are allowed the Volterra theory and series are suitable tools for analytical investigations. The results found — possibly by a computer — by using Volterra series are, however, frequently very complex, and difficult to interpret. In the special case where no memory effects are allowed less sophisticated theory can not only be used, but should in fact be preferred, since the results found then give more insight. Information on how the ignored but still present memory effects influence the distortion behavior is then lost, however.

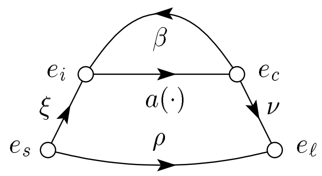

Figure 4: Signal-flow graph describing the nonlinear superposition model. Where $a(\cdot) = \sum_n a_n(\cdot)^n$ is a bijective nonlinear transfer function.

The goal for this section is to derive expressions for the first-, second-, and third-order transfers of the nonlinear function $A_t(\cdot)$, see Figure 3. This information can be used together with the distortion figures of merits in following section, to analyze the linearity performance. Even more useful is the possibility to find the demands on the transfers as dictated by the system specification. This is the basis for amplifier synthesis.

If the system behave weakly nonlinearly its transfer can be represented by a power series,

$$e_\ell = A_t(e_s) = \sum_n A_{tn} e_s^n. \tag{1}$$

The amplifier is assumed to behave weakly nonlinearly for the applied signal amplitude and can therefore be accurately described by the first three terms of its Taylor series. In the special case where the feedback network is linear and nonlinearities exclusively appear in the transfer $a(\cdot)$, that is all other transfers $\xi,\ \rho,\ \beta,$ and ν are considered to be linear, the over-all transfers are,

$$\begin{aligned} A_{t1} &= \frac{\nu\xi a_1}{1 - a_1\beta} + \rho, \\ A_{t2} &= \frac{\nu\xi^2 a_2}{(1-a_1\beta)^3}, \text{ and} \\ A_{t3} &= \nu\xi^3 \frac{a_3 + 2a_2^2\beta/(1-a_1\beta)}{(1-a_1\beta)^4}. \end{aligned} \tag{2}$$

DISTORTION

Two sinusoidal signals applied simultaneously at the system input are relevant test signals for systems receiving information coded as narrow-band phase, frequency, or amplitude modulation on a sinusoidal carrier, a radio receiver is one example. The first sinusoidal e_1 signal represent the wanted information carrying signal. The second sinusoidal e_2 model an interfering unwanted signal.

$$e_s = \widehat{e}_1 \cos(\omega_1 t) + \widehat{e}_2 \cos(\omega_2 t), \tag{3}$$

where $\widehat{e}_1$, and $\widehat{e}_2$ are the amplitude of the sinusoidal input signals.

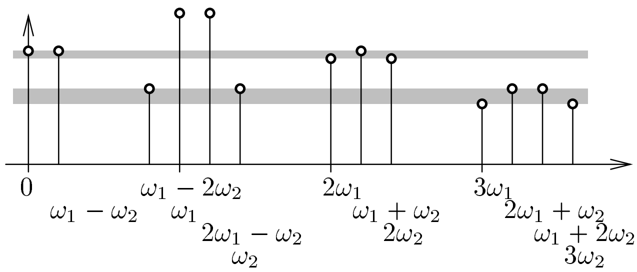

Figure 5: Two tone test, single-sided output signal spectrum.

Nonlinearities in the system give rise to several tones at frequencies different form the two original, see Figure 5. Some of them, called intermodulation products, are particularly troublesome. The magnitude of these products are normalized to the fundamental and called: second-order intermodulation distortion IM_2, and third-order intermodulation distortion IM_3 respectively. If $\widehat{e}_s = \widehat{e}_1 = \widehat{e}_2$ then this result can be used to form:

$$\begin{aligned} \mathrm{IM}_2 &= \widehat{e}_s \left| \frac{A_{t2}}{A_{t1}} \right|, \text{ and} \\ \mathrm{IM}_3 &= \frac{3\widehat{e}_s^2}{4} \left| \frac{A_{t3}}{A_{t1}} \right| \end{aligned} \tag{4}$$

Input-referred intermodulation-distortion intercept points $\mathrm{IP}_{\mathrm{IM2}}$, and $\mathrm{IP}_{\mathrm{IM3}}$.

$$\begin{aligned} \mathrm{IP}_{\mathrm{IM2}} &= \left| \frac{A_{t1}}{A_{t2}} \right|, \text{ and} \\ \mathrm{IP}_{\mathrm{IM3}} &= 2\sqrt{\left| \frac{A_{t1}}{3A_{t3}} \right|} \end{aligned} \tag{5}$$

LOCAL FEEDBACK

The influence of local feedback is known in qualitative terms. Both that global feedback is more effective than local feedback, and that the preceding stages have to handle larger signals when the gain is reduced in the output stage by local feedback is well known [2].

Local feedback can be used to reduce the global loop gain to facilitate high-frequency compensation. The linearity performance deteriorate in general by the same amount as the global loop gain is reduced. On the other hand if the local feedback operate on the amplifying stage that dominate the linearity performance the linearity loss can be small. This is so since the global loop gain loss is compensated by the linearizing effect that the local feedback has on the amplifying stage. The best place for local feedback is at the amplifying stage that dominates the linearity performance, that is the output stage.

This exchange between global and local loop gain with almost no linearity loss is only possible for moderate levels of local loop gain. The linearity performance of the

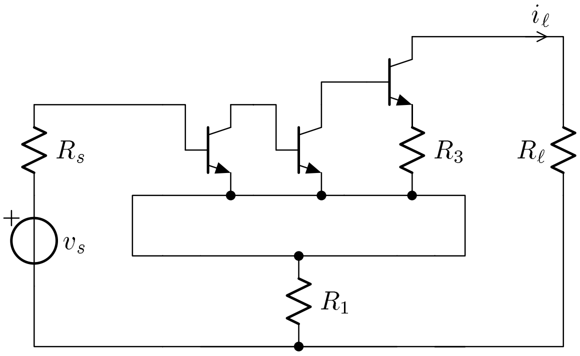

Figure 6: Three-stage transconductance (V→I) amplifier with local series feedback R_3 at the output stage.

amplifier can not benefit from the linearity improvement of the amplifying stage as the local feedback grows strong, since then other amplifying stages will dominate. Two realizations of a transconductance amplifier, one with series and the other with shunt local feedback at the output stage, is depicted in Figures 6 and 1. The local loop gain $a_{e1}\beta_e$ is made stronger as the local series feedback resistor R_3 is increased. Figure 7 demonstrates how no linearity performance is sacrificed at moderate levels of local feedback. At high levels of local loop gain there are a dramatic degradation in linearity, however. The local shunt feedback influence the linearity in an analog way, see Figure 2.

It is worth pointing out that, the linearity of the amplifier is not improved by applying local feedback except when cancellation effects occur, see the peaks in $\mathrm{IP_{IM2}}$ in Figures 7 and 2. Such cancellation effects are seldom exploited in industrial applications since the cancellation effects are very sensitive to changes — temperature, bias currents and voltages for example — in the environment of which the amplifying stages are operating.

DISCUSSION

In the literature various arguments on whether the designer should use local feedback or not are found. Here two of the most justified arguments are presented. Finally the results derived so far is discussed.

The first view is based on the assumption that the number of amplifying stages is fixed by an upper limit due to bandwidth and stability considerations. Local feedback always give lower linearity and should be avoided altogether. The least harmful location for local feedback is the output stage, however [1].

The second view is the assumption that measures should be taken to make the output stage to dominate the distortion. Local feedback might violate the assumption and hinder structured design. If a loop gain reduction is necessary, due to high-frequency compensation difficulties, local feedback should be placed at the intermediate stages

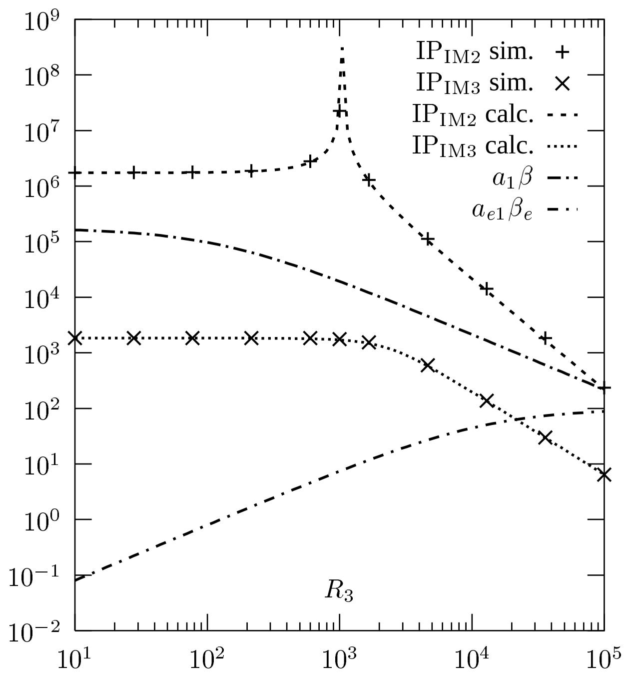

Figure 7: Transconductance amplifier with local series feedback R_3 $[\Omega]$ at the output stage. The input-referred intermodulation-distortion intercept points $\mathrm{IP_{IM2}}$ $[\mathrm{V}]$ and $\mathrm{IP_{IM3}}$ $[\mathrm{V}]$ are chosen to quantify the linearity — large values signify high linearity. $a_1\beta$ $[\cdot]$ is the global loop gain and $a_{e1}\beta_e$ $[\cdot]$ is the local loop gain.

to not violate the orthogonality assumption [3].

The results found so far indicate that no linearity improvements are gained using local feedback, in consensus with [1]. The linearity loss is the smallest if the local feedback is located at the output stage. If the local feedback is very strong, severe linearity loss occur due to intermediate stage contributions. In such case it might be a good idea to reduce the number of amplifying stages.

Local feedback should only be used to relieve high-frequency compensation or biasing difficulties.

REFERENCES

[1] E. H. Nordholt, *Design of High-Performance Negative-Feedback Amplifiers.* PhD thesis, Delft University of Technology, department of Electrical Engineering, the Netherlands, 1980.

[2] M. G. Scroggie, “Distortion in negative feedback amplifiers,” *Wireless Engineer*, vol. 15, p. 91, Feb. 1938.

[3] C. J. M. Verhoeven, A. van Starveren, G. L. E. Monna, and M. H. L. Kouwenhoven, *Structured Electronic Design*. Delft University of Technology, department of Electrical Engineering, the Netherlands, May 2000.