Solid-State Electronics, Vol. 24, pp. 343–366, 1981
Printed in Great Britain.

0038-1101/81/040343-24$02.00/0
Pergamon Press Ltd.

TUNNELING HOT ELECTRON TRANSFER AMPLIFIERS (THETA): AMPLIFIERS OPERATING UP TO THE INFRARED

MORDEHAI HEIBLUM
IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, U.S.A.

(*Received* 16 *June* 1980; *in revised form* 20 *August* 1980)

Abstract—A family of novel three-terminal devices which relies on the transfer of a quasi-monoenergetic hot electron beam through a thin base is described. The devices are similar in principle to the proposed tunneling amplifier by Mead in the early sixties ("Cold Cathode" or "Metal Base" amplifiers). Results are reviewed and the probable reasons for the poor performances are pointed out. It is predicted that, with a proper choice of parameters, metal-base amplifiers can operate as switches, negative resistance devices and continuous amplifiers in the subpicosecond range.

Two subclasses are described: The tunneling emitter (THETA), in the major part of the work, and the nontunneling emitter (BHETA) amplifiers. In the THETA family the metal-oxide-metal-oxide-metal (MOMOM), the MOM-semiconductor (MOMS), and the heterojunctions devices are described. Members of the BHETA family generate quasi-monoenergetic electron beams by injecting electrons by an n^+n^- or a metal-n^- junctions, and include a variety of metals and semiconductor combinations.

Very thin films are required in these devices (oxides ~ 15 Å, metals ~ 100 Å, semiconductors ~ 100 Å). The molecular beam epitaxy technique and lattice matching considerations are required for pinhole free semiconductors and metal films with minimum interface states. Sputter-oxidation methods are needed for thin oxide growth. Systems which combine these features with availability of microfabrication make these devices feasible today.

1. INTRODUCTION

The tremendous reduction in computer size and increase in its speed creates an ongoing effort in search for faster switches. Microwave devices are available up to few tens of GHz, but leave the spectrum of frequencies up to the near i.r.—where lasers are more common—vacant. Lasers operating in the submillimeter regime are scarce, have very low efficiencies, are combersome and not tunable. There are areas of switching, communication, and data processing which are limited by device capabilities, and those in turn have an upper frequency limit which does not exceed 50 GHz.

Novel attempts are now being explored to increase the speed of switching and CW systems employing Josephson devices[1] and field-effect transistors (with submicron gate width)[2] while in the i.r. regime tunable free-electron sources are being tested[3]. Proposals for a very high frequency traveling wave tubes[4], solid state traveling wave devices[5] and semiconductor superlattice devices[6] are being considered.

The success of the point-contact metal-oxide-metal (MOM) junction in the past ten years as a classical mixer and detector at near i.r. and visible frequencies (for a list of references see Ref. [7], triggered Schwarz[8] to propose an MOMOM device as an amplifier in the 10 μm regime. Schwarz revived a 1960 proposal by Mead[9, 10], which came at the time when alternative devices for the bipolar transistor were being sought. In this particular configuration, a biased MOM emits—via tunneling—a fairly monoenergetic electron beam which traverses through a thin metal electrode and surmounts a potential barrier of a second MOM, to be collected by a collector metal electrode. Experimentally Mead found a very small transfer ratio between collector and emitter currents[10]. Other researchers had tried to improve the device performance, including an alteration where the tunneling electrons were ejected into vacuum and collected by an anode (MOMVM), but with limited success (see Section 2.3 and Ref. [7] for references).

A variation of Mead's proposal was fabricated and tested by Spratt *et al.*[11] where the second MOM or the vacuum-anode combination were replaced by a semiconductor (MOMS). Their experiments failed due to pin-holes in the center metal electrode (as was shown by Hall[12] and Lavine and Iannini[13]). More effort to improve the MOMS over the years had met with marginal success, and for a summary and a review of those hot electron transfer devices one should refer to Refs. [14, 15].

The frequency performance of hot-electron devices was estimated by Moll[16] and Atalla and Soshea[17]. Both came to the conclusion that tunnel-emitters are inferior to the bipolar transistor. As will be shown later, their conclusions result from a particular choice of parameters for the tunnel emitter, and the restriction of the current density to 10^3 A/cm^2. A different choice of parameters proves the superiority of the MOM, and the present operation of MOM devices in the near i.r. frequency range is an experimental proof (utilizing the "point-contact" and the "edge" configurations[7]).

In this paper, I propose a novel class of active devices based on hot electron transfer mechanisms which combine the basic ideas of Mead[9] and Spratt *et al.*[11], with the recent success of MOM devices[7]. I elaborate

on the basic pitfalls which prevented a successful operation of these devices in the past and point out what has to be done to make them useful devices.

2. TUNNELING HOT ELECTRON TRANSFER AMPLIFIERS (THETA)

Unlike the bipolar or the field-effect transistors which are "thermal devices", hot electron devices generate hot electron beams by some kind of an "emitter" (via tunneling, for example). The electrons traverse a thin "base" and are collected by a "collector", before thermalizing to the lattice temperature. If the linear dimensions are small, it can be shown that the RC time constants and the transient time can both be remarkably small.

In the next sections I will describe a general tunneling hot electron transfer amplifier (THETA), its expected characteristics, and limiting mechanisms for ideal performance. Early devices and their limitations, followed by proposed changes and novel devices will then be described.

2.1 *Operation principles and characteristics*

Consider hot electron generation via tunneling. Figure 1 describes a general THETA configuration. Since the linear dimensions along x are much smaller than the other dimensions, a one-dimensional model will be fit for the analysis.

The base-collector barrier is designed so that upon biasing (in the polarity shown) a negligible amount of current will flow between base and collector. The barrier between emitter and base is thin enough to allow tunneling. These basic features are presented in the energy diagram of Fig. 1(b). Metals are identified by the Fermi energy (E_F) and oxides or semiconductors by the bottom edge of the conduction band (E_C).

Upon base-emitter biasing, electrons will tunnel from emitter to base with energies near E_F of the emitter (their density decays for energies above E_F and their tunneling probability drops sharply when they are further removed from the top of the potential barrier). In the base they can suffer electron–phonon (e–ph), electron–electron (e–e) and electron–impurity (or defect) collisions, after a characteristic length l_e (which will be noted as the mean free path). Assuming a one-dimensional model, if $d_B < l_e$ and $\Phi_C < E_x \leq eV_{BE}$, where E_x is the normal energy (kinetic energy for motion in the x direction), a substantial fraction of I_E will be collected by the collector, resulting in $I_C = -\alpha I_E$, where α is the transfer ratio (using the convention of $I > 0$ when it flows into the port). Applying V_{CB} in the positive direction (C positive with respect to B) will not affect I_C to the first order, since the quantum mechanical reflection from C-B interface are not strongly dependent on the shape of the potential barrier as long as $E_x > \Phi_C$. Hence, I_C will be constant with any load R_L as long as $\Phi_C/e - V_{CB} \lesssim V_{BE}$. The power gain G is $\alpha^2 R_L/r_{in}$ (as long as $r_{out} \gg R_L$, where r_{in} and r_{out} are the differential slopes of the input and output characteristics), and for $\alpha^2 > r_{in}/R_L$, $G > 1$ (G is limited by $\alpha^2 r_{out}/r_{in}$). If we assume that $\alpha = 0.5 \ldots 1$, the input and output characteristics can be predicted as shown in Fig. 2. Figure 2(a) describes the output characteristics. When $V_{BE} < \Phi_C/e$ (and $V_{CB} > 0$),

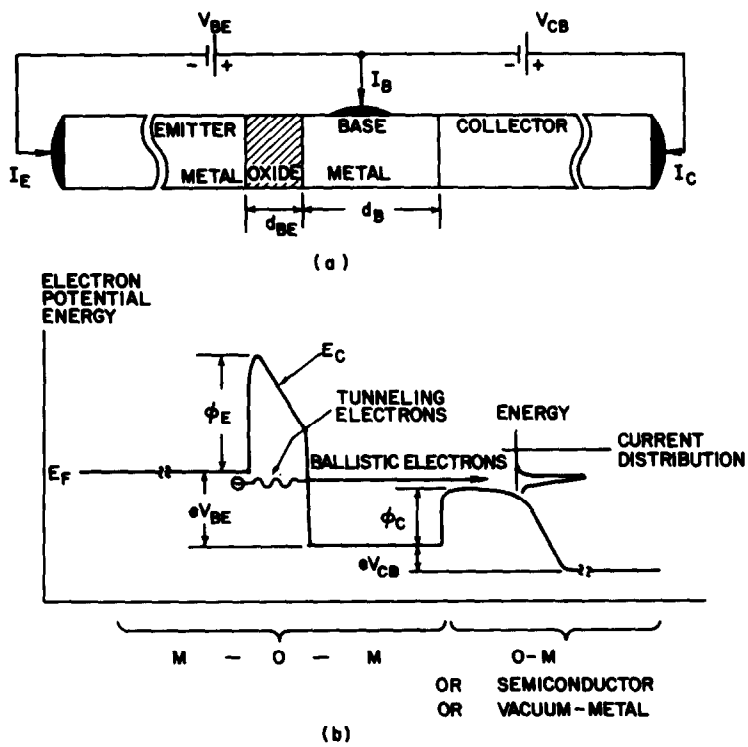

Fig. 1. A general THETA device. (a) Schematic description of an MOM tunneling emitter and a general collector, (b) Generalized energy diagram for MOMOM, MOMS and MOMVM. The horizontal lines are the Fermi levels in emitter, base and collector.

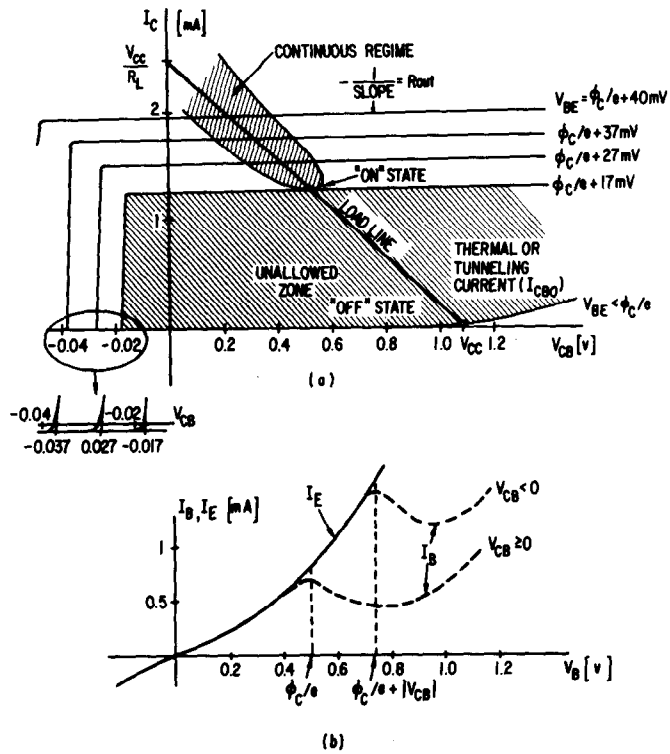

Fig. 2. (a) An idealized output characteristics of a THETA device. In reality the unallowed zone and Rout depend strongly on the monochromacity of the ejected electron beam. Sample values of V_{BE} are presented (assuming $E_x \cong eV_{BE}$). (b) I_B vs V_{BE} exhibits a differential negative resistance with an on-set determined by V_{CB}.

the stream of hot electrons will bounce back from the base-collector interface (the tunneling probability through the collector barrier is $10^{-6} \ldots 10^{-3}$), and thermalize in the base film resulting in $-I_E = I_B$. Biasing the collector positively ($V_{CB} > 0$) will leave the barrier height Φ_C unaffected (neglecting the image correction), and the observed collector current will result only from the tunneling electrons from the base (named I_{CBO}). I_{CBO} can be limited by a proper choice of collector parameters. Note that in this argument, quantum mechanical resonance effects due to multiple reflections from the boundaries are neglected [18, 19].

If the energy spread of the electron current is small, the collector current will suddenly rise when V_{BE} exceeds Φ_C/e (due to the finite spread in energy, the collector current will rise when $eV_{BE} \cong \Phi_C - \delta$, where δ is half the normal energy spread of the electron current). This effect will manifest itself in the output characteristics by an "unallowed zone". Consequently, if the device is connected into a circuit in a "common base configuration" (as shown in Fig. 1(a)), with a load R_L and a supply V_{CC}, a change in the input voltage of $2\delta/e$ will change V_{CB} from close to V_{CC} ("OFF" state in Fig. 2(a)) to $V_{CC} - \alpha|I_E|R_L$, which can be very small and even negative ("ON" state). The input power which is required for "switching" the output can be small (directly proportional to δ^2 which is temperature and parameter sensitive), and the two distinct states of V_{CB} (on both sides of the "unallowed zone"), will qualify the amplifier as a fast switching element with possible time response down to the subpicosecond region.

If one monitors the input characteristics in a "common emitter configuration" (I_B vs V_{BE}), one should expect to observe a similar behavior to the one displayed in Fig. 2(b). If $V_{CB} = 0$; $I_B = I_E$ as long as $V_{BE} < (\Phi_C - \delta)/e$. Over the transition region $\Phi_C - \delta \leq eV_{BE} \leq \Phi_C + \delta$, α will increase from $\sim 10^{-4}$ to ~ 0.5 and I_B $[= (1-\alpha)I_E]$ will drop. A further increase in V_{BE} will decrease α due to a shorter mean free path of the hot electrons in the base and the increase in the energy spread of the tunneling electrons. Hence, in the vicinity of $V_{BE} > (\Phi_C + \delta)/e$, the input characteristics will exhibit differential negative resistance features. If $V_{CB} < 0$, $\Phi_C \cong \Phi_C(0) + e|V_{CB}|$ (for an MOM collector), and the threshold voltage for the differential negative resistance region will increase, as shown in Fig. 2(b). This feature enables us to control the on-set of the differential negative resistance region, or to switch in and out from this region by an independent electrode. At this mode of operation the device can be used as a two-terminal device for the generation of oscillations or amplification.

Attempts to operate similar devices in the sixties failed, mostly due to poor choices of parameters and difficulties in fabrication procedures.

In the next section, I elaborate in some detail on the basic loss mechanisms occuring in current transport in hot electron transfer devices. Understanding them will help us understand proposals for variations and new device configurations.

2.2 *Loss mechanisms*

Figure 3 describes an energy diagram of an MOMOM device. I consider the basic loss mechanisms in that structure, using the dimensions and parameters which

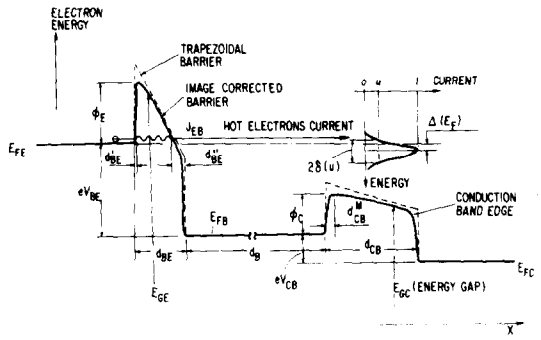

Fig. 3. Energy diagram and current distribution in MOMOM.

researchers have been using up to date. Modifications of these losses applicable in other hot electron devices will be discussed in conjunction with the particular devices.

The transfer ratio α will be equal to unity if the tunneling current distribution is finite in width and scattering and reflections can be avoided. In real structures we will encounter the following phenomena which contribute to a reduction in α: (a) Quantum mechanical energy spread of the tunneling electron current. (b) Energy spread of the tunneling current due to electronic collisions in the conduction band of base-emitter oxide. (c) Energy loss and isotropic randomization in the base metal. (d) Quantum mechanical reflection from the collector oxide-base interface. (e) Transmission through the conduction band of the collector oxide. (f) Quantum mechanical reflections from the collector oxide-metal boundary.

I now discuss in some detail each of the basic loss mechanisms.

(a) *Quantum mechanical energy spread of the tunneling electron current.* The current distribution of electrons which tunnel from one metal into the other through an insulator is quasi-monoenergetic with a peak distribution near E_F. This is a direct result of the diminishing electronic density above the Fermi level (at low enough temperatures) and the reduction in the tunneling probability of the lower energy electrons. The collected current density is $\alpha^* J_{EB}$, where α^* is the internal transfer ratio and J_{EB} is the current density resulting from the electrons flux which traverse from emitter to base. Since the emitter current density J_E is a combination of two components $|J_E| = |J_{EB}| - |J_{BE}|$; one gets $|J_E| < |J_{EB}|$ *always* and in principle, α can *exceed unity* ($\alpha > \alpha^*$). I'll elaborate on this point in Section 3.2.

Most previous works on MOMOM devices involved a symmetrical structure with $\Phi_E = \Phi_C$ and $V_{BE} \geq \Phi_C/e = \Phi_E/e$ which put the base-emitter junction into operation in the Fowler–Nordheim regime[20], as shown in Fig. 3 (in this regime the barrier is triangular and the current component J_{BE} can be neglected).

In the general case, when one considers α, the normal energy distribution of the current is of importance since the base-collector barrier is one dimensional and transmitted electrons must have $E_x > \Phi_C$ to be collected (rather than $eV_{BE} > \Phi_C$)[21]. The current distribution in energy space is given by the general expression[22]

$$J_{EB}(E_x, E_t)\,dE_x\,dE_t = \frac{4\pi m_m e}{h^3} D_E(E_x) f_E(E) \times [1 - f_B(E)]\,dE_x\,dE_t \qquad (1)$$

where $J_{EB}(E_x, E_t)$ is the current distribution in energy space due to electrons tunneling from emitter to base; E_x is the normal energy component (perpendicular to junction plane), E_t is the transverse energy component (in the junction plane), $D_E(E_x)$ is the tunneling transmission coefficient for current J_{EB}, $f_E(E)$ is the Fermi–Dirac occupation function in emitter, $f_B(E)$ is the Fermi–Dirac occupation function in base, m_m is the electron effective mass in the metal (is assumed the same in base and emitter), e is the electron charge.

Expression (1) is a direct consequence of the following assumptions: (a) metals are free-electron metals; (b) tunneling occurs between two parallel planes and through a one-dimensional potential barrier; (c) tunneling is elastic, namely, energy and transverse momentum are conserved in the tunneling process. In (1) and in the following calculations, $m_m = m_0$ where m_0 is the free electron mass and $f_B(E) = f_E(E + eV_{BE})$ will be assumed.

The current which results from electrons tunneling from base to emitter, J_{BE}, can be expressed as in (1), except by exchanging $f_E(E)$ and $f_B(E)$. The net current density of electrons traversing from emitter to base is

$$|J| = |J_{EB}| - |J_{BE}|. \qquad (2)$$

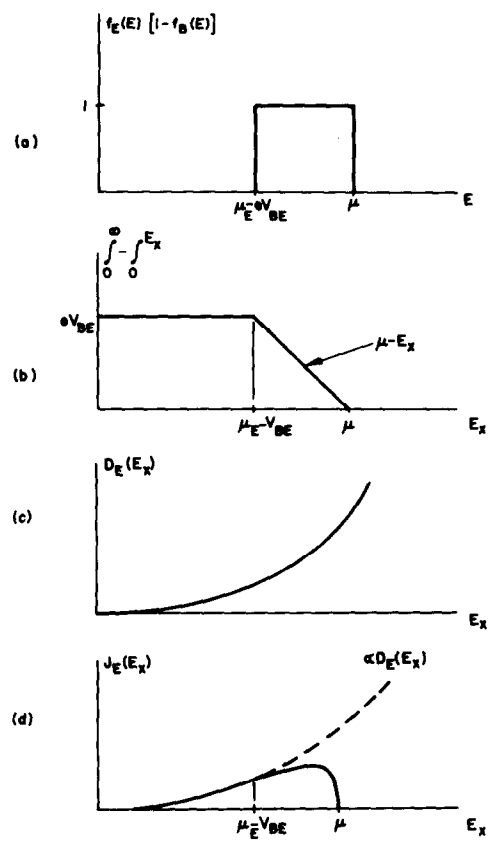

Fig. 4. Steps in evaluating $J_{EB}(E_x)$ at $T = 0°K$, is in eqn (5).

Neglecting the image correction and assuming a rectangular barrier, $D_E(E_x)$ can be calculated using the first order WKB approximation[23],

$$J_{EB}(E)\,dE = \frac{4\pi m_m e}{h^3} f_E(E) \times [1 - f_B(E)]\,dE \int_0^E D_E(E_x)\,dE_x. \tag{6}$$

$$D_E(E_x) = \begin{cases} \exp\left[-\frac{2}{3}\frac{2\sqrt{m_{BE}}}{\hbar} d_{BE} \frac{(\Phi_E - E_x)^{2/3} - (\Phi_E - E_X - eV_{BE})^{3/2}}{eV_{BE}}\right]; & \mu_E - eV_{BE} < E_x < \mu_E - eV_{BE} + \Phi_E \\ \exp\left[-\frac{2}{3}\frac{2\sqrt{2m_{BE}}}{\hbar} d_{BE} \frac{(\Phi_E - E_X)^{3/2}}{eV_{BE}}\right]; & \mu_E - eV_{BE} + \Phi_E < E_x < \mu_E + \Phi_E \\ 1; & \mu_E + \Phi_E < E_x \end{cases} \tag{3}$$

where Φ_E is the barrier height of emitter-base oxide, m_{BE} is the effective mass of tunneling electron in emitter-base oxide, d_{BE} is the emitter-base oxide thickness.

Note that $D_E(E_x) = 1$ is assumed for electrons with E_x above the barrier, and the failure of WKB approximation in the range $E_x \approx \Phi_E$ is ignored. Energies are measured relative to the bottom of the conduction band in the emitter, and μ_E is the chemical potential of the emitter (which will be regarded as the Fermi level here)[24].

Using eqn (1), the "normal distribution" can be calculated,

$$J_{EB}(E_x)\,dE_x = \frac{4\pi m_m}{h^3} D_E(E_x) \int_0^\infty f_E(E)[1 - f_B(E)]\,dE_t = \frac{4\pi m_m}{h^3} D_E(E_x) \times \left\{\int_0^\infty f_E(E)[1 - f_B(E)]\,dE - \int_0^{E_x} f_E(E)[1 - f_B(E)]\,dE\right\} \tag{4}$$

where I have used $E = E_t + E_x$, $dE = dE_t$ and $\int_{E_x}^\infty = \int_0^\infty - \int_0^{E_x}$.

The normal current distribution can be calculated easily at $T = 0°K$,

$$J_{EB}(E_x) = \frac{4\pi m_m e}{h^3} \cdot \begin{cases} V_{BE} D_E(E_x); & E_x < \mu_E - eV_{BE} \\ (\mu_E - E_x) D_E(E_x); & \mu_E - eV_{BE} < E_x < \mu_E \\ 0; & \mu_E < E_x \end{cases} \tag{5}$$

and Fig. 4 shows explicitly the derivation steps. As Fig. 4 describes the normal current distribution is fairly broad with a peak value which is always located below the Fermi level. Young and Müller[25] claimed that surface imperfections cause sufficient energy transfer between normal and transverse energy components, and have demonstrated a fairly good agreement between experimentally measured current distributions emitted from a field emission tip and the *total* energy distribution of the emitted current. Since our model is theoretical, normal current distributions are more appropriate in this analysis, and results for α will reflect the worst case losses resulting from finite energy distribution. No simple approximation can be derived for the total current distribution, and I prefer to leave it in the most general form,

Gomer[26] and Handy[14] give some estimates for the width of the total current distribution in the Fowler–Nordheim regime, which is not adequate for the low voltage and thin oxide cases.

I have carried numerical integrations of eqns (4) and (6) and representative results are demonstrated in Fig. 5. In the analysis $m_{BE} = m_m = m_0$, $V_{BE} \equiv V$, $\Phi_E \equiv \Phi$ and $d_{BE} \equiv d$ had been used. A barrier height $\Phi = 2$ eV is a realistic representation of Al–Al_2O_3–Al junctions, which have been widely used in the past. Table 1 summarizes results for typical total and normal current distributions for few oxide parameters and under different biasing conditions. In Fig. 5 I present the actual shapes of the current distributions in a typical case.

The total energy distributions in a junction with $\Phi = 2$ eV are summarized in Table 1(a) with varying oxide thickness: (a) $d = 10$ Å, (b) $d = 25$ Å and (c) $d = 50$ Å. In all three cases the following basic features are observed: (1) The distribution peak is shifting up in energy when the biasing voltage increases up to $eV = \Phi$; a further increase in V (entering the Fowler–Nordheim regime), causes a shift down in the maxima position (E_{max}). (2) The full width of the energy distribution narrows down when eV approaches a number somewhat above Φ from both sides of the energy scale. A different behavior is exhibited when $d = 50$ Å, and the biasing voltage is near $V = 2$ V. The distribution is very broad and the maxima position is almost 0.06 eV *above* the Fermi level (is noted at the table by a minus sign). (3) Energy distributions are mostly below the Fermi level. (4) Cooling to 77°K enhances the distribution near the Fermi energy and reduces the full width, but mostly for energies above E_F. (5) An increase in the barrier thickness causes the peaks to move closer to E_F and the full width narrows.

A similar behavior is exhibited in Table 1(b) for the normal current distributions. The distributions are wider and the displacement of the distribution peak below the Fermi level is larger. The temperature dependence is weaker, and again, mostly observed at energies above

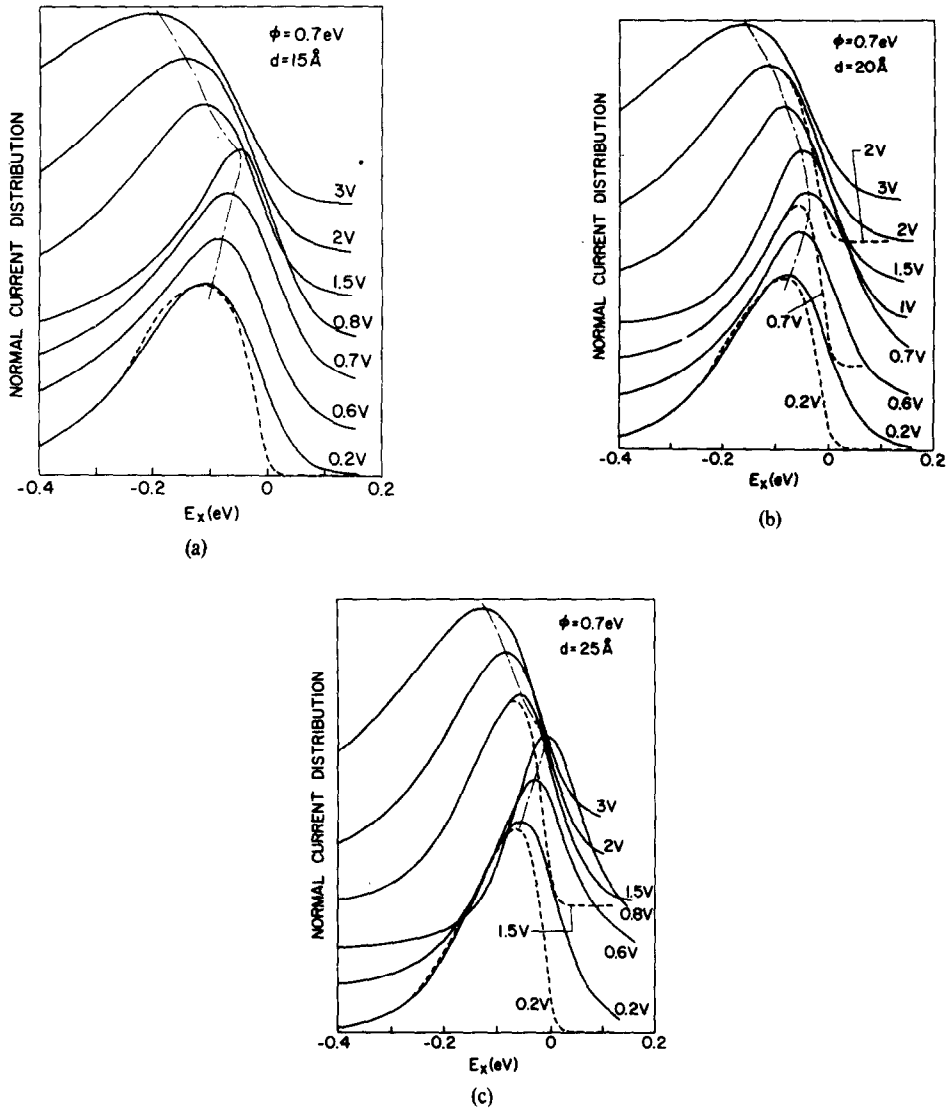

Fig. 5. Normal current distributions for $\Phi = 0.7$ eV. The line ——·——·—— is the locus of the distribution peaks. The line ---- is for $T = 77°K$. Each biasing condition has its own zero level. (a) $d = 15$ Å, (b) $d = 20$ Å and (c) $d = 25$ Å.

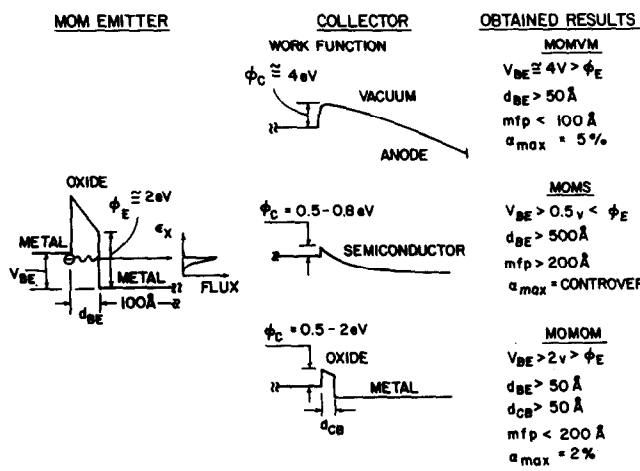

Fig. 6. A summary of the known THETA family: MOMVM, MOMS and MOMOM. Parameters and performances are typical to the works of the sixties.

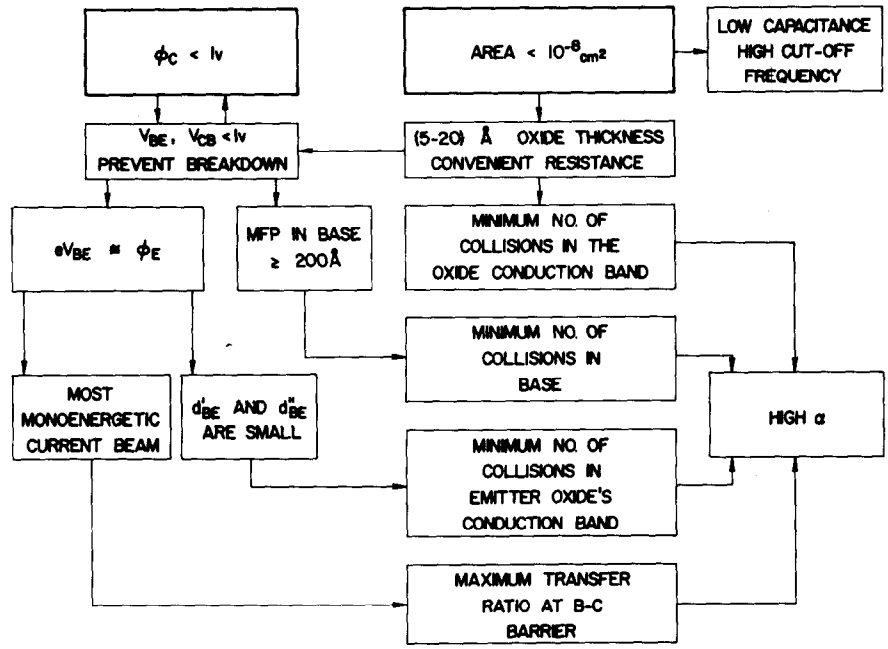

Fig. 7. A summary of parameters and performances of a proposed MOMOM. The two basic requirements are noted by the heavy line blocks and the final results are in the blocks on the right.

E_F. Full energy spreads are about 0.5 eV in the 10 Å, and about 0.25 eV in the 50 Å cases.

In early experiments the parameters chosen were $d_{BE} = 50$ Å, $\Phi_E = 2$ eV and $V_{BE} > 2$ V. Hence, the current density was very small and the area had to be large enough to get useful currents (mm² to cm²). Also, as will be seen soon, hot electrons with high excess energy have a short mean free path which results in a deterioration

Table 1. Summary of peak positions on left, and full width at half maximum on right for two 2 eV and 0.7 eV barrier heights (in eV)

V[v] / d[Å]	0.5		1.5		2		3		5	
10	.07	.3	.06	.29	.05	.26	.056	.24	.075	.34
25	.035	.17	.023	.15	.015	.15	.023	.14	.038	.18
50	.003	.14	-.012	.14	-.058	.26	-.012	.14	.015	.14

(a) ϕ = 2eV: TOTAL CURRENT DISTRIBUTIONS

V[v] / d[Å]	0.5		1.5		2		2.5		3		6	
10	.15	.51	.11	.65	.06	.56	.06	.43	.1	.51	.2	.8
25	.10	.27	.077	.23	.05	.22	.053	.18	.073	.22	.14	.4
50	.03	.17	.017	.17	-.045	.30	.0	.23	.017	.16	.075	.2

(b) ϕ = 2eV: NORMAL CURRENT DISTRIBUTIONS

V[v] / d[Å]	0.2		0.6		0.7		0.8		1.5		2		3	
15	.104	.29	.075	.24	.07	.24	.04	.19	.104	.26	.137	.33	.183	.43
20	.08	.21	.054	.22	.042	.22			.087	.22	.121	.29	.17	.37
25	.054	.19	.31	.19			.008	.16	.05	.22	.081	.18	.012	.30

(c) ϕ = 0.7eV: NORMAL CURRENT DISTRIBUTIONS

of α. Lower barrier height junctions, or ones with thinner oxides, will enable a much higher current flow for a substantially smaller junction area.

Examples of normal current distributions for $\Phi = 0.7$ eV are shown in Fig. 5 and summarized in Table 1(c). Biasing a junction with $d = 15$ Å at 0.2 V will result with a current density of $\sim 12\,\mu A/\mu m^2$. Junctions with areas in the submicron square range are an absolute necessity for rapid signal processing. Working with low-energy electrons will increase their mean free path in the base metal without changing significantly the current distribution. In a typical distribution, $\delta(1/2) \equiv$ half maximum full width $\cong 0.5$ eV and $\Delta(E_F) \equiv E_F - E_{max} \cong 0.05$ eV. Again, minimum $\delta(1/2)$ and $\Delta(E_F)$ are observed for $eV \gtrsim \Phi$ (most pronounced is the case of $d = 25$ Å, $V = 0.8$ eV). Table 1 and Fig. 5 suggest that the collector barrier height should be chosen very close to the emitter barrier height, consequently, when $eV_{BE} \gtrsim \Phi_E \gtrsim \Phi_C$ the narrowest and least displaced (from E_F) current distribution will be achieved, resulting most probably in the highest transfer ratio.

In the above calculations, I assumed that the top of the oxide's valence band energy is at $-\infty$. This assumption is valid for all cases where $E_G \gg \Phi$, where E_G is the energy gap of the oxide. In cases where $E_G \approx 2\Phi$, the above energy distributions are not valid, and in fact, due to the increase in tunneling probability for energies near the valence band, the distributions can be much wider or even exhibit two distinct maxima[27], obviously a situation which we would like to avoid.

(b) *Energy spread of tunneling current due to electronic collisions in the conduction band of the base-emitter oxide.* Under strong biasing conditions ($eV_{BE} > \Phi_E$), the electron spends part of its path in the conduction band of the base-emitter oxide. This part is broken into d'_{BE} and d''_{BE}, as denoted in Fig. 3. d'_{BE} and d''_{BE} depend on the biasing voltage and the image correction via[28]

$$d'_{BE} = d_{BE}\frac{1.2\lambda}{\Phi_{BE}}$$
$$d''_{BE} = d_{BE}\frac{eV_{BE} - \Phi_E + 5.6\lambda}{eV_{BE}} \tag{7}$$

where $\lambda = e^2 \ln 2/8\pi\epsilon d_{BE}$, ϵ is the dielectric constant, and a parabolic approximation for the image correction was assumed. For all practical cases $d'_{BE} = (0.01\text{–}0.03)d_{BE}$, but d''_{BE} is strongly dependent on the parameters chosen. For example, for $\Phi_E = 2$ eV and $d_{BE} = (25\text{–}50)$ Å, d''_{BE} changes between $0.03d_{BE}$ and $0.5d_{BE}$ when the biasing voltage is increased from 2 to 4 V.

While spending time in d'_{BE} and d''_{BE} the electron collides mostly with optical phonons[29] loses energy and changes its momentum. Since the Debye temperature of most oxides is about 1000°K, the loss of energy per collision is ~ 0.1 eV, and we may regard these collisions as almost elastic. Hence, the most dramatic effect of those collisions is the drastic change in the normal energy distribution of the current, which will reduce α considerably. Scattering in the region designated by d'_{BE} is not important since the momentum distribution of the electrons which impinge on the barrier is isotropic to begin with, and those electrons which are scattered back into the emitter are not contributing to the emitter current. But an isotropically scattered electron in the region d''_{BE} will most probably not be collected by the collector. This is easily understood if one looks at an electron travelling normal to the surface with $E = E_{xo}$. After scattering, $E_x = E\cos^2\beta$, where β is the scattering angle relative to the normal. If $E_{xo} - \Phi_C = \Delta E$, the maximum allowed β for collection is $\beta_{max} = \arctan(\Delta E/\Phi_C)^{1/2} \cong (\Delta E/\Phi_C)^{1/2}$ for $\Delta E \ll \Phi_C$. This restricts considerably the acceptance cone, and if we neglect multiple-scattering events (which might direct the electron back into the acceptance cone), we can safely say that only ballistic electrons will be collected.

As a result, the maximum transmission coefficient which is introduced by the oxide for a single energy component, is

$$\alpha_{BE} = \exp\{-d''_{BE}/l_x\} \tag{8}$$

where l_x is the x oriented mean free path, which is defined by

$$l_x = \int_0^1 d(\cos\theta)\int_0^\infty dr p(r) r = \frac{1}{2} l_{e\text{–}ph} \tag{9}$$

where $l_{e\text{–}ph} = \int_0^\infty dr p(r) r$ and $p(r)$ is the probability for the electron to suffer a collision in the region between r and $r + dr$. (Note that semiclassical description here is only qualitative, since $\lambda \approx d$, where λ in the de Broglie wavelength of the electron).

The available data for l_x (and $l_{e\text{–}ph}$) is scarce and numbers are widely distributed. Table 2 presents collected results for Al_2O_3. As one sees, the values vary from 3 to 24 Å and depend strongly on the measuring technique and the process of oxide fabrication. All those results had been obtained on thermally or plasma-grown oxides with thicknesses in the 100 Å range and without any special care to eliminate impurities from the oxides.

For oxides thicknesses in the 100 Å regime and $d''_{BE} \cong 0.3d_{BE}$[34] transmission won't exceed 2×10^{-3} if $l_x \cong 5$ Å.

(c) *Energy loss and isotropic randomization in the base metal.* Scattering of electrons in the base metal results from electron-defect or electron-impurity, electron–phonon and electron–electron collisions.

The electron-neutral impurity scattering is temperature independent and the associated mean free path is proportional to the square of the kinetic energy. When ionized impurities scattering are dominant, scattering times proportional to $T^{3/2}$ are found.

The mean free path between electron–phonon collisions is inversely proportional to the temperature and linearly dependent on the square of the kinetic energy,

$$l_{e\text{–}ph} = \lambda_\sigma (E_K/E_F)^2 \tag{10}$$

where λ_σ is the mean free path of thermalized electrons and E_K is the kinetic energy of the hot electrons. One

Table 2. x oriented mean free path of electrons in the conduction band of Al_2O_3

ℓ_x [Å]	3	4	5	6	24
Reference	30	14	31	32	33

could find λ_σ using the conductivity σ of the metal

$$\lambda_\sigma = \frac{m_m v_{th}}{ne^2}\sigma \tag{11}$$

where v_{th} is the thermal velocity of the conduction electrons and n is their density.

For a more detailed discussion on the dependence of l_{e-ph} on the energy, refer to Section 3.3.

An approximate energy dependence for the mean free path between electron–electron collisions can be estimated from density of states considerations. If one assumes that the interaction is completely inelastic, namely, the hot electron loses the same amount of energy that the thermalized electron gains, the predicted collision time will go down with increase in the hot electron energy. The number of electrons which the hot electron can interact with, in conjunction with the number of states to which the thermalized electron can be excited to, leads to a $(E-E_F)^{-2}$ dependence of the collision time. Since the mean free path is equal to the product of the collision time and the hot electron velocity, one gets

$$l_{e-e} \propto E^{1/2}(E-E_F)^{-2}. \tag{12}$$

This relation is undoubtedly oversimplified, but it provides us with an approximate handy relation in a limited range of energies, and particularly allows to extrapolate l_{e-e} to ranges where data is not available. More detailed theories can be found in Refs. [34–45], with a wide spread in the predicted results. Unfortunately, the experimental results vary as well and are strongly dependent on the method of testing and preparation of the samples.

As was pointed out in Section (b), lattice related scattering which are elastic and isotropic in nature can be regarded as a permanent loss in the collection process, and since electron–electron collision prevents collection too, both scattering can be combined.

The combined mean free path is customarily estimated via

$$l_T = \left[\sum_i (l_{e-i})^{-1}\right]^{-1} \tag{13}$$

where l_T is the combined and l_{e-i} is the mean free path for the particular scattering mechanism. Equation (13) is accurate only if the energy dependence of the scattering mechanisms is similar, an assumption which is valid over a small energy range.

Table 3 presents a collection of published mean free paths for a few metals in the low energy range. As one can see, results of one work differ markedly from another, and only the order of magnitude is significant.

Electrons with excess energy of (3–5) eV have a combined mean free path in the range of (50–100) Å, so that transmission of $\exp(-1)$ or $\exp(-2)$ is an upper bound for a 100 Å thick metal base. Electrons with excess energies in the (0.5–1) eV range might have mean free paths in the (200–700) Å regime, and consequently suffer less total loss.

As will be discussed in Section 3.2, it is desirable to use a thicker base to prevent an output–input coupling (due to the skin-depth and spreading resistance in the base metal), to eliminate pinholes which occur in deposited thin metal films, and to reduce the spreading resistance of the base. This will be possible without reducing the transfer ratio, only if the ballistic electrons are not "too hot".

(d) *Quantum-mechanical reflection from the base-collector oxide interface.* Classically, when the normal energy of the hot electron exceeds the barrier height Φ_C, no reflection will occur. Quantum mechanically, there is always a finite reflection which gets smaller as the excess normal energy above Φ_C increases.

The reflection depends strongly on the gradient of the potential barrier and on the effective mass of the electron on each side of the reflecting interface. The current transmission can be expressed as [55]

$$T = \left(\frac{A_t}{A_i}\right)^2 \frac{m_i^*}{m_t^*} \tag{14}$$

where A is the amplitude of the wave functions, m^* is the effective mass, and i and t stand for media of incidence and transmission, respectively. Crowell *et al.* [55] had calculated T numerically for a metal-semiconductor interface and showed a reduction in T for an increase in the electric field and a decrease in the excess energy above Φ_C. For $m_i^* = m_t^*$ and excess energy (0.1–0.4) eV, a reflection of 25% is inevitable.

The case of a perfectly square barrier presents the highest reflection coefficient. For example, if $\Phi_C = 0.5$ eV and $d = 15$ Å, an excess energy of 0.3 eV will result with a transmission coefficient of 50% while for $d = 10$ Å and the same excess energy $T \cong 90\%$. The detailed shape of the barrier, the field penetration into the electrodes and the Thomas–Fermi screening field will reduce the potential gradient at the boundaries and increase T, up to more than 60% for an excess energy of 0.1 eV. Since $m_i^* > m_t^*$ in the case of metal-oxide or metal-semiconductor interface, T will be greater.

Since the current distribution is not monoenergetic, a numerical integration has to be carried out to get the total T.

Table 3. Accumulation of mean free path data of hot electrons in evaporated thin films, as a function of their excess energy above the Fermi level

Material	Excess Energy [eV]	Mean Free Path [Å]	Reference
Al	0.7	~100	55
	0.7	~200	52
	1	~100	52
	1	60	54
	1.5	~50	52
	1.5-1.85	90	66
	2	150	49
	2.5	$\ell_{e\text{-}e}$~220	53
	2-3	150	32
	3	300	73
	4	70±10	31
	5	$\ell_{e\text{-}e}$~50, ℓ_{l}~250	62
	6	50±10	31
	6-7	75	47
	9	$\ell_{e\text{-}e}$~510, $\ell_{e\text{-}ph}$~130	63
Pd	1	$\ell_{e\text{-}e}$~650, $\ell_{e\text{-}ph}$~110	55
	1.1	170	57,61
Cr	2.8	35	59
Cu	0.55	430	36
	0.927	70-110	69
	1.1	50-200	57,61
K	2.2	1000	58
Au	0.8	55±100(holes)	68
	0.85	520±5	70
	0.95	350	60
	1	700	71
	1	$\ell_{e\text{-}e}$~1200, $\ell_{e\text{-}ph}$~406	55
	1.1	740	57,61
	1.1	740	72
	2	$\ell_{e\text{-}e}$~140	64
	2	200	71
	3	$\ell_{e\text{-}e}$~120	64
	3.2	70	59
	4	$\ell_{e\text{-}e}$~50	64
	5	$\ell_{e\text{-}e}$~50	64
	5	70	71
	7	80	46
	7	47	48
	7	57	50
	7-10	180	46
	5.5-10	$\ell_{e\text{-}e}$<45, $\ell_{e\text{-}ph}$~250	62
	12	400-600	67
Ag	0.65	370	36
	1	$\ell_{e\text{-}e}$~1100, $\ell_{e\text{-}ph}$~570	55
	1.1	440	57,61

Table 3 (*Contd*)

Material	Excess Energy [eV]	Mean Free Path [Å]	Reference
	1.9	200-500	51
	5	48	51
	5.5	70	56
	5.5-8	$l_{e\text{-}e}<42$, $l_{e\text{-}ph}\sim400$	62
	12	250-700	67
In	4.27	42	65
	4.9	30.5	65
	5.65	18.5	65

(*e*) *Transmission through the conduction band of the base-collector oxide.* The base-collector oxide can be divided into two regions: $x < d^M_{CB}$ and $x > d^M_{CB}$ (see Fig. 3, $x = 0$ is at the base-collector oxide interface). In the region $x < d^M_{CB}$, a retarding field decelerates the electrons, and if in this region the normal energy due to phonon and impurity drops below Φ_C, the electron will not be collected. On the other hand, if the normal energy drops below Φ_C for $x > d^M_{CB}$, the electron *must* be collected (the energy peak will prevent if from being scattered back). With an increase in the electron's normal energy, the number of generated optical phonons and consequently reflections will rise. Crowel and Sze[74] had shown that reflection from metal-semiconductor interface increases when n varies from 1 to 2, where $E_x - \Phi = nE_{ph}$ and E_{ph} is the optical phonon energy. A further increase in n leads to a saturation in the reflectivity at values: 45% in GaAs and 15% in Si and Ge ($l_{e\text{-}ph}$(GaAs) $\cong$ 15 Å, $l_{e\text{-}ph}$(Si, Ge) $\cong$ 65 Å). In a thin oxide case, d^M_{CB} is usually equal to or smaller than $l_{e\text{-}ph}$ (see Table 1), hence reflections due to isotropic scattering are minimized. If a monoenergetic electron beam could be generated, a narrow window in the range $\Phi_C < E_x < \Phi_C + E_{ph}$ would have a high transmissivity. The only scattering would be by acoustic phonons (neglecting quantum-mechanical reflections).

When the energy distribution of the current, the quantum mechanical reflection and the phonon-scattering are considered simultaneously, maximum transmission will occur at some value of eV_{BE} in the range $\Phi_C \ldots \Phi_C +$ 0.5 eV.

(f) *Quantum-mechanical reflections from the collector oxide–collector metal interface.* Classically, no reflections occur at this boundary. MacCall has treated this problem quite rigorously, modeling the metal as a periodic potential terminated by a potential barrier at its surface[75, 76]. His numerical calculations resulted in reflection and transmission bands which sensitively depend on the incoming energy and the amplitude of the periodic potential. Reflection might vary between a fraction to few tens of percent.

For evaporated amorphous or polycrystalline metals, the applicability of this model is questionable and reflections are probably dependent mostly on m^* in the adjacent oxide. For $m^*_i \cong m^*_i$, reflection can amount only to a fraction of a percent.

For more loss mechanisms refer to Section 3.2(b).

2.3 *The THETA family*

In the preceding sections, most of the analysis was centered around the MOMOM configuration, which is only one member of the three known members of the THETA family. Creation of the hot electron beam is similar in all three configurations. It is based on an MOM device with a thin metal counter electrode and an oxide with a barrier height of about 2 eV (which is the case of Al–Al_2O_3–Al system which had been utilized extensively). Each member of the family will be briefly discussed in light of previous achievements, which are summarized in Fig. 6.

MOMVM (MOM-vacuum-metal). In this configuration electrons have to surmount the potential barrier at a metal(base)-vacuum interface (which is the workfunction of the base metal). Collection is done by a positively biased anode situated away from the base.

Employing stable metals with workfunctions close to 4 eV (e.g. $\Psi_{Cu} = 4.4$ eV, $\Psi_{Ag} = 4.3$ eV, $\Psi_{Au} = 4.3$ eV, $\Psi_{Al} = 4.5$ eV, etc., while reactive metals have workfunctions in the vicinity of 2.5 eV), with conjunction with Al–Al_2O_3 emitter leads to operation in a biasing regime where $eV_{BE} > \Phi_E$, namely, the Fowler–Nordheim regime. Table 1 gives the full width at half maximum for the normal current distributions with $\delta(1/2) = (0.3$–$0.4)$ eV and $\Delta(E_F) = (0.1$–$0.3)$ eV. Increasing biasing voltage increases $\delta(1/2)$ and $\Delta(E_F)$.

For energies in excess of 4 eV, l_e is smaller than 100 Å (in all known cases), and the reduction in α is substantial. Moreover, the emitter-oxide thickness has to be large enough to sustain the large V_{BE} ($d_{BE} \gtrsim 50$ Å). As a result d''_{BE} is large and scattering events increase $\delta(1/2)$ and $\Delta(E_F)$ even further.

The collection region on the other hand, is free of electron scattering mechanisms. Reflection from vacuum-collector boundary is minimal due to the small potential gradient (no discontinuity in the dielectric constant and effective mass).

Table 4 summarizes prior art. As one can see α_{max} is poor. In Ref. [33] the Al metal was coated with Ba which has a workfunction of $\sim$2.5 eV. This resulted in a

Table 4. Accumulation of parameters and performances of past MOMVM's

E-B Oxide Type	E-B Oxide Thickness [Å]	Base Metal	Base Thickness [Å]	α_{max}	Reference
Al_2O_3	100	Al	100	10^{-3}	30
Al_2O_3	75	Al	100	$3x10^{-4}$	77,78
Al_2O_3	67	Al,Al-Ba	100	$<10^{-4}$,$5x10^{-2}$	33
Al_2O_3	900	Au	80	$1.8x10^{-2}$	79

significant initial improvement, which decayed with time after the Ba evaporation. In Ref. [79] the hot electrons were thermally emitted due to the large d_{BE} and a biasing voltage of ~ 13 V. Note that the current distribution peak was at ~ 4.5 eV, namely, $\Delta(E_F) \cong 8.5$ eV!

MOMS (MOM-semiconductor). An intimate contact between the base and a semiconductor results in a much lower barrier height than the workfunction of the base metal. Most barrier heights are in the range of (0.5–1) eV[80]. As a result, V_{BE} doesn't exceed 1 V and the mean free path of electrons in the base metal is larger than 100 Å.

The transmission into the semiconductors is less than unity due to quantum mechanical reflections and elastic scattering from optical phonons. Hall had argued that due to the short mean free path of hot electrons in semiconductors ($l_{e-ph} \cong 50$ Å)[81], an electron will suffer 10–20 collisions in the depletion layer, and 50% of the incoming electrons will be backscattered[12]. Crowel and Sze included the influence of the electric field in the depletion layer (but neglected the quantum mechanical reflections), and predicted 10% reflection in Si and less than 45% in GaAs for an electric field of 3×10^5 V/cm[55, 74]. Quantum mechanical reflections for an excess energy of 0.1 eV above Φ_C were estimated as 10% in GaAs, 2% in Si, and 10% in Ge[55]. The overall predicted reflection is minimal in Ge.

When the electric field increases, quantum mechanical reflection will increase. Φ_C will decrease (due to image correction), increasing the probability of multiple phonon creation and backscattering.

The MOMS was first proposed and fabricated by Spratt *et al.*[11]. Their overall transmission coefficient was more than 90%, which they accounted for by assuming $l_e(\text{Al}) \cong 1000$ Å. Hall[12] and Lavine *et al.*[13] criticized their interpretation and proposed that a direct injection of electrons into the semiconductor—which occurs via pinholes in the base metal—is responsible for the large α. The problem of growing thin films without pinholes is an outstanding problem and I'll return to it in the next section. Tarng and Wehner had demonstrated this problem by sputtering Mo on top of different metals[82]. Using Auger spectroscopy combined with sputter-etching, they had shown that the first six monolayers of Mo on W substrate uniformly cover the W without pinholes. When W was replaed by Al, the first twelve monolayers covered only $\sim 9.5\%$ of the Al area. Using Au as a substrate led to $\sim 40\%$ coverage after twelve monolayers and Cu substrate to $\sim 90\%$ coverage. The excellent coverage of W was attributed to the similar atomic structures of Mo and W.

Antula had checked the pinhole formation of thin evaporated metals on amorphous Al_2O_3 via electron transmission microscopy[52]. He found that 40 Å Al films cover only 75% of the Al_2O_3 area while 70 Å Al films cover 88%. Au films cover 88% at a thickness of 85 Å and 93% of 100 Å.

These two examples demonstrate the importance of the materials which are in contact, in pinholes formation. Deposition conditions like, temperature of substrate, rate of deposition, and the degree of vacuum are of great importance too, as recent works with molecular beam epitaxy demonstrate.

Table 5 summarizes reported results in the literature on the performance of MOMS devices. In Ref. [83], special precautions were taken to avoid pinholes by using a mercury ball as an emitter, and in Ref. [54] by depositing a very thick base (800 Å).

At this point, I would still regard the high transfer ratios which had been reported for the MOMS as questionable.

MOMOM (MOM-oxide-metal). The MOMOM was the first proposed device (chronologically) among the hot electron devices[9]. A second MOM junction, separated by a thin metal base from the emitter, serves as a collector.

The very narrow base-collector oxide minimizes electron–phonon collisions. Even if one accounts for a large potential gradient at the base-collector boundary, the overall transmission is similar to or better than in the MOMS device.

In the next section, a small area, small Φ_C device, will be described. Prior art utilized mostly a symmetrical structure with both Φ_E and Φ_C in the neighborhood of 2 eV (summarized in Table 6). In order to support voltages on the order of $V_{BE} > 2$ V thick oxides had been grown (or deposited) and the overall performances were poor. In Refs. [52, 85], the author accounted for the current transport through the pinholes in the base metal. He claimed to have an Al_2O_3 barrier with $\Phi_C = 0.7$ eV,

Table 5. Accumulation of parameters and performances of past MOMS's

Emitter Metal	E-B Oxide Type	E-B Oxide Thickness [Å]	Base Metal	Base Thickness [Å]	Collector Type	α_{max}	Reference
Au	Al_2O_3	not given	Al	100	n-Ge	0.9	11
Hg	SiO	40	Au	100	n-Si	0.5	83
Pb	Al_2O_3	~30	Al	800	n-Ge	10^{-3}	54
Au	Al_2O_3	not given	Al	150,800	n-Si	10^{-2},10^{-4}	84

which he achieved by depositing Al in an O_2 environment. The remarkable increase in α which he had observed might be due to the low barrier height and the long l_e in the base metal.

Frequency performance. In 1963, Moll[16] and Atalla and Soshea[17] published two almost identical papers in which they have compared the frequency performance of several hot electron devices with that of an *npn* transistor. They concluded that, in general, a tunnel emitter device will perform poorly at high frequencies. In their model, they used the parameters $\Phi_E = 1$ eV, $d_{BE} =$ 20 Å and $I_E = 10^3 a/cm^2$ which led them theoretically to predict an input time constant ($\tau_{in} = r_{in}C_{in}$) on the order of 10^{-10} sec or larger[87].

The recent frequency performance which the point-contact and the evaporated MOM structures have exhibited proved the contrary. Experimentally, small area tunnel junctions have a frequency response up to at least the near i.r. (For recent review articles refer to Refs. [88, 89, 7].) These results are drastically different from those predicted in Refs. [16, 17] and they can be accounted for by the following features of the "new MOMs": (a) the area is in the submicron square range, (b) the tunneling impedance is tens to several hundred ohms, (c) the oxide thickness is on the order of 5–10 Å, (d) due to the image lowering effects and the "non-bulk" behavior of the few monolayer oxide, the barrier-height of the oxide is probably low, (e) current densities on the order of $10^6 a/cm^2$ are routinely achieved without any damage to the junction[90]. As a result, the calculated τ_{in} can be as low as 10^{-13} sec, a range which is not accessible by any other electronic device[91]. (See later, eqn (20) and Figs. 8 and 9.).

A question which should be raised is the validity of the lumped circuit RC-time constant concept at these high frequencies. Sue and Gustafson[92] had argued that in finite structures the known non-radiative surface plasmon modes are radiative and can be coupled into, from free space without using an evanescent wave type approach of coupling. In contrast to the antenna coupling model—where the collected high frequency voltage is applied across the junction—in the above model the surface plasmon modes propagate into the junction (between the two metals which are separated by roughly 10 Å), and hence are not "shorted" by the capacitive load. Efficient coupling is predicted[92] up to the very near i.r. even if the classical time constant is an order of magnitude or so greater.

The only quantum-mechanical treatment which deals with the interaction of radiation with an MOM junction was published recently by Tucker and Millea[94] based on the original approach taken by Tien and Gordon[93]. In their approach, the applied radiation modulates the Fermi levels of both metals and the electrons change energy periodically, with frequency ω. If the electrons are treated using the wave formulation, they will have non-vanishing probabilities of being at energies greater than E_F by a multiple of $\hbar\omega$ and with amplitudes which

Table 6. Accumulation of parameters and performances of past MOMOM's

Emitter Metal	E-B Oxide	E-B Oxide Thickness [Å]	Base Metal [Å]	Base Thickness [Å]	B-C Oxide	B-C Oxide Thickness [Å]	Collector Metal	α_{max}	Reference
Al	Al_2O_3	70	Al	100	SiO	100	Al	0.1	10
Al	Al_2O_3	40	Al	200	Al_2O_3	50	Al	0.02 (77°K)	30
Al	Al_2O_3	33	Al	150	Al_2O_3	33	Al	0.01 (77°K)	32
Al	Al_2O_3	100	Al	130,150	Al_2O_3	100	Al	.26,.2	52,85

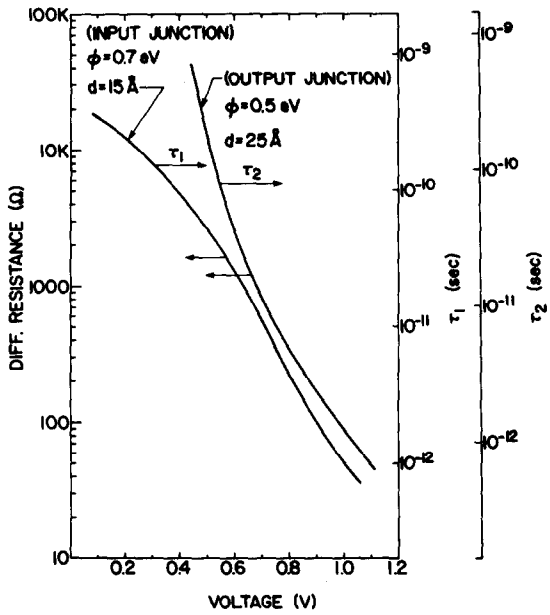

Fig. 8. Dynamic resistances and time constants of the input and outputjunctions in the model MOMOM.

are dependent on the modulating voltage. It was shown[94] that the impedance of the junction drops at higher frequencies and the nonlinearity of the I–V characteristic decreases. A frequency limit of "classical operation" is arrived when $\hbar\omega$ approaches the barrier height.

The MOMVM is expected to be limited by the transit time through the vacuum space, which is about 10^{-10} sec for a length of 0.1 mm and $V_{CB} = 10$ V.

The MOMS will be limited by the transit time through the depletion layer of the Schottky barrier collector. Recent experimental results had demonstrated classical mixing and detection of radiation up to a wave length of 70 μm[95–97]. Theoretical prediction by van der Ziel[98] predicts a minimum wavelength of 5.5 μm for a doping level of 3×10^{19} cm^{-3}, which is substantiated by Tsang and Schwarz who employed another theoretical approach[99].

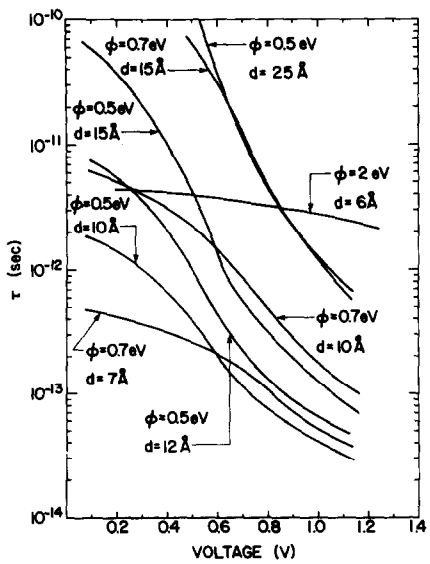

Fig. 9. Time constants of different MOM's as a function of biasing. The superiority of "low Φ-thick d" junctions are demonstrated by comparison with the case $\Phi = 2$ eV and $d = 6$ Å.

The MOMOM will be limited by either: transit time of the electrons through the base, barrier height[100], optical properties of the metals[101] or input and output RC time constants. A proper choice of materials and device dimensions can push the upper frequency limit to $1/t_{base}$, where t_{base} is the transit time through the base. For a 100 Å metal base and electron velocities of 10^8 cm/sec, we may expect $t_{base} \cong 10^{-14}$ sec.

3. PROPOSED "THETA" DEVICES

The failure of MOM-collector devices to exhibit a high transfer ratio has led to a decreased interest and finally, to a complete stop in research activities in the late sixties.

In this section I wish to refine the MOM-collector devices and to propose new configurations which rely upon the generation of a monoenergetic hot electron beam via tunneling and a rapid collection by a high impedance collector.

The problems which we are facing are by no means simple. They could be divided into two major categories: (a) Achieving a high enough α at d.c. operation. (b) Maximizing the input and output coupling efficiencies of high frequency signals and α at these frequencies.

In the next few sections I describe the main idea, predict performances, stress difficulties and propose practical configurations.

3.1 *Consequences of small area devices*

The motivation for working with ultra-small area devices arises from the operating frequency requirements. But the small area MOM's are a new breed of devices with properties which are described in the previous section and Ref. [90].

The past MOM-collector devices were large area devices. Consequently, thick oxides were grown in order to reduce the tunneling probability, and devices would have an impedance in the tens of ohms to kilohm range (spreading resistance, or lead resistance, would be comparatively negligible). Little attention was paid to the collector oxide barrier height, and Al–Al_2O_3-metal collectors had been utilized mostly. As a result, electrons in the base metal were "very hot" (> 2 eV) and suffered many collisions which together with losses in the thick oxides reduced α considerably.

Devices with areas in the submicron square range operate with oxide thickness of less than 10 Å and still have a desirable impedance. A low collector oxide barrier height will allow a long mean free path of hot electrons in the base metal. The consequences of utilizing very small area devices can be summarized as follows: A very thin oxide ($\lesssim$ 10 Å) will suffice to give tens to hundreds ohms impedances, which in turn reduces the losses in the oxide. The thin oxide will limit V_{BE} (electric field is limited by 10^7 V/cm, to prevent breakdown), but large V_{BE} is not needed due to the low Φ_C. A small V_{BE}, in addition (and in particular $eV_{BE} < \Phi_E$), will

prevent an operation in the Fowler–Nordheim regime and as a result, d'_{BE} and d''_{BE} will be negligibly small (Fig. 3), leading to a further reduction in the number of phonon scattering. High collector impedance is easily achieved by an oxide thickness of about 25 Å, consequently d^{M}_{CB} is very small and the number of collisions in that region is less than one. The block diagram in Fig. 7 summarizes the most important consequences which will result from the two basic requirements, if fulfilled (blocks of heavy lines). Another basic requirement is $E_{GE} > 2\Phi_E$ (which is usually fulfilled), otherwise the top of the valence band of the oxide will lead to a broader current distribution[27], and a reduction in α.

3.2 *The MOMOM*

As pointed out above, small area and very thin oxides are an "absolute must" for an acceptable performance of the THETA family. I will estimate the d.c. and high frequency performances of the proposed MOMOM in the sections below.

(a) *Predicted performance.* The overall operation characteristics of the MOMOM are based on the following:

Transfer ratio. The transfer ratio $\alpha = I^*_C/I_E$ is calculated using a cascade model of transmission,

$$J^*_C(E_x, E)\,dE_x\,dE = J_{EB}(E_x)T_B(E)D_C(E_x)\,dE_x\,dE \tag{15}$$

where $T_B(E)$ and $D_C(E_x)$ are the transmission coefficients through the base and collector respectively.

$$I^*_C = 162\int_0^\infty [1-f_C(E)]T_B(E)f_E(E) \times [1-f_B(E)]\,dE\int_0^E d_E(E''_x)D_C(E'_x)\,dE_x. \tag{16}$$

In (16) the current is in $[\mu A/\mu m^2]$ and I^*_C stands for the transmitted current into the collector, so that $I_C = I^*_C + I_{CBO}$. I_{CBO} is the tunneling current between collector and base when the emitter is not connected. It is evaluated similarly to I_E using the substitution $V_{BE}\to V_{CB}$, $d_E\to d_C$, $\Phi_E\to\Phi_C$ and $f_E\to f_C$. Based on eqn (10)

$$T_B(E) = \exp\left\{-d_B\frac{(E-E_{FB})^2}{L_0}\right\} \tag{17}$$

where L_0 is a characteristic constant of the base metal[102], and E_{FB} is the Fermi level in the base. Note that $f_C(E) = f_E(E + eV_{BE} + eV_{CB})$, $E'_x = E_x + eV_{CB}$, and $E''_x = E_x + eV_{CB} + eV_{BE}$. I_E is given by (3)–(6) and $|I_B| = |I_E| - |I_C|$.

Input and output resistances. The d.c. resistances are evaluated using

$$R_{in} = \frac{V_{BE}}{I_E}, \qquad R_{out} = \frac{V_{CB}}{I_C} \tag{18}$$

while the differential resistances are defined as

$$r_{in} = \frac{\partial V_{BE}}{\partial I_E}, \qquad r_{out} = \frac{\partial V_{CB}}{\partial I_C}. \tag{19}$$

For example, r_{in} is evaluated using:

$$r_{in}^{-1} = 162\int_0^\infty f_E(E)\left\{[1-f_B(E)]\int_0^E D'_E(E_x)\,dE_x - f'_B(E)\int_0^E D_E(E_x)\,dE_x\right\}dE \tag{20}$$

where $D'_E = \partial D_E/\partial V_{BE}$ and $f'_B = \partial f_B/\partial V_{BE}$. r_{out} is calculated from the output characteristics. The time constants are evaluated using $\tau = rC$ where $C = \epsilon A/d$ where ϵ is the dielectric constant (which is taken as that of free space due to the extreme thinness of the oxide), and A is the junction area which is assumed here to be 1 μm^2 in the present calculations.

Let's review the requirements from an MOMOM amplifier. (a) $r_{out} \gg r_{in}$ to achieve a power gain, (b) τ_{in} and τ_{out} as small as possible for high frequency performance, and (c) maximizing α. These requirements can be translated into design using the following reasoning: Figure 5 and Table 1 suggest that the most monoenergetic electron beam is supplied by a "low Φ_E-thick d_{BE}" combination (rather than a "high Φ_E-thin d_{BE}") and the half maximum—full width is minimized when $eV_{BE} \cong \Phi_E$ which leads directly to $\Phi_C < \Phi_E$. The combination "low Φ_E-thick d_{BE}" will minimize the time constants too. Because of requirement (a), $d_{CB} > d_{BE}$ and for maximizing l_e (and α), $\Phi_{CB} \sim 0.5$ eV is necessary. Because of (b), the oxides' thickness should be in the (10–25) Å region. For model calculations I've chosen the following parameters: $\Phi_E = 0.7$ eV, $d_{BE} = 15$ Å, $\Phi_C = 0.5$ eV, $d_{CB} = 25$ Å, and Figs. 8–12 exhibit different qualities of this MOMOM amplifier with a chosen cross section of 1 μm^2 resulting from evaluating eqns (15)–(20). The properties of few "high Φ_E-thin d_{BE}" junctions are plotted too for comparison.

The dynamic resistance of *BE* and *CB* junctions with their respective time constants are plotted in Fig. 8. Figure 11 describes the behavior of α for the model junction. As one can see, α reaches its maximum at $V_{BE} \cong 0.8$ V which leads to $r_{in} \cong 100\,\Omega$ and $\tau_{in} \cong 4\times10^{-12}$ sec in Fig. 8. If one assumes $L_0 \cong 200$ ÅeV^2 and (from Fig. 11) $\alpha_{max} \cong 0.67$; one gets $r_{out} \gtrsim 225\,\Omega$ for a power gain $G \gtrsim 1$. Hence, $\tau_{out} \cong 2\times10^{-12}$ sec, and the upper operational frequency is limited by τ_{in} to about 200 GHz.

The transfer ratio for the model device is predicted in Fig. 11 for few chosen L_0's. I've plotted two cases of "high Φ_E-thin d_{BE}" for comparison purposes, which clearly demonstrates their inferiority.

The output characteristics are presented in Fig. 10, using $L_0 = 500$ ÅeV^2. The transferred current is separated from the tunneling current I_{CBO}. At low V_{CB} (< 0.4 V), r_{out} is bounded by $r^*_{out} \equiv \partial V_{CB}/\partial I^*_C$ where $I^*_C =$

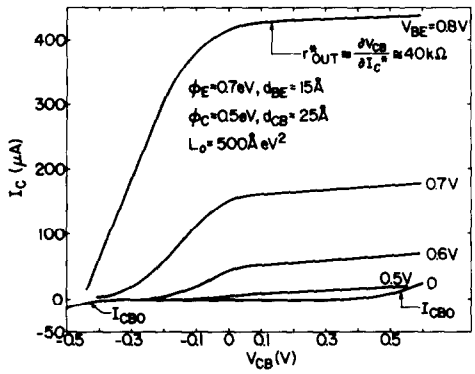

Fig. 10. Output characteristics of the model MOMOM. The region $V_{BE} < 0.7$ V could be regarded as the "OFF" region while $V_{BE} > 0.8$ V as the "ON" region. Note that I_{CBO} is negligible for $V_{CB} < 0.4$ V.

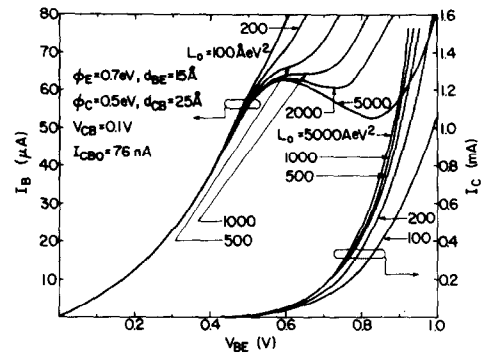

Fig. 12. I_B vs V_{BE} and I_C vs V_{BE} in the model MOMOM. Differential negative input resistance in the common emitter configuration is exhibited for $L_0 > 1000$ Å(eV)². The transconductance increases with the increase of L_0 and V_{BE}.

αI_E, and at higher V_{CB}, the dynamic resistance of the output junction ($\partial V_{CB}/\partial I_{CBO}$) will bound the output resistance. In this particular example, for $V_{CB} \cong 0.4$ V both resistances are about 40 KΩ, which is twice the maximum of r_{out} attainable.

The transconductance[103] changes markedly from less than $5 \times 10^{-4} a/v$ at the threshold value of V_{BE} ($eV_{BE} \cong \Phi_C$) to more than $25 \times 10^{-4} a/v$ for $eV_{BE} > \Phi_C$, as is directly observable in Figs. 10 and 12. As was discussed in Section 2.1, this is highly desirable in switching devices, since a small swing in V_{BE} near threshold will produce a large change in I_C.

As I suggested before, Fig. 12 describes that the MOMOM may exhibit a differential negative resistance at the input for some choice of parameters. One won't find the negative resistance feature in "high Φ_E-thin d_{BE}" structure due to the wide energy spread of the emitter current. In the model structure, the negative resistance is strongly dependent on the mean free path in the base metal, and as one can see it disappears at $L_0 = 1000$ ÅeV². Such long mean free path might be achieved in impurity and defect free, epitaxially grown thin metal films (and possibly at low temperatures). If the negative resistance feature is undesirable, it could be easily eliminated by a different choice of Φ_E, d_{BE} as mentioned above[104].

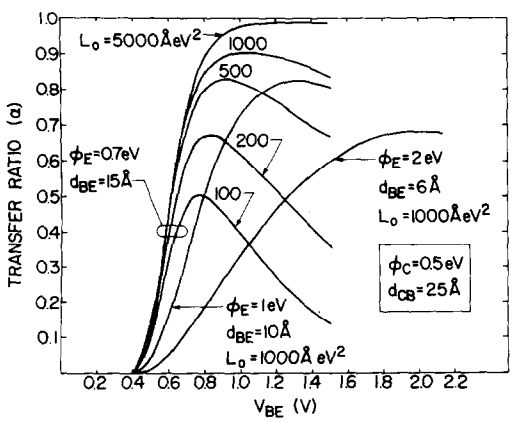

Fig. 11. Transfer characteristics of the model MOMOM with L_0 as a parameter. A comparison with a "high Φ_E-thin d_{BE}" input junction is demonstrated.

The above model will perform well down to the 10 psec range (100 GHz). If the frequency limit is extended further, the time constant of the collector has to be reduced. Figure 9 shows a different choice of parameters: $\Phi_E = 0.7$ eV, $d_{BE} = 7$ Å, $\Phi_C = 0.5$ eV, $d_{CB} = 12$ Å. In a 1 μm² device, $r_{in} \cong 2\,\Omega$ for $V_{BE} = 0.8$ V and $r_{out}(\min) \cong 15\,\Omega$ for $V_{CB} = 0.6$ V. This combination leads to $\alpha = 0.43$ when $L_0 = 200$ ÅeV² and consequently to $G \cong 1.3$ down to the range of 5×10^{-13} sec. If $L_0 = 1000$ ÅeV², $G \cong 2$ up to the upper frequency mentioned above. One should keep in mind that shrinking the area below 1 μm² will increase r_{in} and r_{out} to any convenient value without affecting the time constants.

As I mentioned in Section 2.2, α in principle could exceed 1. I considered the effect of heating up the base electrode without changing the temperature of emitter and collector[105]. T_{base} was allowed to change in the range (300–1000)°K and $T_{emitter} = T_{collector}$ in the range (4.2–300)°K. For $L_0 = 500$ ÅeV² and $V_{BE} = 0.8$ V, α changed in the range 0.83691–0.84169, which is a very small effect. This effect might be more pronounced if the electron bath in the base could be substantially "heated up" without affecting the electron temperature in the emitter and collector electrodes.

(b) *Practical difficulties.* The predicted performances of the MOMOM device were based on an ideal structure. By that I mean: a trapezoidal potential barrier, a pure elastic tunneling process, metal electrodes which are perfect conductors, smooth and parallel interfaces, and negligible quantum mechanical reflections when $E_x > \Phi_C$. Obviously, a real structure does not possess all these merits. I would like to describe a few of the more difficult problems which one might expect to face during an actual fabrication and test process.

Emitter region. The oxide region of the emitter is the thinnest layer in the device, yet is required to support a potential difference of ~1 V and supply a monoenergetic electron beam. A few difficulties are:

Non-tunneling current. Due to the thinness of the oxide transfer of electrons via pinholes and metallic microbridges can become dominant.

When few monolayers of a native oxide are grown thermally on top of a metal electrode, pinholes, due to the oxide-metal lattice mismatch, are most probably in-

evitable. Gray explained properties of thin oxide MOM's by invoking the conduction mechanism through microbridges in the oxide barrier[106]. However, it would be difficult to explain the operation of the point-contact MOM's by this mechanism, because of the very small area at the contact and the simple process of making the contact. Since evaporated thin oxide MOM's—where special care is exercised in oxide growth process—exhibit similar high frequency characteristics[7], I would tend to believe that in both kinds of junctions, microbridge conduction is not a dominant effect[107]. The validity of the tunnel model can be checked by looking for the super-conductive gap when tunneling from a normal metal into a superconductive one. If current transport through the oxide is via metallic microbridges the ejected current distribution will not be monoenergetic. It will have the same form as the density of states in the emitter), and α will be limited by $1-\Phi_C/eV_{BE}$.

Inelastic tunneling. The tunneling current can be enhanced in an undesirable way if electrons are allowed to interact with resonance centers in the oxide during the tunneling process. Then, the tunneling process can be looked upon as a two-step process with energy loss $\hbar\omega$ (where ω is the resonant frequency of the centers), and enhanced overall tunneling probability. Electrons can interact with vibrational modes of trapped molecules in the oxide, with adsorbed molecules in the oxide-metal interface, with phonons at the metal, with surface plasmons at the oxide-metal boundary, and more. Characteristically, this phenomenon can take place only when $eV_{\text{bias}} \geq \hbar\omega$[111]. For the pioneering work of Lamb and Jaklevic refer to Refs. [112, 113].

In an inelastic process, the ejected current into the base is not monoenergetic with an energy peak near E_{FE}, but is rather peaked near $eV_{BE}-\hbar\omega$. In a process which involves an inelastic tunneling path, α will be limited by the ratio of the elastic component to the total current.

Clearly, resonant centers are undesirable.

Oxide's impurities and traps. Imperfections in the oxides and incorporation of impurities will lead to trapping centers and fluctuations in the barrier height. Trapping sites will reduce the collected current and similarity to the case of charged impurities, the trapped charge will modify the barrier height. This is of vast importance in the collector region where small fluctuations in Φ_C can reduce α considerably.

Base region. The base is an isolation screen between input and output and is required to supply an appropriate potential difference to the emitter and collector. Due to its thinness, few problems arise:

Base pinholes. This subject was discussed in some detail in Section 2.3.

Base spreading resistance. If α is close to unity and r_{out} is very large, the resistance of the base metal is not a major problem since $I_B \cong 0$. Since most probably this won't be the case, the finite spreading resistance of the base will lead to a negative feedback between collector and emitter, and make V_{BE} current dependent.

The actual spreading resistance depends on the thickness of the base, the quality of the deposited metal and the actual geometry of the base. Obviously, to minimize its effects it should be much smaller than r_{in}, and for $r_{\text{in}} \cong 50\,\Omega$ it is not a minor task. Finding a metal with a long mean free path for hot electrons will alleviate this problem, since the base could be made thicker without causing a reduction in α. For example, if the specific resistance of the metal is $3\,\mu\Omega$ cm and its thickness is 200 Å, and the area is $(1\,\mu\text{m})^2$, the base spreading resistance is less than $1.5\,\Omega$.

Skin effects in base. Similarly to base spreading resistance, high frequency operation will lead to a finite amount of coupling between collector and emitter.

Mean free path and Φ_C. If an MOMOM is fabricated proceeding from the emitter toward the collector, the metal base is required to (a) have a long mean free path (mfp) for hot electrons, and (b) have a low barrier height Φ_C with its native oxide. These two requirements are not necessarily compatible. Other processes of fabrication might be needed, for example, the base could be constructed out of two layers: the wider with a long mfp and the shorter with a low Φ_C with its native oxide.

Collector region. The main requirement from the collector is $\Phi_C \cong 0.5$ eV. Few published results on metal-oxide barrier heights suggest that mostly $\Phi \sim 2$ eV. Consequently, a major problem remains.

In addition to these difficulties, I'd like to mention few general problems:

Edge effects. A standard configuration of an MOM junction exploits two metal strips overlapping each other. As is known, most of the current is contributed by the overlapping edges[7, 114]. This mechanism of current transport is not necessarily via tunneling and its cause is not known. Masking the edges minimizes the effect[7, 114], but might be a rather difficult task in very small area junctions.

Interface states. An interface between dissimilar materials will have interface states which might serve as trapping centers and reduce the transfer ratio. Cleanliness of each interface and best lattice match between the heterojunction materials will minimize those effects.

Geometrical realization of signal coupling. When frequencies are high, the junction can be coupled into by terminating a waveguide which guides the energy. Since the active area of the junction is very small, some coupling mechanism is necessary to maximize the input signal, while leaving the output circuit uncoupled from the input signal (a "concentrator" or some kind of antenna is necessary[115]).

In the next section, geometrical realizations in light of the main requirements will be discussed.

(c) *Proposed configurations.* An amplifier which operates in the millimeter or optical regime should have the following features: (a) compatability with conventional circuitry, like strip lines and integrated optical circuits, (b) high input and output coupling efficiencies, and (c) good insulation between input and output circuits.

In the following, I'm proposing few integrated configurations of the MOMOM device which can be adapted to the i.r. and visible regimes.

The basic structure is the "Edge MOM" which was developed recently[7]. The fabrication process utilizes

standard optical lithography techniques and achieves effective area of tunneling which is comparable to that of the point contact configuration. One form of the device is described in Fig. 13 (top). An overlap between a 1 μm metal strip and an oxidized edge of a thin metal layer (~ 100 Å thick) forms a junction with an area of tunneling on the order of 10^{-10} cm^2 [117]. The configuration at the bottom of Fig. 13 describes an MOMOM configuration which is a hybrid between "small" (10^{-10} cm^2) and "large" (10^{-8} cm^2) area junctions. This proposal involves a combination of in plane and out of plane radiation coupling schemes. Light can be coupled via a horn shaped dielectric waveguide [118] into the emitter junction. In the focusing process the electric field will intensify as the light propagates into the junction. The base electrode screens the output from the input signal, so that the propagating field along the long wire antenna can result only from a reexcitation by electrons which traverse the base-collector oxide. The long wire antenna is only one means of antenna coupling into free space or the substrate. Figure 14 describes a more advanced configuration where input and output signals are guided into and out from two Edge MOM devices. Cross cuts through the two waveguides in the device are shown also in Fig. 14. The input and output coupling is done via guiding in horn shaped waveguides. The output waveguide replaces the antenna which was previously utilized and enables inplane signal processing. In both examples the output loads should have a relatively high impedance to satisfy the gain requirements. In the antenna case (Fig. 13), some matching mechanism might be necessary, and in the waveguide case (Fig. 14), a high quality factor cavity which in turn is coupled into a waveguide will increase the load impedance substantially. The later configuration is more desirable but its realization is more difficult.

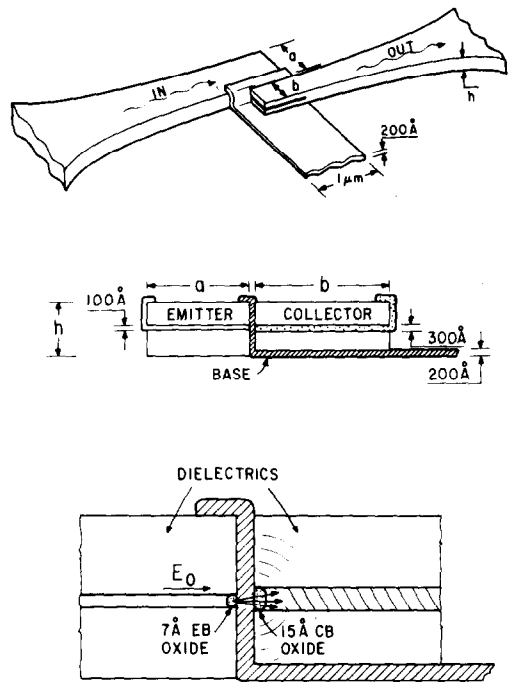

Fig. 14. A "double edge MOMOM". Signal processing is in the plane of the substrate. Typical dimensions are prescribed.

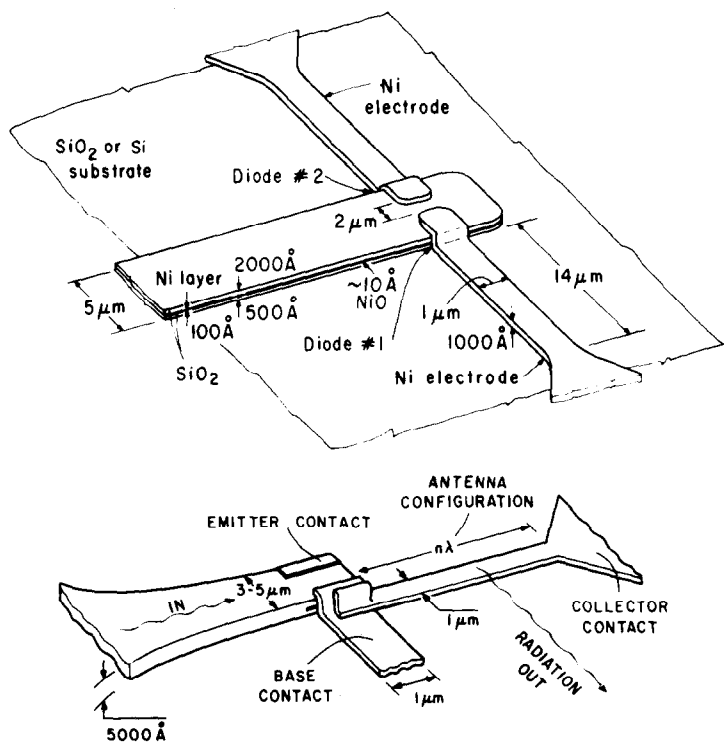

Fig. 13. Utilization of "Edge MOM" device in an MOMOM configuration. Two Edge diodes are described at top, with a tunneling area of 10^{-10} cm^{2c}. In this example, the junctions are symmetric (N_i–N_iO–N_i). For details refer to Ref. [7]. A "hybrid MOM" is shown at the bottom. Light is guided into an edge MOM (emitter) and coupled out via a large area (collector) MOM and a long-wire antenna.

At microwave frequencies, conventional techniques can be adopted with high coupling efficiencies, as was recently demonstrated by Slayman[119]. Coupling of ~72 GHz signal into a "large area" MOM which was imbedded inside a waveguide, had led to sensitivities as high as demonstrated by Schottky barriers.

Exploiting electron beam lithography "sandwich structure" configurations can be fabricated with areas which are comparable with those of point contact configurations. Using molecular beam epitaxy (MBE) techniques single crystal metals can now be grown, with presumably a longer mean free path for hot electrons and higher conductivity. Using MBE techniques, non-native oxides, semiconductors and insulators in crystalline forms can replace the native barriers in MOM's, enabling one to achieve impurity free barriers with a controlled barrier height.

These techniques are available today, and make the MOMOM device an extremely attractive device in frequency regions where no other active device is available.

3.3 *The MOMS*

In the MOMS the collector is a Schottky barrier. Barrier height values in metal-semiconductor junctions are well tabulated[80], and a proper Φ_C can always be achieved. The main disadvantage of the MOMS in comparison to the MOMOM is the large thickness of the depletion layer. This in turn will reduce the transfer ratio and limit the upper frequency of operation (this problem might be alleviated somewhat as discussed in Section 5).

The frequency performance was discussed in Section 2.3, and we may expect amplification up to at least 10 μm wavelength. In conjunction with semiconductor-metal-semiconductor (SMS) structures research, Crowel and Sze looked into the backscattering of hot electrons from the collector's depletion layer[120]. Taking into account the generation and absorption of optical phonons, one can write

$$\frac{1}{\lambda_x}=\frac{1}{\lambda_{ax}}+\frac{1}{\lambda_{gx}} \tag{21}$$

where λ_x is the x oriented mean free path and λ_{ax} and λ_{gx} stand for the associated mean free path with absorption and generation of phonons by the hot electrons. The collision time associated with phonon absorption

$$\frac{1}{\tau_a}=A(E_K+\hbar\Omega)^{1/2}N \tag{22}$$

where E_K is the electron kinetic energy, $\hbar\Omega$ is the phonon energy, $N=(e^{x_0}-1)^{-1}$ is the number of phonons where $x_0=\hbar\Omega/k_BT$, and A is a proportionality constant which depends on the matrix element responsible for the absorption process. Similarly,

$$\frac{1}{\tau_g}=A(E_K-\hbar\Omega)^{1/2}(N+1) \tag{23}$$

where τ_g is the collision time associated with phonon generation. The energy terms come about from a simple density of state arguments (see also Section 2.2). Since $v_x=(2E_{Kx}/m^*)^{1/2}$ where v_x and E_{Kx} are the x oriented velocity and kinetic energy respectively (assuming ellipsoidal bands),

$$\frac{1}{\lambda_x}=\frac{1}{L}\left[\left(\frac{E_K+\hbar\Omega}{E_{Kx}}\right)^{1/2}N+\left(\frac{E_K-\hbar\Omega}{E_{Kx}}\right)^{1/2}(N+1)\right] \tag{24}$$

where L is the mean free path associated with generation of phonons. Crowel and Sze[120] calculated λ_{ax}, λ_{gx} and the transmission coefficient in an SMS structure for a Si and GaAs collectors. Assuming a Maxwellian distribution in the transverse direction ($E_K-E_{Kx}=k_BT$) they found that at low E_{Kx}, λ_{ax} limits λ_x and at higher energies λ_{gx} is the limiting parameter. Neglecting quantum mechanical reflections and assuming e^{-E_e/k_BT} type current distribution, where E_e is the minimum total energy of the electrons ejected from the emitter, they found that the transmission is fairly constant with the increase of E_e in Si (50–80%) and in GaAs (20–50%), depending on the phonon energies.

Even though the ejected current distribution in the MOMS case is different, we can estimate the transfer ratio due to electrons backscattering to be around 50%.

Quantum mechanical reflections diminish fast for an increase in E_x above Φ_C. For small m^* in the semiconductor, reflections of less than 20% are predicted for $E_x-\Phi_C>0.1$ eV. For more information the reader should refer to Crowel and Sze[121].

The above estimations will change with a different choice of model parameters. I would expect to get similar transfer ratios for the MOMS and MOMOM.

Practical configurations can involve schemes similar to the suggested ones in the previous section. Note, that due to the present lack of knowledge on the deposition of single crystal semiconductors on metals and oxides, fabrication should preferably proceed from collector toward emitter.

3.4 *A heterojunction device*

The above devices combine metals, semiconductors and insulators in a layered form. These structures are inherently difficult to fabricate without interface problems which result from lattice mismatch and the exposure to the ambient in between layers deposition. In today's state of the art very little is known about epitaxial growth of semiconductors on metals, metals on oxides, oxides on metals etc., and much less is known on the quality of the interfaces and ways to improve them.

The recent advancement in molecular beam epitaxy systems (MBE)[122–124] led to a great advancement in the deposition capability of very thin epitaxial semiconductor films. Moreover, the MBE system serves primarily now as a research tool and *in situ* surface structures, interface quality and the like are investigated with much more ease routinely. A device which is composed entirely of semiconductors with similar lattice parameters is highly desirable.

Heterojunctions of GaAs–$Ga_{1-x}Al_xAs$ and GaAs–Ge are two examples of fairly well studied combinations. Let's focus our attention on the first combination. The lattice constant of GaAs is 5.654 Å while that of AlAs ($x = 1$) is 5.661 Å. Upon increasing the Al content, $Ga_{1-x}Al_xAs$ changes from a direct (at $x \cong 0.32$) to an indirect gap semiconductor and its energy gap varies from 1.43 to 2.1 eV[125, 126]. Alferov *et al.*[127] concluded that in GaAs–$Ga_{1-x}Al_xAs$ heterojunction, the discontinuity in the valence band is very small, whereas that in the conduction band is close to the difference between the energy gaps of the heterojunction components, varying from 0 to 0.4 eV as the Al concentration increases up to about 80%[127, 128]. The conduction band discontinuity in an n–n heterojunction is described in Fig. 15(a). When the GaAs and $Ga_{1-x}Al_xAs$ layers are thinner than the depletion and accumulation layers the conduction band takes the form described in Fig. 15(b). On the basis of this band model, a heterojunction device is possible with controlled barrier heights which could function similarly to the MOMOM amplifier. Figure 16 describes an example of the proposed device. Tunneling occurs between GaAs layers through a potential barrier provided by a thinner layer of $Ga_{1-x}Al_xAs$. The Al content is chosen in the emitter and collector barriers to provide barrier heights of 0.4 and 0.2 eV respectively (or there about). $d_{BE} < d_{CB}$ is chosen to provide $r_{out} \gg r_{in}$. Emitter and collector are terminated by a heavily doped GaAs and a metal contact, a combination which minimizes voltage drop across the ohmic contact. For small ohmic contact and low spreading resistance in the base, the base region has to be heavily doped, which in turn, reduces the mean free path of hot electrons in it due to impurity scattering. Heavy doping in the base can be achieved by heavily doping the GaAlAs layers (modulation doping[129]. Electrons from the conduction band of the GaAlAs will "spill over" into the base which can be viewed as a quantum well, thus reducing the spreading resistance and the ohmic contact without degrading the mean free path of the hot electrons ejected from the emitter[130]. The collector and emitter materials will have also a sufficient carrier concentration (coming from the donors in the GaAlAs), and most of the potential will

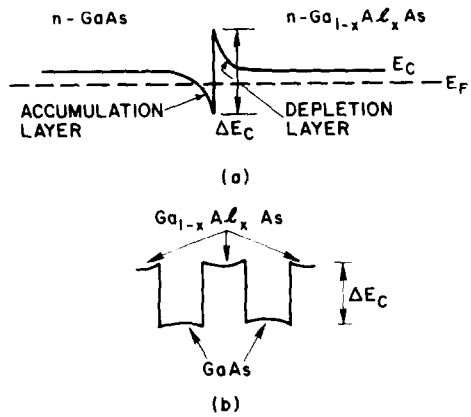

Fig. 15. (a) The conduction band edge in $Ga_{1-x}Al_xAs$–GaAs (N^+n^-) heterojunction. (b) A similar description for very thin layers combination.

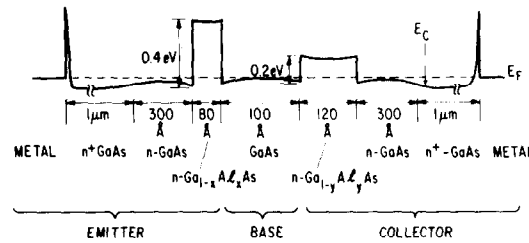

Fig. 16. A heterojunction GaAs–GaAlAs transfer amplifier with sample parameters. Ohmic contacts on emitter and collector sides are provided by a heavily doped Schottky barrier (the contact to the base is omitted).

drop across the GaAlAs, with a resultant band diagram like the one shown in Fig. 3. Note that even a carrier concentration of 10^{20} cm^{-3} in the base will result with base spreading resistance of a few hundred ohms. Consequently, R_{in} has to be in the kilohm range which in turn will increase the input time constant.

Since the tunneling effective mass in GaAlAs is about 0.1 m_0[131] similar tunneling current densities to the ones obtained in MOM's can be obtained for barrier thicknesses up to 150 Å or more. Current density and energy distribution are determined by the term m^*d, and the location of E_F above the conduction band edge. For small $E_{EF} - E_{CE}$, narrow energy distributions are expected. If $\Phi < E_G/2$ we may expect electrical characteristics which are similar to those of the MOMOM (with thicker epitaxial films which are grown with superior quality utilizing the MBE technique.).

One shouldn't ignore quantum-mechanical resonances in the case that $eV_{BE} < \Phi_C$. A double tunneling barrier like that was considered theoretically by Davis and Hosack[18] and experimentally tested by Chang *et al.*[133]. Resonances had been detected in the current transmission and a slight differential negative resistance was observed. When $eV_{BE} > \Phi_C$ (our working conditions), these effects are expected to be small.

In estimating the upper frequency limit of the device which is described in Fig. 16, one may substitute an average drift velocity of 10^7 cm/sec and total travel distance of about 900 Å, resulting in a transit time of $\sim 10^{-12}$ sec. The two n-regions in emitter and collector were chosen to be very narrow to minimize the transit time. As noted above, the base spreading resistance might ultimately limit the upper frequency of operation.

One may expect to find the narrow collector and emitter regions "over flooded" by electrons which spill over from the n^+ regions. If those regions are too narrow they might lead to a significant reduction and even elimination of the GaAlAs–GaAs interface potential barriers. For example, if the GaAs(n^+) layers are 5×10^{17} cm^{-3} doped, we may expect an accumulation layer in the GaAs layer, beyond the GaAs (n^+): GaAs boundary, as thick as (3–5)λ where λ is the Debye distance which in this case amounts to 53 Å[132]. Layers' thickness on the order of 300 Å are a compromise.

A practical problem is to make a low ohmic contact to the base. It is obvious that nonalloyed ohmic contacts have to be utilized, since in the alloying process diffusion of the top metal contact occurs down to a depth of a few

thousand Ångstroms. Low resistance nonalloyed ohmic contact can be achieved by heavily doping the base region in the contact region[134, 135].

3.5 *Non tunneling transfer devices*

In contrast with the above, where all THETA members employed tunneling, I would like to propose a few novel devices which generate the hot electron beam via ballistic electron transfer (this class will be named "BHETA" for "Ballistic Hot Electrons Transfer Amplifiers"). One known and studied example in this family is the Semiconductor-Metal-Semiconductor device (SMS). I'll briefly review its properties and the relevant difficulties in the next section. Then, a brief description of few possible novel alternatives will follow.

(a) *Semiconductor-Metal-Semiconductor (SMS).* Atalla and Kahng[136] and Gepper[137] had proposed the SMS amplifier in 1962. The most extensive work on this structure was done by Crowel and Sze[138]. Actual current transfer ratios which had been observed were about 0.3 (for Si–Au–Ge, Si–Ag–Ge, GaP–Au–Ge and GaP–Ag–Ge). No high frequency tests had been performed, but the predicted usable operational frequency was ~10 GHz[139], which reflected the state of the art then.

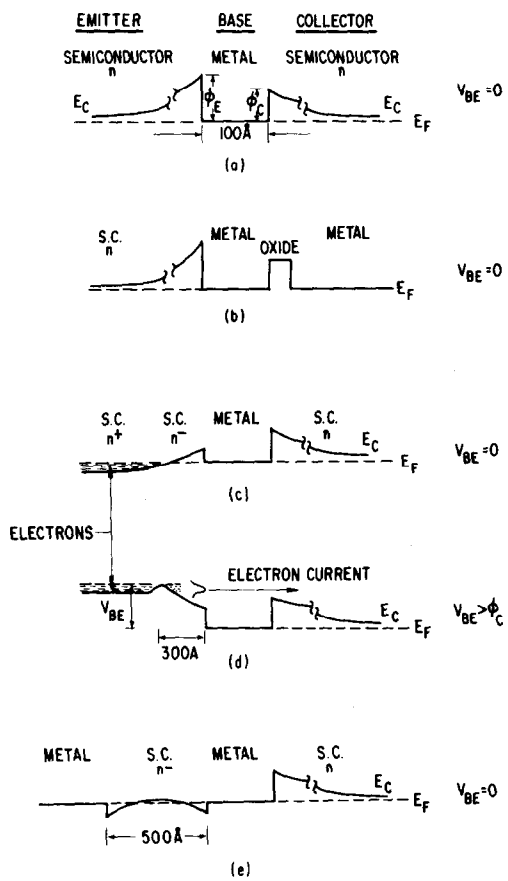

Fig. 17. Non-tunneling emitters hot electron amplifiers (BHETA). (a) The well known SMS. If $\Phi_E < \Phi_C$, I_B vs V_{BE} can exhibit a dynamic negative resistance. (b) An alternative SMOM, with similar features to the SMS. (c) Unbiased and (d) biased n^+n^--metal-semiconductor. Electrons which are emitted from the n^+n^--interface proceed ballistically to the collectors. (e) Similarly to (c), an unbiased MSMS structure. Emitter and base form a non-rectifying contact with an n^- semiconductor thin layer.

A simplified energy diagram is presented in Fig. 17(a). The barrier heights are chosen so that $\Phi_E > \Phi_C$. When $V_{BE} > 0$, electrons are thermally ejected above Φ_E, ballistically cross base metal and collected by the collector. In this structure the ejected electron beam has a Maxwellian distribution with a peak value at Φ_E. The transfer ratio is influenced by factors similar to those occurring in the MOMS device.

Frequency performance is determined by the total transient time from emitter to collector. Since the depletion layers' thickness is no less than ~1000 Å, one couldn't expect operation below the picosecond range. (Thinner depletion layers will permit tunneling, which is unacceptable).

Devices had been fabricated in the past by evaporating a thin base on a flat crystalline semiconductor and contacting the emitter at the top by lowering a "point contact" crystalline semiconductor. The metal films were low quality and unavoidable oxides formed on both sides of base metals. Recent progress in the deposition of metals on semiconductors in single crystal forms[140] raises hopes of utilizing the process of depositing semiconductors on top of metals, in a single crystal form. Thus an improved SMS structure seems feasible.

(b) *Semiconductor-Metal-Oxide-Metal (SMOM).* This configuration is described in Fig. 17(b). The ejecting contact is similar to that of the SMS, but the collector is an MOM. High impedance collector is easily achieved with an oxide thickness in the range of 25 Å, which still has a very short collector transit time. Metals have to be chosen so that $\Phi_C < \Phi_E \cong 0.7$ eV, which involves, again, a search for a low barrier height metal-oxide interface.

(c) *Ballistic emitter* (n^+n^-)*—Metal-Semiconductor.* The structure described in Fig. 17(c) and (d) is similar to the SMS structure. This proposal intends to overcome the long transient and charging times in the emitter. Changing the collector into an MOM junction will increase speed even further.

The emitter is constructed out of an n^+n^- semiconductors combination (the n^+ material doesn't necessarily have to be the same as the n^- semiconductor). The n^+ layer is terminated by an ohmic contact and supplies the ejected carriers, while the n^- region supports the potential difference supplied by the battery V_{BE}. The n^- layer is a wide gap material and is chosen thin enough so that (a) all carriers emitted from the emitter come from the n^+n^- boundary with a total energy eV_{BE} with respect to the base, and (b) these electrons will continue through the n^- region without suffering electron and phonon collisions, and will maintain their quasi-monoenergetic energy distribution[141]. The input I–V characteristic results from the ballistic nature of the current and is determined solely by the background doping in the n^- layer, space charge effects, and the n^+n^- boundary. A simplified treatment was published recently[142–144].

Due to the required thinness of the n^- layer, two important points should be considered: (a) Transfer ratio will increase only for $eV_{BE} > \Phi_C$ (see Fig. 17(d)). For

$\Phi_C \sim 0.5$ eV, strong electric fields ($\sim 10^5$ V/cm) and high current densities will persist in the emitter region under normal operation. This might restrict the operation mode of the device to short pulse operation. (b) Electrons from the n^+ region will most certainly modify the energy diagram in the vicinity of the base metal and suppress Φ_E. One should expect to have a differential negative resistance in the input as long as $\Phi_E < \Phi_C$, provided the temperature is low enough and the width of the ejected current distribution is not too wide.

If one assumes a space charge limited current which obeys Child's law, $J_{EB} = 44/9(2e/m^*)^{1/2}\epsilon V_{BE}^{3/2}/d_{BE}^2$, the equivalent dynamic resistance of the n^- layer $r_{in} \equiv \partial V_{BE}/\partial I_E \cong 10\,\Omega$ for GaAs ($\epsilon = 11.5\epsilon_0$, $m^* = 0.07\,m_0'$, $V_{BE} = 0.1$ V, $d_{BE} \cong 300$ Å and area 1 μm^2). The transient time through the emitter region can be estimated via $\tau = 3(m^*/2e)^{1/2} d_{BE}/V_{BE}^{1/2}$ which is in the low 10^{-13} due to the small effective mass in GaAs. The input time constant is in the same range. Due to the very small input resistance, one should minimize: the spreading resistance of the n^+ layer in the emitter, the base spreading resistance, and ohmic contact resistances (note that $r_{in}(DC) = 2/3 r_{in} \cong 7\,\Omega$ which constrains the spreading resistances even further).

(d) *Ballistic emitter (Metal-Semiconductor)–Metal-Semiconductor (MSMS).* Another ballistic emitter device is described in Fig. 17(e). The metal: n^+ combination in the emitter is replaced by a metal: n^- combination which forms an ohmic contact upon contact (for example, metal–CdS or metal–InAs combinations). A similiar nonrectifying contact is formed with the base metal. By eliminating the n^+ layer one might hope to reduce the emitter spreading resistance.

The n^- semiconductor (wide gap) should be thick enough to prevent degeneracy due to electrons diffusing from both contacts, but not too thick so that the ballistic behavior will be preserved. Biasing, again, should be in the range $eV_{BE} > \Phi_C$ and the device might be restricted to a short pulse operation mode.

The Schottky barrier which forms the collector can be replaced by a low-barrier MOM, a change which will push the upper frequency limit upwards.

4. CONCLUSIONS

The constant search for high frequency devices involves mostly the miniaturization of existing devices. The recent success of metal-oxide-metal junctions as i.r. classical detectors and mixers, revives a new interest in tunnel emitter—metal base devices and analogous structures.

In this paper I've reviewed the basic properties of MOMOM's and MOMS's and have shown that, provided few difficulties could be surmount, they might play an important role in the future of high speed devices, up to frequencies in the 10^{13} Hz range. These amplifiers can be utilized in CW amplification, in switching modes, and negative resistance amplification and frequency generation modes.

The weakest link in our knowledge today are suitable input and output coupling schemes which guarantee a high ratio between the load and the input impedances.

Exploiting similar band diagrams, I've proposed a double heterojunction amplifier which relies in its fabrication on the capabilities of the MBE technique.

Brief descriptions of non-tunneling emitter hot electron devices, which can operate at frequencies as high as the 10^{13} Hz regime, are described later.

Fabrication of these devices (THETA and BHETA), requires state of the art of technology, like: MBE techniques, new oxidation techniques, sub-microfabrication, new waveguiding and cavity designs and more. A joint effort by material and device researchers is necessary for progress in this field.

Acknowledgement—I wish to thank M. Nathan for his valuable suggestions and for reading the manuscript.

REFERENCES

1. J. Matisoo, *Proc. Low Temp. Phys. Conf. LT*15, Grenoble (1978).
2. H. Fukui, *Proc. IEEE.* University/Industry/Government Microelectronic Symposium (1979).
3. S. F. Jacobs, M. Sargent III and M. O. Schully (Editors), *Physics of Quantum Electronics*, Vol. 5 (1978).
4. Y. W. Chan, *Phys. Rev. Lett.* **42**, 92 (1979); H. Schwartz, *Phys. Rev. Lett.* **42**, 1141 (1979).
5. A. Yariv and D. R. Armstrong, *J. Appl. Phys.* **44**, 1664 (1973); A. Yariv and C. C. Shih, *Optics Commun.* **24**, 233 (1978); M. M. Schoucri, *J. Appl. Phys.* **50**, 702 (1979).
6. L. L. Chang and L. Esaki, Proc. of. of the Yamada Conf. on Electronic Properties of 2-d Systems, Japan. *Surface Science* **98**, 70 (1980).
7. M. Heiblum, S. Y. Wang, J. R. Whinnery and T. K. Gustafson, *IEEE J. Quant. Electron.* **QE-14**, 159 (1978); M. Heiblum, Ph.D Thesis, U.C. Berkeley (1978).
8. S. E. Schwarz, *IEEE J. Quant. Electron.* **QE-10**, 62 (1974).
9. C. A. Mead, *Proc. IRE* **48**, 359 (1960).
10. C. A. Mead. *J. Appl. Phys.* **32**, 646 (1961).
11. J. P. Spratt, R. F. Schwarz and W. M. Kane, *Phys. Rev. Lett.* **6**, 341 (1961).
12. R. N. Hall, *Solid-St. Electron.* **3**, 820 (1961).
13. J. M. Lavine and A. A. Iannini, *Solid-St. Electron.* **5**, 109 and 273 (1962).
14. R. H. Handy, *J. Appl. Phys.* **37**, 4620 (1966).
15. J. M. Lavine, *SCP and Solid-St. Tech.* 17 (1969).
16. J. L. Moll, *IEEE Trans. Electron. Dev.* **ED-10**, 299 (1963).
17. M. M. Atalla and R. W. Soshea, *Solid-St. Electron.* **6**, 245 (1963).
18. R. H. Davis and H. H. Hosack, *J. Appl. Phys.* **34**, 864 (1963).
19. S. Takaba, K. Yasui and S. Kaneda, *Appl. Phys. Lett.* **31**, 636 (1977).
20. R. H. Fowler and L. W. Nordheim, *Proc. Royal Soc.* **A119**, 73 (1928); L. W. Nordheim, *Proc. Royal Soc.* **A121**, 626 (1928).
21. If one calculates the total energy distribution, an "acceptance cone" with a limiting angle with respect to the normal Θ_{max}, which fulfills $E\cos^2\Theta_{max} = \Phi_C$ has to be considered.
22. C. B. Duke, *Tunneling in Solids.* Academic Press, New York (1969).
23. K. H. Gundlach and J. G. Simmons, *Thin Solid Films* **4**, 61 (1969).
24. The above calculation is by no means accurate, but serves the purpose of giving the feeling for the energy distribution of the tunnel current. Even if other factors will be included, it is difficult to predict that they will agree better with experiments, due to unknown factors like: fluctuations in oxide thickness, metal-oxide boundary, effective mass of tunneling electrons, unknown barrier shape and height, the validity of the image correction concept, and the dielectric constant of the thin oxide which is applicable to the tunneling process.

25. R. D. Young and E. W. Müller, *Phys. Rev.* **113**, 115 (1959).
26. R. Gomer, *Field Emission and Field Ionization.* Harvard University Press, Cambridge (1961).
27. K. H. Gundlak, *J. Appl. Phys.* **44**, 5005 (1973).
28. J. G. Simmons, *J. Appl. Phys.* **34**, 1793 (1963).
29. Collision with acoustical phonons is neglected due to the small probability of collision. The oxide is assumed to be free of impurities and any lattice imperfections.
30. O. L. Nelson and D. E. Anderson, *J. Appl. Phys.* **37**, 66 (1966).
31. R. E. Collins and L. W. Davis, *Solid-St. Electron.* **7**, 445 (1964).
32. E. E. Huber, Jr., F. E. Johnston, Jr. and C. K. Kirk, Jr., *J. Appl. Phys.* **39**, 5104 (1968).
33. H. Hanter and W. A. Feibelman, *J. Appl. Phys.* **33**, 3580 (1962).
34. This is a representative case of $\Phi_c \cong 2$ eV and $eV_{BE} \cong 4$ eV.
35. J. J. Quinn, *Phys. Rev.* **126**, 1453 (1962).
36. J. J. Quinn, *Appl. Phys. Lett.* **2**, 167 (1963).
37. K. Motizuki and M. Sparks, *J. Phys. Soc. Japan* **19**, 486 (1964).
38. S. L. Adler, *Phys. Rev.* **130**, 1654 (1963).
39. S. L. Adler, *Phys. Rev.* **141**, 814 (1966).
40. C. A. Bates, *Tech. Rep. No.* 216-1, Stanford Electronics Lab. (1963).
41. R. H. Ritchie and J. C. Ashley, *J. Phys. Chem. Solids* **26**, 1689 (1965).
42. R. W. Davis, *Phys. Rev.* **181**, 1118 (1969).
43. L. Kleinman, *Phys. Rev. B* **3**, 2982 (1971).
44. C. A. Kubkonen and H. Smith, *Phys. Rev. B* **8**, 4601 (1973).
45. Y. B. Levinson, *Solid-St. Electron.* **21**, 923 (1978).
46. C. A. Mead, *Phys. Rev. Lett.* **8**, 56 (1962).
47. R. E. Collins and L. W. Davis, *Appl. Phys. Lett.* **2**, 213 (1963).
48. E. D. Savage and D. E. Anderson, *J. Appl. Phys.* **38**, 3245 (1967).
49. H. Kanter, *J. Appl. Phys.* **34**, 3629 (1963).
50. H. Kanter, *J. Appl. Phys.* **34**, 12 (1963).
51. J. S. Escher, P. E. Gregory and T. J. Malony, *J. Vac. Sci. Tech.* **16**, 1394 (1979).
52. J. Antula, *J. Appl. Phys.* **43**, 1830 (1972).
53. W. Gericke and W. Ludwig, *Phys. Stat. Solid (a)* **19**, 583 (1973).
54. C. Guinet, *Appl. Phys. Lett.* **25**, 600 (1974).
55. S. M. Sze, C. R. Crowel, G. P. Carey and E. E. LaBate, *J. Appl. Phys.* **37**, 2690 (1966).
56. C. N. Berglund and W. E. Spicer, *Phys. Rev.* **136**, A1044 (1964).
57. C. R. Crowell, W. G. Spitzer, L. E. Howarth and E. E. LaBate, *Phys. Rev.* **127**, 2006 (1962).
58. H. Thomas, *Z. Physik* **147**, 395 (1957).
59. G. A. Katrich and O. G. Sarbei, *Soviet Phys. Solid-St.* **3**, 1181 (1961).
60. R. W. Soshea and R. C. Lucas, *Phys. Rev.* **138**, A1182 (1965).
61. R. N. Stuart, F. Wooten and W. E. Spicer, *Phys. Rev. Lett.* **10**, 7 (1963).
62. H. Kanter, *Phys. Rev.* **B1**, 522 (1970).
63. R. N. Stuart and F. Wooten, *Phys. Rev.* **156**, 364 (1967).
64. W. F. Kralikowski and W. E. Spicer, *Phys. Rev.* **B1**, 478 (1970).
65. J. Feisner, P. Roboz and P. B. Barna, *Phys. Stat. Solid (a)* **4**, K187 (1971).
66. L. Eckertova, J. Dvorak and J. Boclk, *Thin Solid Films* **36**, 171 (1976).
67. J. Raby and A. Collins, *Thin Solid Films* **20**, 227 (1974).
68. C. R. Crowell, W. G. Spitzer and H. G. White, *Appl. Phys. Lett.* **5**, 215 (1964).
69. L. B. Leder, M. E. Lasser and D. C. Rudolph, *Appl. Phys. Lett.* **1**, 3 (1962).
70. C. R. Crowell and S. M. Sze, *Phys. Rev. Lett.* **15**, 659 (1965).
71. S. M. Sze, J. L. Moll and T. Sugardo, *Solid-St. Electron.* **1**, 509 (1964).
72. W. G. Spitzer, C. R. Crowell and M. M. Atolla, *Phys. Rev. Lett.* **8**, 57 (1962).
73. M. Geddo, A. Stella and G. Soncini, *Phys. Stat. Solid (a)* **47**, 223 (1978).
74. C. R. Crowell and S. M. Sze, *Solid-St. Commun.* **8**, 673 (1965).
75. L. A. MacColl, *Phys. Rev.* **56**, 699 (1939).
76. L. A. MacColl, *Bell. Syst. Tech. J.* **30**, 888 (1951).
77. J. Cohen, *J. Appl. Phys.* **33**, 1777 (1962).
78. J. Cohen, *Appl. Phys. Lett.* **1**, 61 (1962).
79. W. Hess and R. Johannes, *Brit. J. Appl. Phys.* **14**, 287 (1963).
80. For example, Φ(Al–Si) ≅ 0.6 eV, Φ(Au–Si) ≅ 0.8 eV, Φ(W–Si) = 0.65 eV, Φ(Au–GaAs) = 0.9 eV. For more examples refer to A. G. Milnes and D. L. Feucht, *Heterojunctions and Metal-Semiconductors Junctions*, Chap. 6. Academic Press, New York (1972).
81. From Ref. [74], $l_{e\text{-}ph}$(GaAs) = 15 Å, $l_{e\text{-}ph}$(Si, Ge) = 65 Å.
82. M. L. Tarng and G. K. Wehner, *J. Appl. Phys.* **43**, 2268 (1972).
83. D. Kahng, *Proc. IRE* **50**, 1534 (1962).
84. W. Ludwig, G. V. Reinsperger and W. Gericke, *Thin Solid Films* **36**, 415 (1975).
85. J. Antula, *Thin Solid Films* **13**, 93 (1972).
86. Mead (Ref. [10]) claims to have $\alpha \lesssim 10\%$. Repetitions of similar experiments resulted with much lower α (Refs. [30, 32]). It is likely that Mead's high current transfer resulted by some other transfer mechanisms rather than by hot electrons emission and collection.
87. Their calculations are based on the Fowler–Nordheim formula which is applicable only at the high bias regime. For thin oxides and low biasing cases, eqn (6) has to be used.
88. D. J. E. Knight and P. T. Woods, *J. Phys.* **E9**, 898 (1976).
89. J. J. Jimenez and R. Petersen, *Infrared Phys.* **17**, 541 (1977).
90. K. C. Liu, C. Davis, Jr. and A. Javan, *Phys. Rev. Lett.* **43**, 785 (1979).
91. It is likely that Refs. [16, 17] considerably reduced the enthusiasm of researchers in the field, which was quite flourishing in the sixties.
92. D. P. Sue and T. K. Gustafson, *Appl. Phys. Lett.* **31**, 71 (1977).
93. P. K. Tien and J. P. Gordon, *Phys. Rev.* **129**, 647 (1963).
94. J. R. Tucker and M. F. Millea, *Appl. Phys. Lett.* **33**, 611 (1978); J. R. Tucker, *IEEE J. Quant. Electron.* **QE-15**, 1234 (1979).
95. H. R. Feterman, B. J. Clifton, P. E. Tannenwald, C. D. Parker and H. Penfield, *IEEE Trans. Microwave Theory Tech.* **MTT-22**, 1013 (1974).
96. K. Mizuno, R. Kuwahara and S. Ono, *Appl. Phys. Lett.* **26**, 605 (1975).
97. D. T. Hodges and M. McColl, *Appl. Phys. Lett.* **30**, 5 (1977).
98. A. van der Ziel, *J. Appl. Phys.* **47**, 2059 (1976).
99. D. W. Tsang and S. E. Schwarz, *J. Appl. Phys.* **50**, 3459 (1979).
100. When $\hbar\omega > \Phi$, internal photoemission starts taking place. Also Tucker–Millea theory[94] predicts the failure of the classical model,
101. When metal deviates from a perfect conductor, thermal losses start taking place and gain is reduced via a feedback from output to input through the base metal.
102. L_0 has the units of $[\text{Å}(eV)^2]$. Physically, it is the mean free path of hot electrons with excess energy of 1 eV above the Fermi level.
103. The transconductance is defined $g_m \equiv \partial I_C / \partial V_{BE}$.
104. It should be pointed out that a convenient utilization of I_B vs V_{BE} curve is possible in the common emitter configuration. If V_{CE} is kept constant while V_{BE} changes, V_{CB} will decrease with the increasement of V_{BE}, resulting in a less profound negative resistance.
105. This might be possible only with vacuum spaces replacing the oxides.
106. K. E. Gray, *Solid-St. Commun.* **13**, 1787 (1973).
107. An oxidation process which I've adapted employes a

simultaneous sputter-etching and oxidation utilizing an argon-oxygen plasma in a RF sputtering system. The process is time independent when sputter-etching and oxidation proceed with equal rates. The resulted oxides are pinhole free and very reproducable. For more detail see Refs. [108, 109].
108. J. H. Greiner, *J. Appl. Phys.* **42**, 5151 (1971).
109. J. H. Greiner, *J. Appl. Phys.* **45**, 32 (1974).
110. If one electrode becomes easily superconductive (like Pb or Nb), the energy gap in it can be observed when the normal metal is biased negatively with respect to the superconductive one.
111. Only if $eV_{bias} > \hbar\omega$ the tunneling electron can lose energy $\hbar\omega$ and still find an empty level in the collecting electrode (namely, still have $E > E_F$).
112. R. C. Jaklevic and J. Lamb, *Phys. Rev. Lett.* **17**, 1139 (1966).
113. J. Lamb and R. C. Jaklevic, *Phys. Rev.* **165**, 821 (1968).
114. S. R. Pollack and C. E. Morris, *J. Appl. Phys.* **35**, 1503 (1964).
115. The second configuration is adapted from Ref. [7]. Another variation is described in Ref. [116].
116. Y. Yasuoka, M. Heiblum and T. K. Gustafson, *Appl. Phys. Lett.* **34**, 823 (1979).
117. Exploiting advanced lithography techniques, the width of the metal strip can be reduced to 1000 Å or less.
118. R. K. Winn and J. H. Harris, *IEEE Trans. Microwave Theory Tech.* **MIT-23**, 92 (1975).
119. C. Slayman, Ph.D Thesis, Univ. California, Berkeley (1979).
120. C. R. Crowel and S. M. Sze, *Solid-St. Electron.* **8**, 979 (1965).
121. C. R. Crowel and S. M. Sze, *J. Appl. Phys.* **37**, 2683 (1966).
122. A. Y. Cho and J. R. Arthur, *Prog. Solid-St. Chem*, Vol. 10, Part 3, 157 (1975).
123. L. L. Chang and R. Ludeke, *Epitaxial Growth* (Edited by J. W. Matthews), Part A, p. 37. Academic Press, New York (1975).
124. A. Cho, *J. Voc. Sci. Technol.* **16**, 275 (1979).
125. H. C. Casey, Jr. and M. B. Panish, *J. Appl. Phys.* **40**, 4910 (1969).
126. The commonly used expression today for the direct gap variation is $E_G(x) = 1.424 + 1.266x + 0.26x^2$ and can be found in the following reference: O. Berolo and J. C. Woodley, *Can. J. Phys.* **49**, 1335 (1971).
127. Zh. I. Alferov and O. A. Ninua, *Sov. Phys.-Semicond.* **4**, 519 (1970).
128. R. Dingle, A. C. Gossard and W. Wiegmann, *Phys. Rev. Lett.* **34**, 1327 (1975), with the conclusion, $\Delta E_C = (0.85\text{–}0.88)\Delta E_G$ where $\Delta E_G = E_G(Ga_{1-x}Al_xAs) - E_G(GaAs)$.
129. R. Dingle, H. L. Stormer, A. C. Gossard and W. Wiegmann, *Appl. Phys. Lett.* **33**, 665 (1978). Note that for a doping level of $10^{18}\,cm^{-3}$, the average spacing between donor sites is ~ 100 Å, which is larger than the width of the GaAlAs layers. We should expect potential fluctuations in these regions which will alter the barrier heights.
130. Abrupt doping profiles are possible only by using the MBE technique, which is a low temperature process and minimizes diffusion.
131. D. J. Stukel and R. N. Euwema, *Phys. Rev.* **188**, 1193 (1969); Q. H. F. Vreben, *J. Phys. Chem. Sol.* **29**, 129 (1968).
132. This is a result of a numerical analysis which was suggested by P. Price.
133. L. L. Chang, L. Esaki and R. Tsu, *Appl. Phys. Lett.* **24**, 593 (1974).
134. Ohmic contacts with specific contact resistivity of $10^{-6}\,\Omega\,cm^2$ were achieved. P. A. Barnes and A. Y. Cho, *Appl. Phys. Lett.* **33**, 651 (1978); J. V. DiLorenzo, W. C. Nichaus and A. Y. Cho, *J. Appl. Phys.* 951 (1979).
135. By epitaxially growing Ge on top of GaAs and evaporating the metal electrode on it, ohmic contacts with specific resistivity of $5 \times 10^{-7}\,\Omega\,cm^2$ were reported. R. Stoll, C. E. C. Wood, K. Boards and L. F. Eastman, *Electron. Lett.* **15**, 800 (1979).
136. M. M. Atalla and D. Kahng, *A New Hot Electron Triode Structure with Semiconductor-Metal Emitter.* IRE–AIEE, Solid State Device Research Conference (1962).
137. D. V. Geppert, *Proc. Inst. Radio Engrs* **50**, 1527 (1962).
138. C. R. Crowel and S. M. Sze, In *Physics of Thin Films* (Edited by G. Hass and R. E. Thun), Vol. 4, p. 325. Academic Press, New York (1967).
139. S. M. Sze and H. K. Gummel, *Solid-St. Electron.* **9**, 751 (1966).
140. A. Y. Cho and P. D. Dernier, *J. Appl. Phys.* **49**, 3328 (1978).
141. The electrons' minimum energy and their distribution peak is at energy eV_{BE}. Toward higher energies their density decays in a Maxwellian fashion.
142. M. S. Shur and L. F. Eastman, *IEEE Trans. Electron. Dev.* **ED-26**, 1677 (1979).
143. T. J. Maloney, *IEEE Electron. Device Lett.* **EDL-1**, 54 (1980).
144. J. J. Rosenberg, E. J. Yoffa and M. I. Nathan, submitted for publication.