Power Dissipation in Information Processing

Problems of removing the heat generated in logical operations limit the speed of computers.

Robert W. Keyes

The speed of computers has increased by several orders of magnitude in the past two decades. Nevertheless, even more powerful information processing machines are needed (*1*). Information processing machines are based upon the embodiment of logical functions in some kind of physical phenomenon. Many methods of using physical phenomena to perform logical functions have been proposed, and some of these proposed embodiments have been constructed and shown to work. The quest for faster machines leads to questions such as, What are the limits to what the diverse physical methods for performing logical operations can do? How shall they be compared? One way to compare different kinds of logic is to look at the performance that has been obtained from them by serious effort. However, most physical schemes have never been the object of any significant amount of developmental work, and neglected methods will be at a disadvantage in terms of performance achieved. Obviously, such a comparison may be a reflection of the state of technology for using the method rather than of any inherent property or potential of the method.

Of course, the usefulness of a physical method of performing logical processing operations depends on many varied factors. However, it appears that a serious limitation on the performance of modern electrical information processing machines is thermal in nature—the power dissipated in the machine will cause its temperature to rise to unacceptable levels, if development continues along present lines (*2*). It might be expected that this kind of limitation would be susceptible to quantitative physical analysis, and a framework that permits it to be evaluated for different physical embodiments of logic is desired.

It should be emphasized here that the most common, powerful, and useful system for carrying out automated information processing is the general purpose digital computer, and that logical processing is considered in this context. A number of features characterize the environment provided by a general purpose computer (*3*). There are many logic elements. The elements process a sequence of information. An element must perform an information processing act, then, very shortly, be ready to process information again without regard to the contents of the previous act. Also, each part of the information sequence passes successively through many elements. The signals must be standardized at each step to prevent their degradation in traversing a long series of logical stages. The standardization also insures interconnectability; it allows the output of any element to be used as the input of any other element. Such standardized binary signals are familiar in biology as a manifestation of the "all-or-none law" (*4*). Efficient and rapid communication between elements is required. Rapid communication means that the distance between elements must be small to reduce delays in the interconnections, since signals can travel no faster than the velocity of light and frequently travel much more slowly. Efficient communication means that energy is transferred efficiently from element to interconnection to element.

The author is a staff member of the IBM Thomas J. Watson Research Center, Post Office Box 218, Yorktown Heights, New York.

Dissipation in Logic

The problem of the heating of logical processors arises in the first place because information processing is dissipative. Von Neumann recognized the dissipative nature of information processing at an early stage in the development of computers and calculated the energy which is dissipated "per elementary act of information, that is, per elementary decision of a two-way alternative." According to von Neumann (*5*), the "thermodynamical minimum of energy per elementary act of information" is $kT \log_e 2 = 3 \times 10^{-14}$ erg (here k is the Boltzmann constant and T is the absolute temperature, about 300°K), and the human brain dissipates 25×10^{-10} watt per neuron, or 3×10^{-3} erg per "binary act." R. W. Landauer (*6*) explained the nature of the dissipation in detail and considered some particular problems quantitatively. He found that several lines of thought suggest that information processing dissipates power.

The first, perhaps, is the irreversible nature of information processing in a general purpose computer (*6*). It is not in general possible to reconstruct the input of logical functions from the output. The logical irreversibility implies physical irreversibility and dissipation. Although in principle one can conceive of preserving a complete history of the computer operations, so that they could all be reversed and all of the energy used could be eventually recovered, this is not practical in a general purpose computer. In such a computer, logical processing stages perform their function, pass the information on to a next stage, and are restored to a state in which they are ready to process information again, losing their memory of what was done in the preceding step. The logical function AND is illustrated in Fig. 1; it may be seen that the output contains less information than the inputs.

A second contact between logical processing and physics arises by way of measurement theory (*6*). The theory of measurement has been discussed by many authors, most notably Brillouin (*7*). From the measurement point of view, a logical processing element is regarded as something that measures the status of preceding elements. Brillouin shows that measurements require the dissipation of an amount of energy of the order of kT.

A third contact with physical theory

arises from the analogy with life processes that has been suggested by many authors. The role of energy flow in life processes has been emphasized most recently and effectively, perhaps, by Morowitz (*8*). The point is that a living system is an ordered state, or a state of low entropy. The system, if left to itself, quickly decays to a disordered state containing no information, but the ordered state can be maintained by a flow of energy through the system. Similarly, the logical elements of modern digital computers contain information only by virtue of being in a nonequilibrium, dissipative state, and the processor contains information because of the flow of energy through it. In fact, the most favorable state of a system through which energy is flowing is not in general the most disordered state but a state of lower entropy and higher order.

The quantitative consequence of the foregoing considerations is that simple physical ideas such as those of Brillouin and Landauer predict minimum dissipations of the order of kT per logical event. Landauer has emphasized the epistemological significance of this finding: In a real world in which a finite amount of energy is available, procedures involving arbitrarily large numbers of steps are not executable (*9*). The dissipations found are, in fact, greater than kT by many orders of magnitude. The reason for logical elements' dissipating energies so many orders of magnitude greater than kT per operation stems from additional requirements that must be placed on general purpose computer elements.

The first of these properties is nonlinearity. Logical functions are basically nonlinear functions—for example, the AND function shown in Fig. 1. Therefore, the embodiment of logical processing elements in terms of physical phenomena requires nonlinear physical phenomena. A. W. Lo has made the need for nonlinearity in a large information processing system much more clear. Lo's basic point is that digital signals must be standardized; a zero must look like a zero and a 1 must look like a 1, regardless of their source (*3*). In its application to electrical signals, this idea means, for example, that, if it is intended that a voltage V_1 represents a 1, then a circuit that receives a voltage V_1' less than V_1 must transmit a voltage that is closer to V_1 than V_1' is. Otherwise, there would be a steady deterioration of voltage level as information passes from stage to stage through the processor. When it is realized that a similar statement must apply to voltages near V_0, that are intended to represent a zero, it is clear that the standardization requires a nonlinear response. The relationships are illustrated in Fig. 2.

X = A and B

A	B	X
0	0	0
0	1	0
1	0	0
1	1	1

Fig. 1. A logical function, illustrating that the output X is not a linear function of the inputs A and B.

The significance of these observations is that physical systems respond linearly to small signals, and the need for physical nonlinearity requires that large signals be used. Large signals mean, however, that the power is high. Although it cannot be rigorously argued that all of the power is necessarily dissipated, the earlier arguments strongly suggest (and experience amply confirms the inference) that a substantial fraction of the power is, in fact, converted to heat. Thus a large amount of heat is created. The dissipation of heat is an especially serious problem in large, fast computers. A large computer contains a great many logical elements that, in a fast machine, must be packed very close together in order that the delay time involved in the transmission of signals from one element to another be short. The heat is eventually carried out of the computer by some fluid. The higher the speed is, the closer the elements must be packed, the higher the density of heat dissipation is, and the more difficult it becomes to carry heat away from the element, transfer it to a fluid, and carry it away.

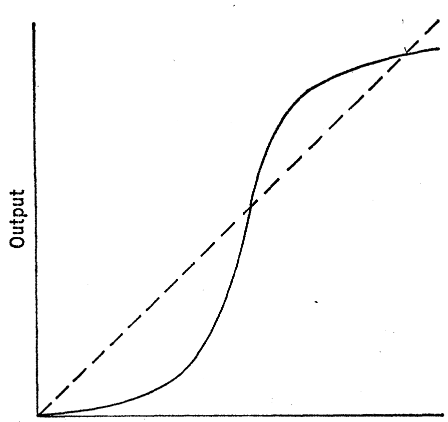

Fig. 2. A nonlinear physical response, illustrating the way in which nonlinearity can restore signal levels (*3*).

The problem would be alleviated if higher-speed elements dissipated less power, but this has not turned out to be the case in high-speed logical processing elements of kinds now known. In fact, in many kinds of physical embodiments of logical elements, exactly the opposite is the case: the power dissipation increases as the speed of the logical elements is increased.

Another consequence of the high degree of excitation needed to obtain nonlinearity in computer elements is that thermal fluctuations are not important. The power spectrum of thermal fluctuations, or thermal "noise," has a uniform density kT per unit of bandwidth. A signal that clearly outweighs the fluctuation must have a spectral power density several times as large, and, since the bandwidth must be something like the reciprocal of the operating time of an element, the signal energy per operation is several kT. The large signals needed to achieve nonlinearity will involve energies greater than kT by many orders of magnitude.

Another aspect of the general purpose computer that is important to an understanding of the power dissipation is the necessity for efficient communication between elements. This requirement manifests itself in different ways in different physical embodiments of logic.

These general observations can be translated into semiquantitative calculations when the scale of nonlinearity can be recognized for particular physical phenomena. Two characteristic parameters of logical elements are relevant to the discussion. The first is speed, or the reciprocal of the time delay between the arrival of an input at an element and the arrival of the result of the processing at the next stage. Higher speed or smaller delay is an objective of the development of logical circuits and devices. A limit on the speed of logical devices is set by the observation that, if a device is to have a delay t_D, then whatever particles or waves represent the signal in the device must be able to traverse the device in a time less than t_D. Thus, as devices are made faster they are also made smaller. Another contribution to delay arises from the fact that signals are processed by many logical elements successively and the time taken for a signal to pass from one logical element to another is part of the total logical delay. Again, if the delay per element is to be less than t_D, then the propagation delay between elements must be less than t_D. Thus, as elements are made faster, the average spacing between them must be reduced and the packing density increased.

The second important characteristic parameter of a logical element is its power dissipation. As suggested above, the power dissipation of logical elements must be relatively high because the attainment of nonlinearity requires that large signals be used. The point here is that, as elements become faster —that is, as they are made smaller and are packed more densely—the problem of removing the heat dissipated in them becomes more and more serious.

Thus the following general line of attack yields a comparison of the thermal limits on the performance of logical elements based on different types of physical phenomena. The basic nonlinearity in the physical phenomenon involved, and thus the minimum required signal level, is identified. The power level can be calculated from the signal level and from other properties and constraints, usually involving the need for controlled communication between elements, of the physical system. Then the signal velocity involved leads to a relation between the size of a logical element, or the space it may occupy, and its speed. Combining the estimate of minimum power with the relation of size to speed provides a quantitative expression of the thermal problem that will be encountered. Supplementing this picture with a knowledge of methods of heat removal permits an estimate of the point, if any, at which the thermal problem becomes insoluble.

Naturally, the greater the extent to which the implementation of the general program outlined above can be based upon fundamental principles, the more convincing and useful it is. However, the exact way in which it can be applied and the degree of detail vary from case to case. In addition, it is apparent that the values of many material properties are involved in quantitative evaluations of the problems and limits described. Physical theory does not now provide means for relating these material properties to fundamental principles. Therefore, it is occasionally necessary to make educated guesses as to the most favorable values (from the viewpoint of improving logical speed) of certain material properties that are consistent with, or suggested by, qualitative physical reasoning, or to fall back on the most favorable experimental determinations of these material properties.

Electrical Logic

The most important physical embodiment of logic to consider in detail is "electrical logic," in which signals are represented by electrical quantities and are carried from circuit to circuit by electrical connectors. All modern computers are based on electrical (transistorized) logic, and the associated technology is so firmly established that no competitive technology would be attractive unless it offered very large potential advantages by comparison. The basic function of electrical signals is to change the potential of electrons in some spatial regions with respect to those in other regions. Now, electrons are dispersed in energy by an amount of approximately kT by thermal agitation. Electrical voltages that are small relative to kT/q are thus just a small perturbation on the steady energy distribution of the electrons and produce only linear effects (q is the charge of an electron; $kT/q = 0.025$ volt at 300°K). Nonlinear effects can be produced by voltages that are large relative to kT/q.

This scale of nonlinearity is most perfectly exemplified by the ideal *p-n* junction, in which the current depends on the voltage as

$$i \propto [\exp(qV/kT) - 1] \quad (1)$$

The junction characteristic as expressed by Eq. 1 is widely accepted as a maximum electrical nonlinearity (*10*). The scale is valid even in such a completely different case as superconductive Josephson tunneling, where the large nonlinear effects occur at a voltage something like the energy gap of the superconductor, a voltage greater than the temperature itself (*11*). It has been suggested that the *i-V* characteristics of tunnel diodes may show greater nonlinearity than is represented in Eq. 1, but only marginally greater nonlinearity has, in fact, been found experimentally (*12*). The (kT/q) voltage scale of nonlinearity is further reinforced by a theorem (*13*) which shows that greater electrical nonlinearity must be accompanied by enhanced shot noise, which apparently has not been observed in solids. A similar scale even seems to be applicable in biology: neuron voltages are a few times (kT/q). Thus, various lines of thought agree in establishing (kT/q) as a practical minimum voltage scale for the production of nonlinear electrical effects. The word *scale* implies that the attainment of the very large nonlinearities needed for reliable logical operation in the presence of such disturbing influences as cross talk, environmental fluctuations, and component variability will require that voltages a great many times (kT/q) be used in real circuits. Questions such as the necessary degree of reliability and the acceptable amount of component variability lie outside the realm of quantitative physical science, and the only statement that can be made about the voltage is that it must be large relative to (kT/q). This same uncertainty, attributable to the existence of compromises with other factors not encompassed in physical theory, pervades all of the situations to be analyzed.

Another constraint on electrical logic stems from the fact that circuits are interconnected with assemblies of conductors that must be regarded as transmission lines. Efficient exchange of signal energy between the logical devices and the lines requires that the devices have impedances comparable to those of the lines, which must be close to Z_0, the impedance of free space, which is equal to $4\pi/c$ (c = the velocity of light) in a Gaussian system of units, or 377 ohms.

Knowing the voltage level and the impedance level allows us to immediately write down the power level

$$P = (kT/q)^2/Z_0 = 2 \times 10^{-6} \text{ watt} \quad (2)$$

(see *14*). As noted above, the ideal scale of nonlinearity can be found in *p-n* junctions, and *p-n* junction transistors are indeed the basis of most logical devices. Equation 2 represents a basic minimum power level that must be multiplied by a large factor to obtain the power level of useful logical circuits. The origin of the large factor, which is 10^4 to 10^5 for the fastest transistorized logic circuits, lies mostly in the need for voltages 10 to 100 times (kT/q) to provide large nonlinearity. Another important conclusion that can be drawn from Eq. 2 is that the power dissipated is independent of the speed of the circuit. At first sight, this conclusion seems to conflict with the circuit designers' practice of trading speed and power. Indeed speed can be traded for power, with a given kind of transistor, in a way that may be written Pt_D = constant. Progress in transistor technology has, however, reduced the value of the "constant," and the independence of power and speed should be compared with the survey of the power dissipation of logical devices shown in Fig. 3, which is historical as regards transistors in the sense that improved transistor technologies have caused a steady motion through the region designated "transistors" (*Tran*) in Fig. 3 in the direction of lower power and higher speeds (toward the lower left). The point is that the power dissipation of the fastest circuits has remained nearly constant at around 100 milliwatts, although their speed has increased by three orders of magnitude since introduction of the transistor into computers.

A disadvantage of electrical logic is the fact that fast transistors must be very small. The response of an electrical element to applied signals is limited by the velocity with which electrons can travel in solids. For transistors, it turns out that the relation between a linear dimension d and the delay time is roughly $t_D = d/s_0$, where s_0 is dimensionally a velocity and has a value of 10^6 cm/sec (*2*). Thus progress toward higher speeds has meant decreasing transistor sizes. The power density in the device has become very high, of the order of 10^4 watt/cm^2 for the fastest transistors. Heat cannot be transferred to fluids at such a high density; it is allowed to flow away from the device by conduction through a solid, spreading out in three dimensions, until its density is low enough to permit transfer to a fluid. In the solid, it encounters thermal spreading resistance. Now, the thermal spreading resistance is inversely proportional to a linear dimension d having the form $R_{th}=c_2/2Kd$. Here c_2 is a dimensionless constant that depends on the exact geometrical configuration of the circuit elements and K is the thermal conductivity of the solid. The temperature rise of the device (ΔT) is equal to the thermal resistance times the power P, or

$$\Delta T = c_2P/2Kd = c_2P/2Ks_0t_D \quad (3)$$

If ΔT has a maximum value, then t_D has a minimum. The following values are one reasonable estimate of the parameters in Eq. 3: $c_2 = 0.05$, $P = 0.1$ watt, K (for silicon) $= 1.5$ watt/cm deg, $s_0 = 10^6$ cm/sec, and ΔT (maximum) $= 60°C$. These values imply $t_D > 28 \times 10^{-12}$ sec. Some time of this order, then, seems to be a limit on how far t_D may be reduced by continued development of silicon circuitry along present lines (*2*).

Optical Logic

Elements based on purely optical phenomena would not suffer from the limitation set by the velocity of electronic motion in solids. The ability to make elements small is an important characteristic of the present state of technology, and many aspects of this technology can be applied to the construction of either optical or electrical elements. Therefore, at any given state of technology in which elements can be made to a certain size, it might be expected that optical elements would be far faster than electrical elements because of the greater velocity of optical signals. In fact, the fastest electronic phenomena known are the very short pulses emitted by mode-locked lasers (*15*). This high inherent speed of optical phenomena has given rise to interest in digital "optical logic" as a way to circumvent the limitations of electrical logic (*16*). It is appropriate, therefore, to examine the thermal problems of digital optical logic.

In a search for sources of nonlinearity in optical phenomena that might be used as the basis of an optical logic system, three such effects were recognized (*17*). The three basic phenomena and the power levels associated with them are as follows.

The first kind of nonlinearity is photochromism or the bleachable absorber —that is, a two-level system in which incident light causes extensive repopulation of absorbing centers, thereby changing the optical properties of the system. The change might, for example, simply be used as an optical switch, in which one light beam changes the transparency of a medium with respect to another beam, or it might be used in a more complex fashion, in which light in one mode changes the threshold of a system for lasing in another mode by its effect on electronic populations.

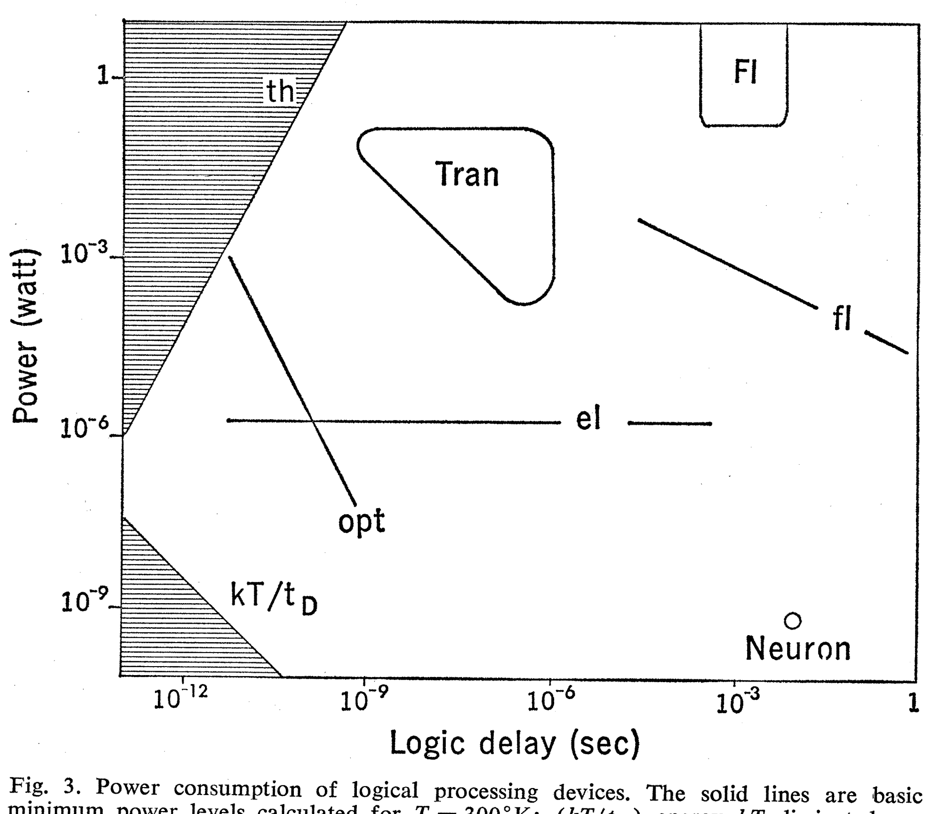

Fig. 3. Power consumption of logical processing devices. The solid lines are basic minimum power levels calculated for $T = 300°K$: (kT/t_D,) energy kT dissipated per logical operation, corresponding to a power level kT/t_D; (*opt*) optical logic (Eq. 5); (*el*) electrical logic (Eq. 2); (*fl*) fluid (pneumatic) logic (Eq. 13); (*th*) the thermal limit (Eq. 14) with $c_1 = 10^{10}$ cm/sec, $m = 10$, and $Q = 100$ watt/cm^2. The performance of real logical devices is expected to be a few orders of magnitude larger than the basic minimum levels, as illustrated by the examples shown: (*Tran*) transistors; (*Fl*) fluid elements; (*Neuron*) [after von Neumann (*5*)] a nerve cell with its processes. The inaccessible regions, above the thermal transfer limit and below the kT limit, are shaded.

Bleaching effects occur when the optical intensity I (in photons per square centimeter per second) satisfies the condition

$$I >> (\sigma t_1)^{-1} \qquad (4)$$

(see *18*). Here σ is the absorption cross section of the centers and t_1 is their lifetime. Equation 4 identifies the basic scale of nonlinearity of the bleaching phenomenon. It states that the rate of absorption of photons per nonexcited center must be large relative to the rate of decay per excited center, so that the light can maintain most of the centers in the excited state. The logic delay time t_D must be larger than t_1. Therefore, many photons are needed per logic delay time for each center involved. The other requirement of a computer system—that the location of beams of light be controlled and that they be conducted from one logical element to another—can be satisfied by confining the beams in optical wave guides, which must, however, have a cross-sectional area of the order of the square of the wavelength λ. An effective system requires that the centers completely absorb the light in a wave guide, so a number of centers that is large compared to λ^2/σ is required. Now σ is related to T_2 (*19*), a time of the order of the inverse line width, and t_{rad} (*20*), the radiative lifetime of the centers, in a way that can be expressed by

$$\lambda^2/\sigma = 2t_{rad}/3T_2$$

and σ and t_{rad} are also related to p, the dipole matrix element of the transition (*19*). A large value of p—namely, the electronic charge times the Bohr radius a_0—is used here in calculating the minimum power of logic and yields a minimum value for t_{rad} of 10^{-6} second. Since, for fast logic, the lifetime of the excited state must satisfy the relation $t_1 << 10^{-6}$ second, the processes that determine t_1 must be nonradiative. Furthermore, T_2 is certainly less than t_1 (and t_D), so, for fast logic, $(\lambda^2/\sigma) >> 1$.

The basic power level of photochromic optical logic, when t_{rad} is expressed in terms of the dipole moment qa_0, is

$$P >> (\lambda^2/\sigma)\ (h\nu/t_D) = \hbar^2\lambda^2 c/8\pi^2 q^2 a_0^2 t_D^2 \qquad (5)$$

The second kind of optical nonlinearity is based on the nonlinear dielectric susceptibility. The time delay involved in nonlinear dielectric effects, which must be less than the logical delay t_D, is the time required for light to traverse an interaction length for nonlinear effects. This interaction length is inversely proportional to the strength of the electric field and, also, to a nonlinear dielectric constant (*21*). The magnitude of a large nonlinear dielectric constant is something like the reciprocal of the internal fields in crystals, or q/a_0^2. Although this value for the nonlinear dielectric constant is used in calculating power levels, somewhat larger values of the nonlinear dielectric constant are known in a few materials, and the possibility that still higher values will be found, which will lower the calculated power level, cannot be excluded. The electric field E needed to attain a delay time less than t_D is

$$E > \lambda q/8\pi^2 a_0^2 c t_D \qquad (6)$$

(see *21*). The electric field in the wave determines its power density. The power level is proportional to the square of the electric field, and, since the electric field is inversely proportional to the interaction length or the delay time, the power is inversely proportional to the square of the delay time. The power level of nonlinear dielectric effects, when the cross-sectional area of the light beam is again set equal to the square of the wavelength, is

$$P > \lambda^4 q^2/256\pi^5 a_0^4 c t_D^2 \qquad (7)$$

(see *17*). The third nonlinear effect involves the self-induced transparency caused by $n\pi$ pulses, described by McCall and Hahn (*19*). An $n\pi$ pulse is one in which the electric field of the light in the pulses satisfies the condition

$$(2p/\hbar) \int E dt = n\pi \qquad (8)$$

Culver and Mehran (*22*) have described methods of performing logical functions with such pulses in a system of two-level centers.

The same large value for p, qa_0, is used here in calculating the power of coherent optical logic with $n\pi$ pulses. If the logic delay is t_D, then the integral in Eq. 8 is approximately Et_D. The power level is again inversely proportional to t_D^2, and, for an area λ^2 and $n = 1$, is

$$P = \pi\hbar^2\lambda^2 c/16 q^2 a_0^2 t_D^2 \qquad (9)$$

(see *17*, *22*). Now it should be observed that the power levels of Eqs. 5, 7, and 9 differ essentially only by numerical factors and therefore define a basic power level of optical logic. [The "numerical factor" that separates Eq. 7 from Eq. 5 may be written $(2\pi)^{-1}$ $(\text{Rydberg}/h\nu)^2$, and $h\nu$ is about 2 electron volts in optics.]

Fluid Logic

Another example of a physical embodiment of logic, which has found some favor because of its simplicity, is "fluid logic." The physics of fluid logic has been reviewed by Glaettli (*23*). The scale of nonlinearity is set by the requirement that the Reynolds number R ($= \rho v L/\eta$, where ρ is the density of the fluid, η is its viscosity, v is the velocity, and L is a linear dimension) must be larger than about 10^3. The delay time t_D is of the order of L/v, or, when v is eliminated by introduction of the Reynolds number,

$$t_D = \rho L^2/\eta R \qquad (10)$$

Also, the pressure is of the order of

$$p = \rho v^2 = R^2\eta^2/\rho L^2 \qquad (11)$$

and the power is

$$P = p v L^2 = R^3\eta^3/\rho^2 L \qquad (12)$$

Combining Eqs. 10 and 12 yields

$$P = R^{5/2}\eta^{5/2}/\rho^{3/2} t_D^{1/2} \qquad (13)$$

In fact, pneumatic logic devices are found to use power of the order of 10^3 times that suggested by Eq. 13.

Thermal Problems

In large computers in which a large number of logical elements are closely packed together, the heat dissipated by the logical circuits is carried out of the machine by some fluid. The heat generated by the elements must be transferred to the fluid through an interface, and this can only be done at rates below some maximum rate Q. A knowledge of the power per logical element and of the maximum rate of heat transfer allows calculation of the area per element and, thus, of an average distance between elements. The minimum distance between elements leads to a minimum propagation delay. Experience shows that signals must travel some number m times the average interelement spacing between logical operations. Thus, there is a limit on systems using elements each of which dissipates power P of the form

$$Q(c_1 t_D)^2 > m^2 P \qquad (14)$$

Here c_1 is the velocity of signal propagation. Equation 14 sets a maximum value on the power that can be dissipated by a logical element with a time delay t_D. On the other hand, the preceding discussion has shown the minimum values of power that are needed to operate a logical element with time delay t_D. Obviously, then, the performance of various types of logical elements is limited by the condition that the power required must be less than the power that can be dissipated.

A practical value for Q is the maximum rate of heat transfer in nucleate boiling of fluids that boil around 300°K—about 100 watt/cm^2. In Fig. 3 this value is used in a comparison of Eq. 14 with the power required for operation of logical elements according to Eqs. 2, 5, 7, 9, and 13. The ranges of power actually used are also shown in cases where information is available. Fast electrical logic circuits use more power by four or five orders of magnitude than the basic level given by Eq. 2. If the actual power level of the fastest transistorized logic circuits, 0.1 watt, is substituted in Eq. 14, and if $c_1 = 10^{10}$ cm/sec—less than the velocity of light because the electromagnetic signals travel through regions occupied by materials of high dielectric constant and index of refraction—it is found that the minimum attainable value for t_D is 32×10^{-12} second. This limit turns out to be about the same as the one caused by the thermal spreading resistance.

The power used by fluid logic devices is very high, both according to Eq. 13, in which values of physical parameters for air at pressure of 1 atmosphere and temperature of 300°K have been used in calculating the line *fl* in Fig. 3 ($\eta =$ 180 micropoise, $\rho = 1.3 \times 10^{-3}$ g/cm^3), and in practice. In fact, fluid logic has been introduced into this discussion, not as a serious contender for the basic design of modern large, fast computers, but to illustrate again the large factor that separates the basic minimum calculated power levels from the power levels needed in real logic circuits. Similar large differences between calculation and actuality for optical logic devices must be anticipated.

In any case, Fig. 3 shows that the basic minimum power level of optical logic devices is greater than the basic minimum power level of electrical logic devices at speeds of less than 10^{-10} second. The rapid rise of the minimum power level of the optical logic device with decreasing delay time causes it to encounter the thermal transfer limit at a value of the delay time two orders of magnitude greater than the delay at the corresponding intersection with the basic power level of the electrical logic device (the intersection of *th* and *el*). Of course, the uncertainty of the large dimensionless factors involved in this comparison obscures its quantitative significance, and the power levels of optical logic depend on a view of material properties that, although seemingly reasonable at present, could conceivably be drastically changed by the discovery of new classes of electrooptic materials, but the comparison clearly suggests that optical logic offers no advantage over electrical logic in surmounting the thermal problems that limit logical speed. As mentioned above, optical components offer a technological advantage—they may be larger for a given speed than electrical components—that might be decisive if technological problems prevented electrical components from approaching the thermal limit. This is not the case; electrical logic circuits with a speed within a factor of 10 of the thermal limits can now be made (*24*).

Two qualitative features of the results summarized in Fig. 3 are important. It appears that technological improvement will be more difficult to achieve in optical logic than in electrical logic because rapid increase of power with increase of speed will be encountered in the former. Furthermore, the sensitivity of the delay time to numerical factors that cannot be accurately included in the calculation is weaker for optical logic than for electrical logic. Whatever neglected numerical factor enters into the thermal limit on delay time is involved in the optical case only to the ¼ power, so the sensitivity of the limits to this factor is limited.

The second important qualitative feature is the fact that the power of electrical logic is proportional to the square of the temperature (Eq. 2). The problems of heat dissipation do not become more serious at the same rate, and it appears that the thermal problems can be alleviated by decreasing the temperature in the electrical logic system. On the other hand, temperature is not a parameter in the performance of optical logic, and no fundamental advantage can be gained by cooling.

Summary

Logical processing elements dissipate the energy of the signals on which they operate to heat. Nonlinear physical phenomena are needed to perform logical processing operations. Nonlinear phenomena require large signal levels; therefore, the amount of heat generated is large. Increases in the speed of logical processing elements are attained by making them smaller and thereby increasing the density of power dissipation, and, frequently, also the power itself. The problem of removing the heat created becomes more and more difficult and, at a certain limiting speed that depends on the technology involved, becomes insoluble. Optical logical elements offer no advantage over electrical elements in surmounting this problem. Of course, new concepts and phenomena, not yet known or not yet applied to logical processing, offer hope for the future.

References

1. *New York Times* (29 September 1968).
2. R. W. Keyes, *IEEE (Inst. Elec. Electron. Eng.) Spectrum* **6**, 36 (1969).
3. A. W. Lo, in *Micropower Electronics*, E. Keonjian, Ed. (Macmillan, New York, 1964), pp. 19–39; *IRE (Inst. Radio Eng.) Trans. Electron. Computers* **EC-10**, 416 (1961).
4. See, for example, the article "All-or-none Law" in *Encyclopaedia Britannica*.
5. J. von Neumann, *Theory of Self Reproducing Automata* (Univ. of Illinois Press, Urbana, 1966), p. 66; see also *The Computer and the Brain* (Yale Univ. Press, New Haven, 1958).
6. R. W. Landauer, *IBM (Int. Bus. Mach. Corp.) J. Res. Develop.* **5**, 183 (1961).
7. L. Brillouin, *Science and Information Theory* (Academic Press, New York, ed. 2, 1962).
8. H. J. Morowitz, *Energy Flow in Biology* (Academic Press, New York, 1968).
9. R. W. Landauer, *IEEE (Inst. Elec. Electron. Eng.) Spectrum* **4**, 105 (1967).
10. W. Shockley, *Bell System Tech. J.* **28**, 435 (1949).
11. J. Matisoo, *Proc. IEEE (Inst. Elec. Electron. Eng.)* **55**, 172 (1967).
12. J. Karlovsky, *Solid-State Electron.* **10**, 1109 (1967).
13. W. Bernard and B. Callen, *Rev. Mod. Phys.* **31**, 1017 (1959); J. B. Gunn, *J. Appl. Phys.* **39**, 5357 (1968).
14. R. W. Keyes, *Proc. IRE (Inst. Radio Eng.)* **50**, 2485 (1962).
15. A. J. DeMaria, D. A. Stetser, W. H. Glenn, Jr., *Science* **156**, 1557 (1967).
16. J. T. Tippet, D. A. Berkowitz, L. C. Clapp, C. J. Koester, A. Vanderburch, Jr., Eds., *Optical and Electro-optical Information Processing* (M.I.T. Press, Cambridge, Mass., 1965); W. F. Kosonocky, *IEEE (Inst. Elec. Electron. Eng.) Spectrum* **2**, 181 (1965).
17. R. W. Keyes and J. A. Armstrong, *Appl. Optics* **8**, 2549 (1969).
18. R. W. Keyes, *IBM (Int. Bus. Mach. Corp.) J. Res. Develop.* **7**, 334 (1963); M. Hercher, *Appl. Optics* **2**, 947 (1967).
19. S. L. McCall and E. L. Hahn, *Phys. Rev. Lett.* **18**, 908 (1967).
20. See, for example, F. K. Richtmyer and E. H. Kennard, *Introduction to Modern Physics* (McGraw-Hill, New York, ed. 4, 1947), pp. 282 ff.
21. J. A. Armstrong, N. Bloembergen, J. Ducuing, P. S. Pershan, *Phys. Rev.* **127**, 1918 (1962).
22. W. Culver and F. Mehran, in preparation.
23. H. Glaettli, in *Advances in Computers* (Academic Press, New York, 1963), vol. 4, pp. 169–243.
24. F. H. Dill, A. S. Farber, H. N. Yu, *IEEE (Inst. Elec. Electron. Eng.) J. Solid State Circuits* **SC3**, 160 (1968).