Sustaining and decaying of the on state of niobium dioxide

Gary C. Vezzoli
U.S. Army Armament Research and Development Command, Applied Science Division, Device Physics and Technology Section, Dover, New Jersey 07801
(Received 4 December 1978)

Data are presented showing that the post switching, high-conductance on state of NbO_2 can be sustained by use of a subthreshold high-frequency square wave. However, it cannot be sustained by a triangular cw of the same peak amplitude and frequency. The decay in the latter regime is due to the slower rise and fall rate of the triangular wave and the relation to recombination time. Equations are given to analyze the charging conditions for the sustained or decaying state, and the change in carrier concentration which permits the critical decay is calculated.

Niobium dioxide undergoes an electrical threshold switching transition when addressed by a suitable pulse[1] or continuous wave.[2] The on state can be preserved after the initial set pulse by use of a continuous pure or distorted square wave of high frequency (~1 MHz) and subthreshold amplitude.[3] The simple dual-power-source switching circuit is given elsewhere.[4] The maintenance of the on state using a square wave, which is distorted by the protective diode, is presented in Fig. 1. This figure shows two oscillograms with the upper trace of each giving first the switching set-voltage pulse, followed by the 1-MHz square wave required to sustain the on state. The maximum voltage of the square wave shown in Fig. 1 is equal to the post-switching on-voltage. The lower trace of each oscillogram presents the initial switch-on current (including the overshoot displacement current), followed by the on-current as it responds to the changing cw voltage. The constancy of the peak current should be noted.

Two oscillograms are presented in Fig. 1 to show that the difference in delay times of the set pulse (before switching occurs) does not affect the maintenance of the on state. The post-switching high-conductance state is sustained even though for a finite time (<100 nsec) the cw voltage is beneath the dc holding value. The data of Fig. 1 are interpreted to mean that the fall and rise rates of the cw are sufficiently fast such that critical relaxation and decay do not occur. Such a decay would cause a switch-off because the cw-voltage is much less than the threshold voltage. The peak cw amplitude must be at least equal to the dc holding voltage, and the time below the holding voltage must be restricted to a maximum allowable value in order that the on state be preserved in this manner. A similar effect of the cw maintenance of the on state is observed in chalcogenide glass threshold switching devices,[3] however, with some difference in the current-voltage on-state curves at given frequencies relative to the I-V curve for the on state of NbO_2.

We have in this laboratory for quite some time contended that a field-induced threshold switching semiconductor-to-metal transition requires both a specific value of electric field and of charge

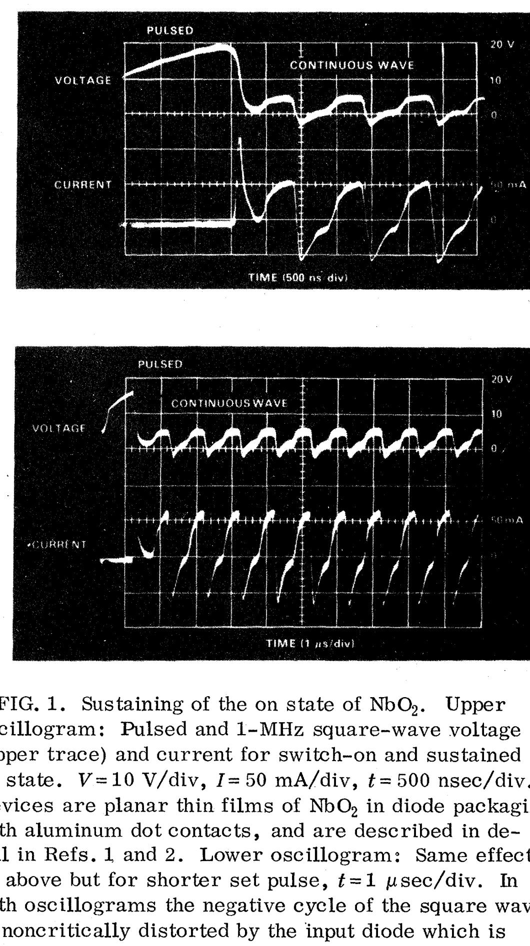

FIG. 1. Sustaining of the on state of NbO_2. Upper oscillogram: Pulsed and 1-MHz square-wave voltage (upper trace) and current for switch-on and sustained on state. $V = 10$ V/div, $I = 50$ mA/div, $t = 500$ nsec/div. Devices are planar thin films of NbO_2 in diode packaging with aluminum dot contacts, and are described in detail in Refs. 1 and 2. Lower oscillogram: Same effect as above but for shorter set pulse, $t = 1$ μsec/div. In both oscillograms the negative cycle of the square wave is noncritically distorted by the input diode which is utilized to protect the two power sources from each other.

density, as well as two types of carriers—one trapped and one free.[5, 6(a)] Contrasting the behavior shown in Fig. 1, if the square-wave source is replaced by a triangular wave of the same amplitude and frequency—thus the only change being the time rate of change of applied voltage—the on state is no longer preserved indefinitely, but undergoes the rapid decay to the off state as shown in Fig. 2. This figure also gives two oscillograms showing the voltage in the upper trace and the corresponding current in the lower trace of each photograph. The set-voltage switching pulses are similar to those in Fig. 1; however, the post-switching 1-MHz triangular cw is incapable of maintaining the on state as attested by the decaying peak current in both oscillograms. The cw current eventually decays to the preswitching off-state value, and the device voltage climbs to the level corresponding to the value before the resistance collapse. Prior to switching, the voltage falls largely across the device as shown by the set pulse. Subsequent to switching, the set voltage falls primarily across the protective load resistor, and in the high-conductance state only the low-level on-voltage falls across the device. Thus in Fig. 2 the unique voltage-charge-time holding conditions which maintained the on state in Fig. 1 are not met in the experiment shown in Fig. 2. This paper endeavors to functionally outline these voltage-charge-time holding conditions.

We assume that the cw-induced on-current will be determined by charging the inherent device capacitance C_0 plus the addition in parallel of nonlinear currents arising from all internal charge transport carriers.[6(a)] Thus

$$I_{on}=C_0\frac{dV}{dt}+\sum_i\frac{dq_i}{dt}.$$

We also assume that the on state of the NbO_2 thin film is dominated by electrons (perhaps undergoing d-d transitions[7] because of overlap of d-orbital wave functions in distorted rutile-type structures), and that q_1 represents a trapped electron, and q_2 a free electron. Then

$$I_{on}=C_0\frac{dV}{dt}+\frac{dq_1}{dt}+\frac{dq_2}{dt}. \quad (1)$$

Since the square and triangular waves differ only in their time derivatives of voltage, and since we wish to express the constant maximum cw current in Fig. 1, we take the time derivative of Eq. (1).

$$\frac{dI_{on}(t)}{dt}=\frac{d}{dt}\left(C\frac{dV}{dt}\right)+\frac{d}{dt}\frac{dq_1}{dt}+\frac{d}{dt}\frac{dq_2}{dt},$$

or by the chain rule,

$$\frac{dI_{on}(t)}{dt}=\frac{d}{dt}\left(C\frac{dV}{dt}\right)+\frac{d}{dt}\left(\frac{dq_1}{dV}\frac{dV}{dt}\right)+\frac{d}{dt}\left(\frac{dq_2}{dV}\frac{dV}{dt}\right). \quad (2)$$

The above equation is utilized in order to express the rate at which the field generates charge.[4-6] The maximum value of measured cw current as a function of time, after the set pulse, must be evaluated. Hence the differentials of Eq. (2) must be summed over every cycle and thus include the contributions of generation and decay. The rationale for using the summation form is given in the Appendix to this paper. Upon differentiating the product $C(t)dV(t)/dt$ and summing we arrive at

$$\frac{dI_{on(max)}}{dt}=\sum\left(C(t)\frac{d^2V}{dt^2}+\frac{dV}{dt}\frac{dC(t)}{dt}+\frac{d}{dt}\frac{dq_1}{dV}\frac{dV}{dt}+\frac{d}{dt}\frac{dq_2}{dV}\frac{dV}{dt}\right). \quad (3)$$

The resulting magnitudes must be evaluated at V_{max} and $t=\tau/4,\ 5\tau/4,\ 9\tau/4,\ldots$. For the square wave we have from Fig. 1

$$\frac{dI_{on(max)}}{dt}=0,$$

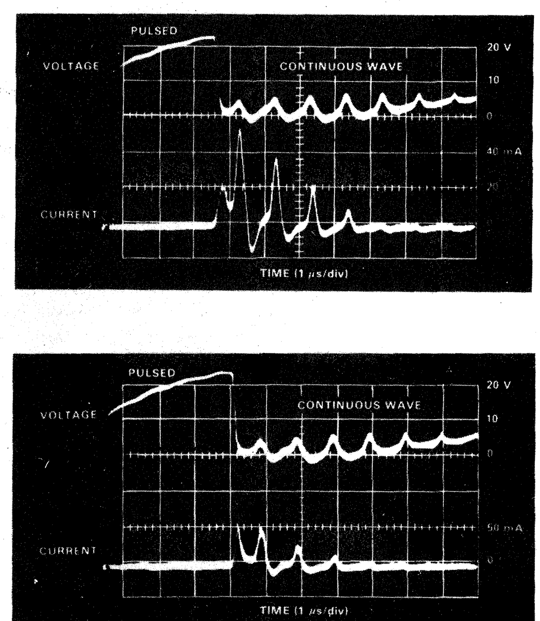

FIG. 2. Decaying of the on state of NbO_2. Upper oscillogram: Pulsed and 1-MHz triangular-wave voltage (upper trace) and current for switch-on and decay of on state. $V=10$ V/div, $I=20$ mA/div, $t=1$ usec/div. Note decaying envelope of on-current. Lower oscillogram: Same effect of decaying on state for longer set pulse, $I=50$ mA/div.

$$\frac{dV}{dt}=\frac{4V}{100\ \text{nsec}}$$

during positive rise time

$$\frac{dV}{dt}=0$$

during constant voltage interval

$$\frac{dV}{dt}=\frac{4V}{50\ \text{nsec}}$$

during positive fall time and negative rise time and $V(t)$ during the negative fall time is an almost linear function of time.

If a pure square wave were substituted for the distorted square wave, the maximum on-current would still be preserved as it is in Fig. 1.[6(b)] Thus, the nonconstant $\dot{V}$ for the negative fall time is not critical, and for an on-state maintaining condition we can write $\dot{V}=4V/(400\ \text{nsec})$ for this interval. Then the second derivative of voltage versus time is essentially zero at every point where $V(t)$ is differentiable, and we have the equation

$$\sum\left[\frac{d}{dt}\left(\frac{dq_1}{dV}+\frac{dq_2}{dV}\right)+\frac{dC(t)}{dt}\right]=0. \tag{4}$$

Any decay which occurs in $I_{\text{on(max)}}$ as a function of time will then according to Eq. (4) be a cummulative effect of the change in the time rate at which the field generates charge. This is because if this rate does not keep pace with or exceed the recombination rate, the on state cannot be sustained. Then, as in Fig. 2, recharging would be too slow to compensate for the allowed rate of decay due to the more slowly falling cw. The on-current then falls off as if $I_{\text{on(max)}}$ were critically damped. Equation (4) thus specifies the condition for the preservation of the on state (provided at least a minimum charge flows).

If we now consider the on-current corresponding to the maximum in the first cycle as an initial current, and consider only the decay or absence of decay due to the next cycle, we have for Fig. 1:

$$\frac{d}{dt}\frac{dq_2}{dV_{\text{net}}}=-\frac{d}{dt}\frac{dq_1}{dV_{\text{net}}}-\frac{dC(t)}{dt}, \tag{5}$$

which is interpreted to mean that the field-induced growth of q_2 occurs directly from the supply of q_1 to sustain the on state. If q_1 and q_2 represent trapped (n_1) and free (n_2) carriers, as one might expect,[8] then Eq. (5) states that if the specified charge balance or equilibrium is not sustained, then the free carriers q_2 will fall back into the traps out of which they had been liberated originally, and the on state will decay into the off state. Thus, if because of decay and a slow charging rate, n_1 does not supply n_2 at a rate equal to or larger than the natural recombination rate of n_2, then a transition must occur to the off state.

For the triangular-wave data in Fig. 2, the voltage falls and rises at a slower rate than for the square wave in Fig. 1 even though the frequencies are the same. For the decay regime the only difference in the treatment is that $\dot{I}_{\text{on(max)}}$ is not zero. We allow $K=dV/dt$ and write

$$\frac{dI_{\text{on(max)}}}{dt}=K\sum\left[\frac{d}{dt}\left(\frac{dn_1}{dV}+\frac{dn_2}{dV}\right)+\frac{dC(t)}{dt}\right].$$

From the decay data of Fig. 2, considering only the first two maxima of I , we have $\Delta I/\Delta t=20\times10^3$ A/sec, $K=2\times10^7$ V/sec. Changing units allows us to write

$$\sum_{t=t(I_{\text{max}_1})}^{t=t(I_{\text{max}_2})}\left[\frac{d}{dt}\left(\frac{dn_1}{dV}+\frac{dn_2}{dV}\right)+\frac{dC(t)}{dt}\right]_{\text{net}}=6\times10^{15}\ \text{electrons/V sec}.$$

Since decay is obviously occurring in Fig. 2, this means that a net recombination into traps or into the valence band is taking place at the rate of about 6×10^{15}/sec (free carriers for every volt decrement in potential). Thus as the amplitude decreases from 4 V to zero in 200 nsec a total of about 5×10^9 net recombinations occur. We know from recovery curve data (Ref. 2) that the distribution carrier lifetime in the on state is about $\tau_{n_1}\approx200$ nsec (as taken from double pulse experiments measuring the reswitching voltage after intervals of voltage interruption). Thus we conclude that during the zero voltage time of 200 nsec at least 5×10^9 carriers undergo net recombination.

In the present treatment we have considered the capacitance C_0 as a variable function of time. We calculate the dielectric relaxation time $\tau_d^{1/2}=\sqrt{RC}$ in the on state as about $\tau_d=80\ \Omega\times2pf\approx10^{-10}$ sec, which is shorter than the rise time at 1-MHz frequency. This indicates that any time constant greater than τ_d should be clearly visible in the experimental data provided proper instrumentation.

To determine the change in the total current density as a function of time one must also consider the current channel cross-sectional area and its time rate of change, and thus evaluate

$$\mathring{J}=\frac{d}{dt}\{I_{\text{on(max)}}(t)/A(t)\}.$$

The treatment given herein must be temperature dependent in dq/dV because the NbO_2 on-state resistance, unlike the chalcogenide glasses, changes considerably as ambient temperature is lowered. Typical Ohmic resistance values are 80–140 Ω for the on state at room temperature and about 400–500 Ω at liquid-nitrogen ambient. Future work will evaluate the recovery curve of the on state at cryogenic and at elevated temperatures to compare it to the extensive temperature-independent recovery data for chalcogenide glasses.

As a further check on the decaying effect of Fig. 2, two more sets of experiments were conducted, one with a 1-MHz ramp function as the cw, and the other with a 0.5–1.5-MHz sinusoidal cw. The ramp function data showed a behavior very similar to the positive cyclic and decaying behavior of Fig. 2. In order to prove that the preservation of the on state in Fig. 1 (relative to the decay in Fig. 2) was not a result of the larger integrated duty cycle of the square wave, a high-frequency sine wave generator was utilized to determine at what frequency the triangular wave would have sustained the on state. The sine wave trace approximated that of the triangular wave very closely, and at a frequency of 1 MHz the sine wave decay data had the appearance of Fig. 2. However, as the frequency was raised such that the rise and decay time would approach that of the square wave, it was found that the on state could be generally sustained as in Fig. 1 at a frequency of about 1.5 MHz. This corresponded to a voltage-rise time of about 150 nsec and a voltage-decrease time of slightly less than 150 nsec (the difference being due to some wave distortion). This is intermediate between the conditions of Figs. 1 and 2, yet sufficient to sustain the on state. As the frequency of the sinusoidal wave was decreased, the expected switch-off occurred. These observations are strong deterrents against any suggestion that switching in NbO_2 may be due to thermal avalanche.

Another observation during these latter sets of experiments was that if the initial cw current, immediately after the switching transition, were not above a threshold value, then the on state was not sustained. However, in such a case the on-current did not show the characteristic envelope decay of Fig. 2. Instead it corresponded to an off-state current of constant peak amplitude, and was a simple response to the cw voltage.

ACKNOWLEDGMENTS

The author is indebted to G. K. Gaule of The Communications Research and Development Command, Ft. Monmouth, N.J., for supplying planar thin-film NbO_2 switching devices that were fabricated by The General Electric Company. Gratitude is also expressed to Stephen DeFeo, Patrick Calella, Neil Albrecht, and John Bera of this laboratory for worthwhile discussions and suggestions, and to L. William Doremus, Chief of the Device Physics and Technology Area of this Command, for encouragement and support to conduct research on NbO_2 devices for lightning-surge-arrestor applications.

APPENDIX

The reason for the summation form becomes apparent if we assume the following.

As the electric field increases, field-induced free carriers are generated at the rate dn_2/dt via the emptying of traps or the direct population of the conduction band. Trapped carriers are generated at the rate dn_1/dt and are trapped in time τ_1. The trapped carrier is promoted by the field (or tunnels) into the conduction band in time τ_2, however, recombines back into the valence band at the rate dn_{1RV}/dt. The free carrier recombines back into a trap at the rate dn_{2RT}/dt, and into the valence band at the rate dn_{2RV}/dt. If we then envision the switching transition to occur after most of the available traps are filled, we postulate it then requires a time of the order $\tau_1+\tau_2$.

Now in Figs. 1 and 2 most of the traps are filled at time t_0 because in both figures the set pulse has initiated a bonafide switch. Thus an initial free-carrier density or number of free carriers N_{2i} is specified. The on-voltage pulse then drops to zero, and the accompanying current decay specifies a new number of free carriers, N_2. The cw voltage then rises but in the case of Fig. 2 at a rate such that

$$\frac{dn_2}{dt} < \frac{dn_{2RT}}{dt} < \frac{dn_{2RV}}{dt}.$$

Then at $t=\tau/4$ when the cw voltage has reached its peak value a new number of free carriers $N_{2(t=\tau/4)}$ are conducting. We can then write

$$N_{2(t=\tau/4)}=N_2-\left(\left|\frac{dn_{2RV}}{dt}\right|+\left|\frac{dn_{2RT}}{dt}\right|-\left|\frac{dn_2}{dt}\right|\right)(\tau/4). \quad \text{(A1)}$$

The voltage then decreases to zero and the number of trapped carriers decreases at the rate $-dn_1/dt-|dn_{1RV}/dt|_{\text{net}}$ because the falling voltage decreases the generation rate, plus decay takes place at the natural recombination rate. Thus at the end of the first half-cycle of the cw current the number of free carriers becomes

$$N_{2(t=\tau/2)}=N_2-\left(\left|\frac{dn_{2RV}}{dt}\right|+\left|\frac{dn_{2RT}}{dt}\right|-\left|\frac{dn_2}{dt}\right|\right)(\tau/4)-\frac{dn_{2d}}{dt}(\tau/4), \quad \text{(A2)}$$

where dn_{2d}/dt equals the decrease in free carriers as a function of time during the decreasing voltage cycle and includes natural recombination.

The above process now continues for the increasing and decreasing negative cycle. It is thus seen that if the regime is continued there is a value of dn_{2d}/dt such that $N_2(t)$ will be less at $t=5\tau/4$ than at $t=\tau/4$. The net effect must be cumulative. Thus the summation sign is used in order to sum the differentials which are approximated by the measured data. The current is then directly proportional to the free-carrier population through $\vec{J} = I/A = Ne\mu\vec{E}/\text{Vol.}$[9]

The same treatment as above could be applied to the number of filled trapping sites at swtiching, N_{1S}. Under the condition

$$\frac{dn_1}{dt} < \frac{dn_{1RV}}{dt},$$

it can be shown that as each cycle continues, the number of filled trapping sites, N_1, will decrease. Thus as the increasing field is reapplied, the charging must first fill the appropriate number of traps, and the excess energy can then populate the conduction band. However, as N_1 decreases, this excess energy is less.

To explain Fig. 1, the question must be asked what condition must be specified for zero decay such that N_2 at $t=(m+1)\tau+\tau/4$ to be equal to N_2 at $t=m\tau+\tau/4$? Obviously, the first condition is

$$\left|\frac{dn_2}{dt}\right| > \left(\left|\frac{dn_{2RV}}{dt}\right| + \left|\frac{dn_{2RT}}{dt}\right|\right).$$

From Eq. (A2) to avoid decay in $I_{\text{on(max)cw}}$ we must specify

$$\frac{dn_2}{dt} - \left(\left|\frac{dn_{2RV}}{dt}\right| + \left|\frac{dn_{2RT}}{dt}\right|\right) \geq \left|\frac{dn_{2d}}{dt}\right|.$$

Thus

$$\frac{dn_{2d}}{dt} < \frac{dn_2}{dt}.$$

But

$$\frac{dn_{2d}}{dt} = \frac{dn_2}{dV_d}\frac{dV_d}{dt},$$

thus

$$\frac{dn_2}{dV_d}\frac{dV_d}{dt} < \frac{dn_2}{dV_i}\frac{dV_i}{dt},$$

where subscripts d and i represent decreasing and increasing voltage.

From Fig. 1 it is clear that $dV_d/dt \gg dV_i/dt$. Thus we conclude that $dn_2/dV_d \ll dn_2/dV_i$.

This states that the decrease in free carriers with decreasing electric field is much less than the increase in free carriers with increasing electric field in the Fig. 1 regime where $N_2(t)_{\max}$ is essentially constant. This is because the decrease in voltage (beneath the holding voltage) with respect to time is occurring in a time shorter than the recombination time of the distribution.[10]

---

[1]G. K. Gaule, P. LePlante, S. Levy, and S. Schneider, IEDM Technical Digest 279 (1976).

[2]G. C. Vezzoli, J. Appl. Phys. (to be published).

[3]G. C. Vezzoli, P. J. Walsh, and L. W. Doremus, J. Non-Crystalline Solids 18, 333 (1975).

[4]G. C. Vezzoli, P. Calella, and L. W. Doremus, J. Appl. Phys. 44, 341 (1973).

[5]P. J. Walsh and G. C. Vezzoli, Appl. Phys. Lett. 25, 28 (1974).

[6](a) P. J. Walsh and G. C. Vezzoli, in *Amorphous and Liquid Semi conductors*, edited by J. Stuke and W. Brenig (Taylor and Francis, London, 1974), p. p. 1402. See (Eq. 37) and description. (b) In the case of the pure square wave the discontinuous slope changes in the negative polarity for voltage and current are not observed. These slope changes in Fig. 1 are due to carrier decay caused by the slower decrease in negative voltage as compared to the decreasing positive voltage. This effect is caused by the protective diode; however, the decay is compensated by the very fast rise time in the positive voltage polarity. Hence the net change in maximum current in both positive and negative directions is zero (for the distorted square cw). The observation of the discontinuous slope change in falling current versus time for the negative polarity identifies a critical decay constant specifying a subholding voltage time that causes highly nonlinear *I-V* conditions in the on state. The nonlinear *I-V* under such conditions is clearly observed in Ref. 2. The subholding voltage time is about 200 ns.

[7]J. B. Goodenough, Phys. Rev. 117, 1443 (1960).

[8]Evidence for trapping is given by P. M. Raccah, Bull. Am. Phys. Soc. 21, 302 (1976).

[9]In these experiments the maximum voltage in the negative cycle for the distorted triangular wave is controlled so as to be equal to that for the distorted square CW, or greater than the holding voltage. Thus, the decaying of the on-state is *not* a result of the diode curtailing the triangular CW amplitude.

[10]From Refs. 1 and 2, the interelectrode volume is about 10^{-8} cc for the NbO_2 device. Therefore, carrier recombination in Fig. 2 occurs at 5×10^{17} carriers/CC which approaches the inverse of screening degeneracy conditions in Mott insulators (inducing a transition to the metallic state via interfering with the Coulombic electron-hole attractive potential and releasing carriers into the conduction band). This implies that a time interval of a multiplier factor of 10 larger than the 200 ns utilized in the calculation would cause the carrier recombination rate to offset the conditions required for Mott degeneracy such that total decay to the off-state would result. This is borne out in Fig. 2 showing that after about 3 μsec the current decays to the off-state level.

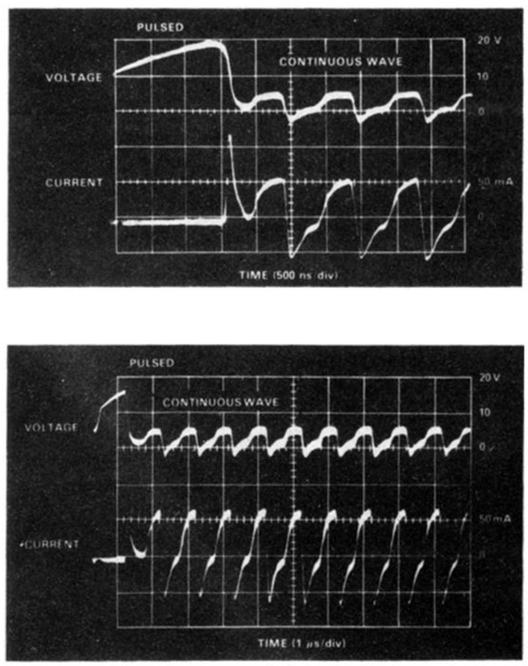

FIG. 1. Sustaining of the on state of NbO_2. Upper oscillogram: Pulsed and 1-MHz square-wave voltage (upper trace) and current for switch-on and sustained on state. $V = 10$ V/div, $I = 50$ mA/div, $t = 500$ nsec/div. Devices are planar thin films of NbO_2 in diode packaging with aluminum dot contacts, and are described in detail in Refs. 1 and 2. Lower oscillogram: Same effect as above but for shorter set pulse, $t = 1$ μsec/div. In both oscillograms the negative cycle of the square wave is noncritically distorted by the input diode which is utilized to protect the two power sources from each other.

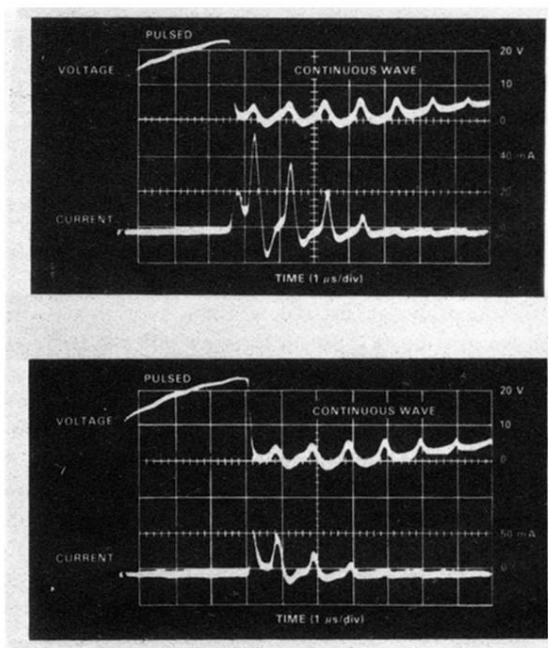

FIG. 2. Decaying of the on state of NbO_2. Upper oscillogram: Pulsed and 1-MHz triangular-wave voltage (upper trace) and current for switch-on and decay of on state. $V = 10$ V/div, $I = 20$ mA/div, $t = 1$ usec/div. Note decaying envelope of on-current. Lower oscillogram: Same effect of decaying on state for longer set pulse, $I = 50$ mA/div.