The Switching Characteristics of a Tunnel Diode

Byung Chan Kim
Department of Physics, Kyungpook National University
(Received September 20, 1965)

ABSTRACT

The switching time, the necessary time for a tunnel diode to change its junction voltage from low to high or from high to low, of a tunnel diode (T.D.) is evaluated by dividing the voltage-current (V-I) characteristic curve of a T.D. into four regions. The switching time for each region is calculated, and it is compared with the value observed, operating the T.D. as a monostable multivibrator. A voltage-pulse-height discriminator is described, and the threshold equation is given. The linearity, pulse resolving ability, and the temperature characteristics are studied. The threshold variation due to the ambient temperature variation is attributed to the peak current variation.

Introduction

When the impurity concentration of a semiconductor is of the order of 10^{15}/c.c., only a small fraction of the states in the conduction band of an n-material are occupied, and also a small fraction of the states in the valence band of a p-material are empty. However, if both materials are heavily doped up to the impurity concentration of the order of 10^{19}/c.c.—10^{20}/c.c., all the states near the bottom of the conduction band of n-material are occupied by electrons and all the states near the top of the valence band of p-material are empty. The energy level diagram for this kind of heavily doped p-n junction is shown in Fig. 1. Normally, the forbidden band effectively isolates the electrons on the two sides of the junction because it presents a potential barrier that the electrons cannot surmount since they do not have sufficient energy. If, however, the junction region is thin enough (100Å~150Å), it follows from quantum mechanical tunneling effect that there is a finite probability that electrons originally on one side of the junction can appear on the other at the same energy. The tunneling current through this kind of p-n junction, when the junction is slightly biased, is given by L. Esaki[1];

$$I=I_{c\to v}-I_{v\to c}=A\int_{E_c}^{E_v}\{f_c(E)-f_v(E)\}Z\rho_c(E)\rho_v(E)dE \qquad (1)$$

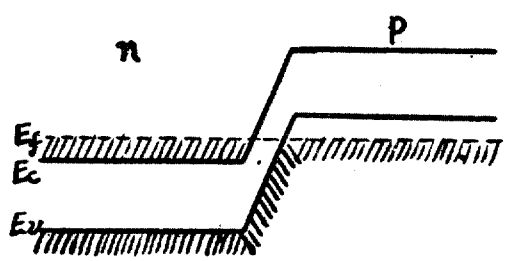

Fig. 1. Energy level diagram for a heavily doped p—n junction

where:

Z : tunneling probability[2]
$f_c(E)$, $f_v(E)$: Fermi-Dirac distribution function
$\rho_c(E)$, $\rho_v(E)$: density of states at energy E
E_c : bottom of conduction band
E_v : top of valence band

Brody et al[3] and Bates[4] calculated the

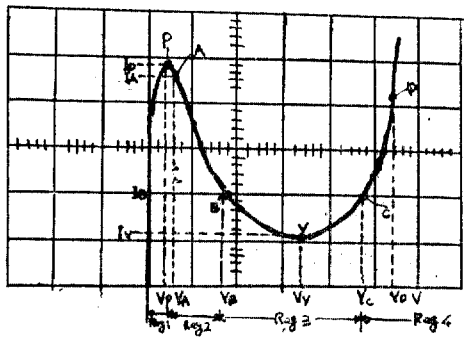

Fig. 2. V-I curve of a T. D.

tunneling current assuming $\rho_c(E)$ and $\rho_v(E)$ are proportional to $(E-E_c)^{\frac{1}{2}}$ and $(E_v-E)^{\frac{1}{2}}$ respectively. The V-I characteristic curve for this device is extremely nonlinear as shown in Fig. 2. The properties and the principles of a T.D.,[5)~7)] its high frequency applications,[8), 9)] its noise performances,[10)] and its application to logic and pulse circuits have been studied by some workers.[11)~13)] Although two parabolic functions could be used for the approximation of the V-I characteristic of a T.D., this gives a poor result at the high junction voltage state. Schuller and Gartner[14)] used a third order polynomial of the form $i=(g/v_o^2)v(v^2-v_o^2)$ for the analytical approximation, but none of the Ge, Si, GaAs,[15)] or InAs[16)] tunnel diodes have this kind of symmetrical V-I characteristics. The linear piecewise approximation will be found in some literature.[12)] But, this approximation cannot give threshold equation when the T.D. is operated as a pulse height discriminator.

Approximation of the V-I curve

Fig. 2 shows the V-I curve of a Si T.D. together with its approximation. The V-I curve is divided into four regions and approximated as follows.

Region I: O→P→A

The T.D. current flows purely by the tunneling process, and it is approximated by a parabolic function,

$$i=I_p\left[1-\left(\frac{v}{V_p}-1\right)^2\right]$$

The end point, A, of this region is defined as $I_A=0.96I_p$, $V_A=1.2V_p$, so that Region II could be approximated by negative resistance.

Region II: A→B

This region is approximated by a negative resistance, and the current flows mainly by the tunneling process. The voltage range of this region is about 150 mV, but this range will be dependent upon the materials the diode is made of. The current through the T.D. will be expressed by,

$$i=I_A-\frac{v-V_A}{R}=I_B+\frac{V_B-v}{R}$$

where, R is the absolute value of the negative resistance. J. B. Lyons, Jr. proposed a circuit for the accurate measurement of negative resistance.[21)]

Region III: B→V→C

In this region, though the excess current is dominant, there are some component due to the minority carrier injection. For a GaAs T.D., Nanavati[22)] gave some physical interpretation for the excess current, and it is discussed by Gold.[23)] The current is approximated by another parabolic function,

$$i=I_V+K[v-V_V]^2$$

where, K is a constant dependent upon the diode, and it could be expressed as $(I_C-I_V)(V_C-V_V)^{-2}$ or $(I_B-I_V)(V_V-V_B)^{-2}$. The boundary point, C, between regions III and IV is defined by $I_C=I_B$ and $V_C-V_V=V_V-V_B$.

Region IV: C→D

In this region the current through the T.D. is mainly due to the minority carrier injection. A positive resistance, R_{CD} defined by $(I_D-I_C)R_{CD}=V_D-V_{CD}$ is used to approximate the T.D. in this region. The position of point D should be dependent upon the final voltage to which the T.D. is driven. A better approximation will be obtained using an exponential function similar to the junction V-I equation for an ordinary

韓國物理學會誌 "새 물리" Vol. 5, No. 2, Dec. 1965

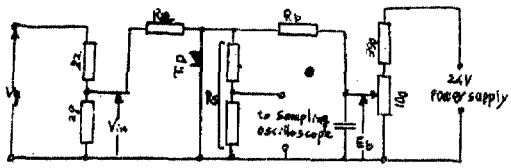

Fig. 3a. Circuit diagram used to observe the switching time of a T. D.

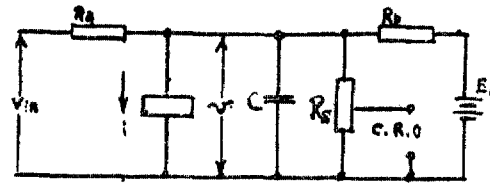

Fig. 3b. Simplified circuit

diode, i.e., $i=I_V+I_S\left[exp\frac{e(v-V_V)}{nkT}-1\right]$. But, this approximation yields a nonlinear differential equation for the calculation of the switching time which cannot be solved by elementary methods. Therefore, the T.D. current in this region is approximated as

$$i=I_C+\frac{v-V_C}{R_{CD}}$$

The Switching Time

Fig. 3a shows the circuit diagram used to observe the switching time of a T.D., and its simplified circuit is shown in Fig. 3b.

It is assumed that the input pulse length is much longer than the T.D. switching time. And also, in the simplified circuit Fig. 3b, the T.D. is represented by a current path in parallel with capacitance C neglecting the small series inductance and resistance together with the variation of the capacitance, C, according to the junction voltage variation.[24] Referring to Fig. 3b, the circuit equation in each region may be written as

$$C\frac{dv}{dt}+i+\frac{v}{R_s}=\frac{E_b-v}{R_b}+\frac{V_{in}-v}{R_a} \quad (2)$$

where,

$$i=I_P\left[1-\left(\frac{v}{V_p}-1\right)^2\right] \text{ in Region I}$$

$$I_A-\frac{v-V_A}{R} \text{ in Region II}$$

$$I_V+K(v-V_V)^2 \text{ in Region III}$$

$$I_C+\frac{v-V_C}{R_{CD}} \text{ in Region IV}$$

putting

$$A=-\frac{I_P}{V_P^2},\quad \frac{1}{R_{e1}}=\frac{2I_P}{V_P}+\frac{1}{R_e},$$

$$\frac{1}{R_e}=\frac{1}{R_a}+\frac{1}{R_b}+\frac{1}{R_s} \quad (3)$$

$$I_{e1}=I_e=\frac{V_{in}}{R_a}+\frac{E_b}{R_b},\quad a_1=-\frac{1}{2AR_{e1}},$$

$$b_1=\sqrt{\frac{I_{e1}}{A}-a_1^2}$$

the time, t_1, necessary for the instantaneous operating point to move from S to A through P will be

$$t_1=\frac{C}{Ab_1}\left[tan^{-1}\frac{V_A-a_1}{b_1}+tan^{-1}\frac{a_1-V_s}{b_1}\right]$$

where V_s is the junction voltage at the stable operating point S. In (3), V_p/I_p shows the incremental resistance of the T.D. at zero bias voltage, and this is comparable to the theoretically calculated value of $(\pi V_p)/(8I_p)$ for a Ge T.D. And the necessary time, t_2, for the instantaneous operating point to move from A to B will be

$$t_2=CR_{e2}\ log\frac{V_B+I_{e2}R_{e2}}{V_A+I_{e2}R_{e2}}$$

where

$$\frac{1}{R_{e2}}=\frac{1}{R}-\frac{1}{R_e},\quad I_{e2}=I_e-\left(I_A+\frac{V_A}{R}\right)$$

In Region III, putting

$$\frac{1}{R_{e3}}=2KV_V+\frac{1}{R_e},\quad I_{e3}=I_e-(KV_V^2+I_V)$$

the solution of (2) will be

$$v=a_3+b_3\ tanh\left[\frac{b_3K}{c}t+tanh^{-1}\frac{V_B-a_3}{b_3}\right]$$

where

$$a_3=\frac{1}{2KR_{e3}},\quad b_3=\sqrt{\frac{I_{e3}}{K}+a_3^2}$$

When the input pulse height is not large enough to drive the T.D. into region IV, the instantaneous operating point will remain in this region until the input pulse

disappears. And junction voltage for this case will be determined by

$$v_{max}=[v]_{t\to\infty}=a_3+b_3$$

In the case where the T.D. will be driven into region IV, the necessary time, t_3, for the operating point to move from B to C through the valley point, V, will be

$$t_3=\frac{C}{b_3K}\left[tanh^{-1}\frac{V_C-a_3}{b_3}-tanh^{-1}\frac{V_B-a_3}{b_3}\right]$$

Also in region IV,

putting $\frac{1}{R_{e4}}=\frac{1}{R_{CD}}+\frac{1}{R_e}$,

$$I_{e4}=I_e+\left(\frac{V_C}{R_{CD}}-1\right)^2$$

the solution of (2) is

$$V=I_{e4}R_{e4}+(V_C-I_{e4}R_{e4})exp\left(-\frac{t}{CR_{e4}}\right) \tag{4}$$

And the final voltage, V_F will be determined by (4) as

$$V_F=[v]_{t\to\infty}=I_{e4}R_{e4}$$

When the input pulse disappears, the T.D. will switch back to its original stable operating point S, where the junction voltage is V_s. And the switch-back time for the instantaneous operating point to move from region IV to the point S through regions III, II, and I will be calculated similarly.

They are

$$t'_4=CR_{e4}\ log\frac{V_F-I'_{e4}R_{e4}}{V_C-I'_{e4}R_{e4}}$$

where

$$I'_{e4}=I_{e4}-\frac{V_{in}}{R_a}$$

$$t'_3=\frac{C}{b'_3K}\left[tan^{-1}\frac{V_C-a_3}{b'_3}-tan^{-1}\frac{V_B-a_3}{b'_3}\right]$$

where

$$b'_3=\sqrt{\frac{I'_{e3}}{K}-a_3}$$

$$I'_{e3}=KV_V^2+I_V-\frac{E_b}{R_b}=\frac{V_{in}}{R_a}-I_{e3}$$

$$t'_2=CR_{e2}\ log\frac{I'_{e2}R_{e2}-V_A}{I'_{e2}R_{e2}-V_B}$$

where

$$I'_{e2}=\frac{V_{in}}{R_a}-I_{e2}$$

$$t'_1=\frac{C}{Ab'_1}\left[tanh^{-1}\frac{V_A-a_1}{b'_1}+tanh^{-1}\frac{a^1-V_s}{b'_1}\right]$$

where

$$b'_1=\sqrt{a_1^2-\frac{I'_{e1}}{A}},\quad I'_{e1}=I_e-\frac{V_{in}}{R_a}$$

These times, t'_4, t'_3 t'_2 and t'_1 are quite dependent upon the biasing condition, E_b and R_b, and are proportional to the junction capacitance, C, when the other circuit conditions are kept constant. Therefore, it is possible to measure the effective junction capacitance if we observe the switch-back time variation which is caused by the connection of external condensers across the T.D. The measured effective junction capacitance by this method was about 100 pF. This value is much greater than the data supplied by the manufacturer as 10 pF in the negative resistance region.

May be the wide junction voltage region, and the neglect of series inductance are the main causes of this difference. In order to measure the actual switching time of a T.D., a fast pulse generator whose

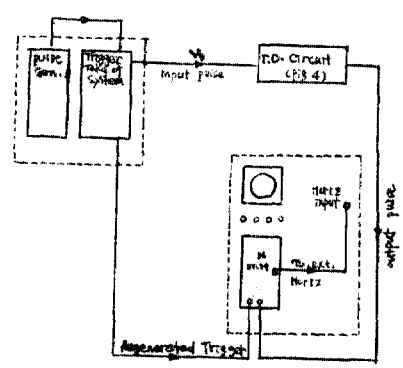

Fig. 3c. Block diagram for the measurement of a T.D. switching time.

Pulse Generator: Tektronix Type 110 pulse Generator and trigger takeoff system (Aek 30049) (rise time≤0.25μS)

N Unit : Tektronix Type N Sampling plug-in Unit (Serial 000256) (rise time≤0.6μS)

Oscilloscope : Tektronix Type 535A (AEK 18098)

Camera : Du Mont Type 302 (AEK 43)

韓國物理學會誌 "새 물리" Vol. 5 No. 2, Dec. 1965

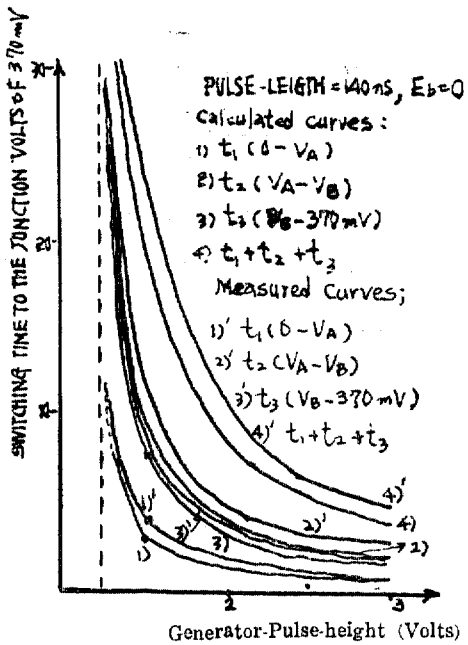

Fig. 4. Generator Pulse-Height vs Switching Time

pulse rise time is less than 0.25 μS and a sampling plug-in unit whose rise time is less than 0.6 μS are employed together with Tektronix Type 535A oscilloscope as shown in the block diagram of Fig. 3c. The calulated switching time is shown in Fig. 4 together with the value observed, and Fig.5 shows the total switchback time. The discrepancy between the value calculated and the value measured may be caused by the approximation of the V-I curve of the T.D., and also by the inaccurate knowledge about the junction capacitance which will be larger at the higher junction voltage than its value at the lower junction voltage. The switching time, when the T.D. is biased by a current generator and is triggered by a current pulse, will be calculated similarly.[20]

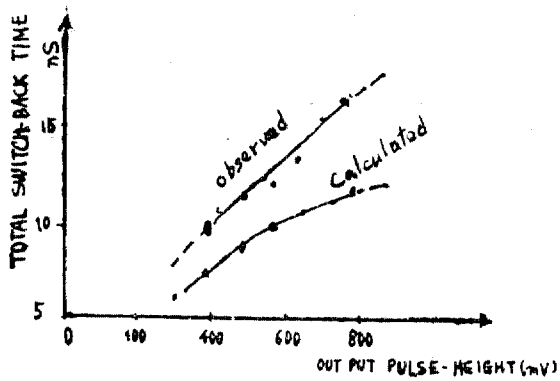

Fig. 5. Output Pulse-Height vs Total Switch-Back Time

Application as a Pulse-Height Discriminator

According to (3), the highest pulse which will not trigger the T.D. will be determined by,

$$b_1=0 \quad \therefore \frac{V_{in}}{R_a}+\frac{E_b}{R_b}=Aa_1^2 \qquad (5)$$

when the input pulse exceeds a certain threshold value which is determined by (5), the T.D. operating point will be shifted to a point in region III or region IV which is determined by the input pulse height. Terefore, this unit can be used as a pulse-height discriminator.

(4 A) Some considerations

Fig. 6 shows an example of a T.D. pulse-height discriminator circuit, and the load lines are shown in Fig. 7. Referring to Fig. 7, the slope of a.c. load line must be small, i.e. R_a, R_b, R_d+Z_d must have large values in order to get a big output pulse. On the other hand, if d.c. load line approaches V-I curve under the high bias current state, the contact point, T, acts as a temporarily stable operating point after the wave reflected from the delay cable arrived, and thereby the output pulse becomes a little longer. The closer d.c. load line approaches

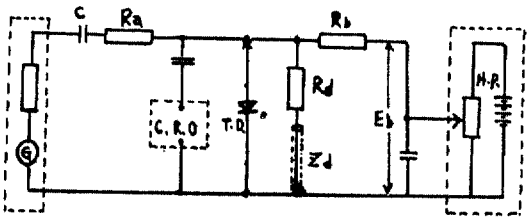

Fig. 6. T. D. Pulse-Height Discriminator

to point T, the longer the output pulse length is. Therefore, the slope of d.c. load line must be big in order to avoid this phenomena under the high bias current. It is also possible to get a kind of staircase output pulse by the proper selection of the resistance and the delay cable,[18] and it is discussed by Steward.[25]

(4 B) The linearity of the discriminator

The linearity of the discriminator is determined by (5), i.e.

$$V_{in}=I_p R_a\left(1+\frac{V_p}{2I_p}\Sigma_i\frac{1}{R_i}\right)^2-\frac{R_a}{R_b}E_b \qquad (6)$$

Where R_i: R_a, R_b and R_d+Z_d

Referring to (6), curve V_{in} vs. E_b is a straight line. In Fig. 8, the V_{in} vs. E_b curves measured are shown together with the theoretical line of (6) for two different values of the coupling condenser c_c, of Fig. 6. Equation (6) and curve 1 of Fig. 8

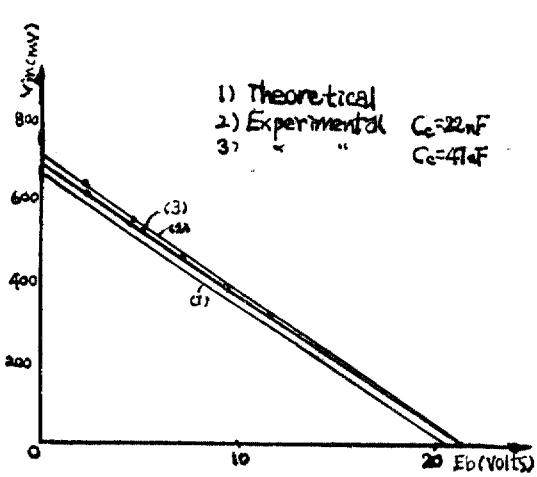

Fig. 8. Linearity of the discriminator

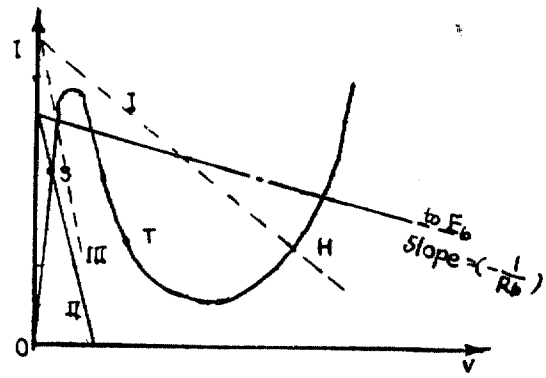

Fig. 7. The Load Lines of Fig. 6 Circuit

Ⅰ :a.c. load line: $-\left(\frac{1}{R_a}+\frac{1}{R_b}+\frac{1}{R_d+Z_d}\right)$

Ⅱ :d.c. 〃 〃 : $-\left(\frac{1}{R_d}+\frac{1}{R_b}\right)$

Ⅲ :a.c. 〃 〃 : $-\left(\frac{1}{R_d}+\frac{1}{R_b}+\frac{1}{R_a}\right)$

after the reflected wave from the delay cable arrives at R_d of Fig. 6

show the highest input pulse which will not trigger the discriminator, and the results measured when 50% of the input pulses triggered the circuit are shown as curves 2 and 3 of Fig. 8. The deviation between curve 1 and curves 2 and 3 may be caused by the different definition of threshold value and the neglected small series inductance. The coupling condenser, C_C, of Fig. 6 causes a little attenuation of the input pulse, and the attenuation will be bigger the shorter the time constant, $C_C R_a$, or the longer the pulse rise time is. This could explain the difference between the slopes of curves 2 and 3. In any case, the linearity between input pulse height and bias voltage is excellent, and the relative variation of the linearity is always less than $\pm 1\%$.

(4 C) The pulse resolving ability

In order to check the pulse resolving ability of the discriminator, the circuit of Fig. 9 was employed. During this measurement, the pulse lengths were 0.4 μS and

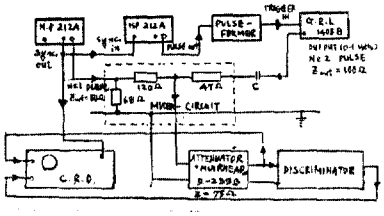

Fig. 9. Block diagram of the circuit of the pulse resolving ability measurement

the pulse rise-times were less than 70 μS, and the positin of No. 2 pulse was adjusted by adjusting the position of the input pulse to the pulseformer block of Fig. 9. Fig. 10 shows the curves of necessary excess pulse height of pulse No. 2(%) vs. pulse space, T_{ps}. If the circuit components are properly selected, the necesssary excess pulse height of pulse No. 2 will remain within $\pm 1\%$ when T_{ps} of Fig. 10 is equal to or less than 0.5 μS. Also T_{ps} will be shorter, shorter the pulse rise time is.

(4 D) Temperature characteristics

Referring to (6), the threshold variation of the discriminator due to the ambient temperature variation could be attributed to either I_p or V_p variation or both. The peak voltage, V_p, and the peak current, I_p, will be the function of Fermi-levels, the density of state functions of both p and n materials, and the tunneling probabilty. Therefore, it is not easy to predict the variation of V_p and I_p according to the ambient temperature variation. But, the actual measurement showed that the peak voltage variation in the ambient temperature range of 13°C—70°C was negligible. Blicher et al.[19] measured the % change of the peak current of a Ge T.D. for different p-region impurity concentration, keeping n-region impurity concentration constant. The result was either positive or negative depending upon the impurity concentration. Therefore, the threshold variation of the discriminator could be attributed to the I_p variation, neglecting the influence of V_p variation. This gives (7) which is derived from (6).

$$V_{in} \simeq R_a \triangle I_p$$

The threshold variation under the same temperature was nearly constant for all the bias voltage range measured, and the mean threshold variation according to the ambient temperature variation is shown in Fig. 11 together with the calculated curve using (7). And also it was observed that the threshold value of the discriminator completely returned to its original value (at 15°C) when the ambient temperature was reduced to 15°C after it had high ambient temperature experience of 55° C for 27 hours and 79°C for 5 hours, without having any residual effect.

Acknowledgement

The author wishes to express his thanks

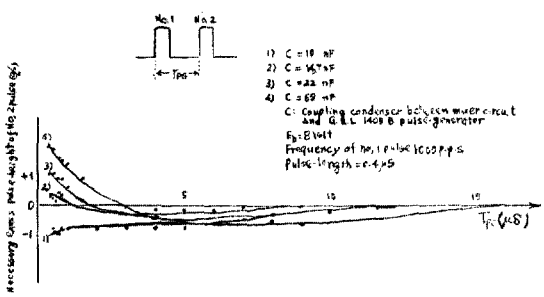

Fig. 10. Pulse Resolving Ability of the Discriminator

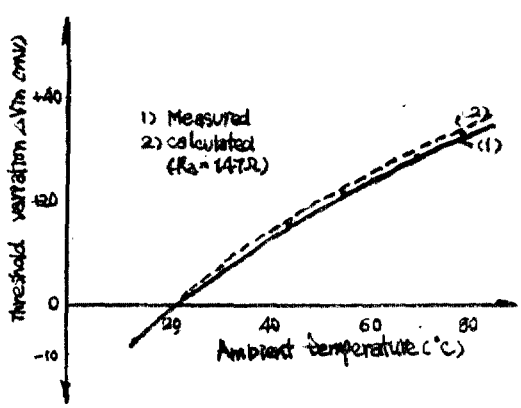

Fig. 11. Temperature Characteristics

to. Mr. G.K. Frölich Hansen and Mr. E. Mose Christiansen for their kind assistance in getting sample tunnel diodes and their constant encouragement and helpful discussions during the course of the work.

References

1) L. Esaki; New Phenomenon in Narrow Germanium p-n Junctions Phys. Rev. **109**, 603. (1958)
2) I.A. Lesk; N. Holonyak Jr., U.S. Davidsohn, M. W. Arons; Ge and Si Tunnel Diodes-Design, Operation and Application, IRE Wescon Convention Record Part 3, **9**, (1959)
3) T.P. Brody, R.H Boyer; The Evaluation of Esaki-Integral and an Approximate Expression for the Tunnel Diode Characteristic, Solid-state Electronics, **2**, 209, (1961)
4) C.W. Bates Jr; Tunneling Current in Esaki Diode. Phys. Rev. **121**, 1070, (1961)
5) R.N. Hall, Tunnel Diode. I.R.E. Transactions on Electron Device. **Ed-7**, 1, (1960)
6) R.A. Pucel; Physical Principles of the Esaki Diode and Some of Its Properties as a Circuit Element. Solid-state Electronics, **1**, 22, (1960)
7) I.A. Lesk, & J.J. Suran; Tunnel Diode Operation and Application. Electrical Engineering, **79**, 270, (1960)
8) H.S. Sommer Jr.; Tunnel Diode as High-Frequency Devices. Proc. I.R.E. **47**, 1201, (1959)
9) G. Dermit; High-Frequency Power in Tunnel Diodes. Proc. IRE., **49**, 1033, (1961)
10) J.J. Tiemann; Shot Noise in Tunnel Diode Amplifiers, Proc. I.R.E., **48**, 1418, (1960)
11) I.W. Janney; Tunnel Diode Applications to Logic and Pulse Circuits. Sandis Corporation, SCTM-375-60-72, (1960)
12) R.H. Bergman, Tunnel Diode Logic Circuits IRE. Transactions on Electronic Computers, **EC-9**, 430 (1960)
13) M.M. Kaufman et al; Tunnel Diode Memory. IEEE International Convention Record **12**, Part 1, 1, (1964)
14) M. Schuller, W.W. Gartner; Large Signal Circuit Theory for Negative-Resistance Diodes, in particular Tunnel Diodes, Proc. I.R.E., **49**, 1268, (1961)
15) N. Holonyak and, I.A. Lesk; Gallium-Arsenide Tunnel Diode. Proc. I.R.E., **48**, 1405, (1960)
16) H.P. Kleinknecht; Indium-Arsenide Tunnel Diode. Solid State Electronics, **2**, 133, (1961)
17) Y. Furukawa; Electrical Characteristics of Esaki Diodes. The Journ. of the Inst. of Elec. Comm. Engrs. of Japan. **43**, 1396, (1960)
18) J. Nagumo and M. Shimura; Self-Osciallation in a Transmission Line with a Tunnel Diode. Proc. I.R.E., **49** 1281, (1961)
19) A. Blicher, R.M. Minton and R. Gliksman; Temperature Dependence of the peak Current of Germanium Tunnel Diode. Proc. I.R.E., **49**, 1428. (1961)
20) B.C. Kim,; The Switiching Time of a Tunnel Diode. Risö Report No. 42, Danish AERE, Denmark. (1962)
21) J.B. Lyons, Jr; Measurement of Negative Resistance of Tunnel Diodes Proc. IEEE **53**, 484, (1965)
22) R.P. Nanavati et al; Execss Current in Gallium Arsenide Tunnel Diodes Proc. IEEE **52**, 869, (1964)
23) R.D. Gold; Comments on Excess Current in Gallium Arsenide Tunnel Diodes Proc. IEEE **53**, 486, (1965)
24) R.P. Nanavati et al; Tunnel Diode Junction Capacitance in the Vicinity of Built-In Voltage Proc. IEEE **51**, 1769. (1963)
25) M. H. Steward et al; Tunnel Diode Loaded by a Shorted Trasmission Line Proc. IRE, **50**, 1830, (1962)