

CIRCUITS FOR THE PRECISE SYNTHESIS OF THE SINE FUNCTION

Indexing terms. Bipolar-transistor circuits, Function generators

Bipolar circuits are described for generating the sine function with high accuracy from an analogue angle input. The technique offers operation over any number of quadrants without switching, high bandwidth and circuit simplicity. The identity of the sine to a summation of tanh terms is utilised, a proof of which is provided.

Methods for generating an analogue signal proportional to the sine of an analogue input include piecewise-linear approximations,¹ polynomial or other continuous-function techniques using analogue multipliers^{2,3} and specially designed translinear circuits.⁴ Invariably, operation is restricted to the range -90° to $+90^\circ$. Servomechanical systems can be designed to extend the angular range, using nonlinear potentiometers and gear trains, with obvious disadvantages.

The technique described here is unique in two ways. First, it provides operation over any desired angular range in a continuous manner ($\pm 360^\circ$ being simply achieved). Second, the output function has an extremely small theoretical error for a circuit of modest complexity, and can be shown to converge exactly on the sine function in the limit. Practical accuracy is determined solely by component quality in most cases, and can be maintained up to high frequencies. A monolithic implementation is essential.

The differential output of a long tailed pair (l.t.p.) of bipolar transistors operating with an emitter supply of I is related to the differential base voltage E_b thus:

$$f(E_b) = \frac{I_{out}}{I} = \tanh \frac{E_b}{2kT/q} \quad (1)$$

This can be used to provide a crude 2-quadrant fit to the sine. With the addition of emitter degeneration a tolerable approximation can be achieved.^{5,6} However, the sine function has zero slope at $\pm M\pi/2$, whereas a circuit based on a single l.t.p. stage always exhibits finite slope at $\pm\pi/2$. For precise applications it is desirable that the zero-slope condition is not only met, but that the output continues to be valid beyond $\pm\pi/2$, at least far enough to cater for overdrive situations and preferably to any desired angular range. This can be achieved by combining the output of several l.t.p. stages in alternating antiphase, as illustrated in Fig. 1a, where five stages are used ($N=5$). All stages receive the input E_b and have an emitter supply of I , and outer stages have base-voltage offsets of nE , where $-m < n < m$ and $m = (N-1)/2$. It is easily shown that the overall transfer function is

$$f(E_b) = \sum_{n=-m}^m (-1)^n \tanh \frac{E_b + nE}{2kT/q} \quad (2)$$

which has N zero crossings, and, with correct choice of offset voltages, can be made to fit the sine function very closely. As N approaches infinity, the fit becomes perfect, as shown in the Appendix.

This is one of a pair of exponential functions which have been found to exhibit a very close fit to the sine, continuous over any practical range within control of the designer. The other function will be described in future correspondence; it has the practical advantage of requiring only one transistor per zero crossing, but as yet no rigorous proof is available. Both functions have been the subject of very extensive numerical investigation since 1967 and techniques have been developed to optimise circuit performance over a chosen angular range.

Fig. 1b illustrates the behaviour of the circuit using dual $n-p-n$ transistors operating at $I = 100 \mu A$ and with $E = 75 mV$ and an ambient temperature of 295 K. The differential-current-to-voltage convertor A has a transresistance of $100 k\Omega$, so the asymptotic value of the function at its output is $\pm 10 V$. Note that the amplitude of the sine function is about 30% of this amount, which is in agreement with both the numerical

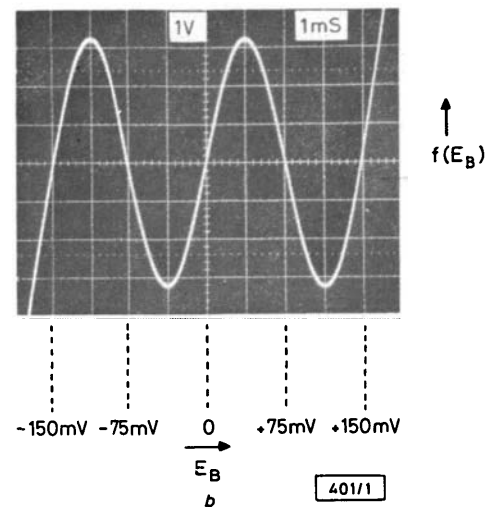
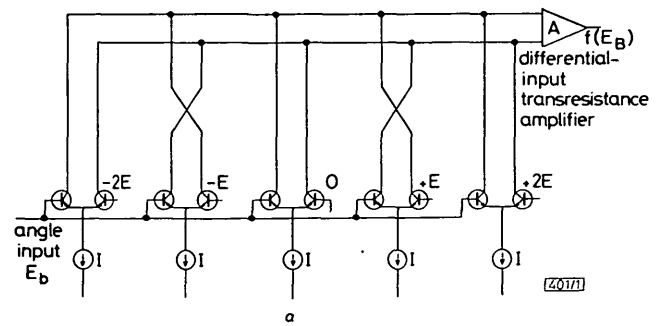


Fig. 1
a Circuit for generating a summation of alternating hyperbolic tangent functions, which closely approximate the sine function
b The output of this circuit for $I = 100 \mu A$, $E = 75 mV$, $T_A = 295 K$

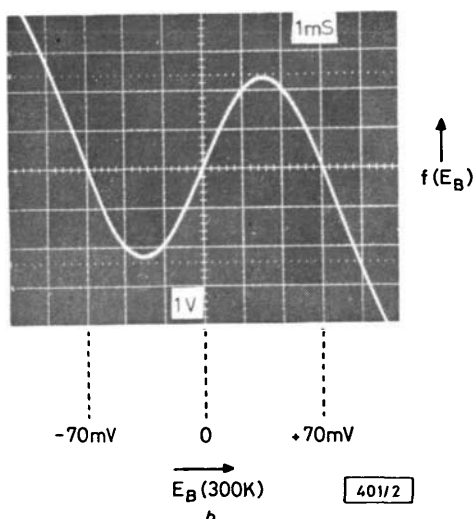
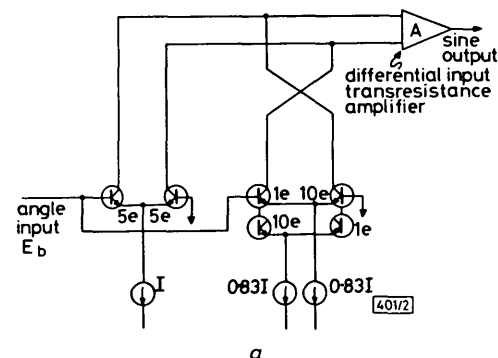


Fig. 2
a A simple implementation of the principle, using emitter area ratios to introduce the offset voltages
b The output of this circuit. In all cases the amplifier A had a transresistance of $100 k\Omega$

analysis and the theoretical result (eqn. 15 of Appendix). For this demonstration the outer offsets were made very slightly greater (2%) than $2E$ to equalise the peaks at $\pm\pi/2$ and $\pm\pi$. As N becomes large exactly-integral offsets can be used. For $N = 9$ the minimum error (computed over the range $-\pi/2$ to $\pi/2$) occurs when $E = 57$ mV (at an ambient of 300 K). The peak departure from an exactly sinusoidal form is then $\pm 0.007\%$ and the r.m.s. error is less than 0.005% . Detailed curves of the error minima are provided in Reference 7.

Temperature stability: The complete design of an ultraprecise sine generator requires a careful appraisal of V_{be} mismatch (which can be mitigated by the use of crossconnected transistor quads, but, in fact, would more likely require the additional use of laser trimming), the errors due to finite beta (largely solved by the use of a slaved current source, as commonly used in precision d.a.c.s) and ohmic emitter resistance (which require operation at low currents, in common with other nonlinear circuits based on the exponential properties of the transistor).

Apart from these somewhat 2nd-order effects, a basic requirement for accurate operation is to maintain the correct value of offset over the temperature range. When N is large the error null is so deep that temperature-stable offsets can be used. In this case, the input E_b needed to achieve a certain angular range is temperature invariant; however, the amplitude of the output function will not be unless special techniques are used to introduce the correct compensation. These complications can be avoided by the use of offset voltages which are exactly proportional to absolute temperature (p.t.a.t.). The stability of angular sensitivity is then achieved by using a temperature-compensated resistive attenuator between the drive source (frequently scaled so that ± 10 V corresponds to $\pm 90^\circ$) and the E_b node, where only a few hundred millivolts are required.

An elegant implementation of the latter approach is shown in Fig. 2a, where unequal emitter areas are used to directly generate the p.t.a.t. offsets, using the well known relationship $\Delta V_{be} = kT/q(\ln \lambda)$, where λ is the area ratio. This circuit needs only two collector regions and four base regions; the integral value of $\lambda = 10$ was chosen for ease of layout and the supply currents were then reoptimised for minimum error. The function generated by this circuit is shown in Fig. 2b. The amplitude at $\pm 90^\circ$ is ± 2.35 V or approximately 36% of the asymptotic value of ± 6.6 V when using the same operating conditions as in the first case.

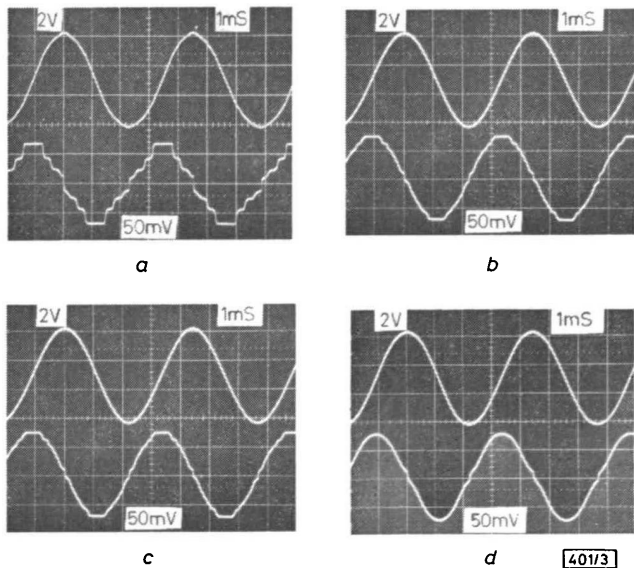


Fig. 3 Comparison of waveform fidelity for four practical sine-function generators; lower waveform is the derivative

- a A commercial monolithic i.c.
- b A laboratory-grade waveform generator
- c The circuit of Meyer *et al.*
- d The proposed circuit

Comparison with alternative techniques: It is illuminating to compare the waveforms generated by circuits based on the prevailing techniques with those generated by the proposed circuit. More revealing is the first derivative, approximated in these experiments by a single-pole highpass filter at 50 times the fundamental frequency. Fig. 3a shows the output of a popular i.c. waveform generator adjusted for minimum distortion; the derivative clearly shows the shortcomings of the piecewise-linear technique. A similar technique is often used in laboratory-grade instruments, of which the waveforms shown in Fig. 3b are typical. The method investigated by Meyer *et al.*⁶ has good accuracy near the zero crossings, but exhibits a step at $\pm 90^\circ$ due to the nonzero slope. Fig. 3c shows results for this type of circuit using the optimum bias conditions given in Reference 6, and a highly accurate triangular drive voltage. Finally, Fig. 3d shows the output of the simple circuit shown in Fig. 2a, from which it is apparent that the waveform and its derivative are essentially identical, evidence of accurate sinusoidal form. Spectral analysis of the waveform shows a total harmonic distortion of less than -75 dB.

Acknowledgment: The author is indebted to Prof. Sir Peter Swinnerton-Dyer for providing a rigorous proof of the connection between the tanh and sine functions, and to Dr. M. Jackson for suggesting improvements in the presentation of this proof, on which the Appendix is based.

Appendix: Numerical analyses suggest that

$$\sum_{n=-m}^m (-1)^n \tanh(x + n\alpha) \rightarrow \beta \sin \frac{\pi x}{\alpha} \quad (3)$$

for large values of m , where α is a constant offset ($= qE/2kT$ in the circuit), x is the input variable ($= qE_b/2kT$) and β is an amplitude-scaling constant. We seek to establish a theoretical basis for this relationship, and find an expression for the amplitude as a function of the offset. Consider the complex function $f(z)$ defined by

$$f(z) = \lim_{m \rightarrow \infty} \sum_{n=-m}^m (-1)^n \tanh(z + n\alpha) \quad (4)$$

for real $\alpha > 0$. Grouping pairs of terms it is apparent that $f(z)$ is analytic except at points of the form $Mj\pi/2 + n\alpha$ (M, n integral, M odd) and has period $j\pi$. Also,

$$f(z + \alpha) + f(z) = \lim_{m \rightarrow \infty} \{(-1)^m \tanh(z + (m+1)\alpha) + (-1)^{-m} \tanh(z + m\alpha)\} \quad (5)$$

and therefore $f(z)$ also has a period 2α . Hence it is a doubly periodic function with periods $j\pi$ and 2α . Moreover, $f(z)$ has just two incongruent poles: a simple pole at $j\pi/2$ with residue 1 and a simple pole at $j\pi/2 + \alpha$ with residue -1 . It has simple zeros at 0 and α .

Now let $\wp(z)$ be the standard Weierstrass function⁸ with periods $j\pi$ and 2α . Then the function

$$\frac{f(z) \wp'(z)}{\wp(z) - \wp(\alpha)} \quad (6)$$

has no poles or zeros, which is confirmed by checking at $z = 0$ and the three halfperiods α , $j\pi/2$ and $j\pi/2 - \alpha$. Hence it is constant with the value (at $z = j\pi/2$) of

$$\frac{\wp''(j\pi/2)}{\wp(j\pi/2) - \wp(\alpha)} \quad (7)$$

Thus we see that

$$f(z) = \frac{\wp''(j\pi/2)}{\wp(j\pi/2) - \wp(\alpha)} \frac{\wp(z) - \wp(\alpha)}{\wp'(z)} \quad (8)$$

It is more illuminating to write this in terms of theta functions. In the notation used in Reference 8, p. 337, we have

$$\omega_1 = \alpha_1, \quad \omega_2 = -(j\pi/2 + \alpha), \quad \omega_3 = j\pi/2$$

and therefore

$$\tau = \omega'/\omega = j\pi/2\alpha, \quad Q = \exp(-\pi^2/2\alpha) \quad (9)$$

Now

$$\wp''(j\pi/2) = 5e_3^2 - e_2^2 - e_1^2$$

and

$$\wp(j\pi/2) - \wp(\alpha) = e_3 - e_1 \quad (10)$$

and, from p. 473, 482 of Reference 8,

$$\frac{\wp(z) - \wp(\alpha)}{\wp'(z)} = \frac{-\alpha\theta_1(u)\theta_2(u)}{\pi\theta_2^2(0)\theta_3(u)\theta_4(u)} \quad (11)$$

where $u = z/2\alpha$. It follows (p. 476) that

$$f(z) = \beta\delta(z) \sin \frac{\pi z}{\alpha} \quad (12)$$

where

$$\beta = \frac{(5e_3^2 - e_2^2 - e_1^2) Q^{\frac{1}{2}}}{(e_1 - e_3)(e_2 - e_3)} \quad (13)$$

(a function of α only) and

$$\delta(z) = \prod_{n=1}^{\infty} \frac{(1 + Q^{4n})^2 - 4Q^{4n} \cos^2(\pi z/\alpha)}{(1 + Q^{4n-2})^2 - 4Q^{4n-2} \cos^2(\pi z/\alpha)} \quad (14)$$

which is a function of both α and z . It can further be shown that, if α is small, the amplitude term β may be approximated by

$$\beta = \frac{4\pi}{\alpha} \exp(-\pi^2/2\alpha) \quad (15)$$

and it is not difficult to show that the factor $\delta(z)$ converges uniformly on 1 for real z as $\alpha \rightarrow 0$, thus proving the original relationship. Eqn. 15 provides a reliable value of the sine-function amplitude for $\alpha < 2$, in practice. Above this, the errors in fitting the function become unacceptable anyway.⁷

BARRIE GILBERT*
Tektronix Inc.
Beaverton, Ore. 97077, USA

26th July 1977

References

- SMITH, G. W., and WOOD, R. C.: 'Principles of analog computation' (McGraw-Hill, 1959)
- SHEINGOLD, D. H.: 'Nonlinear circuits handbook' (Analog Devices Inc. 1974), chap. 2-1
- GILBERT, B.: 'New analog multiplier opens way to powerful function synthesis', *Microelectronics*, 1976, 8, (1)
- GILBERT, B.: 'Translinear circuits: a proposed classification', *Electron. Lett.*, 1975, 11, pp. 14-16
- GREBENE, A. B.: 'Monolithic waveform generation', *IEEE Spectrum*, April 1972, pp. 34-40
- MEYER, R. G., et al.: 'The differential pair as a triangle-sine wave converter', *IEEE J. Solid-State Circuits*, 1976, SC-11, (3)
- GILBERT, B.: 'Precision sine-function synthesis technique' (to be published)
- HANCOCK, H.: 'Lectures on the theory of elliptic functions' (Dover, 1958)

* Formerly with Analog Devices Semiconductor, Wilmington, Mass. USA

MODELLING THE 3rd-ORDER INTERMODULATION-DISTORTION PROPERTIES OF A GaAs F.E.T.

Indexing terms: Intermodulation, Schottky-barrier field-effect transistors

A simple analytical model of the 3rd-order intermodulation-distortion properties of a GaAs field-effect transistor is proposed. The model takes into account the distortion-producing nonlinearities of the Schottky-barrier junction, the transconductance and the drain conductance. Parameters of the model are presented for a GaAs f.e.t. of 0.5 μm gate length, and the model-predicted distortion characteristics are compared with measured data.

Introduction: This letter proposes a simple analytical model of the 3rd-order intermodulation-distortion characteristics of a GaAs microwave f.e.t. and describes these characteristics as a function of the device load admittance. The model should find application in the computer-aided design of low-distortion GaAs f.e.t. amplifiers. Parameters of the model are given for a GaAs f.e.t. of 0.5 μm gate length. The dominant device nonlinearities are identified and the model-predicted characteristics are compared with measured data.

Circuit model: The analytical model developed here is based on a simple unilateral circuit model of a GaAs f.e.t. This circuit model is shown in Fig. 1, together with the load admittance Y_L , the source admittance Y_S , and the 2-tone excitation, which comprises two equal-amplitude signals (i_{ω_1} and i_{ω_2}) at frequencies ω_1 and ω_2 , respectively. In addition to the familiar elements of a simple unilateral circuit model, two current generators (i_A and i_B) are included in Fig. 1. These represent the effects of nonlinearities in the f.e.t. Considered here are the three main nonlinear elements contributing to these effects: the Schottky-barrier junction capacitance at the gate, the gate-voltage-dependent transconductance and the output-voltage-dependent drain conductance. For the purposes of analysis, the contributions of the first two of these nonlinear components are lumped together and are represented by i_A . The output conductance nonlinearity is represented by i_B .

Justification for combining the first two nonlinearities is based on a series of experiments carried out by the authors on a number of different GaAs f.e.t.s. These experiments showed that, for fixed output power and load admittance Y_L at the drain, the 3rd-order intermodulation distortion is practically independent of the source admittance Y_S .

Analysis: The method of analysis used here is based on power-series descriptions of the device nonlinearities, and is similar to the approach used by Khadr and Johnston¹ for the analysis of distortion in f.e.t.s at frequencies around 10 MHz. Considering first the combined nonlinearity due to the Schottky-barrier junction and the transconductance, the current i_T (Fig. 1) as a function of the gate voltage v_i is described by a power series with order-dependent time delays:²

$$i_T(t) = g_{m1}v_i(t) + i_A(t)$$

where

$$i_A(t) = g_{m2}v_i^2(t - \tau_2) + g_{m3}v_i^3(t - \tau_3) + \dots \quad (1)$$

$g_{m1}, g_{m2}, \dots$ are real coefficients and $\tau_2, \tau_3, \dots$ are time delays. For low distortion levels this power series can be truncated after the cubic term in v_i . By substituting the 2-tone gate voltage signal

$$v_i(t) = A(\cos \omega_1 t + \cos \omega_2 t) \quad (2)$$

into eqn. 1 the component of $i_A(t)$ at each of two 3rd-order intermodulation-distortion frequencies $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ is obtained. Using phasor notation, this component can be expressed as

$$I_{A3} = \frac{3}{4}g_{m3}A^3 \angle \gamma$$