Physics – Uspekhi **48** (7) 703 – 712 (2005) © 2005 Uspekhi Fizicheskikh Nauk, Russian Academy of Sciences

REVIEWS OF TOPICAL PROBLEMS PACS numbers: **84.70.+p**, **85.30.–z**, 85.30.Kk

Nanosecond semiconductor diodes for pulsed power switching

I V Grekhov, G A Mesyats

DOI: 10.1070/PU2005v048n07ABEH002471

Contents

1. **Introduction** **703**
2. **General information about semiconductor opening switch** **703**
3. **Drift step recovery diodes** **705**
4. **SOS diodes** **708**
5. **Silicon carbide opening switch** **710**
6. **Conclusions** **711**

References **711**

Abstract. **The development of semiconductor-based nano- and subnanosecond high opening switches is crucial for advancing modern research in experimental physics and radioelectronics, particularly with increasing power (to 10^{10} W) and repetition rate (to 10^4 Hz) of impulse devices. Highlighted in this review are two types of silicon diodes: drift step recovery diodes (DSRDs) and SOS diodes with the attainable current densities and switched-off powers being 10^2 A cm^{-2} and 10^8 W in the former case, and 10^5 A cm^{-2} and 10^{10} W in the latter. The possibility of utilizing not only monocrystalline silicon (as in DSRDs and SOS diodes) for the base material but also monocrystalline silicon carbide is examined.**

1. Introduction

Nano- and subnanosecond electric pulses with a peak power ranging from megawatts to terawatts are employed in a number of state-of-the-art technology areas like relativistic microwave electronics, ultrabroadband radar, electromagnetic countermeasures, electromagnetic compatibility investigation of complex systems, underground radar, laser and accelerator power suppliers, etc. High-power short pulses are also employed in several branches of contemporary experimental physics, for instance, in the field of controlled nuclear fusion and other large-scale physical experiments [1].

There are two approaches to high-power nanosecond pulse generation, which differ by the way in which they store energy: accumulation in capacitive storage devices (low-inductance capacitors and forming lines) with subsequent energy transfer to the load via a closing switch, and accumulation in the magnetic field of the inductive current circuit; in the latter case, transferring the energy to the load requires effecting nanosecond breaking of a high current. The latter method is of paramount importance for high-power pulsed technology, because the stored energy density in inductive storage devices is one and a half to two orders of magnitude higher than in capacitive devices, the cost of storage devices is significantly lower, and, which is also of significance, the pulsed voltage across the load at current breaking may be substantially higher than the voltage at intermediate stages of pulse formation. However, rapid high-current interruption, which involves breaking currents that range into kiloamperes for pulsed megavolt voltages, is a substantially more complicated task than rapid closing.

At the stage of laboratory experiments, this problem is commonly solved with the aid of plasma breakers with nano- or microsecond pumping, injection thyratrons, and exploding wires. For practical applications, however, especially in industrial technologies, these circuit components cannot be used — primarily owing to the short service life of the breakers, actuation instability, and the impossibility of using them in pulse-periodic modes.

In principle, for these purposes it would be advisable to take advantage of semiconductor devices with their practically unlimited service life and high stability, but the combination of nanosecond breaking time and gigawatt pulsed power is a highly nontrivial task for semiconductor devices.

2. General information about semiconductor opening switch

Of conventional commercially produced devices, current breaking for a time on the order of 10 ns can be carried out in special-type field-effect transistors. This transistor is, in essence, a power integrated circuit comprising hundreds of thousands of microtransistors with a characteristic dimension of 10 – 15 μm, operating in parallel. The operating voltage of the device is several hundred volts and the current is dozens of

I V Grekhov A F Ioffe Physical-Technical Institute,
Russian Academy of Sciences,
ul. Politekhnicheskaya 26, 194021 St.-Petersburg, Russian Federation
Tel./Fax (7-812) 247 91 23
E-mail: grekhov@mail.ioffe.ru
G A Mesyats P N Lebedev Physics Institute,
Russian Academy of Sciences,
Leninskii prosp. 53, 119991 Moscow, Russian Federation
Tel. (7-095) 135 24 30. Fax (7-095) 952 50 81
E-mail: mesyats@nsc.gpi.ru

Received 28 January 2005
Uspekhi Fizicheskikh Nauk **175** (7) 735 – 744 (2005)
Translated by E N Ragozin; edited by A Radzig

amperes, so that to produce a pulse with a power of, say, 50 MW, the breaker should consist of 10^4 transistors. Owing to the obvious complexity and high cost of these systems, the question of their development has not even been discussed.

The simplest semiconductor opening switch is a conventional p^+nn^+ diode. When a current passes through it in the conducting direction, the weakly doped n-base is filled with electron – hole plasma due to the injection of electrons and holes across the potential barriers of the n^+n and p^+n junctions. A reverse current pulse is then allowed to pass through the diode (the minus on the p^+-contact); in this case, the holes are drawn out of the plasma by an external field via the p^+-contacts, and the electrons via n^+-contacts. While hole concentration at the p^+n junction exceeds the equilibrium one, the diode carries a direct current limited by the load resistance — this is the phase of a high reverse conduction (HRC). Then, a space charge region (SCR) begins to form, whose boundary displaces from the p^+n junction to the n-base, the voltage across the device increases, and the current in the circuit is lowered — this is the phase of reverse resistance recovery (RRR). The problem amounts, in essence, to making this process fast enough.

A subnanosecond semiconductor breaker was made for the first time as far back as the 1950s: it was the so-called charge-storage diode (CSD) (see, for instance, Refs [2, 3]). This device is extremely simple in design: in a silicon wafer with an n-type conductivity, due to boron diffusion from the surface there forms a p^+n junction and a base region with an abrupt density gradient, i.e., with a strong built-in electric field. In the flow of forward current, the holes injected by this current at a low injection level are retarded by the built-in field near the injector. A rapidly rising pulse of reverse current is then allowed to pass through the diode, and the stored holes are almost completely expelled at the stage of high reverse conduction, after which the current through the diode is abruptly, in $10^{-9} - 10^{-10}$ s, interrupted to go over to the load connected to the diode in parallel. The CSD is indisputably an extremely simple opening switch with a very fast response, but the avalanche breakdown voltage of such a p^+n junction with a heavily doped base lies within the 10 – 50-V range and the working current amounts to hundreds of milliamperes, which is absolutely insufficient for producing high-power high-voltage pulse generators.

Generally speaking, a conventional high-power high-voltage semiconductor diode also constitutes a opening switch in the forward-to-reverse bias switching; in this case, the switched-off power of a single device may be very high — several megawatts — but under ordinary conditions the duration of the current breaking process lies in the microsecond range rather than the nanosecond one. The physics of this process under the conditions of high reverse-current density was considered in detail as early as 1967 in the classical work by Benda and Spenke [4]. Although the electron – hole plasma accumulation in the n-base of the p^+nn^+ structure for direct bias was calculated neglecting all nonlinear effects and the p^+n-junction recovery in the flow of a high reverse current was calculated under several unrealistic assumptions (the constancy of reverse current in time, the field-independent carrier mobility, etc.), the principal physical features of the process were determined quite clearly. First of all, it was revealed that the fall of the carrier concentration at the blocking p^+n junction to the equilibrium value and the onset of space charge region formation in the same place do not result in a sharp decrease in reverse current, when there is

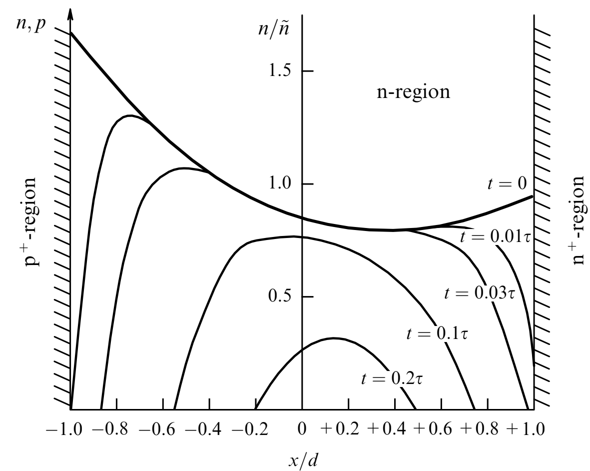

Figure 1. Electron – hole plasma distribution profile in the n-base of a silicon p^+nn^+ structure in the flow of direct forward current j_F $(t = 0)$ and then in the flow of reverse current $j_R = 10 j_F$. The n-base thickness d is equal to the ambipolar diffusion length $L = \sqrt{D\tau}$, where D is the ambipolar diffusion coefficient, and τ is the carrier lifetime in the n-base for a high injection level; $\tilde{n}$ is the average plasma concentration.

a region at the SCR boundary filled with electron – hole plasma; the character of reverse-current decrease is controlled by the plasma 'resolution' in precisely this region. The calculated plasma distribution profile for a high level of injection in the n-base of the silicon p^+nn^+ structure is plotted in Fig. 1, assuming the ratio of forward and reverse current densities $j_R / j_F = 10$. One can see that the plasma 'reservoir' at the SCR boundary (the left part of the structure) exists for a long time to determine the slow decrease in reverse current and lengthen the turning-off process. The asymmetric plasma distribution in the flow of forward current and the high rate of recovery at the left boundary, towards which the external field draws out the holes, stem from the fact that the hole mobility in silicon is three times lower than the electron mobility: $\mu_n = 3\mu_p$.

Figure 2 displays a simplified picture of the motion of plasma fronts in the flow of reverse current upon the emergence of an SCR, which corresponds to the examination in Ref. [4]. The fronts are assumed to be abrupt here, and the

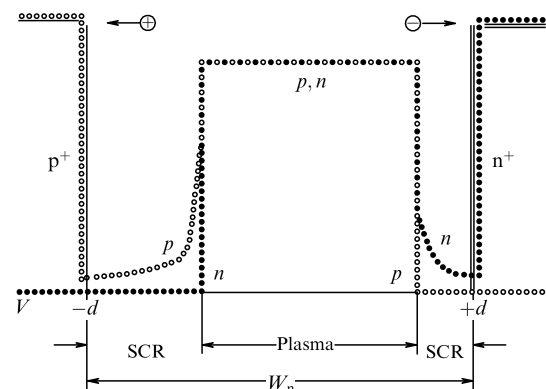

Figure 2. Schematic of the carrier distribution and formation of plasma fronts in the n-base of thickness W_n in the flow of a high reverse current $j_R = 10 j_F$ through the p^+nn^+ structure. SCR stands for the space charge region.

plasma concentration is assumed to be constant along the coordinate. Usually, the average plasma concentration $\tilde{n}$ is rather high ($10^{16}-10^{17}$ cm^{-3}) and its relaxation time for neutrality violation is short (10^{-12} s), and therefore the transfer of holes leftwards and electrons rightwards are rigidly interrelated processes. It was shown that under these conditions the velocity of travel of the left boundary is given by

$$V_1 = \frac{\mu_n}{\mu_n + \mu_p} \frac{j_R}{q\tilde{n}} ,$$

and of the right boundary is defined as

$$V_r = \frac{\mu_p}{\mu_n + \mu_p} \frac{j_R}{q\tilde{n}} ,$$

so that in a silicon diode, where $\mu_n = 3\mu_p$, the left boundary travels three times faster. When the reverse current density $j_R \gg qn_0V_S$ (n_0 is the equilibrium electron concentration in the n-base, and V_S is their saturation velocity), the field in the SCR on the right and on the left is controlled by the traveling carrier charge, i.e., depends on the current density.

In principle, this level of understanding of high-power diode recovery for a high density of reverse current allowed the formulation of the basic principles underlying the design of a high-power nanosecond diode breaker even at that time, but this was not done. The reason most probably lay with the insufficient level of development of high-power semiconductor pulse technology at that time as a whole; in this sense, the work of Benda and Spenke [4] significantly outstripped its time.

3. Drift step recovery diodes

Purposeful work in the quest for a high-power diode nanosecond breaker was carried out in the A F Ioffe Physical-Technical Institute, Russian Academy of Sciences (PTI RAS) in the early 1980s. The impetus was given by the findings of the work by Grekhov et al. [5], who studied the feasibility of developing a high-voltage power CSD and showed, in particular, that the HRC phase duration increases and the RRR phase duration decreases to a value shorter than 0.1 μs as the diffused p^+n-junction depth increases (Fig. 3). It is pertinent to note here that the diodes in this work were manufactured by technology which had been specially developed in the USSR [6] for the power semiconductor industry.

The main feature of the technology consists in the fact that the deep p^+n junctions are created through combined diffusion of boron and aluminium from their oxides in the atmosphere. The surface concentration of Al in this process has a strictly defined value of $(5-7) \times 10^{16}$ cm^{-3}, and therefore the diffusion layer consists of two regions: a heavily doped ($\sim 10^{19}$ cm^{-3}) 'boric' p^+-region 10–20 μm deep, and a lengthy (80–120 μm), relatively weakly doped 'aluminium' region with a smoothly weakening gradient of impurity concentration. The Shockley–Read lifetime τ_n of minority carriers in this region, which shortens with increasing majority p-carrier concentration as

$$\tau_n = \tau_{n0} \left(1 + \frac{p}{p_k}\right)^{-1} ,$$

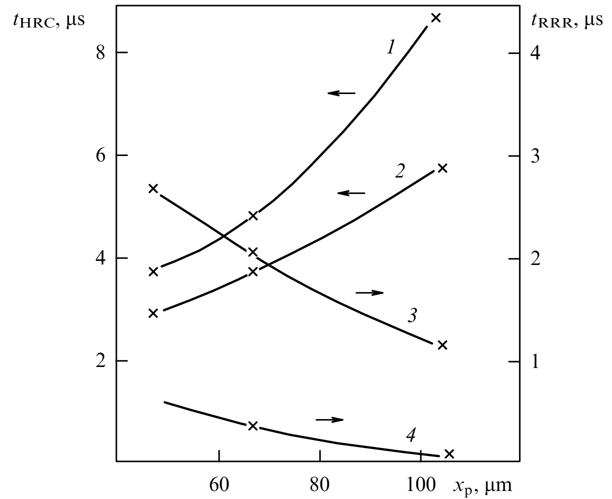

Figure 3. Dependences of t_{HRC} (*1*, *2*) and t_{RRR} (*3*, *4*) on the pn-junction depth x_p. The parameters of the diodes are as follows: base resistivity $\rho_{W_n} = 50\ \Omega$ cm, base thickness $W_n = 100$ μm, and carrier lifetime $\tau_p = 20$ μs. Measurement mode: the densities of forward and reverse currents are equal, $j_F = j_R$; the pulse length $t_F = 100$ μs; *1*, *3* — $j_F = 1$ A cm^{-2}; *2*, *4* — $j_F = 10$ A cm^{-2}, the reverse voltage $U_R = 10$ V.

according to Landsberg and Kousik [7], is on the order of ten microseconds (here, τ_{n0} is the lifetime in a weakly doped material, and $p_k \approx 10^{16}$ cm^{-3} is the experimentally determined value of the boundary number density). Therefore, when a forward current flows through this p^+pnn^+ diode, the p-region turns out to be 'flooded' with electron–hole plasma. In the switching there occurs a rapid lowering of plasma concentration at the p^+p junction; however, unlike the situation with an abrupt p^+n junction, this does not result in SCR formation, because the current is effected by the majority carriers of the p-layer. The plasma front travels through the p-layer towards the pn junction, and only when this front approaches the junction does the SCR begin to form and the reverse current begin to drop. Therefore, according to Grekhov et al. [5], the enhancement of p^+pn-junction depth leads to increasing the HRC phase duration and shortening the RRR phase duration, because a significant portion of the charge turns out to have been withdrawn from the diode by the instant of SCR formation. This is precisely the p^+p-junction design subsequently employed in all high-power nanosecond diode breakers.

As noted above, the presence of electron–hole plasma at the boundary of the expanding SCR slows down the expansion process, i.e., it lowers the build-up of the voltage across the diode and delays the current decrease. Therefore, according to modern views, the recovery should proceed in such a way that the plasma front traveling through the p-region from the p^+p junction to the pn junction and the front traveling through the n-base from the n^+n junction to the pn junction should meet precisely in the pn-junction plane. In this and only this case would the flow of reverse current and the SCR expansion proceed due to the fast motion of majority carriers alone in the opposite directions from the pn junction.

However, this is hard to accomplish exclusively with design approaches. For instance, when designing a device with an operating voltage of, say, 1.7 kV on the basis of n-type conduction silicon commonly utilized in high-power devices, the maximum SCR width and, hence, the n-base thickness

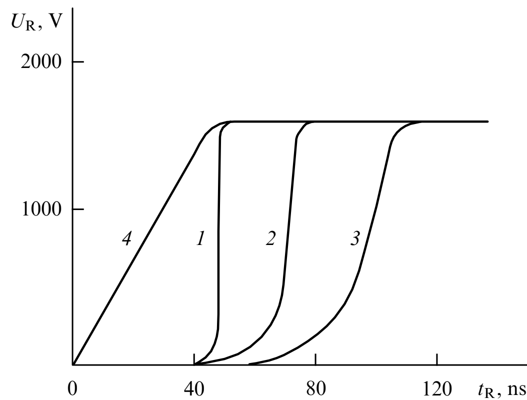

Figure 4. Voltage rise rate across the diode in the course of recovery after the flow of forward current. Mode parameters are: forward current density $j_F = 15$ A cm^{-2}, and reverse current density $j_R = 100$ A cm^{-2}. Curve *4* depicts shape of the pulse of reverse voltage U_R ($j_F = 0$) applied. Forward current pulse duration is as follows: *1* — $t_F = 0.4$ μs, *2* — $t_F = 0.8$ μs, and *3* — $t_F = 1.2$ μs.

should be greater than 140 μm, and the thickness of a p-region made by diffusion technique cannot be greater than 100 – 120 μm. Then, for a more or less uniform plasma distribution over the p- and n-regions of the device, the fronts will meet in the n-base (because the front velocity in the p-region is three times higher), and the current breaking will be rather slow. Grekhov et al. [8, 9] showed for the first time that the nanosecond current interruption in a silicon p^+pnn^+ diode may be effected if the pulse duration of forward current is made short enough for the greater portion of the total amount of injected plasma to be concentrated in the p-region.

The main experimental result of these works is illustrated in Fig. 4. A forward current pulse I_F with an amplitude of 3 A and a duration t_F adjustable in the 0.4 – 1.2 μs range was allowed to pass through the sample, and then a pulse of reverse voltage was applied, which built up to 1.7 kV in 40 ns (curve *4*, $I_F = 0$). It is clearly seen that, as t_F decreases, the rise time of the voltage across the diode shortens to ~ 2 ns for $t_F = 400$ ns. The processes occurring in this case are schematically displayed in Fig. 5 borrowed from Ref. [10]. The p^+pnn^+ structures investigated (see Fig. 5a) were made by combined diffusion of Al and B in n-Si with a donor concentration of 10^{14} cm^{-3}; the pn-junction depth was 120 μm, the p^+-layer thickness was 50 μm, the n-base thickness was 200 μm, and the working area was 0.3 cm^2. The n^+-region was made by phosphorus diffusion to a depth of 50 μm. The plasma distribution profile under pumping by a short (400 ns) current pulse is plotted with a dashed line in Fig. 5a. In the vicinity of the p^+-layer there forms a thin layer of electron – hole plasma with a concentration on the order of 10^{17} cm^{-3}, whose thickness increases owing to ambipolar diffusion as $\sqrt{Dt_F}$, where D is the ambipolar diffusion coefficient. In front of this layer there forms a concentration wave whose front moves rapidly towards the n^+-layer under the conditions of bipolar drift (for a concentration of $\sim 10^{15}$ cm^{-3}).

A formation of strongly nonuniform distribution results: the major part ($\sim 75\%$) of the plasma is concentrated in the p-layer. When a rapidly rising pulse of reverse current passes through the device, the front of the concentration wave travels rapidly in the opposite direction, from the n^+-layer to the pn junction (Fig. 5b). The plasma concentration near

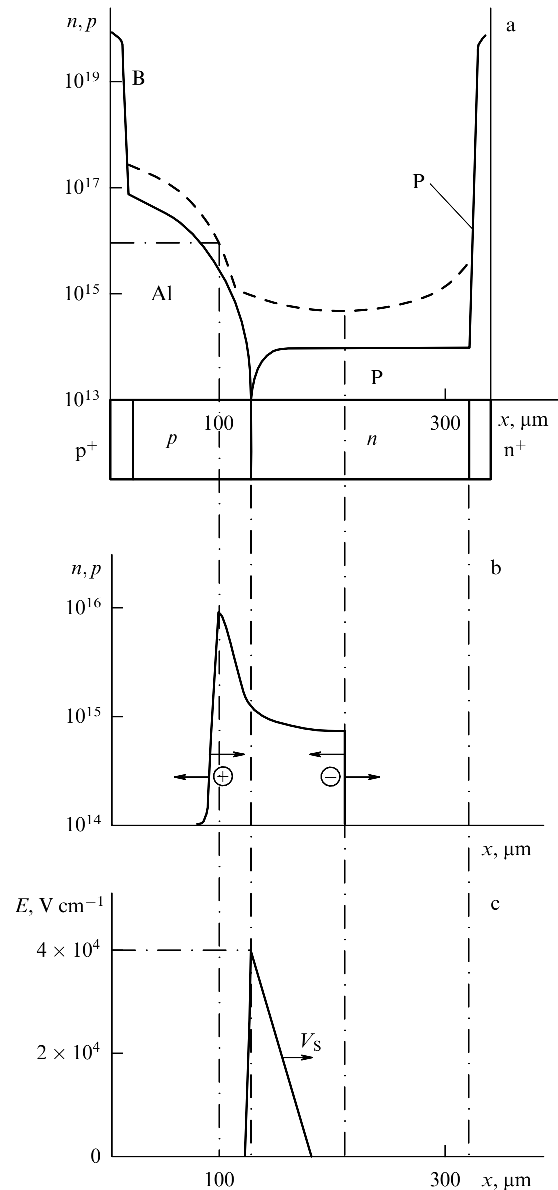

Figure 5. (a) Semiconductor structure design; the dashed line shows the plasma distribution upon the passage of a short forward current pulse. (b) Motion of plasma fronts in the passage of a reverse current pulse. (c) SCR formation upon collapse of the fronts.

the p^+p junction is simultaneously lowered owing to transfer of the holes leftwards, and the concentration front produced travels through the p-layer to the right towards the pn junction. The forward-to-reverse current pulse parameter ratios for a specific p^+pnn^+-structure design are so selected that the fronts meet in the vicinity of the pn junction. From this moment there is no plasma in the diode, and the flow of reverse current is effected due to the motion of majority carriers in the opposite directions from the pn junction (Fig. 5c). As this takes place, an SCR forms near the pn junction, the voltage across the diode increases sharply, and the current through the diode breaks. The rate of this process is defined primarily by the velocity of travel of the SCR boundary through the n-base. That is why the mode parameters are commonly selected in such a way as to ensure that the reverse current density j_R reaches a value $j_S \approx qn_0V_S$ at the instant of time when the fronts meet and that the boundary early in the breaking process travels with the

ultimate possible — saturated — velocity $V_S = 10^7$ cm s^{-1}. This proposition is quite significant: the electrical neutrality in the n-layer behind the trailing edge of the concentration wave is violated for $j_R > j_S$, which is responsible for the occurrence of a 'pedestal' in the voltage pulse being formed and for a rise in switching losses, and for $j_R \ll j_S$ the power being switched diminishes and the current breaking time t_{off} lengthens. In pulsed systems one usually has $j_R \approx 0.8 j_S$ and the ultimately short switching time is $t_{off} \approx W_{SCR}/V_S$, where W_{SCR} is the width of the space charge region for an operating voltage U_b across the device. In particular, for the diodes investigated in Refs [8, 9] it was found that $U_b = 1.7$ kV, $W_{SCR} = 150$ μm, $t_{off} \geqslant 1.5$ ns, and $j_R \approx 130$ A cm^{-2}.

The plasma accumulation in the flow of current through the diode has been studied in sufficient detail. The fundamental process which limits the ultimate possible plasma concentration in silicon is the Auger recombination [11, 12], which is responsible for a sharp decrease in $\tau_{n,p}$ with increasing concentration: from $\sim 7 \times 10^{-5}$ s at $n \approx 10^{17}$ cm^{-3} to 10^{-9} s at $n \approx 10^{19}$ cm^{-3}. An even more significant limitation is the fact that silicon p^+p and n^+n homojunctions are far from being ideal injectors of holes and electrons. With an increase in current density, the injection coefficient goes down owing to the escape of nonequilibrium carriers through potential barriers to the heavily doped p^+- and n^+-layers with a very high recombination rate. The flux through the barrier is approximately proportional to the plasma concentration squared, which limits the possibility of increasing the amount of stored plasma in the diode by increasing the forward current density [13]. This density normally amounts to 30–50 A cm^{-2}, and the charge accumulated in the plasma for a pulse length of, say, 400 ns is equal to $(10-15) \times 10^{-6}$ C cm^{-2}. Then, for a linear rise in reverse current pulse, its requisite working density (~ 200 A cm^{-2}) will be reached in 100–150 ns, after which the current will break because the plasma would have been completely removed from the diode. In actual practice, owing to various kinds of charge losses, this time would be substantially shorter. The above device is referred to as a drift step recovery diode (DSRD).

The working voltage of DSRDs normally ranges from 500 to 1700 V, which corresponds to the doping level n_0 of initial silicon from 10^{15} cm^{-3} to 10^{14} cm^{-3} and to the limiting speed of response from ~ 0.6 to ~ 2 ns; a higher level of n_0 hampers the process of obtaining deep diffused pn junctions with a low surface dopant concentration, and for $n_0 < 10^{14}$ cm^{-3} the voltage drop across the n-base upon passage of the trailing edge of a concentration wave becomes too large. It was found that, despite the relatively low working voltage of single devices, developing high-voltage — for hundreds of kilovolts — assemblies on their basis is not a serious problem. Since the high voltage is applied to the device only during current breaking, i.e., for several nanoseconds, during which the surface breakdown does not manage to develop, there is no need either for voltage splitters or for a special design of the edge profile of the device, which significantly simplifies the assembly design.

Of fundamental importance to the generation of high-voltage nanosecond pulses is the recovery synchronism of a large number of diodes connected in series. Naturally, the amplitude and duration of the forward current pulse, as well as the rate of reverse-current pulse rise, are automatically strictly equal for all diodes of the assembly, but the total amount of plasma introduced by the forward current pulse into the p- and n-regions, as well as the form of plasma distribution, may, in principle, vary from diode to diode owing to the scatter of the carrier lifetimes in the p- and n-layers. This may result in a recovery time mismatch of diodes and the overall process of current interruption would slow down. It turned out, however, that the technological processes elaborated furnish a sufficient reproducibility of these parameters to obtain a nanosecond current breaking in the assembly, because the duration of the forward current pulse (hundreds of nanoseconds) is much shorter than the average lifetime $\tau_{n,p}$ of nonequilibrium carriers in the p- and n-layers, and the p^+p-junction injection coefficient is practically equal for all devices.

The amplitude of reverse current pulse in a single DSRD element may be very high, because the simplicity of the technology permits making devices around silicon wafers of any industrially utilized diameter (up to 125 mm). However, experiments show that the duration of a current interruption process lengthens as the diameter increases in wafers with diameters of 25 mm and larger. This is hypothesized to be due to the skin effect, but no deliberate investigation of this problem has been undertaken. The working area of a device on a wafer 25 mm in diameter is equal to ~ 4 cm^2, i.e., the pulse amplitude at $j_R = 200$ A cm^{-2} amounts to 800 A, and the pulsed power is equal to 1.2 MW at an operating voltage of 1.5 kV. It is possible to built high-power nanosecond pulse generators around these devices, since they are easily connected in series and in parallel. The advent of DSRDs in 1983–1985 led to radical changes in high-power semiconductor pulsed technology: the generators of nanosecond pulses with a power ranging into the dozens of megawatts, operating at frequencies of hundreds of hertz, came to be quite an ordinary equipment. For instance, Efanov et al. [14] described a pulse generator with a power of 64 MW (80 kV, 0.8 kA), a pulse rise time of 0.8 ns, and a repetition rate of 1 kHz. Figure 6 shows a generator developed in the PTI RAS for experiments in cleaning biogas by a pulsed corona discharge [15]. The operating frequency of the DSRD assemblies may, in principle, be very high, because the next cycle can commence practically immediately after the passage of forward and reverse current pulses (i.e., ~ 500 ns after the beginning of a cycle). In reality, the frequency capacity is determined by thermal limitations in forward and reverse current pulse formers, where power transistors are commonly employed as switches.

In some application areas, specifically in laser engineering, the delay between the control pulse and the high-power nanosecond pulse should be as short as possible. In DSRD-based devices, this delay is determined by the total duration of forward and reverse current pulses and cannot be shorter than 200–300 ns, its main part being the duration of the forward current pulse. If a semiconductor structure is prepared in which the plasma distribution nonuniformity required for a sharp interruption of reverse current is effected for a direct forward current, the delay will be equal to the duration of reverse current pulse and may be shortened to 15–20 ns. Such structures were obtained by decreasing the injection coefficient of the n^+n junction in the p^+pnn^+ structure [16] by a strictly controllable lowering of the doping level of the n^+-layer region from which electron injection occurs. In this structure, the plasma concentration at the n^+n junction in the flow of forward current is much lower than at the p^+p junction, and in the passage of the reverse current pulse the plasma front is first formed at the n^+n junction and not at the

Figure 6. Generator of high-power nanosecond pulses. The pulsed voltage is 70 kV, the pulse current 3 kA, the rise time 3 – 5 ns, the pulse energy 2 J, and the repetition rate may be as high as 2 kHz.

p^+p junction. As indicated by experiments, in this diode with inverse recovery order (inverse recovery diode, IRD), the working forward current density is substantially lower than in a DSRD, and the stored amount of plasma is somewhat smaller. However, the optimal density of reverse current should have the same magnitude as in DSRDs, and therefore the rise time of reverse current should not exceed 15 – 20 ns, which defines the pulse delay time.

4. SOS diodes

As shown in the previous section, the working current density in drift step recovery diodes cannot, in principle, exceed 200 – 300 A cm^{-2}, while the skin effect limits the possibility of increasing the working diode area above $\sim 4\ cm^2$. That is why to develop DSRD-based breakers of the gigawatt power range with the interrupted current ranging into the dozens of kiloamperes and the working voltage running into the hundreds of kilovolts requires connecting a very large number of devices in parallel and in series. Estimates show that the cost and complexity of such systems become unrealistically high.

A breakthrough into the gigawatt power range was made in 1992 – 1993: at the Institute of Electrophysics, Ural Branch of the Russian Academy of Sciences (IEP UB RAS) it was experimentally established that a sharp current breaking is also observed for very high forward and reverse current densities (one – two orders of magnitude higher than the densities optimal for the DSRD mode) in a certain range of current densities and pulse lengths [17], its mechanism clearly being different from that in the DSRD. Subsequent experiments and calculations [18 – 23] enabled construction of the physical picture of this phenomenon, which the authors termed the silicon opening switch (SOS) effect.

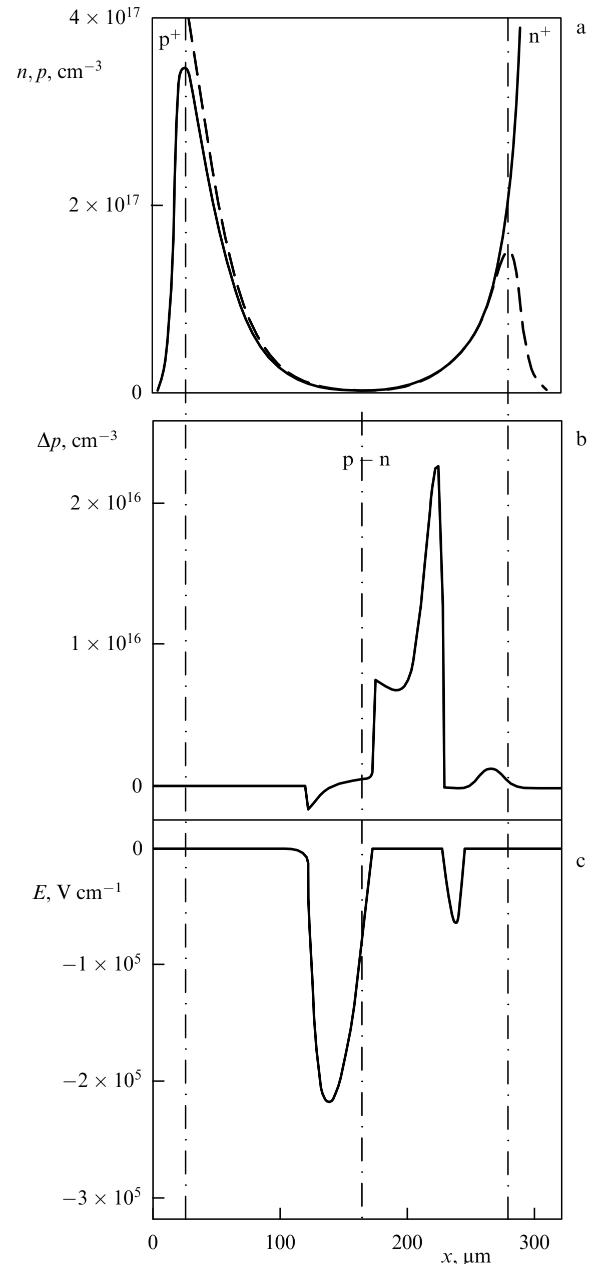

Figure 7. Calculated distributions of plasma concentration and field density in an SOS diode. (a) Electron n (solid line) and hole p (dashed line) concentration distributions at the end of a forward current pulse for $j_F = 0.8$ kA cm^{-2} and a length of 360 ns. (b, c) Distributions of excess-hole concentration $\Delta p = p - N_a$ and field intensity in the breaking of reverse current pulse $j_{R,\max} = 4.2$ kA cm^{-2} at the instant of highest voltage across the structure.

Physical-mathematical simulations of the SOS process involved simultaneous numerical solution of the Kirchhoff equation for the electrical circuit with an SOS diode, the continuity equations for electrons and holes in the diode structure, and the Poisson equation. Shown as an illustration in Fig. 7 are the calculated parameters of the SOS process in

the pumping and recovery of an assembly consisting of 160 diode p^+pnn^+ structures with an area of 0.36 cm^2, a pn-junction depth of 165 μm, an n-base thickness of $\sim$ 65 μm, and a donor concentration of 10^{14} cm^{-3} in the base [22]. The diffusion-layer fabrication technology was similar to that described in the previous section. The load resistance was equal to 200 Ω. The calculated plasma distribution in the device at the end of pumping by a short ($\sim$ 360 ns) pulse of forward current ($j_F = 0.8$ kA cm^{-2}) is depicted in Fig. 7a; Fig. 7b shows the location of the plasma fronts, and Fig. 7c the field distribution in the breaking of the current with a density of 4.2 kA cm^{-2} and a rise time of 35 ns.

It is clearly seen that, owing to the high density of forward current, the plasma concentration introduced by the bipolar drift into the central part of the diode is higher than in the DSRD process. The steeply rising pulse of reverse current forms steep plasma fronts in the p- and n-layers, traveling towards each other, the front in the p-region traveling with a significantly higher velocity. The flux density of holes carried out by the field from the plasma across the left boundary, $p = j_R(qV_s)^{-1}$, continuously increases with increasing current. The plasma front, while traveling to the right through the diffusion p-layer, passes through regions with a continuously diminishing concentration N_a of acceptor dopant. For $p > N_a$, the space charge of uncompensated free holes produces an electric field determined by the concentration difference of the holes moving with a saturation velocity and the immobile acceptors. The field intensity rises steeply, and the width of a space charge region increases as the plasma boundary changes its position.

At this stage, the voltage across the diode rises rapidly and the current goes over to the load connected in parallel to the diode. Naturally, the lowering of current through the diode lowers the flux density of holes in the SCR (in a time on the order of the flight time $\tau = W_{SCR} V_S^{-1} \sim 0.2$ ns), but simultaneously the acceptor concentration is also lowered, because the boundary displaces towards the pn junction; this retards the decrease in the field intensity in the SCR. Calculations show that the field in the SCR reaches the threshold of impact ionization in silicon ($> 2 \times 10^5$ V cm^{-1}) for a lowering of current by 30–40%, resulting in the emergence of an electron component of current in the SCR, which decreases the front propagation velocity

$$V_f = \frac{j_R}{q\Delta p}\left(\frac{b}{b+1} - \frac{j_n}{j_R}\right),$$

where $b = \mu_n/\mu_p$, j_n is the electron current density, and Δp is the excess hole concentration at the pulse front. A fundamentally important feature of the SOS process is the fact that all of the above described processes take place in a rather heavily doped p-region; unlike the DSRD process, the pn junction and the weakly doped n-base remain 'flooded' with a high-density electron–hole plasma and barely participate in the current breaking. The second important feature of the SOS effect is the fact that, owing to the impact ionization in the SCR, the charge extracted from the diode may be substantially larger than the charge 'pumped' by the forward current pulse.

Numerical simulations showed that the main effect on the dynamics of current breaking is exerted by the initial plasma distribution profile in the diode and the dopant distribution profile in the p-layer. It was experimentally confirmed that the shorter the forward current pulse (i.e., the greater the amount of plasma contained in the narrow p^+p-junction region) and the weaker the dopant concentration gradient in the p-layer (i.e., the deeper the pn junction), the faster the current interruption process. In particular, for an ultradeep pn junction (180–200 μm) and a short (dozens of nanoseconds) reverse current pulse, the current breaking takes place in a time shorter than a nanosecond [24].

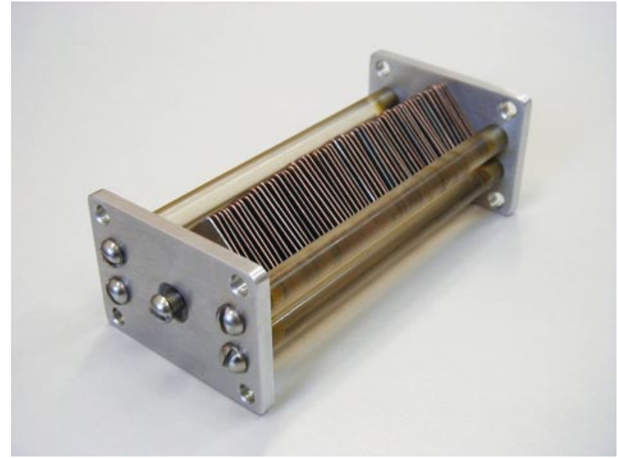

Figure 8. Typical design of an SOS-diode assembly with a pulsed voltage of 180 kV and an interrupted current of $\sim$ 4 kA.

Calculations and experiments indicated that the highest voltage across the p^+pnn^+ structure in the course of current interruption is normally equal to 750–850 V and a large number of the diodes are to be connected in series for generating high-voltage pulses. Specially designed investigations revealed that the possible technological scatter in the p-region width does not entail a significant difference in the voltage across the diodes: although the SCR formation in structures with a higher pn-junction depth begins later, this region broadens faster than for a smaller junction depth, and the difference in voltages at the concluding stage of the process is small ($\sim$ 4%). That is why SOS diodes may be connected in series without employing external voltage dividers. A typical SOS-diode assembly design is illustrated in Fig. 8, and the main characteristics of the assemblies produced are collected in Table 1 [25].

Table 1. Parameters of SOS diodes elaborated at the Institute of Electrophysics, UB RAS.

Parameter	Value
Operating voltage	60–1000 kV
Number of series-connected structures	80–320
Structure area	0.25–4 cm^2
Forward current density	0.4–2 kA cm^{-2}
Broken current density	2–10 kA cm^{-2}
Forward current duration	40–600 ns
Reverse current duration	15–150 ns
Current breaking time	0.5–10 ns
Power dissipated in transformer oil (steady-state operation)	50–500 W
Length/mass	50–220 mm/0.05–0.6 kg

There also exist devices which harness the effect of subnanosecond current breaking, which were developed for the formation of pulses several nanoseconds long. For a short pumping time, they break a current as high as 2 kA in 500–800 ps.

Figure 9. S-5N generator. The output voltage is 400 – 1000 kV, the current through the load is 1 – 3 kA, and the pulse duration (FWHM) is 30 – 50 ns. Pulse repetition rate: 300 Hz (constantly), 900 Hz (a 30-s long pulse packet). Average input power: 18 kW (constantly), 45 kW (a 30-s long pulse packet). Overall dimensions of the body: $3.5 \times 1.4 \times 1.1$ m; mass: ~ 2800 kg.

Investigation and operation of the elaborated SOS diodes as parts of different pulsed generators revealed their extremely high reliability and the capacity to withstand many-fold overloads in current and voltage; in this case, the pulse repetition rate may be brought up to 10^4 Hz. Purposeful bench tests were carried out to intentionally put the devices out of action. It turned out that increasing the current density and the current injection rate by an order of magnitude (from 5 to 50 kA cm^{-2}) results in only an increase in energy losses at the pumping stage and a lowering of opening switch operation efficiency without putting the devices out of action. In this case, the structures operate as an active resistance, which limits the pumping current, because the base modulation at these current densities is accompanied by the emergence of high forward voltages. Attempts to put the SOS diode out of action with the aid of a high operating voltage (a device with an operating voltage of 120 kV was placed in a generator with an output voltage of 450 kV) showed that the SOS diode operated as a voltage limiter (the pulse amplitude did not exceed 150 kV) in the current interruption and in doing this it consumed energy from the pump capacitor. Numerical simulations of this operating mode revealed a sharp intensity increase of carrier avalanche multiplication in the region with the electric field and a corresponding lowering of structure resistance at the stage of current breaking.

A large number of high-power pulse generators have been developed on the basis of SOS diodes for different areas of modern engineering. Shown by way of example in Fig. 9 is a generator with a pulsed power of ~ 3 GW and a voltage up to 1 MV, developed in the IEP, UB RAS. Prior to the advent of SOS diodes, this level of parameters in semiconductor nanosecond pulsed technology appeared to be absolutely impossible. The possibilities of employing SOS diodes in high-power nanosecond pulse generators are significantly increased when use is made of magnetic switches and magnetic compressors. That is why such generators have received wide acceptance [25].

5. Silicon carbide opening switch

All the high-power opening switches described above were made employing the base material of all semiconductor electronics — monocrystalline silicon. However, the last 5 – 6 years saw the opening-up of a realistic possibility of developing high-power semiconductor electronics and, in particular, high-power pulsed technology on the basis of high-quality monocrystalline silicon carbide (SiC).

The energy gap of, for instance, the 4H-SiC polytype is equal to 3.24 eV, i.e., it is significantly broader than for Si (1.12 eV), and therefore the limiting working temperature, which is limited by the rate of thermal carrier generation, is approximately three times higher for SiC devices than for silicon devices ($\sim 600\,^\circ$C instead of $200\,^\circ$C). The critical avalanche breakdown field in SiC is an order of magnitude higher than in Si, and the electron saturated velocity is approximately two-fold and the thermal conductivity is approximately three-fold higher than in silicon. This complex of advantages in principle enables a significant improvement in operation speed, power, and reliability of all the devices of high-power electronics. Extensive studies in this area are now underway, and to date it has been experimentally established that the SiC analogues of all silicon devices of high-power electronics without exception are possible to develop.

The feasibility of developing an SiC analogue of a silicon DSRD (more precisely, IRD, see Section 3) was first demonstrated in Refs [26, 27]. 4H-SiC $\mathrm{p^+n_0n^+}$- and $\mathrm{p^+p_0n^+}$-structures were grown by chemical vapor deposition on $\mathrm{n^+}$-substrates at a temperature of $1500\,^\circ$C. The 40-μm thick n-type base region possessed a dopant (nitrogen) concentration $n_0 = (3-5) \times 10^{14}$ $\mathrm{cm^{-3}}$, and the 12-μm p-type region had an aluminium concentration of 8×10^{14} $\mathrm{cm^{-3}}$; the diodes were 0.6 mm in diameter. The oscilloscope traces of recovery for a constant forward current of 0.4 A are shown in Fig. 10.

Observed in the $\mathrm{p^+p_0n^+}$-structure is a sharp subnanosecond breaking of reverse current which rises to an amplitude of -0.8 A in 10 ns; in the $\mathrm{n^+n_0p^+}$-structure, the breaking process lasts for ~ 12 ns. Generally speaking, these results were predictable, because a fast interruption of reverse current in epitaxial structures with abrupt block $\mathrm{p^+n_0}$ and $\mathrm{n^+n_0}$ junctions and a very large mobility difference of electrons and holes ($\mu_\mathrm{n}/\mu_\mathrm{p} \approx 7$ in 4H-SiC) is possible only in the version in which the block junction draws an electron current. Numerical simulations of recovery [28] showed that the trailing plasma front in a $\mathrm{p^+n_0n^+}$-structure with abrupt junctions is formed (as in silicon IRDs) at the $\mathrm{p^+p_0}$ junction and arrives at the blocking $\mathrm{n^+p_0}$ junction earlier than the plasma concentration vanishes there. Upon arrival of the front, the plasma is no longer contained in the device, and the reverse current flows only due to the motion of majority carriers away from the $\mathrm{n^+p_0}$ junction, resulting in a rapid SCR formation and current breaking.

Simple estimates show that the characteristics of an SiC-based breaking switch may be quite good. For instance, a single $\mathrm{p^+p_0n^+}$-diode with a breakdown voltage of 10 kV (with a base thickness $W_\mathrm{p0} \approx 100$ μm, and a donor concentration $N_\mathrm{d} = 7 \times 10^{14}$ $\mathrm{cm^{-3}}$) will possess an ultimate current breaking time $t_\mathrm{off} \approx 1$ ns and a reverse current density $j_\mathrm{R} \approx 10^3$ A $\mathrm{cm^{-2}}$.

In principle, SOS breakers can also be fabricated on the basis of SiC. To do this, when growing the $\mathrm{p^+pn_0n^+}$- or $\mathrm{n^+np_0p^+}$-structures it is required to ensure the desired dopant

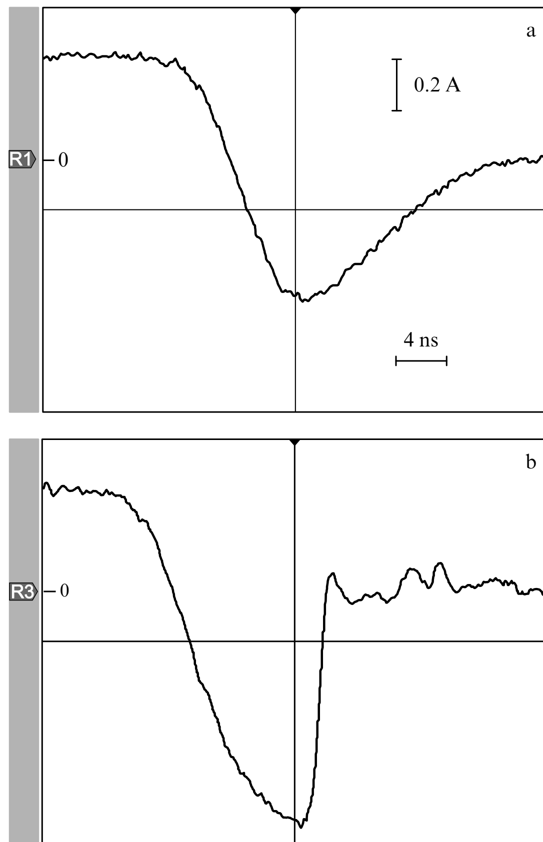

Figure 10. Oscilloscope traces of recovery in the pumping of silicon carbide p^+nn_0 (a) and $p^+p_0n^+$ (b) structures by a direct current of 0.4 A. The current zero is indicated with the mark ▶.

concentration gradient in the p- or n-layer by programmed variation of the dopant content in the gas flow in the growing of these layers. This process is rather complicated, and work on the development of the SiC analogues of a silicon SOS diode has not been undertaken up to the present time.

6. Conclusions

Two main methods are employed for generating high-power pulses. In the first one, the electric energy is accumulated in a capacitive storage device to be switched later to a load with the aid of a closing switch. The second technique consists in magnetic energy being stored in an inductive storage device, and then switched to the load after an abrupt current interruption. In the inductive storage, the energy density exceeds the energy density of capacitive storage by more than an order of magnitude, permitting a substantially more compact design of pulsed generators. However, the absence of a reliable and high-break switch did not allow making inductive energy-storage generators for a long time.

Our review gives a description of the operation of semiconductor diode opening switches intended for high-power nanosecond pulsed technology. Of special interest are two types of silicon diode breakers: a DSRD and an SOS diode. The former was proposed and elaborated at the A F Ioffe Physical-Technical Institute, RAS, and the latter at the Institute of Electrophysics, UB RAS. DSRDs make it possible to switch powers ranging into the hundreds of megawatts in nano- and subnanosecond times for a current density on the order of 10^2 A cm^{-2}. SOS diodes enable switching power as high as several gigawatts in an equally short time for a current density exceeding 10^3 A cm^{-2}.

Prior to the development of high-power semiconductor interrupters, in order to break current, advantage was taken of thin exploding wires, plasma interrupters, injection thyratrons, tacitrons, etc. However, all these devices exhibit long current breaking times, a short service life, and a low pulse repetition rate. The advent of DSRDs and SOS diodes marked a full-scale revolution in high-power nanosecond pulse technology, because it enabled development of the generators of high-power nano- and subnanosecond pulses with a voltage as high as 10^3-10^6 V, a pulse repetition rate up to 10^4 Hz, and a virtually unlimited service life.

The elaboration of high-power nanosecond pulse generators with inductive storage and semiconductor opening switches lent a strong impetus to the works on relativistic microwave electronics, ultrabroadband radar, generators for electromagnetic countermeasures, laser power suppliers, electron accelerators, pulsed X-ray units, and so forth.

References

1. Mesyats G A *Generirovanie Moshchnykh Nanosekundnykh Impul'sov* (Generation of High-Power Nanosecond Pulses) (Moscow: Sov. Radio, 1974)
2. Boff A, Moll J, Shen R, in *3rd Intern. Solid-State Circuits Conf.: Digest of Technical Papers* (New York: Lewis Winner, 1960) p. 50
3. Eremin S A, Mokeev O K, Nosov Yu R *Poluprovodnikovye Pribory s Nakopleniem Zaryada i Ikh Primenenie* (Semiconductor Devices with Charge Storage and Their Applications) (Moscow: Sov. Radio, 1966)
4. Benda H, Spenke E *Proc. IEEE* **55** 1331 (1967)
5. Grekhov I V, Geifman E M, Kostina L S *Zh. Tekh. Fiz.* **53** 726 (1983) [*Sov. Tech. Phys.* **28** 459 (1983)]
6. Grekhov I V et al. "Sposob sozdaniya istochnika diffuzii alyuminiya v kremnii" ("Method for making a source of aluminium diffusion in silicon"), USSR Author's Certificate No. 176989 (Priority data July 6, 1964)
7. Landsberg P T, Kousik G S *J. Appl. Phys.* **56** 1696 (1984)
8. Grekhov I V et al. *Pis'ma Zh. Tekh. Fiz.* **9** 435 (1983) [*Sov. Tech. Phys. Lett.* **9** 188 (1983)]
9. Grekhov I V et al. *Solid-State Electron.* **28** 597 (1985)
10. Grekhov I V *Izv. Ross. Akad. Nauk. Ser. Energetika* (1) 53 (2000)
11. Huldt L *Phys. Status Solidi A* **8** 435 (1971)
12. Haug A *Solid State Commun.* **25** 477 (1978)
13. Howard N R, Johnson G V *Solid-State Electron.* **8** 275 (1965)
14. Efanov V N et al., in *11th IEEE Intern. Pulsed Power Conf., Baltimore, Maryland, USA, June 29–July 2, 1997: Digest of Technical Papers* Vol. 2 (Eds G Cooperstein, I Vitkovitsky) (New York: IEEE, 1997) p. 985
15. Pemen A J M et al. *Rev. Sci. Instrum.* **74** 4361 (2003)
16. Grekhov I, in *11th IEEE Intern. Pulsed Power Conf., Baltimore, Maryland, USA, June 29–July 2, 1997*, Book of Abstracts, p. 0.10.8
17. Kotov Yu A et al. *Dokl. Ross. Akad. Nauk* **330** 315 (1993) [*Phys. Dokl.* **38** 229 (1993)]
18. Darznek S A, Rukin S N, Tsyranov S N *Zh. Tekh. Fiz.* **70** (4) 56 (2000) [*Tech. Phys.* **45** 436 (2000)]
19. Rukin S N *Prib. Tekh. Eksp.* (4) 5 (1999) [*Instrum. Exp. Tech.* **42** 439 (1999)]
20. Darznek S A et al., in *BEAMS '96: Proc. of the 11th Intern. Conf. on High Power Particle Beams, Prague, Czech Republic, June 10–14, 1996* Vol. 2 (Eds K Jungwirth, J M Ullschmied) (Prague: Institute of Plasma Physics Acad. of Sci. of the Czech Republic, 1996) p. 1241
21. Darznek S A, Mesyats G A, Rukin S N *Zh. Tekh. Fiz.* **67** (10) 64 (1997) [*Tech. Phys.* **42** 1170 (1997)]
22. Rukin S N, Tsyranov S N *Pis'ma Zh. Tekh. Fiz.* **30** (1) 43 (2004) [*Tech. Phys. Lett.* **30** 19 (2004)]

23. Grekhov I, Mesyats G A *IEEE Trans. Plasma Sci.* **PS-28** 1540 (2000)
24. Lyubutin S K et al. *Dokl. Ross. Akad. Nauk* **360** 477 (1998) [*Phys. Dokl.* **43** 349 (1998)]
25. Mesyats G A *Impul'snaya Energetika i Elektronika* (Pulsed Power Engineering and Electronics) (Moscow: Nauka, 2004)
26. Grekhov I V et al. *Pis'ma Zh. Tekh. Fiz.* **28** (13) 24 (2002) [*Tech. Phys. Lett.* **28** 544 (2002)]
27. Grekhov I V et al. *Solid-State Electron.* **47** 1769 (2003)
28. Grekhov I V et al. *Fiz. Tekh. Poluprovodn.* **37** 1148 (2003) [*Semicond.* **37** 1123 (2003)]