Compact Modeling of BJT Self-Heating in SPICE

David T. Zweidinger, Sang-Gug Lee, *Member, IEEE*, and Robert M. Fox, *Senior Member, IEEE*

Abstract—**Self-heating effects in bipolar junction transistors have been incorporated into PSpice dc and ac analyses. The effects are intrinsic to the operation of the transistor, and are treated within the device model, avoiding the need for thermal subcircuits. A physical thermal impedance model is provided, which allows prediction of thermal impedance for devices with rectangular emitters from device geometry. A simple approximation is used to predict thermal frequency response. The predictive model can be overridden by measured thermal model parameters. The modifications made to the PSpice code are presented, along with some discussion of implementation alternatives. An example simulation is presented demonstrating the significance of thermal effects in a typical circuit. Run-time comparisons show the modified code is about half the speed of unmodified PSpice, mostly because of slower convergence. It is likely that this performance can be improved with suggested implementation changes.**

I. Introduction

Many workers have pointed out that thermal feedback can significantly affect the operation of bipolar junction transistors (BJT's) [1]–[7]. For example, for a BJT with thermal spreading impedance of 200 K/W and Early Voltage $V_A = 150$ V, the output resistance drops below one-half that predicted by the Gummel–Poon model for all currents greater than 500 μA when the base is driven by a voltage source [4]. As such, it is somewhat surprising that self-heating effects are not routinely included in circuit simulation models. Several current trends make such thermal effects increasingly important for circuit simulation. One is the scaling of transistor geometries, which increases both current densities and thermal spreading impedances. Next is the increasing reliance on computer aids for circuit design, which, coupled with a trend toward concurrent engineering, leads to a need for more predictive simulation. The final trend is the use of dielectrics for device isolation; since SiO_2 is a poor thermal conductor, such isolation technologies greatly increase thermal spreading impedance.

Actually, there are several thermal feedback mechanisms in IC's, which can usually be considered separately [3], [8]. One mechanism is the rise in the overall chip temperature due to the total power dissipated on the chip. This temperature rise is controlled by the chip-to-package and package-to-ambient thermal impedances. This global heating operates over a long time-scale (milliseconds to minutes) and couples all of the devices on a chip. Global heating can generally be reduced by careful packaging and heat-sinking. It should be straightforward to include this effect in SPICE using an overall iteration loop on temperature; however, global heating is neglected in the present work.

For large-area devices, or those dissipating large power, there can be direct coupling between the heat dissipated in one device and the temperature of other devices. This mechanism is strongly affected by the circuit layout. Fukahori and Gray [9] added a finite-element heat-flow solution to a circuit simulator, which they used to compare several layouts of a 741-type op-amp circuit. They showed that thermal coupling from the output stage back to the input could profoundly affect the circuit's gain characteristics. They also demonstrated how careful layout exploiting symmetry can mitigate these effects.

In contrast to other thermal feedback mechanisms, direct heating of a transistor by its own power dissipation cannot be eliminated through scaling or layout. This mechanism can cause substantial errors in modeling even without high power dissipation. As pointed out in [10], for modest power dissipation, the temperature rise is confined within the transistor itself. Thus this effect can be simulated by allowing each transistor to be modeled at its own temperature, controlled by its thermal spreading impedance and its power dissipation. The thermal spreading impedance which controls the temperature rise can be predicted from the transistor's geometry.

This paper describes how such predictive modeling of local thermal heating can be implemented in the dc and ac analyses in PSpice [11]. The philosophy adopted was to enhance the accuracy of predictive BJT modeling by incorporating self-heating into the BJT model with minimal inconvenience to the software user and with minimal changes to the software code.

Prior circuit simulators with self-heating [5], [6] have used empirical models for thermal impedance, usually based on a single-pole RC thermal equivalent circuit. Such equivalent circuits underestimate the range of frequencies over which thermal feedback is significant. These models are further limited in that they are necessarily empirical; no way is provided to predict the thermal behavior without building and measuring a device. With recent trends toward concurrent engineering, designers commonly need

Manuscript received July 2, 1992; revised January 26, 1993. This work was supported by the Semiconductor Research Corporation under Contract 91-SP-087. This paper was recommended by Associate Editor D. Scharfetter.

D. Zweidinger and R. Fox are with the VLSI TCAD Group, Department of Electrical Engineering, University of Florida, Gainesville, FL 32611-2044.

S.-G. Lee was with the VLSI TCAD Group, University of Florida, and is now with Harris Corporation Semiconductor Sector, Melbourne, FL 32902-0883.

IEEE Log Number 9208605.

0278-0070/93$03.00 © 1993 IEEE

to design circuits using technologies which are not yet characterized. Furthermore, thermal measurements are rather difficult to perform accurately.

In the implementation presented here, a hierarchy of thermal-impedance models is provided [12], allowing the user to trade accuracy for complexity. One model level uses a simple predictive model for thermal impedance [7], whose accuracy has been demonstrated for a wide range of rectangular emitter geometries. The model is not compatible with subcircuit implementation, as it does not have an electrical equivalent circuit. The parameters of the predictive model can be partially or completely overridden by empirical parameters.

The thermal impedance model has been verified through single-transistor measurements [7], [12], [13]. These measurements include common-emitter dc output characteristics, output resistance, and reverse transconductance, plus the ac output impedance. These measurements were consistent with the new model and differed significantly from results of models ignoring self-heating.

Thermal modifications to transient analysis have not yet been implemented. It was shown in [14] that thermal effects caused only about a 4% error in the rise-time prediction for a typical digital circuit. In [13], it was shown that much more significant modeling errors due to self-heating are common in analog circuits. For this reason, the small-signal ac and operating-point analyses were the first to be implemented. Transient analysis using the predictive thermal impedance model will be difficult, as this model does not have an equivalent-circuit representation that could be easily implemented in PSpice.

This paper demonstrates how self-heating can be implemented in PSpice. Section II describes the needed changes to the modeling. Section III gives details of the changes to the algorithm. Section IV presents some example results and provides benchmark comparisons to the unmodified simulator. The limitations of the present implementation and suggestions for further research are presented in Section V and Section VI is the conclusion.

II. Thermal Modeling

A. Thermal Spreading Impedance Modeling

The thermal impedance model, derived in [7], is the frequency-domain equivalent of the time-domain derivation of [10], and is based on the same assumptions. All of the heat is assumed to be generated uniformly in the rectangular volume of the collector space–charge region (SCR). An adiabatic surface is assumed, which can be modeled using an image source placed above the surface. In the time domain, the response to a point impulse of heat, measured a distance r away in a uniform medium, is

$$T(r, t) = \frac{\kappa}{8K(\pi\kappa t)^{3/2}} \exp\left(-\frac{r^2}{4\kappa t}\right) \tag{1}$$

where K is the thermal conductivity and κ is the thermal diffusivity. The Laplace transform of (1) gives the thermal impedance

$$Z_{TH}(r, s) = \frac{1}{4\pi K r} \exp\left(-r\sqrt{\frac{s}{\kappa}}\right) \tag{2}$$

For $s = 0$, this equation gives the thermal resistance R_{TH} for the point source. An expression for R_{TH} at any point $r' = (x', y', z')$ in the semiconductor due to the heat generated in the collector SCR can be derived by integrating R_{TH} over the collector SCR and its image. For a rectangular vertical BJT the integral is

$$R_{TH}(W, L, D, H, r') = \frac{1}{4\pi K} \int_0^W \int_0^L \left[\int_D^{(D+H)} \frac{dx\,dy\,dz}{r} + \int_{-D}^{-(D+H)} \frac{dx\,dy\,dz}{r} \right] \tag{3}$$

where $r = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}$, W and L are the width and length of the emitter, D is the collector–base junction depth, and H is the SCR width. The result in (2) can be approximated by the following [7], [13]:

$$R_{TH} = \frac{1}{4\pi K \sqrt{WL} \cdot f_1 \cdot f_2} \tag{4}$$

where

$$f_1(d, h) = (0.058d + 0.14)h + 0.34d + 0.28 \tag{5}$$

$$f_2(a) = 0.98 + 0.043a - 6.9 \cdot 10^{-4}a^2 + 3.9 \cdot 10^{-6}a^3 \tag{6}$$

$$d = \frac{D}{\sqrt{WL}} \tag{7}$$

$$a = \frac{W}{L}, \quad a > 1 \tag{8}$$

$$h = \frac{\sqrt{\dfrac{2\epsilon(V_{CB} + V_0)}{qN_{EPI}}}}{\sqrt{WL}} \tag{9}$$

N_{EPI} is the epitaxial doping density in atoms/cm^3, V_{CB} is the collector-base junction voltage, and V_0 is the collector-junction built-in potential. This approximation has been found experimentally to estimate R_{TH} well for a wide variety of rectangular emitter geometries. Note by comparing (2) and (4) that the same R_{TH} could be computed by replacing the collector SCR with a point heat source placed a distance r_{eff} below the emitter, where r_{eff} is defined as

$$r_{eff} = \frac{1}{2\pi \cdot K \cdot R_{TH}} = 2\sqrt{W \cdot L} \cdot f_1 \cdot f_2. \tag{10}$$

This r_{eff} can be used in (2) to provide a simple estimate

for Z_{TH} for all frequencies as

$$Z_{TH} = R_{TH} \cdot \exp\left[-r_{eff}(1+j)\sqrt{\frac{\omega}{2\kappa}}\right] \tag{11}$$

Equations (10) and (11) can also be used to model Z_{TH} based on a measured value of R_{TH}.

Fig. 1 shows Z_{TH} for a typical device as computed both using the simple point-source model (11) and using the Fourier transform of the impulse response from [10]. Not surprisingly, the point-source representation overestimates the phase of Z_{TH}. However, the magnitude of Z_{TH} is predicted well, and shows that the errors in phase coincide with small $|Z_{TH}|$, so that the overall error is small. Note that $|Z_{TH}|$ varies over a much wider frequency range than would be predicted by a single-pole thermal equivalent circuit, in which $|Z_{TH}|$ would decay in one decade of frequency.

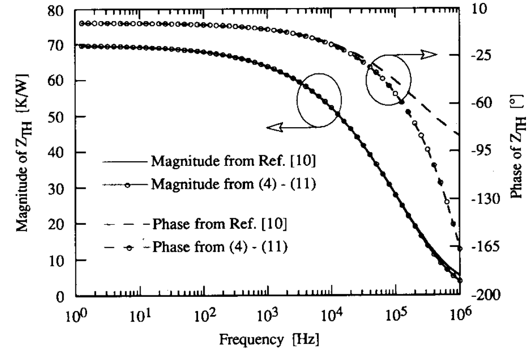


Fig. 1. Magnitude and phase of the thermal spreading impedance, Z_{TH}. Circled values were computed using (4)–(11); uncircled values were computed using the Fourier transform of the more detailed model in [10]. The device is assumed to have emitter width and length of 23 μm, base-junction depth of 3 μm, and collector space-charge region thickness of 2.7 μm.

B. Large-Signal Thermal Modeling

Modeling of thermal effects on large-signal behavior requires that each device be modeled at its own temperature. As implemented in PSpice [11], the Gummel–Poon equations for are I_C and I_B are

$$I_C = \text{area}\left(\frac{I_{be}}{K_{qb}} - \frac{I_{bc}}{K_{qb}} - \frac{I_{bc}}{\beta_R} - I_{bcn}\right) \tag{12}$$

$$I_B = \text{area}\left(\frac{I_{be}}{\beta_F} + I_{ben} + \frac{I_{bc}}{\beta_R} + I_{bcn}\right) \tag{13}$$

where I_{be} is the forward diffusion current, I_{ben} is the non-ideal base–emitter current, I_{bc} is the reverse diffusion current, I_{bcn} is the non-ideal base-collector current, β_F and β_R are the forward and reverse current gains, and K_{qb} is the base–charge factor defined by

$$K_{qb} = \frac{1 + \left[1 + 4 \cdot \left(\frac{I_{be}}{IKF} + \frac{I_{bc}}{IKR}\right)\right]^{NK}}{2 \cdot \left(1 - \frac{V_{bc}}{VAF} - \frac{V_{be}}{VAR}\right)} \tag{14}$$

where IKF, IKR, VAF, and VAR are the forward and reverse knee currents and Early voltages, and NK is a high-current slope factor (set to 0.5 in standard SPICE). Usually, self-heating is only significant in the forward-active region, in which I_{bc} and I_{bcn} can be neglected. The forward diffusion current I_{be} is given by

$$I_{be} = I_S \cdot \left[\exp\left(\frac{V_{BE}}{NF \cdot V_T}\right) - 1\right] \tag{15}$$

and the non-ideal base–emitter current I_{ben}, is given by

$$I_{ben} = I_{SE} \cdot \left[\exp\left(\frac{V_{BE}}{NE \cdot V_T}\right) - 1\right] \tag{16}$$

where NF and NE are emission coefficients for the ideal and non-ideal components and I_S and I_{SE} are transport and leakage saturation currents, respectively. When the temperature T is above the nominal temperature T_0, the temperature-dependent parameters are computed in PSpice as follows:

$$V_T = \frac{k \cdot T}{q} \tag{17}$$

$$\beta_F = \beta_{F0}\left(\frac{T}{T_0}\right)^{XTB} \tag{18}$$

$$I_S = \text{area} \cdot IS \cdot \exp\left[\frac{EG}{V_T}\left(\frac{T}{T_0} - 1\right)\right]\left(\frac{T}{T_0}\right)^{XTI} \tag{19}$$

$$I_{SE} = \text{area} \cdot ISE \cdot \exp\left[\frac{EG}{V_T \cdot NE}\left(\frac{T}{T_0} - 1\right)\right]\left(\frac{T}{T_0}\right)^{XTI/NE}\left(\frac{T}{T_0}\right)^{-XTB} \tag{20}$$

where EG is the band-gap, and XTB and XTI are temperature exponents for β_F and I_S.

In addition to the temperature-dependent currents, it is necessary to find the temperature derivatives of the currents for use in the Newton–Raphson iterations. The frac-

tional temperature coefficient of a current I is defined by

$$T_{CF}(I) = \frac{1}{I} \cdot \frac{dI}{dT} \tag{21}$$

To find closed-form expressions for the temperature coefficients of the collector and base currents in the forward-active region, this definition was applied to (12) and (13), with I_{bc} and I_{bcn} neglected. This gives for the fractional temperature coefficient of collector current

$$T_{CF}(I_C) = T_{CF}(I_{be}) \left[1 - \frac{I_C}{IKF} 2K_{q1} NK \cdot \left(1 + 4\frac{I_{be}}{IKF}\right)^{(NK-1)} \right] \tag{22}$$

where

$$T_{CF}(I_{be}) = \frac{1}{T}\left(XTI - \frac{V_{BE}}{NF \cdot V_T} + \frac{EG}{V_T}\right) \tag{23}$$

and K_{q1} is the inverse of the factor in parentheses in (14). The fractional temperature coefficient of the base current is

$$T_{CF}(I_B) = \frac{1}{T\beta_F I_B}\left[I_{be}\left(XTI - XTB + \frac{EG - V_{BE}/NF}{V_T}\right) + \frac{\beta_F I_{ben}}{NE}\left(XTI - NE \cdot XTB + \frac{EG - V_{BE}}{V_T}\right)\right] \tag{24}$$

Fig. 2 shows the variation of these temperature coefficients with V_{BE}. Over most of the range, $T_{CF}(I_{be})$, given by (23), dominates. Equation (23) is commonly used for $T_{CF}(I_C)$. Above the knee current (IKF), the temperature coefficient approaches $T_{CF}(I_{be})(1 - NK)$, which is typically about one-half of $T_{CF}(I_{be})$. The behavior of $T_{CF}(I_B)$ is also complicated; at low currents, I_{ben} dominates, and $T_{CF}(I_B)$ is controlled by the second term in (24). At higher V_{BE}, I_{be} dominates, and the first term in (24) controls $T_{CF}(I_B)$. Note that the temperature coefficient of the current gain β is just $T_{CF}(I_C) - T_{CF}(I_B)$, and is predicted to become negative at high V_{BE}. Many of the parameters that control these temperature coefficients are difficult to extract accurately. Furthermore, it is not certain that the variations of these parameters over a large range of temperatures are correctly modeled by (18)–(20). Further research in this area is needed.

C. Small-Signal Modeling

It is necessary to correct the admittance matrices in both the large- and small-signal analyses to reflect self-heating. The BJT common-emitter y-parameters can be corrected to include thermal feedback by computing the derivatives of the terminal currents with respect to the base and collector voltages while including temperature variations, yielding [2]

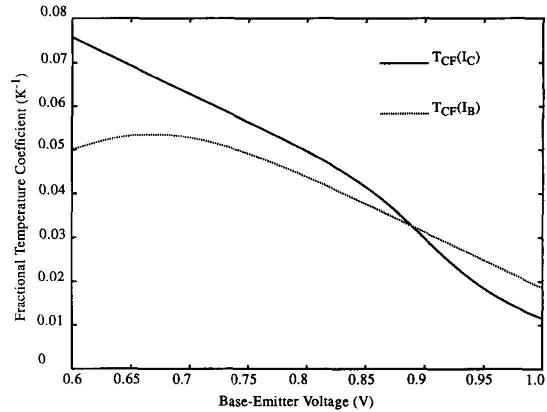


Fig. 2. Typical variation of the fractional temperature coefficients of collector and base currents with base-emitter voltage, as computed in PSpice. Key device parameters were $BF = 220$, $XTB = 1.7$, $XTI = 3$, $NF = 1$, $NE = 2$, $NK = 0.5$, $ISE = 6.9 \cdot 10^{-15}$ A, $EG = 1.11$ eV, $VAF = 150$ V, $VAR = 5$ V, $ISS = 12.3 \cdot 10^{-18}$ A, $IKF = 44$ mA, and $T = 300$ K.

$$y_{ij} = \frac{y_{ijE} + T_{CF}(I_i) Z_{TH} I_i I_j}{1 - T_{CF}(I_i) Z_{TH} P} \tag{25}$$

where i and j are 1 for the base and 2 for the collector, respectively, y_{ijE} is the uncorrected electrical y-parameter, $T_{CF}(I_i)$ is the fractional temperature coefficient of the base or collector current, Z_{TH} is the thermal spreading impedance, and P is the dissipated power in the device. In the dc analysis, the y-parameters are replaced by their dc equivalents, the g-parameters, and Z_{TH} is replaced by R_{TH}. It is easy to show [13] that usually y_{11} and y_{21} are substantially affected by self-heating only if the denominator is large, which requires substantial power dissipation. On the other hand, because of the typically small values of g_{12E} and g_{22E}, y_{12} and y_{22} can be strongly affected by self-heating even with low power.

III. PSpice Implementation

The thermal models described above were implemented in the PSpice simulator, using the Device Equations options, which allows convenient access to well-documented source code for several key model-evaluation routines. While this approach was successful, it did lead to some limitations. For example, the Device Equations option does not allow access to the convergence and time-step control algorithms, nor is it possible to change the circuit topology (no new nodes can be added). Thus the temperature was not used as a state variable, but was computed as needed from the power dissipation. This reduces the storage requirements for the simulator but may have degraded convergence.

A. DC Analysis

To include thermal effects, the PSpice dc analysis subroutine required modification in four ways. First, when

the predictive model is in use, the V_{BC}-dependent value of the normalized SCR width h is computed from (9) and used in (4). Second, the temperature dependences in (17)–(20) were incorporated into the current computations for each device. Third, the conductances were modified to include thermal effects using (25) before the conductance matrix is updated. Finally, temperature limiting was added to improve convergence.

To implement the three-level thermal-impedance model, the following device parameters were added to the model specification: R_{TH}, C_{TH}, N_{EPI}, W_{TH}, L_{TH}, and D_{TH}. These parameters are defaulted to zero, and non-zero values of the parameters determine which thermal model is used. If N_{EPI} is specified without R_{TH}, R_{TH} is computed from the predictive model in (4)–(9) and the thermal-model flag is set. If R_{TH} is specified, the thermal-model flag is also set and Z_{TH} is computed using (10) and (11) unless C_{TH} is also specified, in which case the specified values are used in $Z_{TH} = R_{TH} + j\, C_{TH}$.

Fig. 3 shows a flowchart of the modified dc analysis routine. The thermal model flag discussed above is used throughout the routine to steer around the code for the thermal effects when they are not desired.

On the first pass, all the devices are initialized in the forward-active region. In subsequent iterations, the terminal currents and conductances from the previous iteration are retrieved, along with the voltages resulting from that iteration. In the unmodified PSpice routine, the new currents are linearly extrapolated from the old currents and conductances and the new voltages, and if the extrapolated currents agree with the old ones, the convergence flag is set and the routine terminates. This can cause the program to converge without updating the temperature. During dc sweeps with closely spaced points, this premature convergence can occur for several successive points, leading to the accumulation of substantial error in the device temperature. The potential problem was avoided by skipping the linear-extrapolation convergence check in the first iteration for each point in a dc sweep and for all iterations in Q-point calculations. This forces the temperature to be updated.

If the terminal currents have not converged, the local temperature is then estimated using $T = P \cdot R_{TH} + T_0$, where T_0 is the ambient temperature and P is the power, computed from $P = I_C \cdot V_{CE} + I_B \cdot V_{BE}$. Because the first few iterations can compute dc values of current and voltage much greater than the final operating-point, the dissipated power and local temperature computations can yield unrealistically high values. Unless the temperature is limited, PSpice can diverge due to thermal runaway in these cases. Several temperature-limiting approaches were tried; none gave foolproof convergence for all circuits. The most stable temperature-limiting algorithm clips the device temperature at a fixed amount (typically 50°C) above the ambient temperature. Using the new device temperature (limited or not), all temperature-dependent quantities are updated using the standard PSpice temperature dependences, (17)–(20). These temperature-modified parameters are used to compute the new currents and conductances, using (12)–(16). This leads to new estimates of the power dissipation and local temperature. The temperature coefficients are then computed from (22) to (24) and saved for use in the ac analysis.

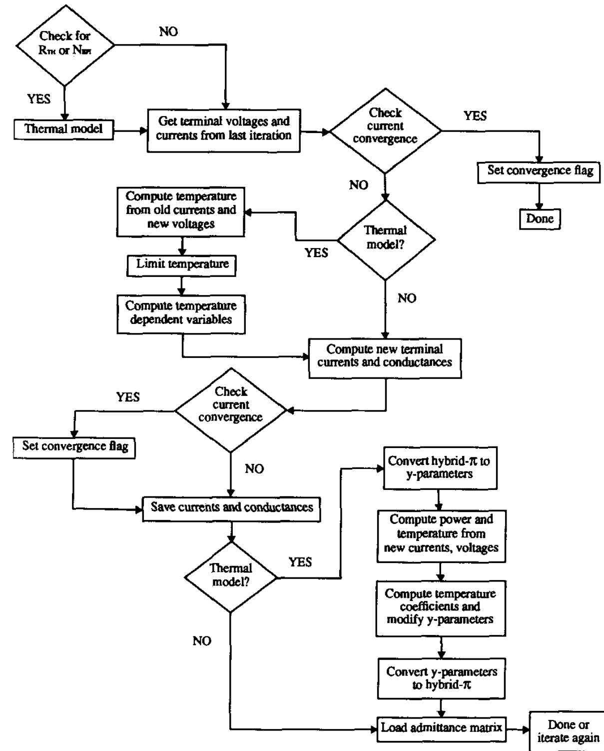


Fig. 3. Flowchart of PSpice dc analysis routine, modified to model thermal effects.

The next step is to convert the dc hybrid-π parameters to the two-port g-parameters using the following conversion equations:

$$g_{11E} = g_\pi + g_\mu \quad (26)$$

$$g_{12E} = -g_\mu \quad (27)$$

$$g_{21E} = g_m - g_\mu \quad (28)$$

$$g_{22E} = g_o + g_\mu \quad (29)$$

where the conductances g_π, g_μ, g_m, and g_o are the conventional hybrid-π parameters [15]. These equations apply to the intrinsic part of the dc model, internal to all parasitic resistances.

The dc equivalent of (25) is then used to compute the thermally modified two-port g-parameters, which are then converted back into hybrid-π form and loaded into the conductance matrix. Although the g-parameters could be loaded directly into the matrix without converting back to hybrid-π parameters, this would require conversion to g-parameters even when a thermal model is not in use.

One of the goals in this work was to minimize the computational overhead when thermal modeling is not required.

B. AC Analysis

The modifications to the ac analysis routine are similar to the conductance-matrix corrections in the dc analysis. Depending on the model level specified, Z_{TH} is computed from either the supplied R_{TH} and C_{TH} or from the R_{TH} computed in the dc analysis routine and (10), (11). In the ac analysis the hybrid-π parameters are converted into y-parameters using

$$y_{11E} = g_\pi + g_\mu + j\omega(C_{xcb} + C_{be} + C_{bc}) \quad (30)$$

$$y_{12E} = -g_\mu + j\omega(C_{xcb} + C_{bc}) \quad (31)$$

$$y_{21E} = g_m - g_\mu + j(xg_m - \omega C_{bc}) \quad (32)$$

$$y_{22E} = g_o + g_\mu + j\omega \cdot C_{bc} \quad (33)$$

where C_{be} and C_{bc} are the base-emitter and base-collector capacitances respectively, xg_m is the excess phase term, C_{xcb} is a base-emitter transcapacitor controlled by the base-collector voltage, and the other terms are standard hybrid-π parameters.

The electrical y-parameters are then modified using (25), and using the temperature coefficients from the dc operating-point analysis. The thermally modified y-parameters are converted back to hybrid-π parameters before they are loaded back into the device admittance matrix.

V. Results

A. Circuit Examples

To demonstrate the thermally modified simulator, a 741 operational amplifier circuit [15] was analyzed. For simplicity, the overload-protection circuits were removed. Compared with some other circuits, the 741 is only modestly affected by self-heating; the circuit was chosen because it provides a familiar benchmark. Some of the transistor model parameters used in the simulation are given in Table I. The n-p-n model was extracted from a 4-μm by 10-μm vertical n-p-n. Since a similar p-n-p device was not available for parameter extraction and the predictive Z_{TH} model does not apply to lateral transistors, a reasonable set of p-n-p device parameters was assumed.

Fig. 4 shows plots of dc sweeps through the active region of the amplifier using the normal and the modified versions of PSpice. The gain error is due primarily to a drop in the output resistances of the gain and load transistors caused by the self-heating. Thermal feedback also caused the increased nonlinearity.

Fig. 5 shows magnitude and phase plots of the output impedance for the second gain stage of the 741 (with the compensation capacitor removed) simulated with and without self-heating. Note the phase anomalies near 1 MHz. Since these errors occur near the unity-gain fre-

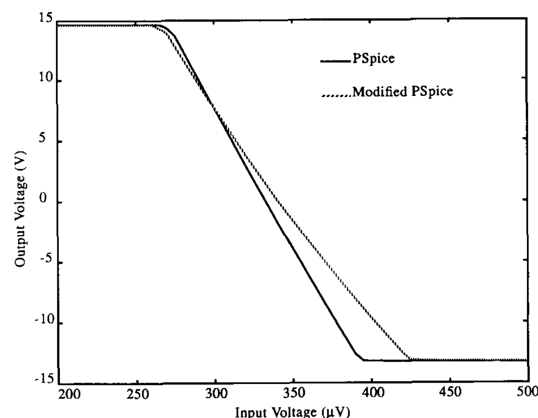
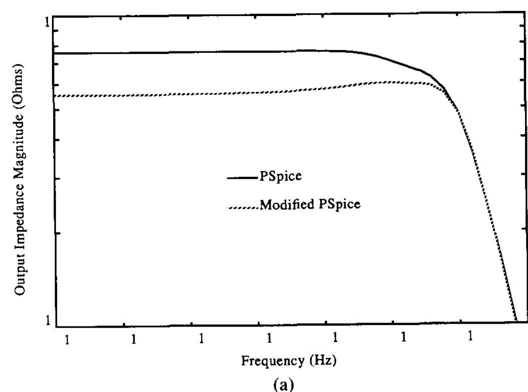
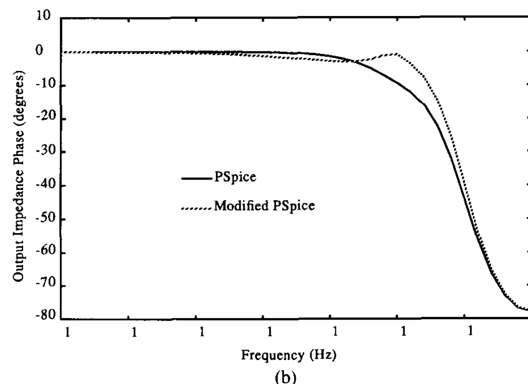


Fig. 4. Input/output curves for a 741 op amp, excluding overload-protection circuits. The dc sweeps were simulated with unmodified PSpice and with the modified version including self-heating.



(a)



(b)

Fig. 5. (a) Magnitude and (b) phase of the output impedance of the second stage of an uncompensated 741 op amp.

TABLE I
Transistor Model Parameters Used For 741 Op Amp Simulation

Parameter	NPN Value	PNP Value	Units
R_{TH}	290	100	K/W
VAF	150	60	V
BF	227.4	80	
IS	28.9E-18	28.9E-18	A
TF	25E-12	150E-12	s

TABLE II
BENCHMARKS FOR A 741 OP AMP Q-POINT, DC, AND AC ANALYSES. ANALYSES WERE RUN ON AN 8-MHZ PC AT

Analysis	Simulator	Number of Iterations	Iteration Increase (%)	Time (s)	Time/Iteration (s)	Time Increase (%)
Q-point	Normal	17	65	8.62	0.51	120
	Modified	28		18.9	0.68	
DC	Normal	362	16	99.64	0.28	97
	Modified	420		196.4	0.47	
AC	Normal	41 Points		26.53	0.64	47
	Modified			39.11	0.95	

quency of the amplifier, they could cause errors in the compensation of such circuits.

B. Benchmarks

Adding thermal modeling increases the simulation time as shown in Table II. Some of the increase in the time per iteration is attributable to the limited access to the data structures in PSpice. New global data structures had to be created to allow sharing new device variables (R_{TH}, temperature, etc.) among routines; accessing these added to the simulation time. The added computations for the predictive Z_{TH} model, the temperature dependences, and the admittance modifications account for the rest of the increase in the time per iteration. The average time per iteration in dc analysis is less than that in Q-point analysis because of use of fast linear extrapolation in the dc sweep after the first iteration for each point. Most of the increase in computation time results from increases in the number of iterations needed to find a solution. The rate of convergence could probably have been improved if access had been available to the convergence-control algorithms.

V. DISCUSSION

The approach adopted in this work was to make the minimal amount of change to "standard" PSpice as possible, in the hope of making the modified software accessible to users. As such, only the most significant mechanisms, those which commonly have significant effects, were included in the modeling. For example, the temperature derivatives given in (22)–(24) are only valid in the forward-active region, where transistors are most sensitive to self-heating. Expressions for the derivatives valid in all regions would be very complex. Also, the only parameters whose temperature variations are updated in the device model are those given in (17)–(20). It is possible to model the temperature variations of many other parameters, such as parasitic resistances and capacitances, transit times, etc., although the temperature dependences of these parameters are relatively weak. Extending the modeling to cover all regions and all parameters would make the simulations more self-consistent and more accurate. However, for most applications the added accuracy is probably not worth the increased complexity and longer run-times.

Another simplification in this implementation is its neglect of inter-device thermal coupling. Such coupling is small for a large class of circuits, but it can be caused by chip-to-package-to-ambient thermal impedance or, with sufficient power dissipation, by direct thermal interaction between devices. Including this coupling would be a useful extension of the present work. In that case, thermal-equivalent subcircuits would probably provide the most convenient representation. The cost of such modeling in terms of user inconvenience and run time is likely to be substantial.

Use of the PSpice Device Models option required a number of compromises. The inability to modify any of the data structures led to substantial run-time overhead and precluded output of many of the temperature-modified parameters in the operating-point report. This option also provided only limited access to the convergence-control routines, which could otherwise have been optimized to handle self-heating (for example, by allowing thermal resistance to be turned on gradually).

The fixed data structures also precluded treating temperature as a state variable. Treating temperature as an implicit variable avoids increasing the dimensionality of the matrices and thus reduces storage requirements and (presumably) the computation time per iteration. However, it is possible that the rate or robustness of the convergence would be improved if temperature were used as a state variable.

Several areas deserve future study. The most obvious is the addition of thermal effects into transient analysis. Modeling self-heating in the time domain presents a significant problem: the impulse response typically is significant over more than four decades of time. Thus the thermal response cannot be modeled accurately using a small number of time constants, as with RC thermal equivalent circuits. Equation (1), which is the time-domain equivalent of the point-source thermal frequency response used in this work, models such behavior simply and fairly accurately. Unfortunately, direct use of this expression would require evaluation of a convolution integral, which is inefficient for this application. What is needed is a simple way to approximate the convolution integral by storing information over many time-scales. This will require modification of the PSpice time-step-control algorithms.

Another area where more work is needed is thermal

modeling of non-rectangular emitter geometries (especially multi-stripe emitters) and of trench-isolated and silicon-on-insulator (SOI) transistor structures. The thermal conductivity of SiO_2 is about one percent that of silicon. Because of the resulting high thermal impedances, even FET's in SOI technologies can have substantial self-heating effects, which will eventually need to be modeled for circuit simulation.

VI. Conclusion

The work described above demonstrates the feasibility and utility of incorporating self-heating effects in ac and dc circuit simulation using physically based modeling. Self-heating can be treated as intrinsic to the BJT device operation, avoiding the need for a thermal subcircuit for the transistor. While the increased accuracy is significant for many applications, it comes at some cost in terms of run-time and robustness of convergence. Further study of implementation algorithms may allow improvement of these characteristics.

References

[1] J. J. Sparks, "Voltage feedback and thermal resistance in junction transistors," *Proc. IEEE*, vol. 46, pp. 1305–1306, June 1958.

[2] O. Mueller, "Internal thermal feedback in four-poles especially in transistors," *Proc. IEEE*, vol. 52, pp. 924–930, Aug. 1964.

[3] G. Meijer, "The current dependency of the output conductance of voltage-driven bipolar transistors," *IEEE J. Sol. St. Circuits*, vol. SC-12, pp. 428–429, Aug. 1977.

[4] W. F. Davis and M. L. Lidke, "The effect of thermal feedback within the bipolar transistor on Gummel-Poon model accuracy," Motorola Internal Rep. May 1988.

[5] R. S. Vogelsong and C. Brzezinski, "Extending spice for electrothermal simulation," in *Proc. IEEE Custom Integrated Circuits Conf.*, pp. 21.4.1–21.4.4, May 1989.

[6] P. C. Munro and F.-Q. Ye, "Simulating the current mirror with a self-heating BJT model," *IEEE J. Solid-State Circuits*, vol. 26, pp. 1321–1324, Sept. 1991.

[7] R. M. Fox and S.-G. Lee, "Scalable small-signal model for BJT self-heating," *IEEE Electron Dev. Lett.*, vol. 12, pp. 649–651, Dec. 1991.

[8] R. M. Fox and S.-G. Lee, "Thermal parameter extraction for bipolar circuit modelling," *Electron. Lett.*, vol. 27, pp. 1719–1720, Sept. 1991.

[9] K. Fukahori and P. R. Gray, "Computer simulation of integrated circuits in the presence of electrothermal interaction," *IEEE J. Solid-State Circuits*, vol. SC-11, pp. 834–846, Dec. 1976.

[10] R. C. Joy and E. S. Schlig, "Thermal properties of very fast transistors," *IEEE Trans. Electron. Devices*, vol. ED-17, pp. 586–594, Aug. 1970.

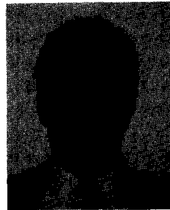
[11] *PSpice Users Manual, version 4.05*, Irvine, CA: MicroSim Corp., January 1991.

[12] D. T. Zweidinger, R. M. Fox, and S.-G. Lee, "Predictive modeling of thermal effects in BJTs," in *Proc. NUPAD IV*, pp. 219–224, June 1992.

[13] R. M. Fox and S.-G. Lee, "Predictive modeling of thermal effects in BJTs," in *Proc. IEEE Bipolar Circuits and Technology Meeting*, pp. 89-92, Sept. 1991.

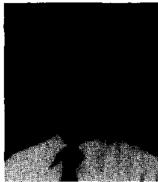
[14] R. T. Dennison and K. M. Walter, "Local thermal effects in high performance bipolar circuits/devices," in *Proc. IEEE Bipolar Circuits and Technology Meeting*, pp. 164–167, 1989.

[15] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*. 2nd Ed. New York: Wiley, 1984.



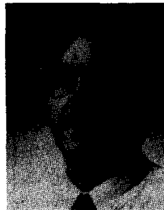
David T. Zweidinger received the B.S.E.E. degree in 1990 and is currently working toward the M.S.E.E. and Ph.D. degrees at the University of Florida, Gainesville, Florida.

His interests include electronic circuits, semiconductor device modeling, and computer-aided design.



Sang-Gug Lee received the B.S. degree in electronic engineering from KyungPook National University, TaeGu, Korea, in 1981, and the M.E. and Ph.D. degrees in electrical engineering, both from the University of Florida, Gainesville, in 1989 and 1992, respectively.

From 1983 to 1986, he was an Instructor in the Department of Electronic Engineering at the Naval Academy, JinHae, Korea. Since 1992, he has been with Harris Semiconductor Company, Melbourne, FL, working on analog IC design. His current research interests are focused on device modeling, characterization, and RF circuit design.



Robert M. Fox received the B.S. degree in physics from the University of Notre Dame, Notre Dame, IN, in 1972 and the M.S. and Ph.D. degrees in electrical engineering from Auburn University, Auburn, AL, in 1981 and 1986, respectively.

He has been on the Electrical Engineering faculty of the University of Florida since 1986. His current rank is Associate Professor. His research interests center on circuit design for advanced silicon technologies.

Dr. Fox is a member of the Audio Engineering Society, ASEE, Phi Kappa Phi, and Eta Kappa Nu.