Solid-State Electronics Vol. 25, No. 9, pp. 933–942, 1982
Printed in Great Britain.

0038–1101/82/090933–10$03.00/0
© 1982 Pergamon Press Ltd.

ON THE THEORY OF THE THERMIONIC EMISSION TRANSISTOR II. TET AS AN ELEMENT OF LOGIC CIRCUITS

SERGE LURYI and RUDOLPH F. KAZARINOV
Bell Laboratories, Murray Hill, NJ 07974, U.S.A.

(*Received* 12 *January* 1982; *in revised form* 22 *February* 1982)

Abstract—The theory of the Thermionic Emission Transistor proposed by us earlier is further developed. It is shown that the structure can be significantly simplified with only a marginal loss in its performance. The simplification consists in eliminating the need for a built-in planar doped triangular potential barrier. The current-voltage characteristics, the voltage gain, and the transconductance are calculated for exemplary TET structures. The proposed simplification makes it feasible to incorporate devices of complementary carrier type on a single chip thus forming the basis for CMOS-like logic. An important advantage of the TET in this application is the low required level of the supply voltage, viz. approx. 0.5 V, which virtually eliminates all the parasitic bipolar effects. The transfer characteristic is calculated for a basic inverter gate. For a numerical example of an inverter pair of total area 10 μm × 10 μm and period 0.6 μm of the surface electrode pattern the estimated inverter delay is 5 ps for fan out of 1.

1. INTRODUCTION

In a previous paper[1] we proposed a novel device called thermionic emission transistor (TET). Operation of this device is based on gate controlled charge injection in unipolar semiconductor structures. The current in the TET flows perpendicular to the semiconductor surface over a planar potential barrier controlled by an interdigitated set of electrodes at the surface. It is shown in [1] that the gate delay of the TET is limited by the time of flight of electrons across the structure and can be in the picosecond range with the gate voltage swing required to switch the output current at room temperature being of order 0.2 V.

In the present work we develop a theory for a radically simplified version of the TET which, nevertheless, is not significantly worse than the original version in terms of the estimated device performance. This simplification allows us to propose a TET-based complementary logic for integrated circuits similar to the well-known CMOS logic[2]. Inasmuch as the TET operates at low voltages (it will be shown that a complementary inverter can operate at a supply level of order half the band gap) we can expect no parasitic bipolar effects.

Let us briefly review the original TET design, see Fig. 1. The device contains an intrinsic layer grown epitaxialy on a heavily doped n^+ substrate. In the process of growth by molecular beam epitaxy (MBE) a thin p^+ layer is built in the intrinsic layer. Acceptors in the p^+ layer are completely ionized forming a sheet of fixed negative charge. This charge sheet gives rise to a potential barrier of triangular shape[3, 4]. The n^+ substrate forms one of the terminals of the device called the cathode. The other two terminals, the anode and the gate are arranged in a periodic pattern of stripes on the surface. Every other stripe represents a metallized n^+ contact or a silicide layer and these stripes are connected on one side to a metallic pad which is the anode terminal. The alternate stripes connected on the other side form an MOS gate terminal.

This design appears to contain two key elements. First, the interdigitated gate-anode surface electrode structure, whose purpose is to allow an independent control of the anode current by the gate voltage. As shown in [1], the lateral inhomogeneity of the electric field averages out exponentially with distance from the surface. The characteristic length λ of the inhomogeneous region is related to the period d of the surface electrodes as $\lambda = d/2\pi$. The second key element of the original design is the planar-doped barrier, which serves as a charge injector. If the built-in charge sheet is located sufficiently far from the surface (far compared to λ), then the barrier height with respect to the substrate (and, therefore, the magnitude of thermionic current) is determined both by the anode voltage V_A and the gate voltage V_G in the following combination:

$$\bar{V} = A\,V_A + G\,V_G \tag{1}$$

where $A = S_A/S$ and $G = S_G/S$ with S_A, S_G being the anode and gate areas, respectively, and $S = S_A + S_G$. The ratio G/A determines the voltage gain which must be greater than unity.

The simplification proposed in this paper consists in doing away with the planar-doped barrier. The question arises, how much transconductance do we have to sacrifice in this case and whether we can achieve any voltage gain at all for a reasonable value of the gate-to-anode area ratio. At first sight it appears that we have to pay a stiff penalty for giving up the built-in charge sheet, since we no longer have a planar *equipotential* barrier whose height is controlled by $\bar{V}$. Instead, the potential profile becomes essentially two-dimensional as shown in Fig. 2, and the thermionic current flows in the vicinity of a saddle point under the anode. Nevertheless, it will be shown that the *IV* characteristics for the simplified version of the TET remain rather similar to those calculated in [1] for the device containing a built-in charge sheet. At low currents we find an exponential dependence

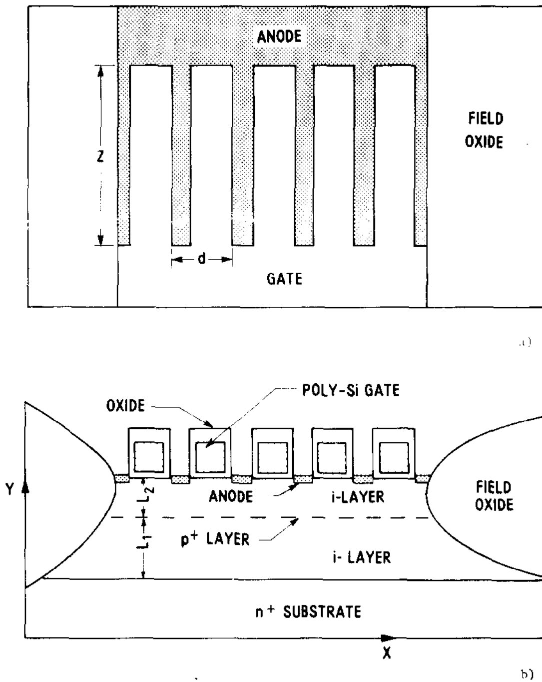


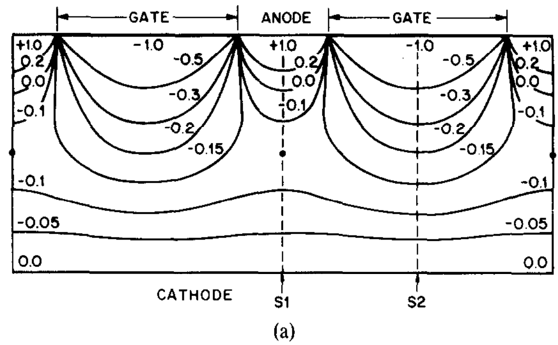
Fig. 1. Schematic diagram of the original version of the TET. The planar-doped barrier shown by the dashed line is eliminated in the present design. (a) Top view. (b) Cross-section in the working area of the device.

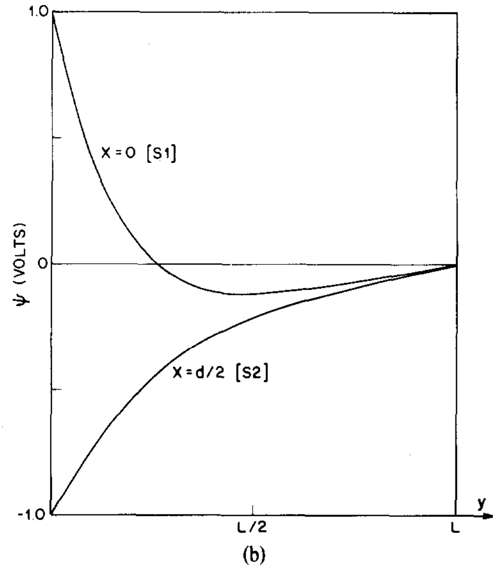
(characteristic of charge injection) of current I on voltage $\tilde{V}$ where $\tilde{V}$ is a linear combination of V_A and V_G like in eqn (1) but with effective weights α and γ in place of the geometrical factors. At higher currents the charge injection regime goes over into a space charge limited mode in which I depends on $\tilde{V}$ in a linear fashion.

The ratio γ/α which now determines the voltage gain is degraded by about 25% compared to the geometric ratio.

With the planar-doped barrier removed it becomes apparent that the TET is conceptually similar to a number of other transistors of the so-called analog type, notably the permeable base transistor (PBT)[5] and the static induction transistor (SIT)[6]. The concept of the analog transistor was first analyzed by Shockley in 1952[7]. The principle of operation of this device is analogous to the vacuum-tube triode in that the current is space-charge limited and is controlled by input electrodes placed between the cathode and the anode.† It is important to distinguish the analog transistor from a conceptually different device also proposed by Shockley in 1952, the junction field effect transistor (JFET)[8], in which the current control is effected through a motion of a depletion region boundary. This distinction is not always made carefully in the assessment of analog transistors, because in some of the proposed versions these

†Although claims to the contrary have been made in literature[5], this analogy goes quite far. The potential barrier in the lightly-doped transit region of an analog transistor near the boundary with the heavily doped source is created by the mobile charge which diffuses from the source just like the space-charge potential of electrons emitted from a vacuum-tube cathode. The similarity ends when one considers the saturation regime of the vacuum device which is due to a fixed cathode work function. In this sense we can say that the source is a cathode with vanishing work function. Because of this the drain current characteristic exhibits no saturation.





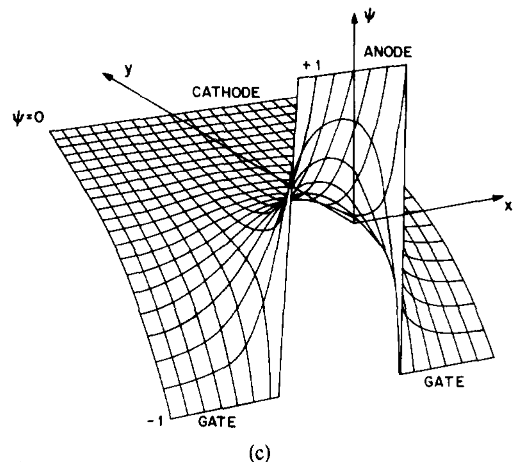


Fig. 2. Potential distribution in the device. (a) Equipotential lines. (b) Variation of the potential along section S1 and S2. (c) Stereometric projection.

devices can be, in fact, JFETs. For example, when the doping concentration in the transit region is large compared to the injected charge density, the SIT and the PBT become equivalent to a vertical JFET and a vertical MESFET, respectively. The conceptual difference between the two types of transistor is in the physical mechanism of current transport. In the analog transistor this is a combination of thermonic and space-charge limited current mechanisms, whereas the JFET and MESFET are basically voltage controlled resistors.

In the PBT the controlling (base) electrodes represent a metallic grid embedded in the body of the semiconductor in the transit region between the cathode and the anode. The PBT represents an improvement over the SIT in which the controlling electrodes are formed by p^+ gate regions (or, generally, by heavily doped regions complementary to the source). As discussed in [5], the metallic base eliminates parasitic bipolar effects such as minority carrier storage, and reduces the base resistance. Theoretically, the PBT with undoped transit region offers the highest possible performance of all analog transistors. However, its practical implementation faces considerable difficulties in growing high-quality material over the metallic grid. In the TET the controlling electrodes are located on the semiconductor surface. The essense of our work is to show that even in this configuration the properly chosen geometry allows one to achieve a sufficient voltage gain.

In our view, neither the FET nor the analog transistor have an advantage over one another in the intrinsic speed of operation, when one considers ideal structures of similar feature size. In both cases the intrinsic delay is limited by the time of flight of carriers from the source to the drain. Strictly speaking, the effective channel length, which determines the transit time is the length of the space-charge accumulation region which coincides with the source-to-drain distance, L, only for long channel devices. Although in analog transistors the channel is usually perpendicular to the semiconductor surface nevertheless the requirement of having a voltage gain always limits the minimum effective channel length by a characteristic lateral dimension of the controlling electrodes. In the case of the PBT, for example, this dimension is the period of the base electrode.

Another important characteristic (especially for VLSI applications) is the required voltage level which crucially determines the power-delay product and thus limits the possible integration density and the speed of operation. According to this figure of merit the analog transistors are clearly not favored if one considers long-channel devices ($L \gg 1\ \mu$m). Indeed, the potential due to space-charge accumulation at a given current density scales as L^2 (assuming saturated velocity) and so does, therefore, the required drain voltage. For submicron dimensions this is no longer a limitation and in this range the analog transistors can operate at the supply level of 0.5 V and lower. In principle, field effect devices, notably the MOSFET, can also operate at the same low voltage level. It should be emphasized that for submicron devices one gains no advantage in speed of operation by going beyond threshold because in the strong inversion regime both the output current and the entire capacitively stored charge are proportional to each other. The real limitation is, therefore, due to insufficient accuracy of the threshold voltage control. It is in this respect that analog transistors can have an important practical advantage. In our view, the only possible advantage of the analog transitor concept is that it can be realized as a bulk device thus minimizing the threshold variation. In the TET the current transport is by charge injection into an undoped material controlled by a potential barrier removed from the surface. Because the density of injected charge in the intrinsic layer greatly exceeds the background doping the latter should not affect the potential distribution. On the other hand, the influence of the uncertainty of charges associated with the oxide is also diminished because the space charge accumulation occurs far from the interface. With state-of-the-art MOS technology one can expect to be able to control the threshold voltage to within a 25 mV margin. This estimate does not include the effect of possible variations in the geometry of the electrodes which may be the only serious source of uncertainty. For the TET concept to be of value in VLSI applications it is imperative that the threshold variation should not exceed 100 mV.

This paper is organized as follows. In Section 2 the electrostatic problem corresponding to the domain shown in Fig. 2 is solved exactly in the low current limit. This gives us an analytic expression for the thermionic current in this limit. In Section 3 we analyze the transition to the space charge limited current mode. The current crowding problem near the anode is investigated in the most extreme limit of $A \ll G$. A complete set of transistor current-voltage characteristics is calculated in Section 4. In Section 5 we consider the possible use of the TET in logic elements of integrated circuits. Operation of the basic inverter formed by two complementary TET devices is analyzed. Inverter transfer characteristics and gate delay are calculated. Our conclusions are summarized in Section 6.

2. TWO-DIMENSIONAL ELECTROSTATIC POTENTIAL AND THERMIONIC EMISSION IN THE LOW CURRENT LIMIT

Depending on the importance of the injected charge in determining the electrostatic field distribution in the intrinsic layer we can distinguish two characteristic modes of the TET corresponding to the low and the high current conditions. At low current levels we can neglect the mobile carrier charge and the problem of finding the electrostatic potential distribution in the device reduces to solving the Laplace equation $\nabla^2\psi(x, y) = 0$ for the domain and the boundary conditions shown in Fig. 2(a). Neglecting the oxide thickness, i.e. assuming that $t_{ox} \ll L\epsilon_{ox}/\epsilon$ where ϵ and ϵ_{ox} are, respectively, the dielectric permittivity of the semiconductor and the oxide, one finds $\psi(x, y)$ in the form of a series [1]:

$$\psi = \bar{V}\left(\frac{L-y}{L}\right) + \frac{2\Delta V}{\pi} \sum_m \frac{1}{m} \sin m\pi A \times \cos(mx/\lambda) \frac{\sinh m(L-y)/\lambda}{\sinh mL/\lambda} \quad (2)$$

where $\Delta V \equiv V_A - V_G$ and $\bar{V}$ is given by eqn (1). It turns out that in the case when $L \gg \lambda = d/2\pi$ one can perform the summation (2) explicitly and obtain the following expression (see Appendix):

$$\psi(x, y) = \bar{V}\left(\frac{L-y}{L}\right) + \Delta V f(x, y) \quad (3)$$

where

$$f(x, y) = G - \frac{1}{\pi}\left\{\arctan\left[\tanh\frac{y}{2\lambda}\tan\frac{\pi G + x/\lambda}{2}\right] + \arctan\left[\tanh\frac{y}{2\lambda}\tan\frac{\pi G - x/\lambda}{2}\right]\right\}. \quad (3a)$$

As discussed in Appendix this solution corresponds to the exact boundary condition on the potential at the surface and an approximate *asymptotic* condition on the field $\mathbf{E}(x, y)$ for $y \gg \lambda$, viz. $E = E_y = \bar{V}/L$.

Figure 2(c) shows a 3-dimensional representation of the surface $\psi(x, y)$ calculated from eqn (3), and Fig. 2(b) the variation of the potential in vertical sections under the gate and the anode. A stereometric view of the carrier energy $(-q\psi)$ profile in the "off" state of the device is shown in Fig. 3. As seen from these figures the surface $\psi(x, y)$ has a saddle point under the anode. The exact location of the saddle point (x_0, y_0) is given by $x_0 = 0$ and

$$\cosh(y_0/\lambda) = \cos\pi A + |2L\Delta V/\bar{V}d|\sin\pi A \quad (4)$$

When the r.h.s. of eqn (4) is less than unity the equation has no solution; physically this means the absence of a saddle point. From eqn (4) it is easily seen that the condition for existence of a saddle point is of the form

$$|2\Delta V/\bar{V}d| > \tan\pi A/2. \quad (5)$$

This condition is satisfied in most cases of practical interest. Moreover in all cases discussed below we find that $y_0 \gtrsim 2\lambda$ and therefore to a good approximation y_0 satisfies

$$y_0 = \lambda\ln|4L\Delta V\sin\pi A/\bar{V}d|. \quad (6)$$

Equation (6) corresponds to taking only the first harmonic in the multipole expansion (2) in the vicinity of the saddle point, viz.

$$\psi(x, y) = \bar{V}\left(\frac{L-y}{L}\right) + \frac{2\Delta V}{\pi}\sin\pi A\cos(x/\lambda)e^{-y/\lambda} \quad (7)$$

The thermionic current can be written as a curvilinear integral

$$I = Z\int_{c0} A^* T^2 e^{\beta\psi(x, y)} dl \quad (8)$$

where A^* is the Richardson constant, Z is the gate width and $\beta \equiv q/kT$. The path of integration c_0 can be rigorously defined in the following way. One first finds the current density distribution $\mathbf{J}(x, y)$. The vector field $\mathbf{J}$ determines an orthogonal curvilinear grid consisting of the zero-flow lines (those perpendicular to $\mathbf{J}$ at every point) and the tangent to $\mathbf{J}$ lines of current. Through each point passes one and only one line of zero flow. The path c_0 corresponds to that line of zero flow which passes through the saddle point. The two-dimensional current distribution will be considered in more detail in Section 3. Here we can proceed in a simpler way by observing that most of the current flows in the immediate vicinity of the saddle point (x_0, y_0). We may, therefore, replace the zero flow line c_0 by a straight line $y = y_0$ which is tangent to c_0 at the saddle point. A straightforward integration using eqns (7) and (8) gives

$$I = (S/d)A^* T^2 \int_0^d e^{\beta\psi(x, y_0)} dx = A^* T^2 S\, e^{\beta\bar{V}y_0/L} I_0\left(\frac{2\beta\Delta V}{\pi}e^{-y_0/\lambda}\right) \quad (9)$$

where I_0 is the Bessel function. Equation (9) together with (6) determines the low-current part of the IV characteristic. Using eqn (6) we can rewrite the argument of the Bessel function in the form $\lambda\beta\bar{V}/L\sin\pi A$. In examples considered below for all realistic currents this argument is of the order of or greater than 2. We can, therefore, use the asymptotic form of the Bessel function. $I_0(x) \sim (2\pi x)^{-1/2}\exp|x|$ with the accuracy of about 5%. In this way we obtain

$$I = A^* T^2 S\left(\frac{\sin\pi A}{\beta V}\frac{L}{d}\right)^{1/2} e^{\beta\bar{V}[1-(y_0/L)+(\lambda/L\sin\pi A)]}. \quad (10)$$

Since, as evident from (6), location of the saddle point y_0 is a slowly varying function of V_A and V_G, the dependence of I and $\bar{V}$ is mainly exponential at low current densities.

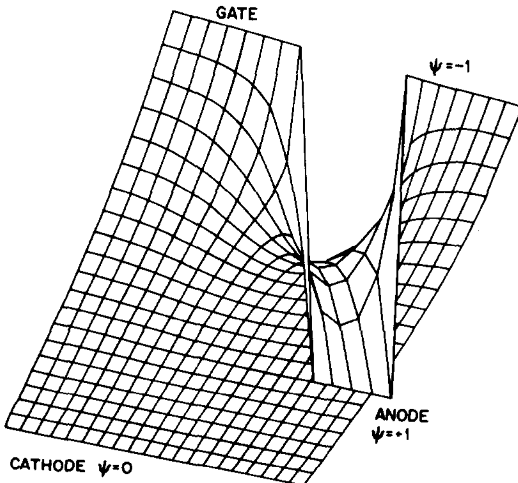


Fig. 3. Stereometric view of the 2-dimensional carrier energy distribution in the intrinsic layer.

3. SPACE CHARGE LIMITED REGIME AND CURRENT CROWDING

At high current densities it is important to take into account the potential $\delta\psi$ associated with the mobile charge drifting with saturated velocity, v_s, on the downhill slope toward the anode. It is easy to take this effect into account in a one-dimensional problem, i.e. when the current density J is uniform in the x direction. One simply integrates Poisson's equation with the charge density $\rho = J/v_s$ along the downhill slope y_0 and finds

$$\delta\psi = \frac{Jy_0^2}{2\epsilon v_s} = \frac{Iy_0^2}{2\epsilon S v_s}. \tag{11}$$

The effect of the mobile charge on the barrier height for thermionic emission reduces to a replacement of the anode voltage V_A by an effective potential $\tilde{V}_A = V_A - \delta\psi$. Physically this corresponds to a screening of the applied field by the mobile charge.

A question arises, how should this be modified in the case when the area of the anode electrodes is less than half of the total area of the device as is the case in the TET. It might appear that the 2-dimensionality of the problem (current crowding near the anode) can qualitatively change the situation, especially in the case when $S_A \ll S$. However, as is shown below even in this extreme case the current crowding problem is not that severe and the potential $\delta\psi$ is modified by a numerical factor less than 2.

To see this we consider the space-charge limited current for a problem of cylindrical symmetry. Suppose that the current results from carriers drifting with saturated velocity from the wall of a cylinder of length Z and radius y_0 to a central electrode of negligible radius. In cylindrical coordinates the Poisson equation is given by

$$\frac{1}{r}\frac{\partial}{\partial r}(rE) = I/2\pi r\epsilon v_s Z. \tag{12}$$

Integrating this equation we find the field $E(r)$ in the form

$$E(r) = I/2\pi\epsilon v_s Z + \frac{\text{const}}{r}. \tag{13}$$

The second term in (13) is due to the charge on the central wire and corresponds to the field (with zero divergence) considered in Section 2. The first term describes the field due to the space charge. This field is finite everywhere. For the cylindrical problem the space charge field is constant because of the velocity saturation. Thus, we find that the screening potential due to the space charge is given by

$$\delta\psi = \frac{Iy_0}{2\pi\epsilon v_s Z} = \frac{Iy_0^2}{\epsilon S v_s} \tag{14}$$

where S is the surface area of the cylinder. From eqns (11) and (14) we see that for the same current density per unit area of the cathode the current crowding near a thread-like anode increases the screening potential by a factor of 2 as compared to the planar geometry. Thus, for a specific geometry of the electrodes we can describe the screening effect by a potential of the form (11) but with a numerical factor $1 < \xi < 2$ whose exact evaluation requires numerical computation.

4. CALCULATION OF THE IV CHARACTERISTICS

The current-voltage characteristics of the device interpolated between the low current and the high current regimes can be obtained by replacing $V_A \to \tilde{V}_A = V_A - \delta\psi$ in the formulas of Section 2 (including expression (6) for y_0). It is convenient to express both I and V_A through a parameter $\tilde{V} = A\tilde{V}_A + GV_G$. This can be done by the following procedure.

For a given V_G and assuming a value of $\tilde{V}$ we define $\Delta\tilde{V}$ by

$$\Delta\tilde{V} = (\tilde{V} - V_G)/A \tag{15}$$

and calculate a self-consistent value for the saddle point $\tilde{y}_0$, viz.

$$\tilde{y}_0 = \lambda \ln|4L\Delta\tilde{V}\sin\pi A/\tilde{V}d|. \tag{16}$$

The current is then given by

$$\frac{I}{A^*T^2S} = \left(\frac{\beta\tilde{V}d}{L\sin\pi A}\right)^{-1/2} e^{\beta\tilde{V}[1-\tilde{y}_0/L+\lambda/L\sin\pi A]} \tag{17}$$

and the screening potential by

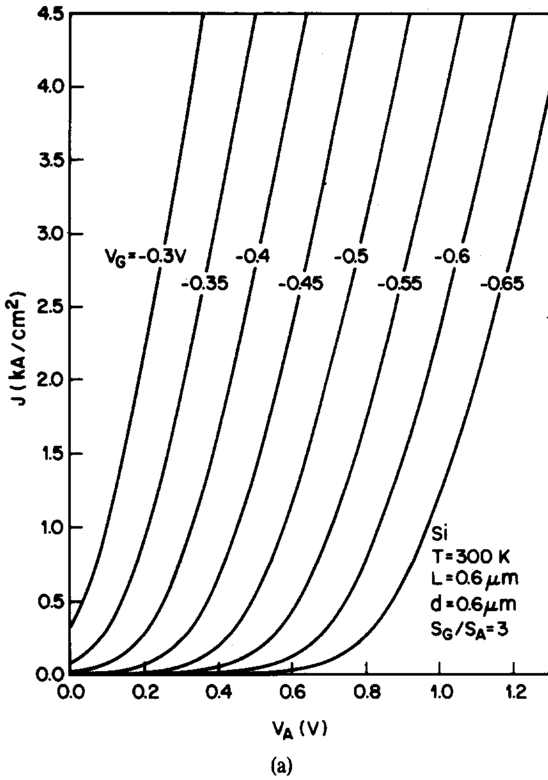
$$\delta\psi = \xi I\tilde{y}_0/2\epsilon v_s S. \tag{18}$$

Finally, the required anode voltage is calculated, viz.

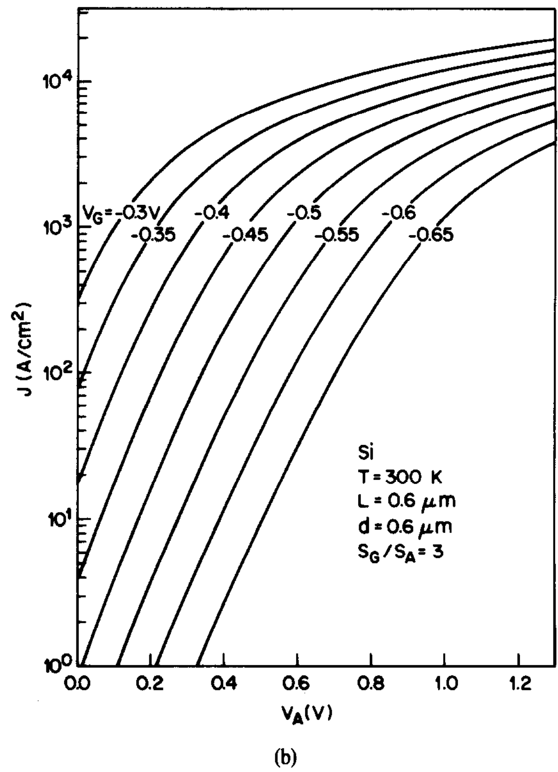
$$V_A = V_G + \Delta\tilde{V} + \delta\psi. \tag{19}$$

Equations (15)–(19) determine the IV characteristics of the TET in a parametric form. These characteristics evaluated for exemplary parameters of the device are represented in Figs. 4 and 5. As seen from these figures, at small currents the dependence on voltage is nearly exponential which corresponds to the thermoionic current. At higher currents the exponential dependence saturates and goes over into a linear law characteristic of a space-charge limited current. It is also seen that curves corresponding to different gate voltages are shifted with respect to each other along the V_A axis. Moreover, the magnitude of this shift is greater than the corresponding difference in V_G. The ratio of shifts ΔV_A and ΔV_G which combined leave the current invariant, determines the voltage gain, K. Recall that in the charge-sheet version of the TET [1] the voltage gain was directly related to the area ratio, $K = S_G/S_A$, whereas in the present version the gain is somewhat smaller than that given by K.

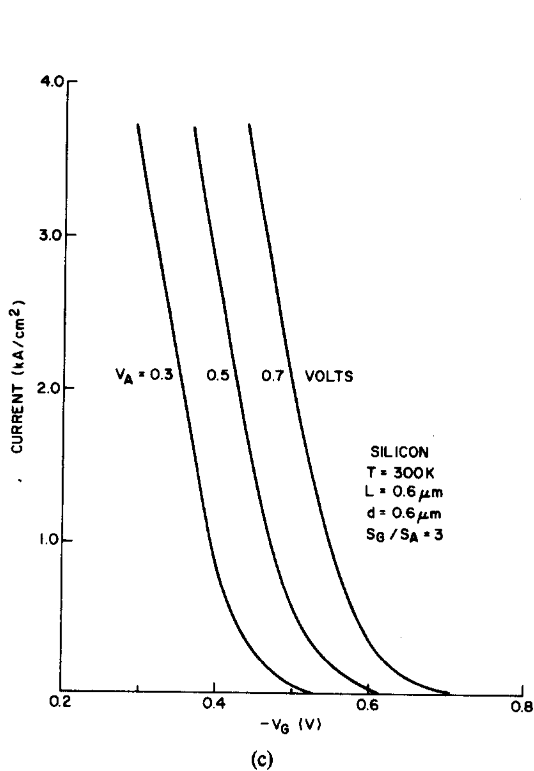
The transconductance g_m depends on both applied voltages V_A and V_G. In the charge injection regime g_m is proportional to the current because of the exponential IV characteristic. For high current the transconductance saturates at a maximum value $g_{m\,sat}$. One can evaluate $g_m(V_A, V_G)$ by a procedure similar to eqns (15)–(19)



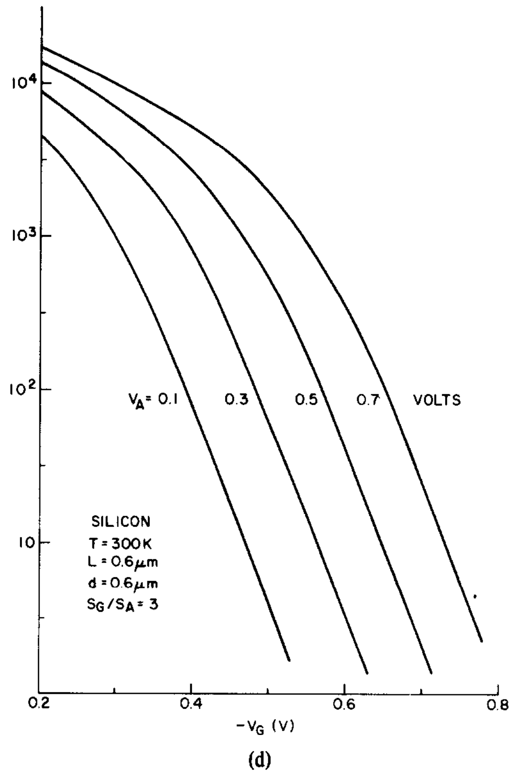
(a)



(b)

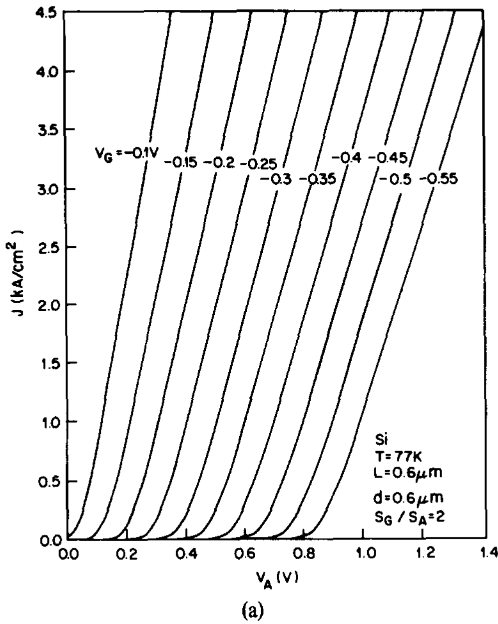


(c)

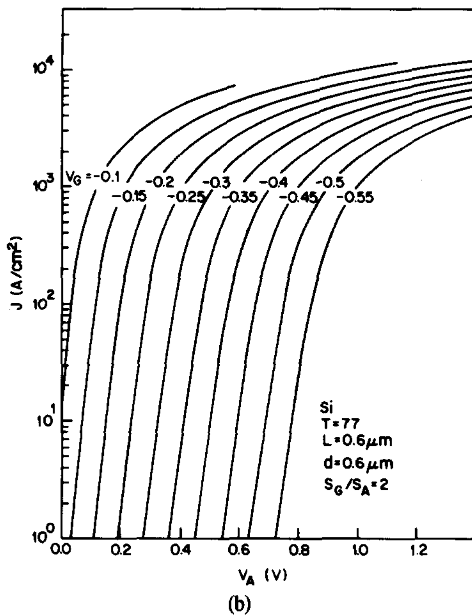


(d)

Fig. 4. A family of *IV* characteristics calculated for a silicon device at room temperature. (a, b) Current vs anode voltage for different V_G. (c, d) Current vs gate voltage for different V_A.



(a)



(b)

Fig. 5. A family of IV characteristics for silicon at $T = 77K$.

above. However, for our purpose it is sufficient to use the following estimate of g_{msat}:

$$g_{msat} = \frac{S_G}{S_A} \frac{2\epsilon v_s S}{\xi y_0^2} \tag{20}$$

which can be derived by using the fact that y_0 is a slowly varying function of the applied voltages. By the same token, the slope of the I vs V_A curves in the high current limit depends only weakly on temperature or V_G since these affect the space charge limited current only through y_0.

As a numerical example, let us consider a device having the total active area $S = 10\ \mu m \times 5\ \mu m$, and the remaining parameters as in Fig. 4, viz. $S_G/S_A = 3$, and $L = d = 0.6\ \mu$. At room temperature, for the anode voltage $V_A = 0.5$ V and the electrostatic potential on the gate $V_G = V_{GS} + V_{bi} = -0.6\ V$ (where V_{bi} is the gate-to-body built-in voltage, i.e. the difference between the work functions of the gate material and the silicon substrate, and V_{GS} is the voltage applied to the gate), the output current equals $I = 20\ \mu A$. When V_G increases 0.3 V the current increases by two orders of magnitude to $I = 2\ mA$, which is comparable to a typical current in the "on" state of a field-effect transistor. Regarding $I = 20\ \mu A$ as an "off" state of the TET we demand that in this state the input voltage should be equal to zero, $V_{GS} = 0$. This implies $V_{bi} = -0.6$ V, which is a typical built-in voltage for a number of silicides, notably $TaSi_2$, on n-type silicon. These materials are especially convenient since they produce approximately equal built-in voltages on silicon substrates of both complementary polarities.

5. PAIR OF COMPLEMENTARY TET DEVICES AS A LOGIC ELEMENT.

In this Section we consider an inverter circuit based on two TET devices with complementary types of conductivity. The circuit is quite analogous to the well-known CMOS FET inverter, see equivalent circuit in Fig. 6. It is important to realise that no bipolar effects should be expected in a TET-based circuit since the device can operate at a supply voltage as low as half the energy gap.

Figure 7 shows schematically the lay-out of one inverter gate. Let us describe the structure in more detail. We start from a p-silicon substrate which serves as the source for the p-type transistor and is at the supply voltage level ($VDD = 0.5\ V$). The source (cathode) of the n-transistor is provided by a deep diffusion of donors in the substrate, and is at the ground (GND) voltage level. An epitaxial intrinsic layer of thickness L serves as a base

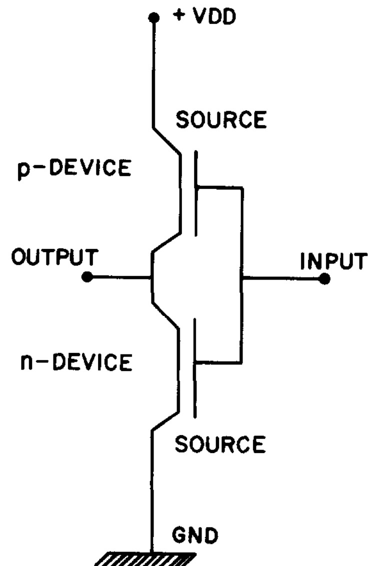


Fig. 6. Equivalent circuit of a complementary inverter.

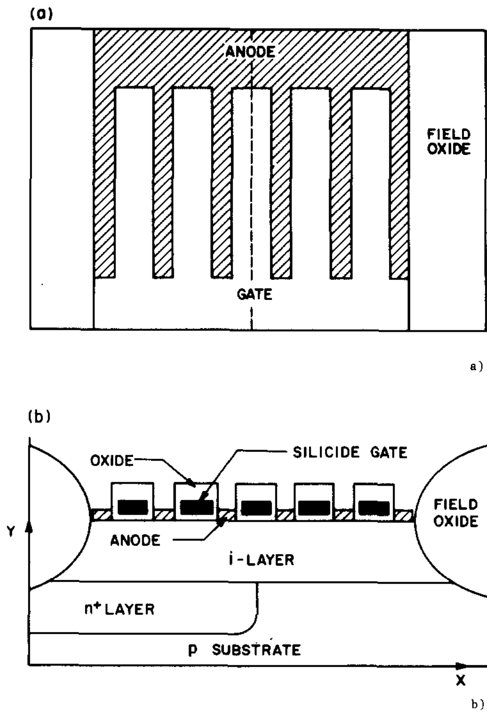


Fig. 7. Schematic layout of the TET inverter. (a) Top view. (b) Cross-section.

for both transistors of the inverter pair. The gate and the anode terminals are arranged as an interdigitated pattern of electrodes on top of the intrinsic layer. The gate electrodes represent an MOS structure (silicide–gate oxide–silicon) common for both transistors. The anode electrodes are in contact with the top of the intrinsic layer. The contact is made ohmic for electrons in the n-type transistor and holes in the p-type device (e.g. by n^+ and p^+ polysilicon drain structures). The nodes are connected and their common potential is the output voltage of the inverter.

The power supply lines to the circuit run entirely underneath the base intrinsic layer with the voltage applied between the n-diffusion region and the p-substrate. The sources of the two transistors are biased by VDD with respect to one another and form a *forward*-biased pn^+ junction. At $VDD = 0.5\ V$ the power dissipation associated with the forward current is negligible even at room temperature. For example, for a diode with $N_D = 3\times10^{20}\,\mathrm{cm}^{-3}$ and $N_A \sim 5\times10^{16}\,\mathrm{cm}^{-3}$ the experimental value of the forward current density at 0.5 V is about $10\,\mathrm{mA/cm}^2$. The only negative consequence of this current is that it draws on the supply battery.

Electrically the circuit represents two variable resistances in series which divide the *VDD* to *GND* voltage depending on the potential on the gate. If an appropriate silicide (e.g. $TaSi_2$) is used for the gate metal then the two devices may be regarded as symmetrical.† Both transistors are "normally off", i.e. either of them is in the "off" state when the gate is at zero voltage with respect to its source. Thus the n-transistor is "off" and the p-transistor is "on" when $V_G = GND$. In this state $V_{out} = VDD$. Conversely, when $V_G = VDD$ then the p-transistor is "off" and $V_{out} = GND$.

The transfer characteristic (V_{out} vs V_{in}) of a TET inverter can be obtained by a graphical construct[2]. As seen from Fig. 6 the anode voltage (source to drain) for the n-transistor equals the output voltage, $V_A^{(n)} = V_{out}$, whereas for the p-transistor $V_A^{(p)} = VDD - V_{out}$. Since the currents in both transistors are equal we can superimpose the plots $I_n(V_{out})$ and $I_p(V_{out})$ as shown in Fig. 8 and read off the transfer characteristic from the intersections of curves corresponding to same gate voltage. Figure 9 shows the characteristic obtained in this way for an inverter composed of an n-transistor with the parameters similar to those used in Fig. 4 and a symmetric complementary device.

It may be instructive to give an approximate analytic derivation of the transfer characteristic. Note that all points of the curve in Fig. 9 correspond to the low current regime, which physically means that the inverter in its steady state draws only a small current. In this regime the current-voltage characteristics are nearly exponential and are well-described by dependencies of the form

$$I = I_n\, e^{\beta(\alpha V_A + \gamma V_G)} \tag{21}$$

where I_n is a constant and α and γ are the inverse ideality factors of the anode and the gate characteristics, respectively. As can be ascertained from Figs. 4 and 5, the ratio γ/α is somewhat smaller than the corresponding geometric ratio, S_G/S_A, e.g. $\gamma/\alpha = 2.5$ for $S_G/S_A = 3$. For the p-type transistor the dependence analogous to (21) is

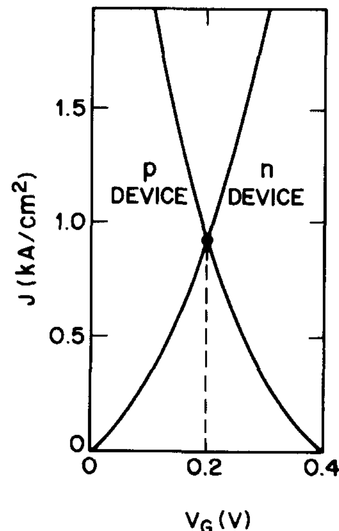


Fig. 8. Graphical superposition of *IV* characteristics for symmetric complementary devices. Device parameters as in Fig. 4 with supply voltage chosen at 0.4 *V*.

†Asymmetry is introduced to a small degree by differences in v_s and A^* for electrons and holes, and to a greater degree by different built-in barriers V_{bi}. The latter can be further adjusted by a judicious choice of the doping levels N_D and N_A or by ion implantation at interface. To this end it may be advantageous to use n-substrate and p^+ diffusion rather than p-substrate and n^+ diffusion.

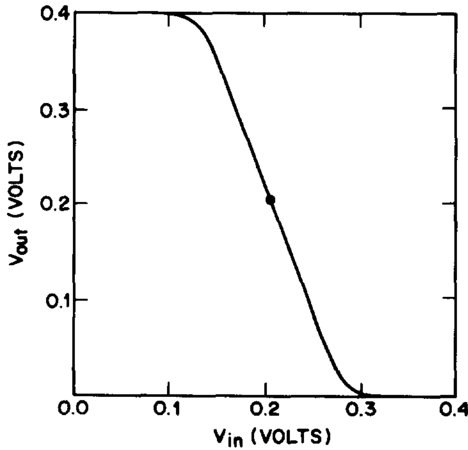


Fig. 9. Inverter transfer characteristic for a numerical example of a symmetric inverter. Device parameters as in Figs. 4 and 8.

obtained by shifting the source voltage, cf. the equivalent circuit in Fig. 6,

$$I = I_p\, e^{\beta[\alpha(VDD-V_A)+\gamma(VDD-V_G)]}. \tag{22}$$

For a symmetric pair of devices one has $I_n = I_p$. Equating the currents we find in this case

$$\alpha V_A + \gamma V_G = (\alpha + \gamma)\, VDD/2. \tag{23}$$

Equation (23) correctly describes the central part of the transfer characteristic including its slope γ/α which, thus, equals the voltage gain K of a single device at low currents. It does not describe the flat portion of the transfer curve. Indeed, for $V_A \lesssim kT/q$ eqn (21) is invalid because it neglects the reverse diode current. Similarly, eqn (22) is invalid when $VDD - V_A \lesssim kT/q$. For a non-symmetric case, $I_p \neq I_n$, the transfer curve will be shifted by the amount

$$\frac{kT}{q}\ln(I_p/I_n) = V_{bi}^{(n)} - V_{bi}^{(p)} \tag{24}$$

where $V_{bi}^{(n)}$ and $V_{bi}^{(p)}$ are the built-in voltages for the n- and the p-transistors, respectively.

Let us now estimate the inverter delay time τ_{in}, given by

$$\tau_{in} = C/g_m \tag{25}$$

where C is the total input capacitance of one inverter stage and g_m is the transconductance of the driving device in its "on" state. The capacitance C equals $2C_1$, where C_1 is the input capacitance of a single transistor. As shown in [1], in the high current limit C_1 is of the form

$$C_1 = \frac{\bar{y}_0}{v_s} g_m + \frac{\epsilon S}{\pi d}\ln ctg\frac{\pi a}{2d} \tag{26}$$

where a is the length of the gap between the gate and anode electrodes. The first term in (26) is associated with the accumulation of mobile charge on the downhill slope in the base and represents the useful gate-to-channel capacitance. The second term represents the gate-to-anode parasitic capacitance and was derived in [1] in a model assuming thin electrodes separated by a small gap. Because of the Miller effect this term enters the delay time with an extra factor of 2. From eqns (25) and (26) and taking the expression (20) for g_m in the high current limit we find

$$\tau_{in} = 2\frac{\bar{y}_0}{v_s}\left[1 + \frac{\xi\bar{y}_0}{\pi d}\frac{S_A}{S_G}\ln ctg\frac{\pi a}{2d}\right]. \tag{27}$$

The first term in (27) corresponds to the time of flight of carriers on the downhill slope of the potential barrier, see Fig. 3. As discussed above, the length of the downhill slope, $\bar{y}_0$ is a slowly varying function of the applied voltages, see eqn (16). According to our numerical calculations in the high current limit the quantity $\bar{y}_0$ stabilizes at approximately $\bar{y}_0 = d/3$. The second term in (27) describes the delay associated with charging the parasitic gate-to-anode capacitance. For $d/a \approx 10$ this delay is typically less than 30% of that described by the first term. Let us note, however, that in reality this parasitic capacitance may prove to be somewhat higher because of the finite thickness of the electrodes neglected in our calculation.

To summarize, for an inverter with the period of surface electrodes $d \sim 0.6\ \mu$m the expected delay is of order $\tau_{in} = 5$ psec. It is important to realize that in a real integrated circuit the speed of operation of the TET-based inverter can be expected to approach by order of magnitude the above "intrinsic" gate delay. Indeed, the current flowing in this inverter during switching is of the order of that in a CMOS FET inverter while the switching voltage is an order of magnitude lower. Accordingly reduced will be the charge associated with all parasitic capacitances such as wiring, interconnect, etc. and the corresponding parasitic delay times. In other words, this means that for reasonable dimensions of the TET, say $S = 10\ \mu\text{m} \times 10\ \mu\text{m}$ as in the above example, the inverter capacitance C will be of the same order as the total parasitic capacitance.

6. CONCLUSIONS

(1) In this paper we developed a theory of the Thermionic Emission Transistor, simplifying the structure proposed earlier [1]. It is shown that in a TET device with a simple base (intrinsic layer sandwiched between a heavily doped substrate and a surface set of interdigitated electrodes) one can achieve a current control which is nearly as efficient as in the original more complicated version of the device containing a planar-doped triangular barrier. To show this we considered the two-dimensional electrostatic potential and the thermionic current at low current levels as well as the space charge limited current in the high injection regime.

(2) On the basis of this theory we calculated the characteristics of the TET structures for several exemplary sets of parameters. It was found that the output

current depends on a linear combination of the applied voltages V_A and V_G, viz. $\alpha V_A + \gamma V_G$. The voltage gain, $K = \gamma/\alpha$, is mainly determined by the area ratio of the surface electrodes, S_G/S_A. We found $\gamma/\alpha \sim 2.5-2.7$ for $S_G/S_A = 3$, and $\gamma/\alpha \sim 1.7$ for $S_G/S_A = 2$. The intrinsic response time of a single device is mainly determined by the time of flight of carriers with saturated velocity across the drift portion of the base. The extent of the drift region equals approximately one third of the period d of the surface electrodes, so that for $d = 0.6\,\mu$m the time of flight is about 2 psec.

(3) The intrinsic layer can serve as a base for both n- and p-type TET devices. This opens an attractive possibility of using these devices in complementary logic circuits, analogous to CMOS. Expecially attractive is the fact that the low supply voltage required by a TET circuit eliminates all parasitic bipolar (latch-up) effects. We calculated the gate delay τ_{in} of a basic complementary inverter and plotted its transfer characteristic for a numerical example. The current which flows in an inverter of total area $10\,\mu\text{m} \times 10\,\mu\text{m}$ during switching is of order $3\,mA$ which is comparable to CMOS current. However, the required switching voltage is about an order of magnitude smaller, viz. $\sim 0.3\,V$. Accordingly reduced will be the charge associated with all parasitic capacitances and the time required to charge them. For numerical example of a device with $d = 0.6\,\mu$m the estimated total inverter delay is $\tau_{in} = 5$ picoseconds.

Acknowledgements—We are grateful to G. E. Smith and S. M. Sze for helpful discussions, and J. E. Iwersen for comments on our earlier work.

REFERENCES

1. R. F. Kazarinov and S. Luryi, *Appl. Phys.* **A28** (1982).
2. C. Mead and L. Conway, *Introduction to VLSI Systems.* Addison-Wesley, Reading, Mass. (1980).
3. R. J. Malik, K. Board, C. E. C. Wood, L. F. Eastman, T. R. AuCoin and R. L. Ross, *Electron. Lett.* **16**, 837 (1980).
4. R. F. Kazarinov and S. Luryi, *Appl. Phys. Lett.* **38**, 810 (1981).
5. C. O. Bozler and G. D. Alley, *IEEE Trans. Electron Dev.* **ED-27**, 1128 (1980).
6. J. I. Nishizawa, T. Terasaki and J. Shibata, *IEEE Trans. Electron Dev.* **ED-22**, 185 (1975); T. Ohmi, *IEEE Trans. Electron Dev.* **ED-27**, 536 (1980).
7. W. Shockley, *Proc. IRE* **40**, 1289 (1952).
8. W. Shockley, *Proc. IRE* **40**, 1365 (1952).

APPENDIX

Electrostatic field and potential distribution in the base of the TET at low injection levels.

Our electrostatic model of the TET base at low injection (i.e., when the injected charge is negligible compared to charges on the electrodes and the substrate) is described by the boundary conditions $\psi(x, L) = 0$ and $\psi(x, 0) = V_A$ if $-Ad/2 < x + md < Ad/2$ for some $m = 0, 1, 2, \ldots$, and $\psi(x, 0) = V_G$ otherwise. For arbitrary ratio of geometric dimensions L/d the solution to Laplace's equation for these boundary conditions is given by the infinite series (2), or

$$\psi(x, y) = \bar{V}\frac{L-y}{L} + \Delta V f(x, y) \qquad \text{(A1)}$$

where

$$f(x, y) = \frac{2}{\pi}\sum_m \frac{1}{m}\sin m\pi A \cos mx/\lambda \frac{\sinh m(L-y)/\lambda}{\sinh mL/\lambda}. \qquad \text{(A2)}$$

Equation (A1) corresponds to the origin $x = 0$ chosen in the middle of an anode electrode. If, instead, one wishes to choose the origin in the middle of a gate elctrode, one has to replace $f(x, y)$ by $g(x, y) = f(x + d/2, y)$, or

$$g(x, y) = -\frac{2}{\pi}\sum_m \frac{1}{m}\sin m\pi G \cos mx/\lambda \frac{\sinh m(L-y)/\lambda}{\sinh mL/\lambda} \qquad \text{(A3)}$$

We are interested in such geometries when $L \gg \lambda$, say $L \sim d$, and in this case the series (A2) admits explicit summation. Physically, such simplification becomes possible because the solution is no longer sensitive to the exact boundary condition on the potential at $y = L$, and is determined instead by an asymptotic *field* condition $\mathbf{E}(x, y) \to (\bar{V}/L)\hat{\mathbf{y}}$. Consider the field in the y-direction,

$$E_y(x, y) = \bar{V}/L + \Delta V f'(x, y) \qquad \text{(A4)}$$

where

$$f'(x, y) = \frac{2}{\pi\lambda}\sum_m \sin m\pi A \cos(mx/\lambda)\frac{\cosh m(L-y)/\lambda}{\sinh mL/\lambda} \qquad \text{(A5)}$$

For $L \gg \lambda$ one has

$$f'(x, y) = \frac{4}{d}\sum_m \sin m\pi A \cos(mx/\lambda) e^{-my/\lambda} \qquad \text{(A6)}$$

which is a geometric progression. After some algebraic manipulations we find

$$f'(x, y) = \frac{1}{d}\left[\frac{\sin(\pi A + x/\lambda)}{\cosh(y/\lambda) - \cos(\pi A + x/\lambda)} + \frac{\sin(\pi A - x/\lambda)}{\cosh(y/\lambda) - \cos(\pi A - x/\lambda)}\right]. \qquad \text{(A7)}$$

In particular, on the line under the middle of an anode electrode the field is given by

$$E_y(x_0, y) = \frac{\bar{V}}{L} + \frac{2\Delta V}{d}\frac{\sin\pi A}{\cosh(y/\lambda) - \cos\pi A} \qquad \text{(A8)}$$

whence we obtain eqn (4) for the location of the saddle point.

We can now calculate the potential $\psi(x, y)$ everywhere in the device by integrating the field E_y given by eqns (A4) and (A7) from the substrate, $y = L$, where $\psi(x, L) = 0$, to a point (x, y) along a straight vertical line, viz.

$$\psi(x, y) = \int_y^L E_y(x, y)\,dy. \qquad \text{(A9)}$$

Integrating, we find an expression of the form (A1) with

$$\pi f(x, y) = \left[\arctan\left(\tanh\frac{y}{2\lambda}\tan\frac{\pi G + x/\lambda}{2}\right) + \arctan\left(\tanh\frac{y}{2\lambda}\tan\frac{\pi G - x/\lambda}{2}\right)\right]_y^L. \qquad \text{(A10)}$$

This expression was used to plot the equipotential lines in Fig. 2 and the stereometric projection of the potential surface $\psi(x, y)$ in Fig. 3.