from *SIAM News*, Volume 31, Number 6

The Design and Production of Microcircuits: Part II

By James Case

Whereas the circuit designers of the future will be obliged to concern themselves with all manner of quantum and classical field effects, as described in Part I of this article (*SIAM News*, May 1998, page 16), those of the relatively recent past had (as many working today still have) only to consider events taking place on directed graphs specified by incidence matrices of the following form:

$$\begin{array}{c} \\ e1 \\ e2 \\ e3 \\ e4 \\ e5 \\ e6 \end{array} \begin{array}{c} \begin{array}{cccc} n1 & n2 & n3 & n4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & -1 & 0 \end{bmatrix} \end{array} = A_0. \tag{1}$$

The rows and columns of A_0 correspond to the edges and nodes of the graph, respectively. Such matrices are always sparse, since each row contains only two nonzero entries, one positive and the other negative. A positive entry in the *ij*th position and a negative one in the *ik*th indicate that edge *i* runs *from* node *j to* node *k*. If the graph is composed of N nodes and B edges, or "branches," then A_0 will be of dimension $B \times N$. The directions of the edges can be specified arbitrarily, but specified they must be.

Node Voltages: Fictitious But Conceptually Useful

It is customary to choose a "reference" node, at which the voltage is required to vanish identically, and to define the voltage at any non-reference node to be the sum of the "voltage drops" across a sequence of edges leading from the reference to the non-reference node. Kirchoff's voltage law (KVL), according to which the voltage drops around any closed loop sum to zero, ensures that the proposed definition is path-independent and therefore unambiguous. If v_0 denotes the N-vector of node voltages, and if Δv_0 is the B-vector of voltage drops across edges, then $A_0 v_0 = \Delta v_0$ by construction.

The fact that the entries in each row of A_0 sum to zero indicates that one of the rows is redundant and can be deleted. We can therefore denote by A the $B \times (N-1)$ matrix obtained from A_0 by deleting the rightmost column, and by v and Δv the vectors obtained by eliminating the final entry in v_0 and by setting the voltage at the reference node equal to zero in Δv_0. Then $Av = \Delta v$. Moreover, if i is a B-vector of "branch currents," then the components of the product $A^T i$ represent the sums of the currents incident on the several nodes of the graph, which must vanish according to Kirchoff's current law (KCL). The equations $A^T i = 0$ and $Av = \Delta v$ therefore represent $B + N - 1$ (scalar) relations among the $2B + N - 1$ unknown components of the vectors v, Δv, and i. Other B independent relations are required to determine uniquely the entire slate of unknowns.

At this point, we have two unrelated sets of equations. The first expresses the conservation of whatever happens to be flowing across the edges of the graph, whether it be water, electrical current, or taxicabs, while the second expresses the recoverability of the (observable) voltage drops across edges from the purely fictitious—but less numerous and conceptually rather useful—node voltages. Still missing are the laws of electricity that relate the voltage drop across an individual edge of the graph to the current flowing through it. The simplest example of such a law is the familiar Ohm's law, which states that the two are directly proportional to one another and applies to ordinary wires operating in the "linear range."

For all its utility, Ohm's law is only an approximate formulation. An abundance of considerations, including the fact that the resistance in a wire is temperature-dependent, complicate the actual relationship. In practice, however, even highly nonlinear devices do exhibit ranges of linearity, and care is frequently exercised not to exceed the limits of those ranges. When the entire system is operating within its linear range, the "branch constitutive relations" (BCRs) can be written in the form $Zi + Y\Delta v = w$, for appropriate matrices Y, Z, and w. We can now write the complete system of equations,

$$\begin{aligned} A^T i &= 0 \\ Av - \Delta v &= 0 \\ Y\Delta v + Zi &= w, \end{aligned} \tag{2}$$

and solve for the $2B + N - 1$ unknown components of the vectors v, Δv, and i. Alternatively, if Av is substituted for Δv in

the BCRs, the system can be rewritten in the form

$$Ai = 0$$
$$YAv + Zi = w, \tag{3}$$

and solved for the $B + N - 1$ unknown components of the vectors i and v, from which the B unknown components of Av are easily recovered. And in certain rare circumstances, it is possible to eliminate i from the system, leaving the $N - 1$ nodal voltages v as the only unknowns in the twice-reduced system

$$A^T YAv = A^T w. \tag{4}$$

Additional Unknowns

This represents a substantial size reduction, since B typically lies well above the lower limit of its potential range $N \le B \le N(N-1)/2$ of constructibility. The latter efficiency is available, however, only in an extremely limited class of problems. Indeed, because both A and Y are typically sparse, while the product YA need not be, even the first reduction may prove counterproductive. In general, due to the presence of nonlinear and dynamical effects, the unknown quantities are functions of time related by functional equations of the form

$$ß(i, v, \ldots, t) = 0. \tag{5}$$

The dots in (5) represent the additional unknowns introduced into the BCRs by circuit elements other than wires. A capacitor, for instance, is a familiar circuit element across which the behavior of the current (i_C) and voltage drop (v_C) is related through an intermediate quantity (q_C) representing charge on the capacitor. The most general form of the capacitor law is $g(q_C, v_C) = 0; i_C = \dot{q}_C$. Inductors are equally familiar circuit elements across which the currents i_L and voltage drops v_L are related through intermediate quantities ϕ_L known as "fluxes." Generally speaking, inductors are governed by equations of the form $h(\phi_L, i_L) = 0; v_L = \dot{\phi}_L$. Hence, the missing quantities in (5) will typically include a smattering of q's and ϕ's, along with their time derivatives. Circuits formed by capacitors, inductors, and resistances, known as LRC circuits, long occupied a prominent position in physics and engineering curricula.

Today, the fastest growing segment of the semiconductor market seems to be for field programmable gate arrays (FPGAs), microchips that contain as many as 25,000 programmable *and reprogrammable* "logic gates." Some in the industry estimate that, by the year 2000, as many as 25% of all chips sold will be of the field programmable variety. An engineer, seated at a PC or workstation and armed with an appropriate CAD program, can display the complete circuitry of such a chip on the screen, with individual layers represented by characteristic colors, as if drawn with crayons on transparent overlays. He or she can then zoom in on any particular part of the circuit, and open or close individual gates merely by clicking on them. When finished, the newly configured design can be exported to a filing cabinet-sized peripheral, not unlike an overgrown laser printer, where the chip is physically altered. The opening or closing of a single gate causes a row of A to appear or disappear. The prevalence of such chips will oblige future circuit analysis packages to exhibit ever increasing flexibility.

Although integrals and time lags have been known to occur, the functional equations (5) governing the behavior of most circuits of interest can be brought into the form

$$f(\dot{x}, x, t) = 0. \tag{6}$$

Often, they can be further simplified to

$$\dot{q}(x) + j(x) - e(t) = 0, \tag{7}$$

the form most natural for simulation programs. In practice, the first term can usually be expanded, via the chain rule, to give

$$C(x)\dot{x} + j(x) - e(t) = 0, \tag{8}$$

where $C(x) = \partial q/\partial x$. This is the form accepted as inputs by many of the more popular of the commercially distributed circuit simulation programs.

Direct and Alternating Current Analyses

On occasion, it is possible to approximate $C(x)$ by a constant matrix C. Even then, however, because C is typically of low rank, (8) cannot be transformed into an explicit ODE. Circuit analysis typically requires the solution of differential–algebraic equations (DAEs), which turn out to be harder to solve than either purely differential or purely algebraic equations. More about that later on.

When a circuit contains oscillators, even a constant excitation $e(t) \equiv e_0$ need not drive the system to a steady state $x(t) \equiv x_0$. On other occasions, however, it is intuitively obvious that one or more such x_0's will exist. The process of finding them, and determining their stability, is known for historical reasons as direct current (DC) analysis. Some variant of Newton's method is typically used to locate potential steady states, and all manner of eigenvalue techniques have been devised for the determination of stability. Because of the size of the problems encountered, even the computation of the Jacobi matrices

$C = C(x_0)$ can be nontrivial. Here, however, the structure of C is to some extent favorable, as summarized by the following schematic:

	Node voltages	Branch voltages	Branch currents	Auxiliary variables
KCLs	O	O	X	O
KVLs	X	X	O	O
BCRs	O	X	X	X

(9)

Here the X's and O's indicate functional dependence and independence, respectively. The display is a reminder that all $B + N - 1$ of the KCLs and KVLs, and usually a few of the BCRs, are both linear and algebraic. Only a BCR can be nonlinear. If the problem formulation includes β auxiliary variables, there will be at most $B + \beta$ nonlinear equations. The first $B + N - 1$ rows of $C(x)$ are independent of x, and need not be recalculated at each successive step of Newton's method, or upon arrival at x_0. Numerical differentiation was used frequently in the past, but symbolic techniques are gaining popularity.

Once a stable steady state x_0, corresponding to a constant excitation e_0, has been found, it is often instructive to augment that excitation with a small periodic perturbation. It usually suffices, when doing so, to consider the linearized system:

$$C\dot{x}(t) + Gx(t) = e_0 + \varepsilon \Re\left[Ke^{i\omega t}\right]. \tag{10}$$

Here $C = C(x_0) = (\partial q/\partial x)(x_0)$, $G = G(x_0) = (\partial j/\partial x)(x_0)$, and K denotes a complex column vector of appropriate length. As usual, the linear operator $\Re$ selects the real part of any complex quantity, and ε denotes a small positive real number.

The resulting process is known as alternating current (AC) analysis, or complex phasor analysis (the name by which the argument of $\Re$ in (10) is commonly known). It is here that linear transform methods, particularly Laplace transform methods, prove their worth. Any complete account of circuit analysis techniques must deal with them at length. Such an account was presented by Roland Freund and Peter Feldmann as a short course at the 1995 SIAM Annual Meeting, and the present exposition borrows heavily from the course notes. Regrettably, those notes remain unpublished.

The Difficulties of Solving DAEs

To determine the range of validity of the results of complex phasor analysis, it often becomes necessary to deal directly with the given system (7). This is straightforward enough when $C(x) = \partial q/\partial x$ is nonsingular, since the system then reduces to a collection of ODEs. These equations must then be solved numerically and the results compared with those obtained by linearization. The full power of numerical integration techniques (particularly the linear multistep techniques) is routinely brought to bear. Stiff systems are frequently encountered, with all the difficulties they entail. Most often, however, $C(x)$ is of less than full rank, meaning that the equations to be solved are DAEs. It isn't hard to see why DAEs are more difficult to solve than either purely algebraic or purely differential equations.

Consider, to that end, the following linear initial value problem (LIVP) with constant coefficients:

$$\begin{aligned} A\dot{x} &= Bx + b(t) \\ x(t_0) &= x_0. \end{aligned} \tag{11}$$

If A is nonsingular, the problem concerns ODEs and is of no further interest. But if A is singular, new difficulties emerge. Most are present even if B is nonsingular.

In that fortuitous event, the $n \times n$ matrix product $B^{-1}A$ can be factored as follows:

$$B^{-1}A = P^{-1}\begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix}P, \tag{12}$$

where C is square and nonsingular, while N is square and nilpotent, meaning that $N^k = 0$ for some positive integer k. The least such k is called the order of N's nilpotence. Next, let

$$\begin{aligned} Px(t) &= \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}, \; Px_0 = \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}, \\ \text{and } PB^{-1}b(t) &= \begin{bmatrix} p(t) \\ q(t) \end{bmatrix}, \end{aligned} \tag{13}$$

where $u(t)$, u_0, and $p(t)$ are vector-valued functions of $0 < m < n$ components, while $v(t)$, v_0, and $q(t)$ furnish the remaining $n - m$ components. In terms of these new variables, the LIVP (11) can be written in the following decoupled form:

$$\begin{aligned} C\dot{u} &= u + p(t), \quad u(t_0) = u_0, \\ N\dot{v} &= v + q(t), \quad v(t_0) = v_0. \end{aligned} \tag{14}$$

The unknown vector function $u(t)$ can now be obtained by the usual methods, since C is invertible, but $v(t)$ requires special treatment. Writing $v(t) = N\dot{v}(t) - q(t)$, differentiating j times with respect to t, and denoting by $f^{(j)}(t)$ the jth time derivative of any $f(t)$, yields

$$v^{(j)}(t) = Nv^{(j+1)}(t) - q^{(j)}(t), \tag{15}$$

for $j = 1, 2, \ldots, k-1$. So, by successive substitution,

$$\begin{aligned} v(t) &= Nv^{1}(t) - q(t) \\ &= N^2 v^{(2)}(t) - Nq^{(1)}(t) - q(t) \\ &= N^3 v^{(3)}(t) - N^2 q^{(2)}(t) - Nq^{(1)}(t) - q(t) \\ &\vdots \\ &= N^k v^{(k)}(t) - N^{k-1} q^{(k-1)}(t) - \\ &\qquad N^{k-2} q^{(k-2)}(t) - \ldots - Nq^{(1)}(t) - q(t), \end{aligned} \tag{16}$$

which, in view of the fact that $N^k = 0$, expresses $v(t)$ as a linear combination of the given $q(t)$ and its first few time derivatives. From (14) it follows that

$$v(t_0) + N^{k-1} q^{(k-1)}(t_0) + N^{k-2} q^{(k-2)}(t_0) + \ldots + Nq^{(1)}(t_0) + q(t_0) = 0, \tag{17}$$

a compatibility condition between $v_0 = v(t_0)$ and $q_0 = q(t_0)$, $q_1 = q^{(1)}(t_0), \ldots, q_{k-1} = q^{(k-1)}(t_0)$.

Equation (17) illustrates two separate difficulties characteristic of DAEs: The forcing term $q(t)$ must possess the required number of derivatives at $t = t_0$, and $v_0 = v(t_0)$ must be as stipulated by those derivatives. Otherwise, the problem has no solution! There is no way to tell, merely by inspection, whether or not problem (11) is well posed. The index k of nilpotency of the matrix N, called the index of the LIVP, furnishes a rough measure of its difficulty. For time-varying linear systems, as for nonlinear systems, there is no guarantee that the index even remains constant! Small wonder that DAEs are among the more active areas of applied mathematical research. The systems that arise in circuit analysis are somewhat special, in that the equations expressing Kirchoff's laws are all linear and algebraic. Accordingly, circuit analysts often spend much time and effort preconditioning particular DAEs by eliminating as many variables as possible by the methods of linear algebra, before tackling the nonlinear and differential parts of the problem. At a guess, they'll be working even harder at it as FPGAs proliferate.

New Mathematical Problems Continue to Arise

Traditional circuit analysis continues, as it has for more than a century, to generate new and interesting mathematical problems. As miniaturization forges ahead, however, new problems arise. Some were discussed in Part I of this article. One that was not is known as "crosstalk" in the case of digital and "interconnect" in the case of analog circuits.

In both cases, the problem has to do with the "electrical length" λ of a signal-conducting "wire." If two such wires run side by side for a distance of more than about $\lambda/4$, the electrical field induced by one will tend to produce spurious signals in the other. If their parallel segments are no longer than about $\lambda/8$, however, the spurious signals will ordinarily be too feeble to detect. The electrical length of a wire has to do with the highest frequency being transmitted through it, and would be infinitessimal if the square pulses of digital circuits were actually square. It is only because their corners are actually rather well rounded, and their Fourier decompositions correspondingly free of ultra-high-frequency sine and cosine waves, that the constraining λ's are as long as they still are.

Circuit analysis is a rapidly evolving discipline. Each passing year renders the "lumped-parameter" methods of yesteryear more inadequate, and the "distributed-parameter" considerations of the future more inescapable. For the moment, rules of thumb like the $\lambda/4$–$\lambda/8$ dichotomy are exceedingly useful, but even they have limits. Challenging problems remain to be solved if accustomed rates of progress are to persist.

James Case is an independent consultant who lives in Baltimore, Maryland.