D.C. AND SMALL-SIGNAL A.C. PROPERTIES
OF SILICON BARITT DIODES

T.G. VAN DE ROER
D.C. AND SMALL-SIGNAL A.C. PROPERTIES OF SILICON BARITT DIODES

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE TECHNISCHE WETENSCHAPPEN AAN DE TECHNISCHE HOGESCHOOL EINDHOVEN, OP GEZAG VAN DE RECTOR MAGNIFICUS, PROF. DR. P. VAN DER LEEDEN, VOOR EEN COMMISSIE AANGEWZEZEN DOOR HET COLLEGE VAN DEKANEN IN HET OPENBAAR TE VERDEIGEN OP DINSDAG 8 NOVEMBER 1977 TE 16.00 UUR

DOOR

THEODORUS GERARDUS VAN DE ROER

GEBOREN TE BRUNSSUM

DRUK: WISO RO HELMIND
DIT PROEFSCHRIFT IS GOEDGEKEURD
DOOR DE PROMOTOREN

Prof. Dr. M.P.H. Weenink
en
Prof. Dr. H. Groendijk
Aan Elly
ACKNOWLEDGEMENTS

The research reported in this thesis was carried out at Eindhoven University of Technology in the Group of Electromagnetic Field and Network Theory. I am grateful to the members of this Group for maintaining a friendly and research-minded atmosphere in which it was a pleasure to work.

An essential contribution was made by J.J.M. Kwaspen who was largely responsible for the preparation and execution of the measurements. He also made most of the drawings for this thesis.

I acknowledge the cooperation with the Group of Electronic Devices who took a continuing interest in the work. In particular the names of C.J.H. Heijnen, Head of the Semiconductor Technology Lab., and M.J. Foolen who made the devices should be mentioned.

The whole project also benefited greatly from the help offered by people from Philips Research Labs, notably M.T. Vlaardingerbroek who stood at its beginning, B.B. van Iperen and H. Tjassens who offered advice in the impedance measurements and L.J.M. Bollen, F.C. Eversteijn, F. Huizinga and H.G. Kock who contributed a great deal to the technological part.

Thanks are also due to F. Sellberg of the Microwave Institute Foundation in Stockholm for making available his calculations.

Last, but not least, I wish to thank Miss Tiny Verhoeven for her able typing work.
"Grau, teurer Freund, ist alle Theorie, 
und grün des Lebens goldner Baum."

Mephisto
## CONTENTS

### I. INTRODUCTION

1. INTRODUCTION

### II. EARLY THEORETICAL MODELS

- II-1. D.C. Characteristics
- II-2. The models of Haus, Statz and Pucel and of Weller
- II-3. The model of Vlaardingerbroek and Van De Roer
- II-4. Scaling laws
- Appendix

### III. EQUATIONS AND RELATIONSHIPS

- III-1. Transport equations
- III-2. Field equations
- III-3. Normalizations

### IV. D.C. THEORY

- IV-1. Introduction
- IV-2. Boundary conditions
- IV-3. The high-field region
- IV-4. The low-field region
- IV-5. Method of solution
- IV-6. Results

### V. A.C. IMPEDANCE

- V-1. Introduction
- V-2. The contact region
- V-3. The diffusion region
- V-4. The drift region
- V-5. Conclusion

### VI. NOISE

- VI-1. Introduction
- VI-2. Shot noise
- VI-3. Thermal noise: the impedance-field method
- VI-4. Conclusion
VII. TECHNOLOGY
   VII-1. Introduction 67
   VII-2. Formation and evaluation of platinum silicide contacts 68
   VII-3. Formation of p-n junctions 69
   VII-4. Further processing 70

VIII. DIAGNOSTIC MEASUREMENTS
   VIII-1. Introduction 73
   VIII-2. Capacitance-voltage measurements 73
   VIII-3. R.F. impedance below punch-through 78
   VIII-4. Current-voltage measurements 79

IX. R.F. IMPEDANCE MEASUREMENTS
   IX-1. The waveguide bridge method 85
   IX-2. Description of the hardware 87
   IX-3. Theory 89
   IX-4. Calibrations 96
   IX-5. Measuring at elevated temperatures 99
   IX-6. Measuring under pulsed bias 100

X. R.F. NOISE MEASUREMENTS
   X-1. Theory 102
   X-2. Experiment 106

XI. RESULTS AND CONCLUSIONS
   XI-1. Introduction 108
   XI-2. P-n-p diode series F 108
   XI-3. M-n-p diode series G 115
   XI-4. M-n-p diode series K 121
   XI-5. Conclusions 127

REFERENCES 129
SUMMARY 135
SAMENVATTING 139
LEVENSBERICHT 143
I. INTRODUCTION

Early in the history of electron devices it was recognized that transit-time effects can have an influence on the behaviour at high frequencies. Early papers by Benham [1] and Müller [2] deal with transit-time effects in vacuum-diodes. The theory was generalized later by Llewellyn and Peterson [3] to tubes with more electrodes.

In 1954, Shockley [4], realizing that transistors become transit-time limited at higher frequencies, explored the possibilities to make two-terminal semiconductor devices having an impedance with a negative real part in some frequency range. He discussed two methods: using the transit time in a constructive way or finding ways to induce a negative differential conductivity in semiconducting materials. Both possibilities have been realized in later years, the first in Impatt and Barritt diodes and the second in Gunn diodes. For a recent review of these devices see [5].

A third possibility, the tunnel effect, was discovered by Esaki [6].

The transit-time device described by Shockley was a p-n-p diode, i.e. a transistor with floating base. If the collector-emitter voltage is raised high enough to fully deplete the base of majority carriers, minority carriers are injected from the emitter and flow towards the collector. The point at which this starts to happen is called punch-through, hence the often used name punch-through diode.

As the injected current is a function of the applied voltage, a modulation of this voltage will also modulate the injected carrier stream. These modulations will travel to the collector in a certain time and due to the finiteness of this transit time the external current modulation will experience a delay with respect to the voltage modulation. For sinusoidal modulation this can be translated into a (frequency-dependent) phase shift, and at those frequencies where the phase shift is between 90 and 270 degrees, the real part of the impedance will be negative.
The structure proposed by Shockley has the disadvantage that the field in the base region is non-uniform, rising from a low value at the emitter to a higher value at the collector, which makes its analysis rather difficult. Besides, velocity modulation occurs which causes Joule losses, as it gives a current component in phase with the field. Now it is well known that the drift velocity in semiconductors saturates at high field strengths and it would be preferable to operate under this condition so that velocity modulation is not possible. Therefore Read, in 1958 [7], proposed to inject carriers from a reverse-biased junction. To accomplish this the field at the junction must be so high that avalanche multiplication of carriers occurs, otherwise no current flow is possible. Now the possibility exists to maintain the field throughout the diode at such a high value that the drift velocity is saturated everywhere.

It took quite an advance in semiconductor technology before, in 1965, the first diodes operating on this principle could be produced. They became known as Impatt diodes (from Impact Avalanche Transit Time).

Meanwhile, experiments [8] showed that transit-time effects in p-n-p or n-p-n structures exist but no negative resistance was found. However, Yoshimura [9] showed theoretically that even with constant mobility (and thus a large in-phase current) a negative resistance is possible. Wright [10,11,12] proposed a n-p-i-n structure which has the advantage that the region of saturated velocity can be larger than in an n-p-n structure. He predicted useful negative resistance and power outputs. A similar structure was proposed by Rüegg [13]. That the operation of these devices was not very well understood at that time is demonstrated by the fact that Rüegg believed his device would show no small-signal negative resistance and therefore would not be self-starting as an oscillator.

In spite of all this activity on the theoretical side it lasted until 1971 before the first experimental realisation of an oscillating punch-through diode was announced [14]. Unlike the proposed devices this was a metal-semiconductor-metal structure, made by polishing a silicon slice down to 12 µm thickness and metalizing it on both sides. Around the same time oscillating p-n-p devices were realized [15], but publication in the
open literature was delayed [16]. Soon after the first publication oscillating p-n-p and M-n-p devices were announced by several laboratories [17,18,19] and the name Baritt diode (from Barrier Injection Transit Time) was coined. An extensive review of their characteristic properties was given by Snapp and Weissglas [20].

Since then steady improvements in power output, efficiency and frequency have been made [21,22,23], but compared to Impatt diodes the Baritt still is a low-power device. Its main advantages seem to be low noise and ease of fabrication. Also it performs well as a self-mixing oscillator [24,25].

A further advantage could be that its negative resistance range is restricted to a frequency band of about one octave. This might seem a disadvantage at first sight but for many applications a broad-band negative resistance is not necessary and even inconvenient, giving rise to oscillations at undesired frequencies.

Whether Baritt diodes will find applications in microwave technology remains an open question. They face a hard competition from Impatt and Gunn diodes and the newly emerging GaAs microwave field effect transistors.

Whereas theories abound, experimental data are relatively scarce. Therefore, in 1972 a program was started in cooperation between the group of Electron Devices and of Electromagnetic Theory at Eindhoven University of Technology comprising the manufacturing of Baritt diodes along with theoretical analysis and measurements of impedance and noise. The author's contributions to the first part of this program, concerning the small-signal properties, are subject of this thesis.

The scope of the present work is to present an analysis of the d.c. and small-signal a.c. properties of Baritt diodes and make a comparison between p-n-p and M-n-p devices. The theoretical part has been kept analytical mostly which made it necessary to introduce a number of approximations. Understanding was its goal rather than obtaining correct numerical values. Nevertheless it has been tried to match theory and experiment as closely as possible, to which end much attention has been
paid to obtaining accurate information about the diode parameters.

The material is ordered as follows: in the next chapter a review will be
given of some of the earlier theoretical models which are eminently
suited to give insight into the characteristic properties of Baritt
diodes. This will make it easier to follow through the next four
chapters where a more elaborate theoretical model will be developed.
These will be followed by chapters discussing the manufacturing technol­
ogy and the measurements. The last chapter will give results of the
measurements, comparison with theory and conclusions.
II. EARLY THEORETICAL MODELS

II-1. D.C. Characteristics

In this chapter some models will be discussed that were proposed shortly after the first experiments to explain the characteristics of Baritt diodes. Although containing a number of rather drastic simplifications they have been found to be well suited to explain qualitatively a number of observed phenomena.

Before tackling the a.c. behaviour, let us start with a review of the d.c. properties. In Fig. II-1 a sketch is given of the physical structure of a Baritt diode. Clearly, it bears a great resemblance to a parallel-plate condensor and we may expect the field and current to be uniformly distributed in the lateral plane. This is important because it allows us to restrict the analysis to one dimension in space which of course is a considerable simplification. Even so the problem is complicated enough. In Fig. II-2 then the charge and field distributions and the energy bands are sketched as a function of the depth coordinate for a p-n-p diode below punch-through. In this situation we can consider the device as consisting of two diodes back-to-back separated by a thin ohmic layer. When the bias voltage is raised the depletion layer of the back-biased diode widens and absorbs the voltage whereas the forward-biased diode is hardly affected. This evidently gives possibilities to probe the impurity concentration by C-V measurements. Also, at high frequencies we may
picture the device as a series circuit of two capacitors and a resistor. This too gives possibilities for diagnostic measurements which will be discussed further in chapter VIII.

In the situation sketched in Fig. II-2 the current is determined by the back-biased diode. In good quality material it is very low and is carried mainly by minority carriers. When the voltage is raised further, eventually the two depletion layers meet, a situation called reach-through or punch-through. Now the current is still low (we do not suppose the peak field is high enough to produce impact multiplication of
carriers) but when we direct our attention to the left-hand junction we see that here a fairly low barrier for holes exist. Holes that have enough energy to overcome this barrier are picked up by the field and swept to the other side. When the voltage is increased further, the barrier is lowered and the hole current increases rapidly, according to the formula [26]:

\[ J = A^*T^2 \exp \left( \frac{V_m}{V_T} \right) \]  

where \( A^* \) is the modified Richardson constant [27] and \( V_T \) is substituted for \( \frac{kT}{q} \). The quantity \( A^*T^2 \) is called the saturation current and is the theoretical limit of the current a p-n junction can supply. Its value, however, is so large (about \( 10^{11} \text{ Am}^{-2} \) at room temperature) that in practice it is never attained.

\[ \text{Fig. II-3. P-n-p diode above punch-through.} \]

\[ \text{a. space charge density:} \]

1. holes, 2. ionized donors, 3. total.

b. electric field.

c. energy band diagram.
The hole current very soon surpasses the electron current and the latter can be neglected for all practical purposes. The hole density now has a spatial dependence as sketched in Fig. II-3a and the corresponding field profile is given in Fig. II-3b. Evidently in the first part of the diode the holes must diffuse against the field and a steep concentration gradient is necessary. Further on the drift velocity increases by the field into saturation and the hole density flattens out.

![Diagram](image)

*Fig. II-4. M-n-p diode above punch-through.*

a. space charge density:
   1. holes, 2. ionized donors, 3. total.

b. electric field.

c. energy band diagram.

When the forward-biased contact is a Schottky-barrier diode, i.e. a rectifying metal-semiconductor junction, the situation is somewhat different. Now an additional barrier exists at the junction [27] which lowers the saturation current. Instead of (II-1) we now have (see Fig. II-4)
Values of $\Phi_h$ of less than 0.2 V have never been observed and since $V_T = 0.025 V$ at room temperature, the saturation current is reduced to values low enough to be realized experimentally. The voltage at which the diode current equals the saturation current is called the flat-band voltage as the energy bands at the junction run horizontally.

One would expect that the current cannot be raised further but this is not true. A new effect comes into play, the Schottky effect. Holes in the vicinity of the junction induce charges in the metal which exert an
attracting force. This can be represented as a potential, the so-called image-force potential, which is sketched in Fig. II-5b. This potential must be added to the electric potential and lowers the barrier. This barrier lowering is determined by the gradient of the electric potential, that is, by the electric field $E_c$ near the junction, which relation can be expressed as [27]:

$$\Delta \phi_h = -\sqrt{\frac{qE_c}{\varepsilon \varepsilon_0}}$$  \hspace{1cm} (II-3)

In practice one always finds a barrier lowering exceeding that given by this expression but still proportional to $E_c^{1/2}$. Not much is known about the physical origins of this effect, but it is suspected that there is a relation with the condition of the metal-semiconductor interface, as a correlation has been found with manufacturing parameters [28].

On the basis of the foregoing considerations we may expect the current-voltage characteristics of p-n-p and M-n-p diodes having the same n-layer width and doping to look like Fig. II-6.

![Current-voltage characteristics of Baritt diodes](image)

Fig. II-6. Current-voltage characteristics of Baritt diodes. 

a. p-n-p, b. M-n-p.

Now that we have an impression of the d.c. behaviour of Baritt diodes, we can turn our attention to their a.c. properties. Clearly, we must distinguish at least two regimes of operation, namely below and above flat-band. For each of these situations a model has been proposed in the literature, which we will now proceed to discuss.
II-2. The models of Haus, Statz and Pucel and of Weller

Shortly after the first announcements of punch-through oscillators Haus, Statz and Pucel [29] published a theory which enabled them to calculate the small-signal impedance and the shot noise. This model divides the diode into two regions (Fig.II-7a): a narrow injection region including the injecting contact and the potential barrier, and a drift region comprising the rest of the n-layer. The behaviour of the injection region is described by eqn. (II-1) and in the drift region the drift velocity is assumed to be saturated everywhere. In view of the foregoing section this is a rather crude approximation. Nevertheless this model has been found to give a good qualitative explanation of a number of phenomena.

![Fig. II-7. Models of Haus, Statz and Pucel (a) and of Weller (b). 1. injection region, 2. drift region.](image)

To calculate the small-signal impedance we split all variables into a d.c. part, with index 0, and a (small) a.c. part, index 1. The a.c. parts have a time dependence $\exp(j\omega t)$. The fact that the a.c. components are small enables us to linearize the equations. The a.c. component of the injected carrier current, $J_{i1}$, is found as the first term of a Taylor-series expansion of Eq. (II-1) or (II-2) around the d.c. operating point:

$$J_{i1} = -\frac{J_0}{V_T} V_{ml}$$ (II-4)

where $J_0$ is given either by (II-1) or by (II-2). To come from (II-4) to a
relation between the a.c. current and the a.c. field at the barrier it is assumed that \( E_1 \) is independent of position between the junction and the barrier (this supposes that the a.c. current in this region is predominantly dielectric displacement current). Then, with \( E_{11} \) the a.c. field at the barrier, one gets:

\[
J_{11} = \frac{J_0 x_m}{V_T} E_{11}
\]  

(II-5)

When one neglects the hole space charge, \( x_m \) can be calculated readily [26]:

\[
x_m = \left( \frac{2eV_m}{qN_D} \right)^\frac{1}{2}
\]

A model for operation above flat-band was given by Weller [30]. It starts from (II-3) and obtains by Taylor-expansion:

\[
J_{11} = \frac{J_0}{2V_T} \left( \frac{q}{4\pi\varepsilon_0 E_0} \right)^\frac{1}{3} E_{11}
\]  

(II-6)

where \( E_{11} \) in this case is the a.c. component of \( E_C \). The drift region now comprises the whole n-layer.

Eqs. (II-5) and (II-6) enable us to find an a.c. boundary condition for the drift region from the d.c. parameters. The analysis of the drift region is the same in both models. In this one-dimensional analysis the total a.c. current \( J_1 \) is not a function of position and equals the external current divided by the diode area. Then, using Poisson's equation, the electric field in the drift region is, following Wright [10]:

\[
E_1(x) = \left( E_{11} - \frac{J_1}{j\omega e} \right) \exp \left( -j\frac{\omega(x-x_1)}{V_s} \right) + \frac{J_1}{j\omega e}
\]

(II-7)

where \( V_s \) is the value of the saturated drift velocity and \( x_1 \) is equal to \( x_m \) below flat-band and zero above. Using the boundary condition \( E_{11} \) can be eliminated and we obtain:
\[ E_1(x) = \frac{J_1}{j\omega C} \left\{ 1 - \frac{1}{1+j\eta_c} \exp(-j\theta) \right\} \]  

where

\[ \theta = \frac{\omega(x-x_i)}{v_s} \]

and the injection parameter \( \eta_c \) is defined as:

\[ \eta_c = \frac{\omega E_{1i}}{J_{1i}} \]

so that its value becomes

\[ \eta_c = \frac{\omega E_i}{J_0 x_m} \]

below flat-band

\[ \eta_c = \frac{4\omega E_i}{J_0} \left( \frac{n_0 E_{1i}}{q} \right)^{\frac{1}{2}} \]

above flat-band

One notes an anomaly in the case of M-n-p diodes. As the current is increased and the flat-band condition is approached, \( \eta_c \) approaches infinity because \( x_m \) goes to zero. Above flat-band, however, \( \eta_c \) starts from zero because of \( E_{1i} = 0 \). This discontinuity can be removed by taking into consideration that the image-force potential is present also below flat-band. It was neglected there because its effect is noticeable only when \( x_m \) becomes very small.

Finding the impedance of the drift region now is easy. The result is

\[ Z_d = \frac{1}{\omega C_d} \left\{ -j + \frac{1}{1+j\eta_c} \cdot \frac{1-\exp(-j\theta_d)}{\theta_d} \right\} \]  

where

\[ C_d = \varepsilon A/(\ell_d-x_i) \]

is the so-called "cold" capacitance of the drift region and

\[ \theta_d = \omega(\ell_d-x_i)/v_s \]

its transit angle.
The first term between brackets in (11-10) is evidently due to the dielectric character of the semiconductor material. The second gives the effect of the modulated charge carrier stream. It contributes not only a resistive part but also a reactive part. This last effect is often described as "electronic capacitance".

Before discussing the impedance further it will be interesting to pause for a moment and have a look at the ratio $\omega E_1 / J_{1c}$ where $J_{1c} = J_1 - j\omega E_1$ is the a.c. charge carrier current. At the beginning of the drift region this ratio is by definition equal to $\eta_c$. Further on we will denote it by $\eta(x)$. From (II-8) it follows that

$$\eta(x) = +j + (\eta_c - j)\exp(j\theta) \quad (\text{II-11})$$

In the complex plane this describes a circle with centre at $+j$ and radius $|\eta_c - j|$, see Fig. II-8.

One sees immediately from this figure that the first part of the diode is dissipative, as here $J_{1c}$ has a component in phase with $E_1$. Only after $\theta = \pi/2$ $J_{1c}$ gets a component in antiphase with $E_1$ so that power is

---

Fig. II-8. Ratio of a.c. electric field and a.c. convection current in the complex plane.
produced. After $\theta \approx 3\pi/2$ dissipation occurs again, so it is desirable to choose $l_d$ such that $\theta_d \approx 3\pi/2$. Furthermore one concludes that it would be preferable to have $n_c$ on the imaginary axis above $+j$. In other words, there should be an inductive relationship between field and carrier current at the injection plane. Then the whole drift region is active and the optimum transit-angle is $\pi$. It is interesting to mention here that Impatt diodes fulfill this condition nearly perfectly.

Now let us take up the discussion of the impedance again. The real part of $Z_d$ is easily obtained from (II-10) as:

$$R_d = \frac{1}{\omega C_d} \cdot \frac{1 - \cos \theta_d + n_c \sin \theta_d}{\theta_d (1 + n_c^2)}$$

(II-12)

Two conclusions can be draw from this expression. First $\sin \theta_d$ must be negative to obtain a negative resistance. The optimum transit-angle is somewhat larger than $3\pi/2$ which corroborates the conclusion from the foregoing discussion. Second, the optimal $n_c$ lies at an intermediate value, between 2 and 3.

---

![Quality factor of drift region as a function of transit angle.](image-url)
In practice often the negative quality factor of a diode is used as a measure for its performance. This is because the possibilities of matching a microwave circuit to the diode are more determined by this $Q$ which is the ratio of $|X_d|$ and $R_d$ than by the absolute value of $R_d$. If we assume that the contribution of the electronic capacitance is small, then $Q$ is simply $(\omega C_d R_d)^{-1}$.

![Graph](image)

**Fig. II-9b. Quality factor of drift region as a function of injection parameter.**

In Fig. II-9a $\omega C_d R_d$ is plotted as a function of $\Theta_d$. For given $l_d$ and $v_s$ this also represents $R_d$ as a function of frequency. In this graph $\eta_c = 2.5$. In Fig. II-9b $\omega C_d R_d$ is plotted against $\eta_c$ for $\Theta_d = 3\pi/2$. From this graph the dependence on $J_0$ may be deduced.

These figures speak for themselves and we won't discuss them further. We merely note that the minimum negative $Q$ that can be obtained is about twenty.

**II-3. The model of Vlaardingerbroek and van de Roer**

The two models discussed before assume the drift velocity to be saturated from the potential barrier onwards. For diodes above flat-band this can be a reasonable approximation when $E_c$ is high, but below flat-band it never is. The field rises from zero in the latter case so that
the carriers must be transported by diffusion mainly. This demands the existence of a carrier density gradient which is not compatible with a saturated drift velocity.

In view of this, Vlaardingerbroek and the author [31] proposed another model which can be considered as an extension of the model of Haus et al. The new model takes account of the fact that the drift velocity first increases linearly with field and saturates only at high field strength. The velocity-field curve is approximated by two straight lines: constant mobility \( \mu \) up to a certain field value \( E_s \) and saturated velocity \( v_s = \mu E_s \) above. Consequently, the drift region now consists of two parts, one where the mobility is constant and one where the drift velocity is saturated. The first of these will be called source region in the following and the second will retain the name drift region. The model in this way combines older theories of Yoshimura [9] and Wright [10]. As Yoshimura showed, the source region can have a small negative resistance itself, but more important, as the new model shows, is that it provides a boundary condition to the drift region favourable for negative resistance.

\[ \text{Inset shows assumed } v-E \text{ characteristic.} \]
This model will now be discussed in some detail, not only because it provides deeper insight into the operation of Barritt diodes but also because the model this thesis is based on is an extension of it. As we will use its derivations rather extensively, paper [31] is attached as an appendix to this chapter. The model is illustrated by Fig. II-10.

In [31] it has been assumed that the boundary condition (II-5) can be applied at a small distance behind the potential barrier. This was necessary because, neglecting diffusion, one obtains zero drift velocity and infinite hole density at the barrier position which, when used to calculate the a.c. impedance, gives unrealistic results, especially at low currents. The applied procedure is thus a crude way of taking account of the fact that the drift velocity is not zero in the potential maximum.

The analysis thus starts at the plane \( x_i > x_m \) where the boundary condition (II-5) is applied. Then from Eqs. (3) and (8) of [31] we can calculate \( n_s \), the value of \( n \) at the plane \( x_s \) where the drift velocity saturates. A slight change of notation has been made to simplify the representation. The symbol \( \alpha \) is substituted for \( \omega /\omega_c \) and \( \Theta_s \) is used for the transit-angle of the source region denoted by \( \omega_c \tau \) in [31]. The result is:

\[
\eta_s = j \left[ 1 + \frac{1}{\alpha} \left( 1 - \frac{J_o + \sigma E_s}{J_o + \sigma E_i} \exp(-j\Theta_s) \right) \right] \\
- \frac{1}{1 + j\eta_c} \cdot \frac{E_i}{E_s} \cdot \frac{J_o + \sigma E_s}{J_o + \sigma E_i} \exp(-j\Theta_s) \right]^{-1}
\]

(II-13)

where \( \eta_c \) follows from (II-9a). The other symbols have the same meaning as in [31].

Clearly, \( \eta_s \) consists of two parts: one due to \( \eta_c \) and one entirely due to the source region. What has been said before about the impossibility of applying Haus' boundary condition at the potential maximum is confirmed here: when \( E_i \) is made zero the contribution of \( \eta_c \) vanishes.
To bring out the significance of (II-13) more clearly we write it in a different form, substituting

\[ \beta = \frac{J_0 + \sigma E_s}{J_0 + \sigma E_s} \]

After some rearrangement we get:

\[ \eta_s = j \frac{E_s}{(n_c - j)E_1 \exp(j\Theta_s)} \]

Apart from the denominator, which is close to one for small currents, this shows a striking resemblance with (II-11) and it turns out that \( \eta \) is moved from the real axis towards the imaginary axis by the transit through the source region. As has already been shown in the preceding section this is beneficial to the negative resistance of the drift region.

When \( \eta_s \) from (II-13) is substituted instead of \( \eta_c \) in (II-10) the impedance of the drift region is obtained. By substituting \( \phi = \arctan \alpha \) and \( \psi = \arctan \eta_c \), the expression for the real part of \( Z_d \) becomes relatively simple. It reads:

\[ R_d = \frac{1}{\omega C_d} \left[ \frac{J_0}{\sigma E_s} (1 + \alpha^2)^{-1} \left\{ \sin(\phi - \Theta_d/2) - \beta \sin(\phi - \Theta_d/2 + \Theta_s) \right\} + \frac{\beta E_1}{E_s} (1 + \eta_c^2)^{-1} \sin(\psi + \Theta_d/2 + \Theta_s) \frac{\sin \Theta_d/2}{\Theta_d/2} \right] \]

The second term in the square brackets is due to the influence of the injecting contact. It has a maximum negative value when \( \Theta_d = \pi \) and \( \psi + \Theta_s = \pi \). These conditions are not difficult to fulfill. Note that the optimum transit angle of the drift region has been reduced to \( \pi \) radians. This is the result of the extra delay produced by the source region.

The first term of (II-15) is due to the source region alone. Since \( \beta \) is greater than one, it contributes a positive resistance unless
\[ |\phi - \theta_s| \approx \pi \] which is a rather improbable situation. Fortunately it stands in proportion to the second term as \( J_0/\alpha E_i \) which can be made a small number.

One notes that when \( J_0 \) goes to zero both components of \( R_d \) become zero, the second one because \( \eta_c \) becomes infinite. This is in accordance with experimental findings.

We thus conclude that the delay introduced by the source region can increase the negative resistance of the drift region. This is beneficial to the total diode resistance, at least when the source region itself does not contribute a large positive resistance. This however is not likely; the impedance of the source region cannot be large, first because its width is small and second because it has a high hole density giving a large conductivity.

II-4. Scaling laws

We conclude this chapter with a few remarks on the influence of various parameters. From the foregoing analysis it appears that the parameters always occur in certain combinations e.g. \( J_0/\alpha E_s, E_i/E_s, \omega \ell_d/v_s \) and \( \omega e/\sigma \). This is true for the drift region and source region, but not completely for the injecting contact. Nevertheless one can state roughly that when \( J_0/N_D, \omega/N_d, \omega \ell_d \) are kept constant, the negative \( Q \) remains the same. So, supposing optimum parameter values are found at a certain frequency, to go to another frequency one has to scale \( J_0 \) and \( N_D \) proportional with frequency and \( \lambda_d \) inversely proportional.
On the theory of punch-through diodes

M.T. Vlaardingerbroek
Philips Research Laboratories, Eindhoven, The Netherlands

Th.G. van de Roer
Eindhoven Technical University, Department of Electrical Engineering, Eindhoven, The Netherlands

(Received 11 September 1972)

An analytical small-signal theory of punch-through diodes is presented in which both the dc and ac hole drift velocity depend on the local electric field. The negative resistance is caused by the velocity and space-charge modulation in the bulk of the n layer, which arise from the interaction of the holes with the electric field. The field dependence of the injection tends to decrease this negative resistance at low current densities.

Recently, much attention has been paid to the theoretical description of punch-through or BARRITT microwave oscillator diodes. Most analytical theories rely on (a) the field-dependent injection of holes by the injecting barrier and (b) the transit-time delay of holes, which makes the phase difference between the ac part of the current induced in the external circuit and the ac diode voltage larger than \( \pi/2 \). Generally it has been assumed that the holes travel at saturated drift velocity throughout the diode. This latter assumption, however, precludes the possibility of velocity modulation due to ac fields and—as is well known from the theory of negative differential resistance in thermionic and semiconductor space-charge-limited diodes—the combined effect of space-charge and velocity modulation can result in an effective negative resistance.

In this letter a model is proposed in which the hole drift velocity, \( v \), is taken to be proportional to the field strength \( E \), for \( E \leq E_a \), the proportionality constant being the mobility \( \mu \). For \( E > E_a \), the hole velocity is assumed to saturate at \( v = v_s \). It will be shown that negative resistance occurs even if the injection conductivity \( \sigma_i \) (the ratio of the ac hole injection current density and the ac field strength near the injecting barrier) is taken to be zero. This is in agreement with transit-time theories of thermionic space-charge-limited diodes. At low current densities the injection mechanism is found to reduce the negative resistance.
We consider the planar structure in the inset of Fig. 1. The $n$ layer, having uniform donor density $N_p$, is fully depleted. The region between the source contact and the potential minimum is swamped with holes so its impedance is negligible. The region between the potential minimum and the plane $x = x_n$, where $E = E_n$, we call the source region; the remainder of the diode is the drift region. Following Ref. 6, we find for the total current density $J(t)$ in the diode

$$J(t) = \varepsilon \frac{\partial E(x, t)}{\partial t} + e p(x, t) v(x, t)$$

$$= \varepsilon \frac{dE(x, t)}{dt} - e N_p v(x, t), \quad (1)$$

where $\varepsilon$ is the dielectric constant and $p$ is the hole density. Use has been made of Poisson's law and $dx/dt = v(x, t)$. It should be noted that the total space charge is the sum of the positive charges of the holes and donors. The total differential in Eq. (1) means that we consider the fields as experienced by a moving hole as a function of time. We assume the dependent variables to consist of a dc and a small ac part. For the dc parts Eq. (1) is

$$J_0 = v_0(x) \left( \varepsilon \frac{dE_0}{dx} - e N_p \right). \quad (2)$$

We introduce a new independent variable, the transit-time $\tau$, defined by $\tau = \int_0^x v_0^{-1}(x') dx'$; furthermore, $\sigma = N_p e \mu$ and $\omega_c = \sigma / \varepsilon$. We solve Eq. (2) for the source region by taking $v_0 = \mu E_0$ and using the boundary condition $E_0 = 0$ when $x = 0$:

$$E_0 = \left[ J_0 / \omega_c \right] \left[ \exp(\omega_c \tau) - 1 \right]. \quad (3)$$

Furthermore, from $x = \int_0^\tau \mu E_0 d\tau'$ we find

$$x \omega_c^2 / \varepsilon \mu J_0 = x'(\tau) = \exp(\omega_c \tau) - \omega_c \tau - 1, \quad (4)$$

which yields the variation of $E_0$ with $x$. We find the end of the source region by substituting $E_0 = E_s$ into Eq. (3) so as to find $\tau = \tau_s$, which can be substituted into Eq. (4) to obtain $x_s$. In practical BARRITT diodes it appears that a hole spends more than half of its transit time in the source region, so that the usual assumption of constant drift velocity is not justified.

With regard to the ac impedance of the source region, the ac part of Eq. (2) is, using $\partial / \partial t = j \omega$, and denoting the ac quantities by the index 1,
This equation is solved by considering \( E_0 E_1 \) as the dependent variable and using Eq. (3). Assuming that the ac field strength is uniform in the region between the source contact and the plane \( x = 0 \), the boundary condition to be used is

\[
\mathcal{J}_1 = \sigma_i E_1; \quad \mathcal{J}_1 = J_0 \left( \frac{2 \epsilon/kTN_0}{\ln(\frac{J_0}{J_0^0})} \right)^{1/2},
\]

where \( \mathcal{J}_1 \) is the conduction current density, \( T \) is the absolute temperature, and \( J_0 \) is the current density at flat-band voltage. In our model, however, \( E_0 = E_1 = 0 \) at \( x = 0 \), so we must apply the boundary condition (6) in a plane \( x = x_1 \) (or \( r = r_1 \)) just beyond the potential minimum at \( x = 0 \) where the diffusion can be neglected. In terms of the total ac current density \( J_1 \), the boundary condition reads

\[
E_1(x_1) = J_1 / (\sigma_i + j \omega \epsilon) .
\]

The solution of Eq. (5) now becomes

\[
\mathcal{E}_1(\tau) = \frac{\omega \epsilon}{j \omega} \left[ 1 + \frac{J_0}{\sigma E_0(\tau)} \left[ \frac{\omega \epsilon}{\omega \epsilon - j \omega} - \left( \exp(\omega \epsilon \tau_i) + \frac{j \omega}{\omega \epsilon - j \omega} \right) \right. \right.
\]
\[
\left. \times \exp[(\omega \epsilon - j \omega)(\tau - \tau_i)] \right] \]
\[
+ \frac{j \omega}{\omega \epsilon} \frac{\sigma}{\sigma_i + j \omega \epsilon} \frac{E_0(\tau_i)}{E_0(\tau)} \exp[(\omega \epsilon - j \omega)(\tau - \tau_i)] \right\} J_1 .
\]

At high current densities, \( \sigma_i \gg \sigma \) so the last term can be neglected and the ac field strength is determined only by space-charge and velocity modulation in the bulk of the source region. At low current densities the injection mechanism, as characterized by the last term in Eq. (8), must be taken into account. From numerical evaluation we found that the result is not critically dependent on the choice of \( \tau_i \) (for \( x_1 \) we normally took values of the order of 0.1 \( \mu \)). It should be noted that the influence of the injection on the field strength \( E \) rapidly decreases for increasing \( \tau \) because of the factor \( E_0(\tau_i)/E_0(\tau) \). This is in contrast to other analytical models, in which the modulation due to the field-dependent injection is maintained throughout the interaction region.1-4

The voltage across the source region \( V_{1s} \) is found from

\[
V_{1s} = \int_{x_1}^{x_f} E_1(x) \, dx = \mu \int_{\tau_1}^{\tau_f} E_0(\tau)E_1(\tau) \, d\tau .
\]

The impedance of the source region \( Z_s \) is found by dividing the result of Eq. (9) by \( J_0 A \), where \( A \) is the diode area. The result is lengthy but straightforward. We
therefore restrict ourselves here to the high-current case \( \sigma_i = \infty \):

\[
Z_s = \frac{\mu J_0}{\sigma A(\omega - j\omega)} \left[ -x'(\tau_e) - \frac{\omega_e}{\omega - j\omega} \frac{\alpha E_e}{J_0} \right] \\
+ \frac{\omega_e^2 \exp(\omega_e \tau_e)}{j\omega(\omega - j\omega)} \left\{ 1 - \exp(-j\omega \tau_e) \right\},
\]

where \( E(x_e) = E_e \) and \( x'(\tau_e) \) is defined in Eq (4).

![Graph](image.png)

**FIG. 1.** Plot of \( Z = Z_s + Z_d \) in the complex plane; \( N_D = 10^{15} \text{ cm}^{-3} \); \( \mu = 450 \text{ cm}^2 \text{ V}^{-1} \text{ sec}^{-1} \); \( v_s = 0.7 \times 10^7 \text{ cm sec}^{-1} \); \( W = 8 \mu \text{m} \); \( A = 3 \times 10^{-4} \text{ cm}^2 \). The full curves are calculated using the appropriate value of \( \sigma_i \). For comparison the dotted curve shows the results obtained neglecting the injection (\( \sigma_i = \infty \)) at low current density. The numbers along the curves denote the frequency in GHz.

With regard to ac impedance of the drift region, the method of calculation is taken from the theory of
IMPATT diodes. The total current in a plane, defined by $\tau > \tau_p$, is

$$J_1 = j\omega e E_1(\tau) + J_{c1}(\tau) \exp(-j\theta),$$  \hspace{1cm} (11)$$

where $\theta = \omega(x - x_0)/v$, and $J_{c1}(x_0)$ is the conduction current density at $x = x_0$ (or $\tau = \tau_p$). The latter current is found by applying Eq. (11) to the plane $x = x_0$, where $\tau = \tau_p$ and $\theta = 0$. The value of $E_1(\tau_p)$ is obtained from Eq. (8). The calculation of the drift region impedance is now straightforward. Again, to avoid the writing of lengthy equations, we only give the result for $\sigma_i \to \infty$ (the limit of high current densities):

$$Z_d = \frac{1}{j\omega C_4} \left( 1 + \frac{J_0}{e E_1(\tau_p)} \frac{1 - \exp((\omega - j\omega)\tau_p)}{\omega - j\omega} \frac{1 - \exp(-j\theta)}{\theta_d} \right),$$  \hspace{1cm} (12)$$

where $C_4 = eA/(w - x_0)$ and $\theta_d = \omega(w - x_0)/\sqrt{v}$.

Equations (10) and (12) together yield the diode impedance $Z = Z_s + Z_d$ for high current densities ($\sigma_i \to \infty$). We have evaluated the corresponding expressions for the general case ($\sigma_i \neq 0$), which hold for all current densities, numerically. Some results are given in Fig. 1, where $Z$ is plotted for various values of the bias current. The results are in reasonable agreement with the experimental results shown in Ref. 7, taking into account the relative uncertainty in $\mu$, $w$, $v_s$, etc. We draw the following conclusions:

(i) Re$Z$ can be negative in more than one frequency region.

(ii) Increasing the current density shifts the negative resistance region towards higher frequencies. Above about 100 A/cm$^2$ the model predicts no useful negative resistance. Experiments showing negative resistance at higher current densities may be explained by the occurrence of avalanche breakdown (IMPATT diode).

(iii) In Fig. 1, one curve shows a plot of $Z$ for low current densities but assuming $\sigma_i \to \infty$ (which means neglecting the injection). The maximum value of the negative resistance is in this case much larger than when using the appropriate value of $\sigma_i$. Apparently the field-dependent injection acts as a damping at low current densities, since in the short range in which the injected ac current influences the field strength [last term of Eq. (8)] the field and the drift velocity are in phase. This conclusion is contrary to what is suggested by a theory in which the electron drift velocity is taken to be either
constant or independent of the ac field strength. At high current densities (> 50 A/cm²) the approximation
\( \sigma_i - \infty \) appears to be valid, which means that the negative resistance finds its origin in the combined effect of velocity and space-charge modulation of the hole current under influence of the ac electric field strength.

(iv) The results of our analytical model are in reasonable agreement with those of numerical calculations.⁸,⁹ For example, the numerical results in Ref. 9 could be reproduced to within 10% for high current densities. At low current densities (< 10 A/cm²) our results are in qualitative agreement with a maximum discrepancy of 1 mho/cm² in the conductance.

The advantage of an analytical theory is that the physical mechanism becomes more clear.

III. EQUATIONS AND RELATIONSHIPS

III-1. Transport equations

Electrons in a semiconductor experience an intensive quantum-mechanical interaction with the crystal lattice, which makes their behaviour quite different from that of free electrons. Ways have been found, however, to avoid the use of quantum-mechanics throughout, notably the concept of quasi-particles. Some quasi-particles encountered in solid-state physics are electrons in the conduction band, holes in the valence band, phonons and photons. A description of these can be found in many textbooks, e.g. [32].

Once having adopted the quasi-particle idea one can consider the collection of electrons and holes in a semiconductor as a gas to which statistical mechanics applies. The state of this gas then is described by distribution functions (one for each particle species). The distribution function $f_h$ of the holes for instance gives the average number of holes in a unit cell in phase space as a function of the space coordinate $r$, the velocity coordinate $\nu$ and time $t$. The macroscopic quantities of interest then can be written as integrals over velocity space, e.g.:

- The hole density $p = \int f_h d^3 \nu$
- The drift velocity $\nu = \frac{1}{p} \int \nu f_h d^3 \nu$
- The thermal energy $W = \frac{1}{p} \int \frac{1}{2} m^* (\nu - \nu)^2 f_h d^3 \nu$
- The heat-flow vector $Q = \frac{1}{p} \int \frac{1}{2} m^* (\nu - \nu)^2 (\nu - \nu) f_h d^3 \nu$

If the distribution function is Maxwellian, $W$ can be interpreted in terms of a carrier temperature: $W = \frac{3}{2} k T$.

To describe the change of the distribution function under the influence of external fields and collisions, Boltzmann's equation is used:
\[ \frac{\partial f_h}{\partial t} + \mathbf{v} \cdot \nabla f_h + \frac{\mathbf{F}}{m^*_h} \nabla_w f_h = \left( \frac{\partial f_h}{\partial t} \right)_c \tag{III-1} \]

where the r.h.s. is a symbolic notation for the influence of collisions. \( \mathbf{F} \) is the external force exerted upon the carriers by electric and magnetic fields and \( m^*_h \) is the hole effective mass, for simplicity assumed to be a scalar.

By integration of the Boltzmann equation multiplied by suitable factors one obtains the higher moments, i.e. transport equations for \( p, v, W \) etc. For a thorough discussion of these derivations, see e.g. [33].

As throughout this work we assume that all quantities are dependent on one space coordinate only, we give here the first three moments in their one-dimensional form:

\[ \frac{\partial p}{\partial t} + \frac{\partial}{\partial x} (pv_x) = \left( \frac{\partial p}{\partial t} \right)_c \tag{III-2a} \]

\[ \frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + \frac{2}{3n^*} \frac{\partial}{\partial x} (pW) - \frac{q}{m^*} E_x = \left( \frac{\partial v_x}{\partial t} \right)_c \tag{III-2b} \]

\[ \frac{\partial W}{\partial t} + v_x \frac{\partial W}{\partial x} + \frac{2}{3} W \frac{\partial v_x}{\partial x} + \frac{\partial Q_x}{\partial x} = \left( \frac{\partial W}{\partial t} \right)_c \tag{III-2c} \]

This hierarchy of equations is never complete since each equation also contains the next unknown in the series. Some way of truncating the series thus has to be found. This problem will be discussed in a while.

In semiconductor device theory it is customary to use the concept of relaxation times to specify the collision terms. A discussion of this concept has been given by Blötekjaer [34].

Using relaxation times means assuming that, when the external fields are taken away, the macroscopic quantities relax to equilibrium values with certain time constants, for instance:
\[
\left( \frac{\partial p}{\partial t} \right)_c = - \frac{p - p_D}{\tau_p} \quad (\text{III-3a})
\]
\[
\left( \frac{\partial v_x}{\partial t} \right)_c = - \frac{v_x}{\tau_m} \quad (\text{III-3b})
\]
\[
\left( \frac{\partial W}{\partial t} \right)_c = - \frac{W - W_L}{\tau_L} + \frac{m v_x^2}{\tau_m} \quad (\text{III-3c})
\]

Usually \( \tau_p \) is called the hole lifetime, \( \tau_m \) the momentum relaxation time and \( \tau_e \) the energy relaxation time. \( W_L \) is the thermal energy corresponding to the temperature \( T_L \) of the crystal lattice:
\[
W_L = \frac{3}{2} kT_L.
\]

A few remarks should be made about these expressions:

- when electron-hole pair creation by impact ionization is present, like in Impatt diodes, a term describing this has to be added to (III-3a). Also thermal generation of carriers is not represented here.

- Eq. (III-3b) expresses the fact that the hole velocity, when it has a drift component, is randomized by collisions. When these collisions are elastic, the energy is conserved, so the thermal energy increases. This is the origin of the second term in the r.h.s. of (III-3c). The first term here describes the transfer of energy to the crystal lattice mainly by inelastic collisions.

- to give the collision terms a more general character the relaxation times often are assumed to be functions of the macroscopic quantities.

A look at the magnitudes of the relaxation times will show us how the transport equations can be simplified. For silicon the orders of magnitude are:

\( \tau_p = 10^{-3} \text{..} 10^{-6} \) sec., \( \tau_m = 10^{-11} \) sec., \( \tau_e = 10^{-12} \) sec.

We are dealing with transit-time devices having transit times in the order of \( 10^{-10} \) sec. This is so short compared to the carrier lifetime.
that the probability for a hole to recombine during transit is negligible. So the r.h.s. of (III-2a) may be put equal to zero.

On the other hand the transit time and signal period are much longer than the momentum and energy relaxation times. Then the \( \frac{\partial}{\partial t} + v \frac{\partial}{\partial x} \) terms in (III-2b,c) can be neglected.

The set (III-2) has thus been simplified considerably. Nevertheless, in semiconductor device theory it is customary to introduce a further simplification. This is the so-called isothermal approximation which consists of neglecting the spatial gradients of \( W \) and \( \dot{Q} \). This at the same time conveniently terminates the hierarchy of moment equations.

Now (III-2b) takes the form:

\[
\dot{v} = \mu E - \frac{D}{p} \frac{\partial p}{\partial x}
\]  (III-4)

The indexes on \( v \) and \( E \) have been dropped and the mobility

\[
\mu = \frac{q \tau_m}{m}
\]

and the diffusion coefficient

\[
D = \frac{2W \tau_m}{3m^2}
\]

have been introduced. Under low-field conditions \( D \) satisfies the Einstein relation: \( D = \mu kT_L/q \).

Now let us try to shed some light on the question of the validity of the isothermal approximation. Assuming that the spatial gradients of \( W \) and \( \dot{Q} \) are small (III-2c) becomes, substituting (III-3c):

\[
W = W_L + \frac{\tau_e}{\tau_m} m^* v^2 - \frac{2}{3} \tau_e W \frac{\partial v}{\partial x}
\]  (III-5)

Now \( \tau_e \) is larger than \( \tau_m \) by a factor of five to ten. In a high-field region where \( v \approx v_s \) and \( \partial v/\partial x \) is small one finds that \( m^* v^2 \) is of the same magnitude as \( W_L \) so that \( W \) can be almost an order of magnitude larger than \( W_L \).
So the isothermal approximation looks rather drastic. Nevertheless its consequences may be less serious than it seems. Let us have a look at the relaxation times.

On physical grounds one would expect \( \tau_e \) and \( \tau_m \) if they can be written as functions of anything, to be functions of \( W \) and \( v \). Then, when \( \partial v / \partial x \) is small, one can write (III-5) as \( W = W(|v|, T_L) \) and consequently also \( \tau_m, e = \tau_m, e(|v|, T_L) \). So, if the proper \( (|v|, T_L) \) dependences are assigned to \( \tau \) and \( D \) the only approximation in (III-4) remains the neglect of spatial variation of \( W \). Using (III-5) the term \( \partial W / \partial x \) in (III-2b) becomes of the form \( v \partial v / \partial x \) and terms of this form have already been neglected.

Unfortunately things are made worse again: in the literature \( \mu \) and \( D \) are always given as functions of \( |E| \) because this is how they actually are measured. The measured dependence for silicon is that they are constant at low fields and decrease at higher fields. The drift velocity approaches a constant value at high fields.

The dependence of drift velocity on electric field has been measured by many authors. Recently Jacoboni et.al. [35] have given an extensive review of the high-field properties of silicon. The variation of \( D \) with \( |E| \) is much less well known. According to Sigmon and Gibbons [36] \( D \) is nearly constant for holes as well as for electrons, but Canali et al. [37] report a strongly decreasing \( D \) in the case of electrons.

In this work we will stick to the convention of specifying \( \mu \) and \( D \) as functions of \( |E| \), mainly because they are given this way in the literature. It may be clear from the foregoing that this is not an entirely satisfactory approach. The consequences are not as serious as one would expect at first sight. Notably in the high-field region of Baritt diodes the drift velocity rises with field but as the saturation velocity is approached the variation of \( v \) becomes smaller. The hole density gradient is small too so that diffusion plays a minor role only and \( v \) depends mainly on \( E \). In this situation it makes only little difference whether one uses \( \mu(|E|) \) or \( \mu(|v|) \) resp. \( D(|E|) \) or \( D(|v|) \).
A situation where serious errors could occur is encountered in the region to the left of the potential maximum. Here field and diffusion act in opposite directions and the velocity remains low whereas $|E|$ can reach appreciable values. This difficulty has been circumvented by keeping $\mu$ and $D$ at their low-field values when $E$ is negative.

The dependences of $\mu$ and $D$ on temperature have already been mentioned briefly. For $\mu$ it is well documented and also reviewed in [35]. For $D$ the Einstein relation has been verified within the accuracy of the measurements.

For the dependence $v(E)$ Canali et.al. [38] give a formula:

$$v = \frac{\mu E}{1 + (\mu E/v_s)^{\beta} \beta^{1/\beta}}$$

(III-6)

where $\mu$ is the low-field mobility and $v_s$ the saturation velocity. Both as well as $\beta$ are functions of temperature. Their values for holes in silicon are given in Table I at three different temperatures.

<table>
<thead>
<tr>
<th>T, °C</th>
<th>$\beta$</th>
<th>$\mu_{m^2}/v_s$</th>
<th>$v_s$, m/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>1.21</td>
<td>0.0450</td>
<td>0.81x10^6</td>
</tr>
<tr>
<td>97</td>
<td>1.25</td>
<td>0.0305</td>
<td>0.79x10^5</td>
</tr>
<tr>
<td>157</td>
<td>1.28</td>
<td>0.0210</td>
<td>0.69x10^5</td>
</tr>
</tbody>
</table>

In the course of the present work it has been found that higher values of $v_s$ than quoted in Table I consistently gave better agreement between theory and experiment. Also its temperature dependence seems to be weaker than indicated here. It should be noted that Canali's experiments did not employ fields higher than 60 kV/cm whereas in Barritt diodes values of 200 kV/cm are reached frequently. Looking at the data given in [38] one finds that they can nearly as well be matched by a curve with $\beta = 1$ and $v_s = 10^5$ m/s. Such a value for $v_s$ is also given by other authors [39].

A point that has not been mentioned yet is the dependence of mobility on doping concentration. It is well known that the low-field
mobility decreases with increasing impurity concentration due to ionized impurity scattering [32]. Caughey and Thomas [40] give the following empirical expression:

\[
\mu = \frac{\mu_{\text{max}} - \mu_{\text{min}}}{1 + (N/N_R)^\alpha} + \mu_{\text{min}}
\]  

with for holes in silicon at room temperature:

\[
\mu_{\text{max}} = 0.0495 \text{ m}^2/\text{Vs}, \quad \mu_{\text{min}} = 0.0048 \text{ m}^2/\text{Vs}, \quad N_R = 6.3 \times 10^{22} \text{ m}^{-3}, \quad \alpha = 0.76.
\]

In view of their connection with \( \nu \) one expects also \( \nu_s \) and \( D \) to depend on concentration. For the low-field case it is not unreasonable to expect that the Einstein relation remains valid so that \( D \) follows \( \mu \). However, data on the combined dependence of \( \nu \) on field, temperature and concentration are not available. Scharfetter and Gummel [41] give a formula for the combined effects of field and doping but without any experimental substantiation.

Therefore we have assumed that the impurity concentration has an effect only on the low-field mobility and that \( N_R \) and \( \alpha \) in (III-7) are independent of temperature.

III-2. Field equations

The complete electromagnetic field in the diode of course is found as a solution of Maxwell's equations where the transport equations are used to find the current term. To do this in three dimensions would be a formidable task, but, as already has been said in Ch. II, it is permissible to treat the whole as a one-dimensional problem. The main objection that can be raised is that we are dealing with a conductive medium so that a kind of skin-effect may occur. It can be made plausible, however, that this effect is small. Suppose that we can define an effective conductivity \( \sigma_{\text{eff}} = q \mu_h p_{\text{av}} \) where \( \mu_h \) is the low-field hole mobility and \( p_{\text{av}} \) is a suitable average of the hole density. For the latter we can take \( J/q \nu_s \) where \( J \) is a typical current density. For a current density of \( 10^6 \text{ A/m}^2 \), which is fairly typical, and a hole mobility of \( 0.05 \text{ m}^2/\text{Vs} \) we find \( \sigma_{\text{eff}} = 0.5 (\Omega\text{m})^{-1} \). At a frequency of
7 GHz we then find a skin depth of 1 cm which is about 100 times a typical diode radius. Even if one takes $\sigma_{\text{eff}}$ ten times higher the skin depth is still 30 times the radius.

Because of the one-dimensionality of the analysis it is not necessary to use the full set of Maxwell's equations. We can replace them with Poisson's equation:

$$\frac{\partial E}{\partial x} = \frac{q}{\varepsilon} (p-n+N_D-N_A)$$

(III-8)

where $p$ is the hole density, $n$ the electron density, $N_D$ the donor density and $N_A$ the acceptor density. Eq. (III-8) is sufficiently general to describe a semiconductor with varying doping density, including p-n junctions. In the present work we will restrict ourselves to a uniformly doped depleted n-type layer for which $n$ and $N_A$ are zero and $N_D$ is a constant. Occasionally the equation will be applied to a p-contact where $N_D$ is zero and $N_A$ is constant.

Differentiating (III-8) with respect to time, substituting (III-2a) and integrating with respect to $x$ yields the relationship

$$J_c + \varepsilon \frac{\partial E}{\partial t} = J(t)$$

(III-9)

where $J_c = qpv$ is the hole current or convection current. In other words the total current is independent of position. This is a more handy relation to use than (III-2a). With the help of (III-4) we find for $J_c$:

$$J_c = qpv(E) - q\frac{\partial p}{\partial x}$$

(III-10)

where $v(E)$ is given by an expression of the form (III-6).

The set (III-8,9,10) will be the basis of the analysis in the following chapters.
III-3. Normalizations

In the course of the analysis to be described in the following chapters it will be handy to make use of reduced, or normalized, quantities. This not only reduces the number of symbols but also makes it easier to estimate the relative importance of various effects. Two sets of normalizing quantities have been used, one of which is appropriate to a diffusion-dominated region and one which is more suitable for regions where diffusion is of secondary importance.

The first set contains the following normalizing quantities:
- voltage: the so-called thermal voltage \( V_T = kT/q \)
- length: a quantity similar to the Debye-length \( \ell_N = \sqrt{\varepsilon V_T/qN_D} \)
- density: the donor-concentration \( N_D \)
- time: an analogue of the dielectric relaxation-time \( \tau_d = \varepsilon/\sigma \)
  where \( \sigma = qv_hN_D \) in our case.

From these all other normalizing quantities can be derived, e.g.:
- field: \( E_N = V_T/\ell_N \)
- velocity: \( v_N = \ell_N/\tau_d \)
- current density: \( J_N = qN_D v_N \)
- impedance: \( Z_N = V_T/J_N A \) where \( A \) is the diode area.
- diffusion constant: \( D_N = \mu h V_T \).

The second set has the same normalizing values for density and time, but now account is taken of the fact that the drift velocity saturates. So the reducing quantities become:
- velocity: the saturated velocity \( v_s \)
- distance: \( \ell_N = \tau_d v_s \)
- field: \( E_N = v_s/\mu \)
- current density: \( J_N = qN_D v_s \)
- voltage: \( V_N = E_N \ell_N \)
- impedance: \( Z_N = V_s/J_N A \)
- diffusion constant: \( D_N = \ell_N^2/\tau_d \)
It is instructive to calculate numerical values of the parameters introduced here. If we take: \( \mu_h = 0.05 \, \text{m}^2/\text{Vs} \), \( v_s = 10^5 \, \text{m/s} \), \( \varepsilon = 10^{-10} \, \text{As/Vm} \), \( T = 290 \, \text{K} \) and \( N_D = 10^{21} \, \text{m}^{-3} \) we get the results summarized in Table II:

### TABLE II

<table>
<thead>
<tr>
<th>Qu.</th>
<th>unit</th>
<th>Set I</th>
<th>Set II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_d )</td>
<td>s</td>
<td>0.13x10^{-10}</td>
<td>0.13x10^{-10}</td>
</tr>
<tr>
<td>( \lambda_N )</td>
<td>m</td>
<td>0.13x10^{-6}</td>
<td>1.3x10^{-6}</td>
</tr>
<tr>
<td>( E_N )</td>
<td>V/m</td>
<td>0.19x10^6</td>
<td>2x10^6</td>
</tr>
<tr>
<td>( V_N )</td>
<td>V</td>
<td>0.025</td>
<td>2.6</td>
</tr>
<tr>
<td>( J_N )</td>
<td>A/m^2</td>
<td>1.6x10^6</td>
<td>16x10^6</td>
</tr>
</tbody>
</table>
IV. D.C. THEORY

IV-1. Introduction

Before the actual realization of operating Baritt diodes, d.c. theories existed only for the space-charge limited diode [42,9], i.e. a diode where the carrier density in the region following the injecting contact is so high that it dominates over the influence of the contact itself. Then the actual nature of the injecting contact is unimportant, provided it supplies enough carriers. The source region in the model of Vlaardingerbroek and the author, with the boundary condition E = 0 is an example of a space-charge limited region.

Soon after the announcement of oscillating MSM diodes [14], a d.c. theory of these diodes was published [43] which took full account of the injecting contact. Here space-charge effects were neglected completely which restricts the validity of the analysis to low current densities. Another paper [26] discussed p-n-p diodes. It considered two regimes of operation: the low-current regime where the injecting contact is dominant, and the high-current regime where the diode can be considered space-charge limited. This paper did not take account of diffusion effects. Baccarani et.al. [44] calculated carrier transport in MSM diodes using the concept of quasi-fermi levels [27] which includes diffusion. They too used simplifications, neglecting the effect of the hole space charge on the electric field and using an approximation for the v-E relationship. Finally, El-Gabaly et.al. [45] performed a numerical analysis of an MSM diode where diffusion, hole space charge and v-E dependence were taken into account and where much attention was paid to the boundary conditions. An interesting conclusion from their work is that the flat-band condition can be reached already at fairly low current densities.

For a full description of the Baritt diode all the above-mentioned effects have to be taken into account but their influence may weigh differently in different regions. The approximate profiles of hole density and field have already been discussed in Ch. II. In Fig. IV-1 they are sketched once more for a p-n-p structure. In an M-n-M diode
the n-region shows qualitatively the same picture, but the hole density at the left-hand junction is lower.

One expects from this figure that in the left part of the diode diffusion will play a predominant role and the mobility will be close to its zero-field value. To the right the non-linearity of the v-E characteristic will be important but diffusion becomes a secondary effect. It will be shown further on in this chapter that for this region a series solution can be found which gives a great saving in computing effort compared to a numerical approach. In the low-field part no analytical solution has been found and here the use of numerical techniques is necessary.

Prior to solving the equation let us say a few words about the boundary conditions.

IV-2. Boundary conditions

As the set (III-8,9,10) leads to a second-order differential equation
two boundary conditions are necessary. Near the reverse-biased junction the drift velocity is close to saturation and it turns out that the relationship between \( p \) and \( E \) is uniquely defined, see Sec. 3 of this chapter. This is equivalent to one boundary condition. The other one must be derived from the properties of the injecting contact. Let us study a p-n junction first.

From Fig. IV-1 it can be seen what the field and the density profiles in the forward-biased junction can be expected to look like. Close to the junction the field and the diffusion counteract each other and we assume that thermal equilibrium reigns, meaning that \( J_c \) is small compared to each of the terms in the r.h.s. of (III-10). Since the p-region is heavily doped we expect that the Pauli exclusion principle comes into play so that we have to use the Fermi distribution function:

\[
    p = \frac{N_v}{1 + \exp(V/V_T)}
\]  

(IV-1)

This is used to obtain a boundary condition in the following way: By differentiating with respect to \( x \) and rearranging (IV-1) yields:

\[
    \frac{dp}{dx} = \frac{p(N_v - p)E}{N_v V_T} \quad (IV-2)
\]

In the p-region Poisson's equation (III-8) becomes:

\[
    \frac{dE}{dx} = \frac{q_e (p - N_A)}{e} \quad (IV-3)
\]

Combining the last two equations we find:

\[
    \frac{dp}{dE} = \frac{\epsilon}{q N_v V_T} \cdot \frac{p(N_v - V_T)}{p - N_A} E \quad (IV-4)
\]

This is easily integrated. As a boundary condition for the p-region we assume that at the left side of this region we have \( E = 0 \) and \( p = N_A \). The result is:

\[
    N_A \ln \frac{N_A}{p} + (N_A - N_v) \ln \frac{N_v - p}{N_A - p} = \frac{\epsilon}{2q V_T} E^2 \quad (IV-5)
\]

This can be used as a boundary condition for the n-region.
In the case of a metal contact the boundary condition is somewhat easier to derive. Below flat-band, if we again assume thermal equilibrium we may put:

\[ p(0) = N_v \exp\left( -\frac{\phi_h}{V_T} \right) \]  

\( \phi_h \) is the barrier for holes going from the metal to the semiconductor. The lowest value found is for a platinum silicide-to-silicon contact for which it is about 0.2 volts.

As soon as the saturation current is reached the Schottky effect becomes operative, as already explained in Ch. II and we can write:

\[ J = J_s \exp\left( \frac{1}{V_T} \cdot \frac{qE(0)}{4\pi e} \right) \]  

\( \beta = J_s \exp\left( \frac{E(0)}{E_s} \right) \)  

Here the dependence of the barrier lowering on the junction field has been represented by a phenomenological proportionality constant as in practice never the theoretical value is found.

IV-3. The High-field Region

From here on we will work with reduced quantities, to be denoted by italics (E). The second set of normalizing parameters of Ch. III is used. Then (III-8,10) become in reduced form:

\[ \frac{dE}{dx} = p + 1 \]  

\[ \frac{dp}{dx} = \frac{p\nu(E) - J}{D(E)} \]  

The dependence of \( \nu \) on \( E \) will be represented by (III-6) with \( \beta = 1 \) which gives:

\[ \nu = \frac{E}{T + E} \]  

\( D \) is kept constant. The magnitude of \( D \) is interesting. With the same data as used at the end of Ch. III one finds \( D = 0.0047 \), so it is a very small parameter. This means, in view of (IV-9) that either a
steep gradient of $p$ exists or that $J$ is very close to $p\nu$. The first situation exists in the region to the left of the potential maximum. In the high-field region with which we are now dealing the second case prevails.

Neglecting diffusion altogether for a moment, we find from (IV-9) with (IV-10):

$$p = J(1+1/E)$$

This suggests that $p$ can be developed in inverse powers of $E$. Now, from (IV-9) $x$ can be eliminated which gives:

$$\frac{dp}{dE} = \frac{p\nu(E) - J}{(p+1)^{D}} \quad (IV-11)$$

If we now substitute

$$p = \sum a_n E^{-n} \quad n = 0, 1, 2, \ldots$$

we find:

$$a_0 = a_1 = J \quad (IV-13a)$$

$$a_n = -D \left\{ (n-1)a_{n-1} + (n-2)a_{n-2} + c_{n-1} + c_{n-2} \right\}, \quad n > 1 \quad (IV-13b)$$

with:

$$c_m = \sum k a_k a_{m-k} \quad k = 1, 2, \ldots, m; \quad m \geq 1 \quad (IV-13c)$$

What has been said in the previous section, is confirmed here: $p$ is a uniquely determined function of $E$.

The next step is to find $E(x)$. To this end we define:

$$\frac{1}{p+1} = \sum b_n E^{-n} \quad n = 0, 1, 2, \ldots \quad (IV-14)$$

which gives:

$$b_0 = \frac{1}{1+a_0} \quad (IV-15a)$$
Finding \( x(E) \) now is a matter of simply integrating (IV-9a):

\[
x(E) = x_{cc} + b_0 (E_{cc} - E) - b_1 \sum_{n=1}^{\infty} \frac{b_n}{2^n} \left( E^{2-n} - E_{cc}^{2-n} \right)
\]

where \( x_{cc} \) and \( E_{cc} \) are the values of \( x \) and \( E \) at the collecting contact. To find \( E_{cc} \), a boundary condition is necessary which has to be obtained from the injecting contact. To formulate it differently: we have found the profile of \( E \), but we don't know its location.

Another integration yields the electric potential:

\[
V(E) = V_{cc} + \frac{1}{2} b_0 (E_{cc}^2 - E^2) + b_1 \sum_{n=1}^{\infty} \frac{b_n}{2^n} \left( E^{2-n} - E_{cc}^{2-n} \right)
\]

where \( V_{cc} \) is the (reduced) potential at the collecting contact.

In (IV-12) the upper limit of the series has been left open. This is done on purpose because the series is non-convergent. With increasing \( n \) the terms first decrease but after some \( n \) increase again. The value \( N \) at which this happens is greater the greater \( E \) is. Apparently we are dealing with an asymptotic series and we must truncate it at the point where the terms start to increase again. The "solution" thus obtained will be a worse approximation the smaller \( E \) is. A definite limit of convergence does not exist. The range of validity of the series approximation depends on what difference one allows between it and the exact solution. As the latter is not known we have to find another criterion. This has been done the following way:

A fairly large number \( N \) of coefficients \( a_n \) is computed, say 30, and the value of \( E \) determined for which

\[
|a_n E^{-N}| < 10^{-7} |a_0 + a_1/E|
\]

It is then assumed that this is the smallest value of \( E \) for which the series represents a valid solution. When computing \( x(E) \) and \( p(E) \) the series are truncated when the last computed term is smaller in magnitude than \( 10^{-4} \) times the sum of the preceding terms. It has been
verified by comparison with a fourth-order Runge-Kutta integration that this gives sufficient accuracy for our purposes. The limit value of $E$ thus found lies between 0.5 and 1, depending on the values of $J$ and $\mathcal{V}$.

IV-4. The low-field region

At low field strengths the series solution of the last section breaks down and no other analytical solution has been found so one has to resort to numerical techniques.

A second-order Runge-Kutta scheme has been tried by Legius [46] and been found to work well. The set of equations (IV-9) is discretized by incrementing $x$ with a step $h$. The iteration integration scheme then is:

$$E_{n+1} = E_n + \frac{1}{2}(K_1 + K_2)$$  \hspace{1cm} (IV-18)

$$p_{n+1} = p_n + \frac{1}{2}(L_1 + L_2)$$

where the $K$ and $L$ are defined by:

$$K_1 = h[p_n + 1]$$  \hspace{1cm} (IV-19)

$$L_1 = h\frac{v(E_n) - J}{(p_n + 1)D(E_n)}$$

$$K_2 = h[p_n + L_1 + 1]$$  \hspace{1cm} (IV-20)

$$L_2 = h\frac{(p_n + L_1)v(E_n + K_1) - J}{(p_n + L_1 + 1)D(E_n + K_1)}$$

This scheme works but with a fixed step it is not very handy. The step has to be chosen small enough that convergence is obtained close to the injecting contact where the gradients of $p$ and $E$ can be very steep. Then it is unnecessarily small for the region adjoining the high-field region. Therefore the step is adapted after each integration step in such a way that the step in $E$ remains approxi-
mately the same. Using (IV-9a) this is done by putting

\[ h_{n+1} = \frac{p_{1} + 1}{p_{n} + 1} h_{1} \]

where \( p_{1} \) and \( h_{1} \) are the starting values of \( p \) and \( h \). The integration is started at the point where the series approximation of the last section breaks down. It is interesting to note that a suitable starting value for \( h \) corresponds to a distance of a few debye-lengths.

IV-5. Method of solution

The Runge-Kutta method is meant to solve initial-value problems. In our case we have a boundary condition at the injecting contact and a prescribed relationship between \( p \) and \( E \) near the other contact. This difficulty is resolved by the following procedure:

The electric field at the collecting contact is not known but we can make an estimate of it by assuming the drift velocity saturated everywhere. This estimate then becomes:

\[ E_{0} = (1+J)[\ell_{d} - x_{m}] \quad (IV-21) \]

where \( \ell_{d} \) and \( x_{m} \) are the reduced values of the diode width and the location of the potential maximum, respectively. The latter can only be guessed but since \( x_{m} \) is small (~0.1µm) its value does not affect the result very much.
Now somewhere near $x = \ell_d$ the field will have the value $E_0$. Let us denote this point by $x_0$ and define an auxiliary coordinate (see Fig. IV-2):

$$y = x_0 - x$$

Integration of the equations (IV-9) now is continued until the boundary condition valid at the injecting contact is satisfied. The value of $y$ at which this occurs gives the value of $x_0$. The place of the collecting contact then is known and the field and density at this point, if desired, can be calculated as well as the diode voltage.

IV-6. Results

![Graph](image)

*Fig. IV-3. Calculated d.c. field profile with field and diffusion components of the current in the low-field region of a p-n-p diode with abrupt p-n junctions. $N_p = 1.6 \times 10^{21} m^{-3}$, $\ell_d = 7.1 \times 10^{-6} m$, $A = 3 \times 10^{-8} m$, $I_{dc} = 30 mA$. 

-45-
To illustrate the method described above a specific example has been calculated. The parameter values have been chosen such that they represent the p-n-p diodes of which further results are given in Ch. XI. In Fig. IV-3 the electric field and the diffusion and drift components of the current are given. Only the low-field part, which

![Diagram of V_{DC} x T, °C]

**Fig. IV-4. D.C. voltage as a function of current and temperature for a p-n-p diode.**

- measured by De Cogan [47].
- calculated.

\[ N_D = 0.625 \times 10^{21} \text{ m}^{-3}, \quad \ell_d = 4.0 \times 10^{-6} \text{ m}, \quad A = 4.15 \times 10^{-8} \text{ m}^2. \]
is the most interesting, is displayed. The figure clearly demonstrates that close to the injecting junction the forces of diffusion and field almost balance each other whereas a little distance behind the zero-field point diffusion has become negligible.

Another example is the following. De Cogan [47] has measured the I-V characteristics of p-n-p diodes with a narrow n-layer at different temperatures. He found that at a certain current density the voltage remains constant within 0.1 percent as the temperature varied between -20 and +75°C. It has been tried to confirm these results by calculations and good agreement has been found, see Fig. IV-4.

In all these calculations abrupt p-n junctions were assumed. Under this condition it was found that the acceptor concentration of the p-contact played no role, as long as it was higher than $10^{23}$ m$^{-3}$. On the other hand the results were quite sensitive to the doping and width of the n-layer. So for uniformly-doped diodes the method of matching the calculated I-V characteristics to the measured ones offer a means of determining the concentration and width of the central layer.

With M-n-p diodes a different situation is encountered. Now the values of $\phi_h$, $J_s$ and $E_s$ have a great influence, so we cannot use this method to obtain information about doping and width. But when the latter two are known with reasonable accuracy we can instead obtain information about the contact parameters. In Ch. VIII it will be explained how the concentration and width can be obtained from other measurements.
V. A.C. IMPEDANCE

V-1. Introduction

In this chapter the a.c. small-signal impedance will be discussed. When one speaks about a.c. impedance one implicitly assumes the presence of time-harmonic currents and voltages. Due to the strong non-linearities connected with transport under the circumstances we consider we cannot expect sinusoidal signals. But in the presence of a steady-state a.c. signal the field quantities will be periodic in time and can be resolved in a set of harmonics. The term a.c. impedance then usually refers to the impedance defined at the fundamental harmonic. This impedance will in general be dependent on the amplitude of the fundamental and on the amplitudes and phases of the higher harmonics.

A small-signal situation exists when the amplitudes of the time-varying parts of all quantities are so small that the influence of the higher harmonics can be neglected. In this case the a.c. signal can be considered as a small perturbation of the d.c. situation and the equations can be linearized with respect to the a.c. components.

The equations we start from are (III-8,9,10). If we combine (III-9) and (III-10), omit n and N_A and convert to reduced quantities, using set 2 of Ch. III, we get:

\[
\frac{\partial E(x,t)}{\partial x} = p(x,t) + 1 \quad (V-1)
\]

\[
J(t) = p(x,t)v(E) - p\frac{\partial p(x,t)}{\partial x} + \frac{\partial E(x,t)}{\partial t} \quad (V-2)
\]

All dependent variables are split into a d.c. component and an a.c. component where the latter is always understood to be small compared to the former. For E this gives e.g.:

\[
E(x,t) = E_0(x) + \text{Re}[E_i(x)\exp(\alpha t)] \quad (V-3)
\]

where \(\alpha t\) has been substituted for the usual \(\omega t\) so that \(\alpha = \omega \tau_d\) just as in Sec. II-3. The other quantities \(J, p\) etc. are split up in the same way.
The small-signal impedance is defined as:

\[ Z = \int_0^\epsilon E_1 \frac{dE_1}{dx} \]  \hspace{1cm} \{V-4\}

If we split Eqs. (V-1,2) into d.c. and a.c. parts, using (V-3) we obtain for the a.c. quantities:

\[ \frac{dE_1}{dx} = p_1 \]  \hspace{1cm} \{V-5a\}

\[ J_1 = p_0 \left( \left. \frac{dv}{dE} \right|_{E_0} \right) E_1 + v_0 p_1 - D \frac{dp_1}{dx} + j\alpha E_1 \]  \hspace{1cm} \{V-5b\}

The a.c. component of \( v(E) \) has been replaced by the first term of a Taylor-series development around \( E_0 \).

Although this is a set of linear differential equations the solution is by no means simple. Of course the equations can be integrated numerically using the methods of the preceding chapter to find the values of the coefficients. In this work a different approach is taken. The diode is divided into regions in each of which an approximate analytical solution is tried. This gives less accurate results but surely provides more insight than a numerical technique. This approach in fact is an extension of the model of Sec. II-3 and has been published earlier [48,49]. Since these publications it has been modified somewhat so for clarity a full account will be given here.

In the model of Sec. II-3 diffusion was neglected altogether. To get some insight into this matter let us take a look at a situation where the d.c. drift velocity is at its saturated value (\( v_0 = 1 \)) but where velocity modulation by diffusion is still possible. Then (V-5) leads to a second-order D.E. with constant coefficients:

\[ J_1 = j\alpha E_1 + \frac{dE_1}{dx} - D \frac{d^2E_1}{dx^2} \]  \hspace{1cm} \{V-6\}

of which the solution can be written as:
\[ E_I = \frac{J_I}{j\alpha} + A_1 e^{|y_1|^2} + A_2 e^{|y_2|^2} \]  \hspace{1cm} (V-7)

with

\[ y_{1,2} = \frac{1}{2D} \left\{ 1 + (1 + 4j\omega D)^{\frac{1}{2}} \right\} \]  \hspace{1cm} (V-8)

Now, since \( D \) is a small number (cf. Sec. IV-3), this can be approximated by:

\[ y_1 = -j\alpha - \alpha^2 D \]  \hspace{1cm} (V-9)

\[ y_2 = +j\alpha + 1/D \]

Apparently there are two propagating waves, one traveling forward (in the direction of carrier drift) and weakly damped and one traveling backward which is heavily damped.

A situation similar to the one analyzed here exists in the high-field region of the diode close to the collecting contact. The forward wave arrives from the direction of the injecting contact and its amplitude will be determined by the conditions prevailing there. The backward wave will be excited at the collecting contact and its amplitude will be determined by the a.c. boundary conditions at this contact. Because of the strong damping of this wave it will travel only a very short distance (\( \Delta x \approx D \)) and will influence neither the forward wave nor the a.c. diode voltage. So we may as well forget about it and assume that its only function is to satisfy the boundary conditions at the collecting contact.

In this analysis the damping of the forward wave is small enough to be negligible. However, in a more realistic picture, as we move from the collecting contact to the left, the drift velocity decreases gradually from its saturated value and velocity modulation will occur. Also diffusion becomes more important in a similar way as it did in the d.c. analysis. Both effects will increase the damping. To allow for this in the analysis we have inserted a region between the potential maximum and the source region of Sec. II-3 where special attention is paid to diffusion.
In the new model we now distinguish three regions, see Fig. V-1:

- the contact region, between the injecting contact and the potential maximum;
- the diffusion region, between the potential maximum and a point where the drift velocity is so high that diffusion becomes negligible;
- the drift region which comprises both source and drift region of the older model. In this region the same $v-E$ relationship as in Sec. IV-3 will be used.

In M-n-p diodes above flat-band the width of the contact region is zero of course. At high currents the field at the contact can rise so high that even the diffusion region may be left out.

The point that separates diffusion region and drift region is expected to be in the neighbourhood of the point where the series development
of Sec. IV-3 breaks down. Since in the a.c. case diffusion plays a somewhat less prominent role than in the d.c. case the transition may be somewhat further to the left. In our model it is specified by its d.c. field value $E_i$. Of course the choice of $E_i$ will influence the results. Using it therefore is only reasonable if it satisfies two demands:
- small changes should not affect the results greatly;
- once a suitable value for $E_i$ has been found in a certain situation, it should be possible to predict what value it must have in different situations (e.g. different temperature, donor concentration).

Further discussion of this point will be postponed until Sec. V-4. Now we will discuss the three regions consecutively.

V-2. The contact region

The model of Haus et.al. [29] for the contact region has three deficiencies:
- it assumes the a.c. field to be uniform which it is not;
- transit-time effects in this region have been neglected;
- the value of $x_m$ is calculated only approximately.

The second of these probably is the least serious as the width of the contact region is small so the transit-time will be small compared to the signal period. The first one may be more serious, especially at high current densities.

The third point concerns d.c. calculations and is easier to improve. Sellberg [50] has computed the d.c. field pattern in a p-n contact region and has found that the results can be represented by a few relatively simple formulas. He used the first set of Sec. III-3 to normalize the variables and found that this leaves only two parameters, namely the reduced d.c. current density $J_{o\delta}$ and the reduced acceptor concentration in the p-region $N_{A\delta}$. For the reduced values of $x_m$, $p_m$ (hole concentration at $x_m$) and barrier potential $V_m$ he gives the following expressions:
One notes that at low current densities the exponential dependence of $J_0$ on $V_m$, used by Haus et al., is a good approximation. At high current densities deviations from this law occur (due to the low value of $J_N$ in this set, $J_{m0}$ can reach values in the order of one).

It is interesting to note that Wright [51], by simple physical arguments, has come to formulas bearing a great similarity to (V-10), viz.:

$$p_{m0}^3 + p_{m0}^2 = 2J_{m0}^2$$

$$\ln J_{m0} = J_{m0} \ln J_{m0} - V_{m0} \delta(J_{m0})$$

where $\delta(J_{m0})$ varies from 1 at low currents to 1.5 at high currents.

Eqs. (V-10) will be used to provide d.c. boundary conditions for the diffusion region. We will retain the boundary conditions of Haus and Weller (cf. Eqs. II-9 and IV-8) for the a.c. case:

$$\eta_c = \frac{\alpha V_I}{J_0 x_m}$$

$$\eta_c = \frac{2\alpha (E, E)}{J_0}$$

The first of these is not quite in agreement with (V-10c) but the difference is small and we neglect it. Wright [51] and others have suggested for an a.c. boundary condition: $\eta_c = \frac{\alpha}{P_m}$.
This supposes that the a.c. convection current at $x_m$ is determined by the field alone which clearly is not true. We therefore prefer (V-11).

To conclude this section we calculate the impedance of the contact region. This becomes:

\[ Z_C = \frac{\eta_C}{G + j\omega \tau_c} \]  

(V-12)

V-3. The diffusion region

In this section the influences of field and diffusion are of equal importance so Eqs. (V-5) should be used in their full complexity. Finding an analytical solution will be a complicated matter, if at all possible, so we will content ourselves with a rather crude approximation, justifiable mainly by the fact that the width of the diffusion region is small. Reduced quantities will be employed using the second set of Sec. III-3.

The d.c. behaviour is analyzed under the assumption that the variation of $p$ can be described by a linear interpolation between $x_m$ and $x_1$, the beginning of the drift region:

\[ p_0 = p_m + \frac{p_m - p_m}{x_m - x_m} (x - x_m) \]  

(V-13)

Consequently, for $E_0$ we have:

\[ E_0 = (1 + p_m) (x - x_m) + \frac{1}{2} \cdot \frac{p_m - p_m}{x_m - x_m} (x - x_m)^2 \]  

(V-14)

By requiring continuity of $E$ and $p$ at the transition to the drift region the value of $x_1$ can be calculated. By definition we have:

\[ E_0(x_1) = E_{\phi} \]

and since in the drift region we will neglect diffusion:

\[ p_\phi = \frac{J_0}{\nu_0(E_{\phi})} \]
which, substituted in (V-14) gives:

\[ x_i = x_m + \frac{2E_i}{2p_m + p_i} \]  

(V-15)

An approximation solution of (V-5) is obtained by replacing \( p_0 (dv/dE) \) and \( v_0 \) by their average values \( p_a \) and \( v_a \). The solution then takes the form:

\[ E_i = \frac{J_i}{f_a p_a} + A_1 \exp(\gamma_1 x) + A_2 \exp(\gamma_2 x) \]  

(V-16)

with

\[ \gamma_{1,2} = \frac{1}{2D} \left\{ v_a \pm \left( v_a^2 + 4D(p_a+j\alpha) \right)^{1/2} \right\} \]  

(V-17)

This has the same form as (V-7,8) but it may not always be allowable to simplify \( \gamma \) the same way since \( v_a^2 \) is not necessarily large compared to \( 4D(p_a+j\alpha) \). Nevertheless it is instructive to carry out this simplification. Then we find for \( v_a^2 \gg 4D(p_a+j\alpha) \)

\[ \gamma_1 = -\frac{p_a}{v_a} + \frac{D(\alpha^2-p_a^2)}{v_a^3} - j\alpha \left( 1 - \frac{2Dp_a}{v_a} \right) \]  

(V-18a)

\[ \gamma_2 = \frac{v_a}{p_a} + \frac{p_a}{v_a} + \frac{D(\alpha^2-p_a^2)}{v_a^3} + j\alpha \left( 1 - \frac{2Dp_a}{v_a} \right) \]  

(V-18b)

The first wave is forward-traveling and damped mainly by velocity modulation, but also by diffusion. The second wave travels backward and is heavily damped by diffusion.

The amplitudes \( A_1 \) and \( A_2 \) of the waves follow from the boundary conditions at \( x_m \) and \( x_i \). Here a difficulty arises. The boundary at \( x_i \) is an artificial one introduced only for the purpose of calculation. In reality it does not exist and a backward wave will not be excited at \( x_i \). Now, our main purpose with this analysis is to find out how the boundary condition provided by the injecting contact is transferred to the drift region. To this end it seems more appropriate to leave out the backward wave. This is not correct.
mathematically of course, but it gives a better picture of the actual physical situation.

With the boundary condition at \( x_m \) derived from (II-9):

\[
J_f = (\eta_c + j\alpha)E_f(x_m)
\]

we obtain for \( E_f(x) \):

\[
E_f = \frac{J_f}{p_a^j + j\alpha} + \frac{\eta_c}{\alpha(p_a^j + j\alpha)}\exp\gamma_f(x - x_m)
\]  \( (V-19) \)

From this we can calculate the boundary condition at \( x_i \) and the impedance of the diffusion region:

\[
\frac{\alpha E_f(x_i)}{J_f} \frac{\eta_i}{1 + j\alpha} = \frac{\alpha}{p_a^j + j\alpha} + \left( \frac{\eta_c}{1 + j\alpha} - \frac{\alpha}{p_a^j + j\alpha} \right)\exp\gamma_f(x_i - x_m)
\]  \( (V-20) \)

\[
Z_i = \frac{1}{p_a^j + j\alpha} + \left( \frac{\eta_c}{1 + j\alpha} - \frac{\alpha}{p_a^j + j\alpha} \right)\exp\gamma_f(x_i - x_m) - 1
\]  \( (V-21) \)

From (V-20) it can be seen that the diffusion region has a screening influence similar to that of the source region discussed in Sec. II-3. When \( |\gamma_f(x_i - x_m)| \) is small \( \eta_i \) approaches \( \eta_c \) so that the boundary condition at \( x_m \) is transferred unchanged to the drift region. In the other case the exponential vanishes and only the first term in the r.h.s. is left which means the boundary condition at \( x_i \) is determined by the diffusion region alone. Looking at (V-18a) one sees that the damping of the forward wave increases with \( p_a \) and thus is highest at the highest current densities, so that there the screening effect is at its strongest. At low currents it does not disappear, however, because of the diffusion term.

A phase shift also exists. As has been discussed already this is beneficial to the negative resistance of the drift region. In the present case the transit angle of the diffusion region is approximately equal to:

\[
\theta_i = \frac{\alpha}{p_a} (x_i - x_m)
\]  \( (V-22) \)
V-4. The drift region

In this region diffusion will be neglected altogether. On the other hand, allowance is made for a general v-E characteristic. This is possible because (V-5) now reduces to a first-order D.E. for which a solution in integral-form can be found. It has been given by McCumber and Chynoweth [52]. A simpler representation is obtained when one converts from the coordinate x to the d.c. transit time \( \tau \) as the independent coordinate. The derivation has been given before [49] and only a condensed version will be given here. Dascalu [53,54] has given a derivation along the same lines for majority-carrier transport.

The (reduced) d.c. transit-time \( \tau(x) \) is defined by:

\[
\tau = \int_{x_L}^{x} \frac{dx}{v_o(E)} \quad [V-23]
\]

Using the d.c. part of Poisson's equation and bearing in mind that \( v \) now depends on \( E \) only one obtains from (V-23):

\[
\chi = \int_{x_L}^{x} \frac{v(E) dE}{v(E)+J_0} \quad [V-24a]
\]

\[
\tau = \int_{x_L}^{x} \frac{dE}{v(E)+J_0} \quad [V-24b]
\]

With (V-23), (V-5) with \( D = 0 \) can be converted to:

\[
J_1 = \left( \frac{J_0}{v_o} \cdot \frac{dV_o}{dE} + j\alpha \right) E_1 + \frac{dE_1}{dt} \quad [V-25]
\]

where \( v_o \) and \( dv_o/dE \) have to be considered as functions of \( \tau \). The last equation is solved by:

\[
E_1(\tau) = J_1 \left( 1 + \frac{J_0}{v_o} \right) \exp(-j\alpha \tau) \left( \frac{\eta_L}{\eta + J_0 + v_o} \cdot \frac{\eta_L}{\eta + J_0 + v_o} \right) + \int_{0}^{\tau} \frac{v_o}{J_0 + v_o} \exp(j\alpha \xi) d\xi \quad [V-26]
\]

where \( v_L = v(E_L) \)

-57-
The impedance of the drift region can also be expressed as an integral over $T$:

$$Z_d = \int_0^T \left\{ \frac{\nu_0}{J_0 + \nu_0} \exp\{-j\omega t\} dt \right\} \left\{ \frac{\eta_\infty}{J_0 + \nu_0} \frac{\eta_\infty}{\alpha(1 + j\omega) \xi} + \int_0^T \frac{\nu_0}{J_0 + \nu_0} \exp(j\omega t) dt \right\} \{V-27\}$$

with $\tau_d = \tau(\xi_d)$

This expression can be worked out in such a way that the unsaturated part of the drift velocity is separated out. The effect of the non-saturated drift velocity can then be discussed. This has been done in [49] and will not be repeated here as its conclusions are the same as those from the model in Sec. II-3.

V-5. Conclusion

With the expressions derived in the foregoing the small-signal impedance can be calculated. As an example in Fig. V-2 the impedance of a p-n-p diode is shown at three different temperatures and three different frequencies. The diode parameters are the same as those of diode F2 discussed in Ch. XI. One notes that at higher frequencies the curves shift to higher currents which is in agreement with the scaling laws derived in Sec. II-4. Also at higher temperatures the curves shift to higher currents. This is found to be a consequence of the decrease of low-field mobility with temperature. Further discussion of numerical results will be postponed until Ch. XI where they will be compared with experimental data.

We just have to say a few words about the field $E_i$ that determines the separation between the diffusion region and the drift region. In Sec. V-1 the conditions which it must satisfy have already been mentioned. Now we want to find an estimate for $E_i$ and a clue as to how it should change when the parameters change. For this purpose let us have a look at the equation for the d.c. current which follows from (III-10):

$$J_0 = p_0 v(E_0) - D \frac{dp_0}{dx}$$
The drift region is characterized by the fact that the second term on the right is small which implies:
\[ \frac{d\rho_0}{dx} \ll J_0 \]
If we take \( \rho_0 = J_0 / \nu_0 \) the magnitude of \( Vd\rho_0 / dx \) can be calculated.
After some manipulation this leads to the condition: \( E^2 \gg D \).

Now let an estimate for \( E_\parallel \) be:
\[ E_\parallel = (kD)^\frac{1}{2} \quad (V-28) \]
where \( k = 10 \) for instance. With the value of 0.0047 for \( D \) calculated in Sec. IV-3 we find \( E_\parallel = 0.22 \) which is somewhat smaller than the field for which the series solution of Sec. IV-3 breaks down. The estimate can be improved by comparison with experiment but Eq. (V-28) shows how it should be adapted when the parameters \( (T,N_D) \) change.

![Graphs showing impedance of a p-n-p diode as a function of d.c. current, temperature, and frequency.](image)

Fig. V-2. Impedance of a p-n-p diode as a function of
d.c. current, temperature and frequency.
\[ N_D=1.6\times10^{21} \text{ m}^{-3}, \quad L_d=7.1\times10^{-6} \text{ m}, \quad A=3\times10^{-8} \text{ m}^2. \]
VI. NOISE

VI-1. Introduction

Under small-signal conditions the noise in a Baritt diode comes from two sources mainly [29,55]:
- the injected current has a shot-noise component;
- the thermal motion of the carriers throughout the diode also produces noise currents.

A third noise source can be the multiplication of carriers when the field at the collecting contact rises to very high values. This is an undesirable effect and care should be taken to avoid it, e.g. by doping the high-field region lower so that the field gradient there is less steep. Here we will consider only the first two noise sources.

We will now proceed by calculating the open-circuit noise voltage, that is, the noise voltage appearing at the diode terminals in a situation where the a.c. current is blocked but the d.c. bias current still can flow. Under small-signal conditions the diode behaves as a linear device for a.c. signals and the noise can be represented by a voltage source in series with the diode impedance. It is assumed that the two noise sources mentioned above are uncorrelated so that their mean-squared noise voltages can be added.

VI-2. Shot noise

The injected shot noise current is given by the well-known formula [56]:

\[ I_s^2 = 2qI_o \Delta f \]  

(VI-1)

This current induces an electric field \( E_s \) at the injecting contact which in its turn modulates the injected d.c. current and also produces a dielectric displacement current. The sum of the injected and the induced currents must be zero because of the open-circuit assumption. This gives a boundary condition which, converting to
reduced quantities with set 2 of Ch. III, becomes:

\[ J_\delta + (\alpha/\eta_C + j\alpha)E_\delta = 0 \]  \hspace{1cm} (VI-2)

where \( J_\delta = I_s/AJ_N \).

This boundary condition was introduced by Haus et al. [29] and applied to their model where the drift velocity is saturated from \( x_m \) onwards (cf. Sec. II-2). We will now introduce it into our a.c. model, developed in Ch. V. When the diode is operating below flat-band the boundary condition will be applied at \( x = x_m \), otherwise at \( x = 0 \).

In the diffusion region the noise field is calculated from (V-5) with \( J_1 = 0 \). We then find the forward wave of (V-16,17) which, with (VI-2), becomes:

\[ E_\delta = -J_\delta \cdot \frac{\eta_c}{\alpha(1+j\eta_c)} \exp y(x-x_m) \]  \hspace{1cm} (VI-3)

From this the boundary condition for the drift region and the noise voltage over the diffusion region can be obtained. The result for the latter is:

\[ V_{\delta d} = -J_\delta \cdot \frac{\eta_c}{\alpha(1+j\eta_c)} \exp y(x_{\Delta}-x_m) \cdot \frac{1}{\gamma_1} \]  \hspace{1cm} (VI-4)

In the drift region we have to solve (V-25) with \( J_1 = 0 \) and \( E_\delta \) at \( x_{\Delta} \) from (VI-3) as a boundary condition. After integration the noise voltage is obtained:

\[ V_{\delta d} = -J_\delta \frac{\eta_c}{\alpha(1+j\eta_c)} \exp y(x_{\Delta}-x_m) \frac{V_{\Delta}}{1+V_{\Delta}} \int_0^{T_d} (J_0 + V_0) \exp(-j\omega t) dt \]  \hspace{1cm} (VI-5)

Summation of (VI-4) and (VI-5) gives the mean-squared shot noise voltage across the diode:

\[ V^2_{\delta} = \frac{2}{J_\delta^2} \left| \frac{V_{\delta d} + V_{\delta c}}{J_\delta} \right|^2 \]  \hspace{1cm} (VI-6)
VI-3. Thermal noise: the impedance-field method

To calculate the thermal noise two methods exist, the Langevin method [57] and the impedance-field method [58]. The equivalence of both has been shown recently [59]. We will use the impedance-field method. It is outlined below where it is assumed from the start that we are dealing with thermal noise in a one-dimensional structure. The impedance-field method itself is more general and can be applied to all distributed noise sources in more dimensions.

Charge carriers in a solid have a random thermal motion. This was mentioned already in Ch. III. For the present it means that there are noise currents in addition to the deterministic currents. If it were possible to connect the planes \(x\) and \(x + \Delta x\) via external leads to a high impedance (see Fig. VI-1a) then in these leads a noise current \(\delta i(t)\) would be induced which by Ramo's theorem [60,61] can be calculated:

\[
\delta i(t) = \frac{1}{\Delta x} \sum_{j=1,2,\ldots,N} q u_{xj}
\]

where \(N\) is the number of carriers between \(x\) and \(x + \Delta x\) and \(u_{xj}\) is the random velocity in \(x\)-direction of the \(j^{th}\) carrier. We can now replace the effect of thermal motion by an equivalent current source \(\delta i(t)\) connected between \(x\) and \(x + \Delta x\) (Fig. VI-1b). This current source will produce an open-circuit voltage \(\delta V_t\) at the diode terminals.

The impedance-field method now consists of two parts:
- finding an expression for the noise current from an analysis of the noise-generating process and
- finding the relation between \(\delta i(t)\) at \(x\) and the open-circuit noise voltage produced by it.

To calculate the mean-squared noise current we assume that the random motions of individual carriers are uncorrelated. Then we can write:

\[
(\delta i(t) \Delta x)^2 = (\sum_{j} q u_{xj})^2 = q^2 N u_x^2
\]
So far we have been working in the time domain. To go to the frequency domain we must calculate the autocorrelation function of $u_x(t)$ and Fourier-transform this. In a model where the carriers undergo collisions at an average time interval $\tau_o$ and where the mean free path is independent of the velocity it is found [58]:

$$\overline{(\delta I(\omega) \Delta x)^2} = 2q^2 N \tau_o \bar{u}_x^2 \Delta f$$  \hspace{1cm} (VI-9)

It can be shown that $\tau_o$ is twice the momentum relaxation time $\tau_m$. Furthermore, if the distribution function is isotropic, one can write

$$\frac{1}{2} \mu x = \frac{1}{3} W$$
where $W$ is the thermal energy defined in Ch. III. Then the factor $\tau_0 u_x^2$ can be written as

$$\tau_0 u_x^2 = \frac{4}{3} \cdot \frac{\tau_m W}{m^*} = 2D$$  \hspace{1cm} (VI-10)

According to the definition of $D$ in Ch. III Eq. (VI-9) then becomes, also substituting $N$:

$$\left(\delta I(\omega)\Delta x\right)^2 = 4q^2 A p(x) U(x) \Delta x \Delta f$$  \hspace{1cm} (VI-11)

where $A$ is the diode area.

---

![Diagram of four-pole used in the impedance-field method.](image)

**Fig. VI-2.** Four-pole used in the impedance-field method.

The next step is to calculate the noise voltage $\delta V(\omega)$ induced by the current $\delta I(\omega)$. With reference to Fig. (VI-2) one can write down a linear relationship:

$$\begin{pmatrix} V_T \\ V_X \end{pmatrix} = \begin{pmatrix} z_{TT} & z_{TX} \\ z_{XT} & z_{XX} \end{pmatrix} \begin{pmatrix} I_T \\ I_X \end{pmatrix}$$  \hspace{1cm} (VI-12)

$Z_{TT}$ of course is the small-signal impedance.

Note that in a device with carrier drift in the $x$-direction this relation is non-reciprocal because different combinations of waves are excited by the currents $I_T$ and $I_X$. 

-64-
Now, from (VI-12), for a current $\delta I$, injected at $X + \Delta X$ and extracted at $X$ we have a terminal voltage:

$$
\delta V_T = \left( Z_{TX, X+\Delta X} - Z_{TX} \right) \delta I = \frac{dZ_{TX}}{dX} \delta I \Delta X \tag{VI-13}
$$

To find the total mean-squared noise voltage we have to divide the diode in elements $\Delta x$ and sum the $\delta V_T^2$ of all the elements. This gives, combining (VI-11) and VI-13):

$$
\overline{V_T^2} = 4q^2 A \int_{\Delta x} d \delta \frac{dZ_{TX}}{dX} \left| p(x) D(x) dx \right|^2 \tag{VI-14}
$$

We proceed now to calculate the impedance field $dZ_{TX}/dX$ for our model of a Barritt diode. From (VI-13) it follows that to do this we must impress a current $\delta I$ between $X$ and $X + \Delta X$, calculate the resulting field (still using the open-circuit assumption) and integrate it to obtain the terminal voltage $\delta V_T$.

When $X$ is in the diffusion region we have:

- to the left of $X$ a backward wave:
  $$
  E_1 = P \exp \gamma_2 (x-X) \tag{VI-15a}
  $$

- to the right of $X + \Delta X$ a forward wave:
  $$
  E_1 = Q \exp \gamma_1 (x-X-\Delta X) \tag{VI-15b}
  $$

- between $X$ and $X + \Delta X$ the complete set of waves:
  $$
  E_1 = \frac{\delta J}{p + j \alpha} + R \exp \gamma_1 (x-X) + S \exp \gamma_2 (x-X-\Delta X) \tag{VI-15c}
  $$

Here $\gamma_1$ and $\gamma_2$ are the same as in Sec. V-3 and

$$
\delta J = \delta I / \Delta X \tag{VI-16}
$$

The amplitude constants are determined by demanding continuity of $E_1$ and $dE_1/dx$ at $X + \Delta X$. In the limit $\Delta X \to 0$ they become:
The values of $R$ and $S$ are of no further importance. The noise voltage over the diffusion region now is:

$$\delta V_{li} = \frac{\Delta X}{p_{a+f_{\alpha}}} \left\{ \frac{\gamma_1 \gamma_2}{\gamma_1 - \gamma_2} \exp\gamma_1(x_i - x_l) - \frac{\gamma_1}{\gamma_1 - \gamma_2} \exp\gamma_2(x_i - x_m) \right\}$$  \hspace{1cm} (VI-18)

The forward wave, when arriving at $x_i$, excites a wave in the drift region which is calculated by solving the homogeneous form of (V-25). This gives for the voltage over the drift region:

$$\delta V_{td} = \frac{\Delta X}{p_{a+f_{\alpha}}} \frac{\gamma_1 \gamma_2}{\gamma_1 - \gamma_2} \frac{V_{d}}{V_{o}} \int_{t_0}^{t_d} (J_o + V_o) \exp(-j\alpha t) \, dt$$  \hspace{1cm} (VI-19)

Substituting the sum of $V_{li}$ and $V_{td}$ in (VI-13), using (VI-16), yields $dZ_{TX}/dx$.

When $X$ is in the drift region the calculation goes along the same lines. It can be shown now that the field to the left of $X$ may be neglected. We can then put $E_1 = 0$ as a boundary condition at $X$ so that between $X$ and $X + \Delta X$ we have:

$$E_1 = \left( \frac{\delta J}{\nu_0} \right)(x - X)$$  \hspace{1cm} (VI-20)

To the right of $X + \Delta X$ we again have the homogeneous solution of (V-25). The final result is:

$$\frac{dZ_{TX}}{dx} = \frac{\exp(j\alpha t)}{J_o + V_o} \int_{t_0}^{t_d} (J_o + V_o) \exp(-j\alpha t) \, dt$$  \hspace{1cm} (VI-21)

VI-4. Conclusion

Although the expressions obtained are rather lengthy their evaluation is straightforward. Numerical results will be given in Ch. XI together with measurements.
VII. TECHNOLOGY

VII-1. Introduction

The manufacturing of Baritt diodes is not a very difficult process compared with the fabrication of, say, klystrons or Impatt diodes. The first operating devices [14] were made by thinning down a silicon wafer to a thickness of 12 µm and then metalizing it on both sides. This is a rather delicate process and it is difficult to maintain a uniform layer thickness. Nowadays it is common practice to start from an epitaxial n-type layer grown on a p+ substrate so that one junction is present already. This way uniform layers of any desired thickness can be made. The disadvantage is that the substrate introduces a series resistance which reduces the already small negative resistance attainable. This effect can be minimized by etching down the substrate to less than 50 µm thickness.

The second junction is made as a p-n junction by a shallow diffusion or an ion implantation, or as a metal-semiconductor rectifying contact by a suitable metalization.

To improve the power output more complicated structures have been made where the central layer has a non-uniform doping profile. The profile is made either in the epitaxial growing process [20] or afterwards by ion implantation [22]:

In our laboratory both M-n-p and p-n-p diodes were made having a uniform n-layer. Starting material were silicon expitaxial slices made at Philips Research Labs. The substrates were p-type, orientation [1.1.1.] and having a resistivity of 0.01 ohmcm. On these n-type epilayers were grown of about 7 µm thickness and 3 Ωcm resistivity. To form Schottky contacts on the n-layers platinum silicide was used. This is known to have a low barrier for holes [62] and therefore one expects a large saturation current. However, our experiments have shown that this expectation is not always fulfilled. To form p-n junctions a shallow boron diffusion from a doped-oxide source was applied.
VII-2. Formation and evaluation of platinum silicide contacts

To make a platinum silicide layer one has to deposit platinum onto the slice and then heat it in a neutral atmosphere. Platinum and silicon then interdiffuse and react to form the intermetallic compound platinum silicide of which four different phases exist: PtSi, Pt₂Si, Pt₁₂Si₅ and Pt₃Si.

The deposition is done in a Randex r.f. sputtering system, Model 2400-6J, equipped with a turbomolecular high-vacuum pump of Leybold, Model 450. It allows three different materials to be sputtered in sequence, as well as sputter etching, without breaking the vacuum. The vacuum chamber is evacuated to 2 x 10⁻⁷ torr after which the sputtering is done in argon at a pressure of 3 x 10⁻² torr.

First the slice is sputter-etched to obtain a clean silicon surface. Then about 200 Å of Pt is deposited. The slice is placed in a furnace and heated in dry nitrogen. In our first experiments a temperature of 550°C was applied for 20 minutes. According to the literature [63] this should be sufficient to form platinum silicide. Measurements on test diodes with an n⁺-substrate (giving single Schottky diodes) showed that the Schottky barriers produced this way were far from ideal. Therefore a heat treatment of 2 hours at 650°C was tried. The layers were analyzed by powder X-ray diffraction and by electrical measurements. A full account of the experiments has been published elsewhere [28]. The results are summarized in Table VII-1.

<table>
<thead>
<tr>
<th>Process</th>
<th>XRD</th>
<th>$\Phi_h (V)$</th>
<th>ideality</th>
<th>reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>av.</td>
<td>max.</td>
<td>factor n</td>
</tr>
<tr>
<td>20 min. 550°C</td>
<td>negative</td>
<td>0.76</td>
<td>0.80</td>
<td>1.14</td>
</tr>
<tr>
<td>2 hrs 650°C</td>
<td>PtₙSi</td>
<td>0.80</td>
<td>0.84</td>
<td>1.05</td>
</tr>
</tbody>
</table>

It was noticed after the heat-treatment at 550°C that the surface colour had darkened which indicates that a reaction has taken place. Keeping in mind that the sputter etching produces a nearly amorphous
silicon surface layer and that also the deposited platinum is amorphous one might speculate that at this temperature only amorphous (or polycrystalline with very small grain size) platinum silicide is formed. This would explain the negative X.R.D. result.

At the higher temperature diffraction lines belonging to all silicides are found but due to the fact that many lines coincide it was not possible to say with certainty which silicides are present. However, it is clear that there is a correlation between the parameters of the sintering process and the electrical properties.

![Graph showing doping profiles in a p-n-p diode before and after a diffusion of 1 h. at 1050 °C.](image)

**Fig. VII-1.** Doping profiles in a p-n-p diode before and after a diffusion of 1 h. at 1050 °C.

---

- Assumed epitaxial profiles before,
- Acceptor profiles after diffusion.

**VII-3. Formation of p-n junctions**

To make a p-n junction a diffusion of a p-dopant, usually boron is necessary. In our laboratory this is done with the so-called Silox process. This is a C.V.D. technique in which a silicon dioxide layer is formed on the slice by a reaction of silane SiH₄ and oxygen at
350°C [64]. To use the oxide as a doping source the silane is mixed with a small volume of diborane $\text{B}_2\text{H}_6$ so that the oxide contains some boron. This boron is subsequently driven into the silicon by heating the slice to 1050°C for 30 min. A disadvantage of the diffusion method is that also the p-dopant in the substrate diffuses out into the n-layer, so that the n-layer width is substantially reduced. To illustrate this in Fig. VII-1 the approximate doping profiles are sketched of a Baritt diode before and after a diffusion of 1 hour at 1050°C. It has been assumed that initially the substrate has an abrupt profile with a concentration of $1.5 \times 10^{25} \text{ m}^{-3}$. The diffusion constant of boron in silicon is taken constant at a value of $0.011 (\mu\text{m})^2/\text{hr}$. The n-doping is $1.5 \times 10^{21} \text{ m}^{-3}$. The figure indicates that a reduction of the layer width of about 1 $\mu\text{m}$ is possible. In the next chapter methods to measure the layer width will be discussed.

A better method is to use ion implantation. In this process boron ions are accelerated by a moderately high voltage and shot into the silicon. The depth they reach is dependent on the acceleration voltage, e.g. 0.2 $\mu\text{m}$ for 20 kV. The crystal lattice of the silicon is damaged by the implantation and has to be annealed at a temperature between 750 and 950°C. At this temperature the outdiffusion from the substrate is much lower so that the reduction of the n-layer width can be restricted to less than 0.5 $\mu\text{m}$.

VII-4. Further processing

After the rectifying junction has been made a contact must be made to it. This has been done by sputtering another 200 $\AA$ of Pt (after sputter etching) immediately followed by 2000 $\AA$ of gold. The gold then is electroplated to a thickness of about 5 $\mu\text{m}$ (Fig. VII-2a).

Using the standard phototesist technique the gold is etched to leave a pattern of circular gold dots (Fig. VII-2b). These are used as a mask to etch out the mesa diodes (Fig. VII-2d). Before this can be done the platinum silicide must be removed by sputter etching since it is not attacked by chemical etchants. Some of the gold is removed too but
Fig. VII-2. Processing steps in the manufacturing of Baritt diodes.

a. metalized silicon wafer.

b. after photo-mask and etching of gold.

c. after backsputtering of platinum silicide.

d. after mesa etching.

e. diode chip mounted in a microwave package.
this is no problem since the gold is much thicker than the PtSi (Fig. VII-2c).

After mesa etching the wafer is scribed and broken into single diode chips which are eutectically bonded onto the goldplated pedestal of a microwave package (Fig. VII-2e). The top of the diode is connected to the package flange with gold wires which are thermocompression-bonded and the package is closed in a neutral atmosphere.

The process described here has been used for the diodes reported in this thesis. Since then several improvements have been made which are of no consequence for the impedance and noise properties but which improve the fabrication yield, the stability and the oscillator power of the diodes.

First, the substrate is thinned down by chemical etching to about 50 \( \mu \text{m} \). This reduces the parasitic series resistance.

Second, the adherence of the metal layer is improved when titanium or chromium is used as an intermediate layer instead of platinum. Also the substrate side is metalized with Pt-Au or Pt-Ti-Au and the chips are soldered onto the pedestal with gold-tin eutectic, melting at 280\(^\circ\)C. This makes the mounting easier and probably also reduces the parasitic resistance. Third, the gold wire, usually 20 \( \mu \text{m} \) in diameter, is replaced by a ribbon of 20 x 100 \( \mu \text{m} \) cross-section. Hereby the series resistance and inductance are reduced.
VIII. DIAGNOSTIC MEASUREMENTS

VIII-1. Introduction

In this chapter some measurements are discussed that can be characterized as diagnostic because they yield, via simple relationships, information about diode parameters that are important in the calculations. The r.f. impedance and noise measurements do not fall in this category and will be discussed in the next chapter. Of course, by comparison with theory it should be possible to draw conclusions about the diode parameters from these measurements too, but the relationships are so complex and involve so many parameters at the same time that this is a rather dangerous thing to do.

The diagnostic measurements we speak about are capacitance-voltage measurements in the "zero"-current regime below punch-through and current-voltage measurements above punch-through. Furthermore a diagnostic interpretation of the r.f. impedance data below punch-through will be discussed.

VIII-2. C.V. measurements

Measurements of capacitance versus voltage is a popular diagnostic tool in semiconductor technology. Used on single diodes it can yield the doping profile among others. This application will be discussed here briefly with special reference to Baritt diodes.

In Fig. VIII-1 a Barritt diode biased below punch-through is sketched. As already mentioned in Ch. II, in this situation we can treat it as two separate diodes connected by a thin ohmic layer. Consider now first the back-biased diode 1. When the edge of the depletion layer shifts an amount $\Delta x_1$ the stored charge and the voltage across the layer change by:

\[
\Delta Q_1 = AqN_D(x_1)\Delta x_1
\]

\[
\Delta V_1 = x_1\Delta E(x_1) = \frac{qN_D(x_1)}{\epsilon} x_1\Delta x_1
\]

(VIII-1)
The differential capacitance then is

\[ C_1 = \frac{\Delta Q_1}{\Delta V_1} = \frac{\varepsilon A}{x_1} \]  
(VIII-2)

This capacitance can be measured by superimposing a small a.c. voltage on the d.c. bias. At low frequencies the capacitance can be influenced by the charging and discharging of slow traps. For this reasons the measurement usually is done at a frequency of 1 MHz.

By differentiating \( C_1^{-2} \) with respect to \( x \) and combining with (VIII-1) one obtains the familiar result

\[ \frac{dC_1^{-2}}{dV_1} = \frac{2}{q\varepsilon A^2 N_d(x_1)} \]  
(VIII-3)

Whereas this expression together with (VIII-2) gives a nice way to measure \( N_d(x) \) of single diodes, the situation in Baritt diodes is more complicated since one can only measure the combined capacitances and voltages of the two depletion layers. For the capacitance one has

\[ C = \frac{\varepsilon A}{x_1 + x_2} \]

and if one assumes uniform doping the voltage change can be written as:
\[ \Delta V = \frac{qN_D}{\epsilon} (x_1 \Delta x_1 - x_2 \Delta x_2) \]

So instead of (VIII-3) one now has

\[ \frac{dC^2}{dV} = \frac{2}{q \epsilon \Lambda^2 N_D} \left( 1 + \frac{2x_2 \Delta x_2 + x_1 \Delta x_2 + x_2 \Delta x_1}{x_1 \Delta x_1 - x_2 \Delta x_2} \right) \]

(VIII-4)

Fig. VIII-2. Apparent doping profile of a uniformly-doped M-n-p Baritt diode as deduced from C-V measurements.

About the correction term (the second term in brackets) the following can be said:

The voltage over the forward-biased diode is determined by its current which is equal to the leakage current passed by the reverse-biased junction. If the latter is a good quality p-n junction, its leakage current is low and nearly constant over a wide voltage range. Then
dx_2 \text{ will be small and the correction term can be approximated by } x_2/x_1 \text{ which is about one at zero bias and becomes smaller at high voltages. So applying Eq. (VIII-3) blindly gives at low bias voltages an apparent donor concentration which is a factor two lower than the real one. At higher voltages the approximation becomes better. This is illustrated by Fig. VIII-2 where the apparent } N_d \text{ is plotted against the apparent } x \text{ (calculated from (VIII-2)) for a Baritt diode that should have uniform doping. The values for doping and width found by the method of the next section are indicated by dotted lines.}

It may be concluded that \( k_d \) is reasonable well approximated by the value of \( x \) where the apparent \( N_d \) starts to rise sharply and that a reasonable estimate for \( N_d \) is obtained at the higher bias voltages. Clearly the C-V method is of only limited value for Baritt diodes. However if one may assume that the depletion capacitances do not change between 1 MHz and the microwave region one can use them as calibrated impedances in the r.f. measurements. For this application it is necessary to take care that the experimental conditions in both cases are such that a comparison is valid. Concerning this point the following remarks can be made:

a. In the r.f. impedance measurements the diode is placed between the broad walls of a rectangular waveguide which is excited in the dominant TE_{01}-mode. The lines of the electric field are then parallel to the axis of the package. To get a comparable situation in the C-V measurements the package has to be mounted between two parallel planes.

b. The dielectric constant should not change between 1 MHz and the microwave region. This is a well established fact [65].

c. The influence of leakage currents on the forward biased diode has already been mentioned but their direct contribution to the diode impedance has been neglected. One can model the influence of this current by (non-linear) resistors parallel to the two depletion capacitances.

The complete equivalent circuit of the diode then becomes that of Fig. VIII-3.

At low frequencies the values of \( R_{1,2} \) are given by the differential
resistances \( \frac{dV}{di} \) of the diodes. For the forward-biased diode 2 this probably remains true up to microwave frequencies since its transit time is very small.

Now, if the value of \( \frac{l}{2\pi RC} \) of one of the diodes is equal to one at a frequency between 1 MHz and the microwave region, this diode will appear as a resistance at 1 MHz and as a capacitance at microwave frequencies. The apparent capacitance at 1 MHz then can no longer be used as a calibration value for the microwave measurements. Especially

\[
\text{Fig. VIII-3. Equivalent circuit of a Baritt diode below punch-through.}
\]

the forward-biased diode will be a possible cause of difficulties in this respect, since its differential resistance decreases strongly with increasing current. To give an idea of what is possible let us make a simple calculation:

the capacitance \( C_2 \) and resistance \( R_2 \) can be calculated as:

\[
C_2 = A \left( \frac{eqN_D}{2(V_D-V)} \right)^{\frac{1}{2}} \quad \text{and} \quad R_2 = \frac{V_T}{I + I_0}
\]

where \( V_d \) is the so-called diffusion voltage, that is the voltage over the depletion layer at zero bias. To calculate \( R \) the familiar diode equation \( I = I_0 (\exp \frac{V}{V_T} - 1) \) has been used.

For our diodes \( A \approx 3 \times 10^{-8} \, \text{m}^2 \), \( N_D \approx 10^{21} \, \text{m}^{-3} \), \( V_d \approx 0,5 \, \text{V} \) and \( I_0 \approx 10^{-10} \, \text{A} \). Then at zero bias \( C_2 \approx 4 \, \text{pF} \) and \( R_2 \approx 2.5 \times 10^8 \, \Omega \) so that at 1 MHz \( \omega R_2 C_2 \approx 6000 \). This is large enough to neglect \( R_2 \), but when \( I_0 \) increases to \( 10^{-6} \, \text{A} \), which is still a small current, \( \omega R_2 C_2 \) becomes in the order of one.

-77-
At microwave frequencies the parallel resistances in Fig. VIII-3 may be neglected even for leaky diodes and a simple series circuit remains whose impedance is:

\[ Z = R + jX = \frac{E_d - x_1 - x_2}{\sigma A} - j \frac{x_1 + x_2}{\omega \varepsilon A} \]  

(VIII-5)

where \( \sigma \) is the conductivity of the material: \( \sigma = q n_e N_d \). Thus if one plots \( Z \) at varying bias in the complex plane a curve like in Fig. VIII-4 results. The part below punch-through gives a straight line with slope: \( \frac{dX}{dR} = \frac{\sigma}{\omega \varepsilon} \).

This expression is remarkable because it does not contain dimensional parameters. At microwave frequencies \( R \) and \( X \) are of comparable magnitudes so that a plot of this kind can be used to obtain the donor concentration. The always present parasitic series resistance does not affect the result since it is constant.

Even more information can be extracted from this plot. Similarly to the case of Impatt diodes [66] the curve shows a sharp kink at the point of punch-through. Here the injected holes start to make their contribution to the impedance which is of a quite different character.
due to the transit-time effect. Now, at punch-through the diode is fully depleted and can be considered as a pure capacitance whereas the only resistance present is the parasitic series resistance due to the semiconductor substrate and the mounting wires (cf. Ch. VII). So from X at the kink the layer width $\delta_d$ and from $R$ the parasitic resistance can be determined.

VIII-4. I-V measurements

In Ch. IV it has already been pointed out that current-voltage measurements can be used to obtain information about the diode parameters. To avoid unnecessary complications it is desirable to keep the diode temperature constant as the current is increased, otherwise the influence of the varying temperature will intermingle with that of the current and a comparison with theory becomes rather difficult.

In Barritt diodes, like in all active microwave semiconductor devices power densities are high and a temperature rise of 100°C is normal. The rise of the diode temperature $T$ can be related to the dissipated power $P$ by:

$$P = C_t \frac{dT}{dt} + \frac{dW}{dt}$$  \hspace{1cm} (VIII-6)

where $C_t$ is the heat capacity of the diode, i.e. the product of the specific heat at constant pressure $C_p$ and the volume. $dW/dt$ is the energy carried away per unit time.

This expression assumes that the temperature is uniform across the diode. In Barritt diodes with their non-uniform field distribution this is not very likely. However, in view of the small layer width and the good heat conductivity of silicon one does not expect the temperature difference to be more than a few degrees.

The heat produced has to flow away through the substrate and the package to the heat sink, a process that is described by the equation:

$$\frac{dW}{dt} = \frac{T-T_o}{R_t}$$  \hspace{1cm} (VIII-7)
where \( T_0 \) is the heat-sink temperature and \( R_t \) is called the heat-flow resistance, measured in Kelvins per Watt. In principle it is temperature dependent but this dependence is weak and usually is neglected.

Eqs. (VIII-6 and 7) show that the thermal behaviour of the diode can be modeled by the parallel-circuit of a thermal capacitance and a heat-flow resistance and the diode temperature rises according to:

\[
T = T_0 + PR_t (1 - \exp(-t/R_t C_t))
\]  

(VIII-8)

To put in some numbers: our diodes typically have a width of 7 µm, an area of \( 3 \times 10^{-8} \) m² and draw 50 mA at 70 V d.c. With the specific heat of 1.62 J/cm³ for silicon the initial temperature rise is 10 K/µsec. A heat flow resistance of 30 K/W gives a final temperature rise of 105 K and a time constant \( R_t C_t \) of 10 μsec.

It is clear from these numbers that to avoid temperature effects one has to use current pulses of less than a microsecond duration and a repetition time much longer than 10 μsec.

As a matter of fact, it is not so much the pulse width that counts but the instant after applying the pulse at which it is possible to take a measurement. In other words the pulse rise time is the quantity to be considered.

In our measurement the bias current is always supplied from a constant-current source. This is first because the current is the important parameter in the theoretical models and second because the slope of the I-V characteristic is rather steep. Now when the current source is pulsed with a short rise time the voltage rise time is determined by the capacitance parallel to the diode.

When working on a microsecond time scale we can represent the diode by a non-linear resistor and the package plus surrounding hardware by a parallel capacitor, Fig. VIII-5. The differential equation for this circuit is:

\[
I_o(t) = C \frac{dV}{dt} + I(V)
\]  

(VIII-9)
It is assumed here that the \( I(V) \) relationship of the diode is an instantaneous one. This is true for time scales longer than the transit-time and shorter than the charging time constants of traps or the carrier lifetime. At 0.1-1 \( \mu \text{sec} \) both conditions are fulfilled. Then (VIII-9) can be integrated directly and gives, with \( I_0(t) \) a step function and initial condition \( V = 0 \), the result:

\[
t = \int_0^V \frac{C_d V}{I_0 - I(V)} \text{d}V
\]

(VIII-10)

\[\text{Fig. VIII-5. Low-frequency equivalent circuit of a mounted Barritt diode above punch-through.}\]

The forms that Baritt diode \( I-V \) characteristics can take have already been sketched in Ch. II, Fig. II-6. To get an idea of what the solution of (VIII-10) can look like let us approximate the characteristics by two, resp. three straight lines, see Fig. VIII-6a:

\[
\begin{align*}
I &= 0, & V &\leq V_{\text{PT}} & \text{(VIII-11a)} \\
I &= \frac{V-V_{\text{PT}}}{R_d}, & V_{\text{PT}} &\leq V \leq V_{\text{FB}} & \text{(VIII-11b)} \\
I &= \frac{V_{\text{FB}}-V_{\text{PT}}}{R_d} + \frac{V-V_{\text{FB}}}{R_d} + \frac{V-V_{\text{FB}}}{R_d}, & V &> V_{\text{FB}} & \text{(VIII-11c)}
\end{align*}
\]

-81-
Here $V_{PT}$ is the punch-through voltage, $V_{FB}$ the flat-band voltage. Then the solution of (VIII-10) becomes:

$$I = 0 \quad \{ \begin{array}{l}
I_o t \\
V = \frac{I_o t}{C}
\end{array} \} \quad t \leq t_o = \frac{CV_{PT}}{I_o} \quad (VIII-12)$$

$$I = I_o \left\{ 1 - \exp \left( -\frac{t - t_o}{CR_{dl}} \right) \right\} \quad t > t_o \quad (VIII-13)$$

$$V = V_{PT} + 1R_{dl}$$

---

**Fig. VIII-6a. Idealized I-V characteristics of Baritt diodes.**

---

--- **p-n-p, --- M-n-p.**
When \( I_0 \) is larger than the saturation current (VIII-13) is valid only up to a time

\[
t_1 = t_0 + CRd_1 \ln \left( \frac{I_0 R^d_1}{I_0 R^d_1 + V^P - V^F_0} \right)
\]

(VIII-14)

After this instant we have:

\[
I = I_0 \left( I_0 - \frac{V^F - V^P}{R^d_1} \right) \exp \left( - \frac{t-t_1}{CR^d_2} \right)
\]

\[
V = V^F + R^d_2 \left( 1 - \frac{V^F - V^P}{R^d_1} \right)
\]

(VIII-15)

---

**Fig. VIII-6b.** Transient behaviour of Baritt diodes.

---

p-n-p, — — — M-n-p.
In Fig. VIII-6b these waveforms are sketched for the case $V_{PT} = 50$ V, $V_{FB} = 55$ V, $R_{d1} = 200$ ohm, $R_{d2} = 1000$ ohm and $C = 100$ pF. One notes that as long as the diode stays below flat-band the final current is reached in a time at which the temperature rise is still acceptable. Above flat-band the settling time is much longer due to the higher differential resistance and a higher temperature rise has to be accepted.

It may be mentioned here that a comparison of pulsed I-V characteristics taken at different temperatures with a d.c. characteristic at room temperature gives a way to determine the thermal resistance [67,68]. We have found that this method works well for M-n-p diodes above flat-band where the I-V characteristics are spaced well apart. In the case of operation below flat-band, especially with p-n-p diodes, the characteristics are very close and the method becomes inaccurate.
IX. R.F. IMPEDANCE MEASUREMENT

IX-1. The waveguide bridge method

There are various ways to measure impedance at microwave frequencies. The classical method is to mount the unknown impedance at the end of a transmission line and to measure the standing wave pattern on the line. A modification of this method uses directional couplers to separate incident and reflected waves and then measures the amplitude and phase relationship between these waves. Both methods become inaccurate when an impedance with a relatively small real part, like a Baritt diode, has to be measured since the standing wave ratio is very high in this case. The accuracy can be improved by measuring the s-parameters of the directional couplers and use these in a computer calculation to correct the measured data. Then one still has to cope with non-linearities and drift in the detectors.

![Diagram of waveguide bridge](image)

*Fig. IX-1. Principle of the waveguide impedance bridge.*

Van Iperen and Tjassens [69] have described a waveguide bridge that allows impedances with a small real part to be measured with great accuracy. Its principle is illustrated by Fig. IX-1. The comparing element is a hybrid-T to whose symmetrical ports a measuring branch...
and a reference branch are connected. The two remaining ports are connected to a signal generator and a null detector, respectively. When the hybrid-T is perfectly symmetric there is no direct coupling from the signal port to the detector port and a null is obtained at the detector when the reflection coefficients of the measuring branch and the reference branch are equal. The reference branch contains a variable attenuator and an adjustable short which together serve to produce a reference reflection coefficient. The measuring branch contains a variable attenuator, a variable phase shifter and the unknown impedance. If the reflection coefficient of the reference branch is known the unknown impedance can be calculated from the attenuation and phase settings in the measuring branch. In practice it is done somewhat differently: the bridge is balanced first with a known impedance. Then, leaving the reference branch unchanged, the unknown impedance is substituted and the bridge is balanced again by adjusting the phase and attenuation in the measuring branch. From the change of the latter two the unknown can be calculated.

The waveguide bridge has several advantages:
- the response characteristic of the detector plays no role,
- the equalization of the two reflection coefficients is done with great accuracy,
- waveguide attenuators and phase shifters have better resolution, stability and accuracy than their coaxial counterparts,
- the bridge is easy to adapt for pulsed measurements.

A disadvantage is that its operation is rather time-consuming due to the many calibrations that have to be done.

For our measurements a bridge based on these principles has been built for the waveguide band of 5.5 to 8 GHz. The same arguments as used in Sec. VIII-4 make it interesting here too to measure with pulsed bias. Modifications therefore have been made that allow pulsed-bias measurements and also the use of elevated temperatures. A description and the theory of operation will be given in the following sections. A more detailed account will be published elsewhere [70].
IX-2. Description of the hardware

The hybrid-T was manufactured with high mechanical precision to assure good symmetry. The direct coupling from the signal port to the detector port was measured to be less than -50 dB.

The bridge signal is generated by a klystron oscillator which is stabilized in frequency by reference to a crystal-controlled source. Due to the stabilization the measuring frequency can only have the values \((n \times 200 \pm 30)\) MHz. This is quite sufficient since the variation of Baritt impedance with frequency is smooth. The oscillator is connected to the hybrid-T via a low-pass filter to eliminate higher harmonics that could obscure the detector minimum.

The reference branch contains a flap attenuator and an adjustable short-circuit plunger. Mechanical stability is the only demand on these components, since their setting should not change during the measurement.

The measuring branch is the most critical part of the bridge. The variable attenuator is a rotary-vane attenuator of Flann Microwave Instruments, Model no. 14/11. The rotary-vane principle assures that its characteristics do not change with time. The scale is marked in 0.01 dB steps below 4 dB attenuation and is easily readable to 0.002 dB. The mechanical stability is such that the resettability is of the same magnitude. Since the manufacturer specifies an accuracy of 0.1 dB which is not sufficient for our purposes, the scale has to be calibrated. This is described in Sec. 4.

A phase shifter of the necessary accuracy (0.5 degrees) is not available commercially, so one was made in the form of a squeeze section. This is a piece of waveguide (in our case 1.10 m long) with a narrow longitudinal slot in the center of both broad walls. By squeezing it together one can change the propagation coefficient and thereby the phase shift. The squeezing is done with a micrometer screw via a lever. The reading of the micrometer is a measure of the phase shift. The calibration of this relationship is described in Sec. 4. The lever principle not only reduces the force on the micrometer but also improves the reading accuracy.
The diode mount is a piece of reduced-height waveguide made by the same techniques as Van Iperen and Tjassens [69] have used. Fig. IX-2 shows a cross-section. Around the diode clamp channels are drilled through which heated silicone oil can be run. In this way the diode temperature can be raised to 125 °C. The temperature is measured with a thermocouple that is brought into contact with the base of the package.

![Fig. IX-2. Cross-section of the reduced-height diode mount.](image)

Between the phase shifter and the diode mount a tapered transition is necessary. A smooth taper, sketched in Fig. IX-3, was made to keep its reflections as low as possible. A provision was made to cool the taper when the mount is heated. Otherwise the heat will spread to the rest of the bridge and the associated expansion of components will cause additional phase shifts. To reduce the heat flow from the mount to the taper the mount flange is machined in such a fashion that only a small part of its area makes contact with the taper.

![Fig. IX-3. Cross-section of the waveguide taper.](image)
For detection of the bridge signal two alternative methods were used. For c.w. signals an unsophisticated spectrum analyzer is employed. When the bias is pulsed the signal is pulse-modulated and using the spectrum analyzer is not possible so it is replaced by a heterodyne receiver. A difficulty arising here is that outside the bias pulse the bridge is unbalanced and a large signal appears which overloads the i.f. amplifier. This gives rise to severe ringing effects that distort the signal during the pulse. To avoid this a PIN-diode switch is inserted between the hybrid-T and the receiver that shuts off the signal outside the bias pulse. The problems associated with the pulse measurements will be discussed in Sec. 6.

Fig. IX-4. The measuring branch viewed as a cascade of two-ports.

1X-3. Theory

In Fig. IX-4 a block diagram of the measuring branch is given. One can view this branch as a cascade of two-ports, transforming the diode impedance to a reflection coefficient at the hybrid port. These two-ports are:
- the attenuator,
- the phase shifter,
- the taper,
- the diode mount,
- the diode package.

Each two-port can be characterized by a set of s-parameters. We will do this for the first three but following the existing custom we will represent the mount and the package by equivalent impedance networks. Accordingly we define intermediate impedances \( Z_d \), \( Z_p \), \( Z_w \) and reflection coefficients \( \Gamma_w \), \( \Gamma_t \), \( \Gamma_p \), \( \Gamma_a \) as shown in Fig. IX-4.
The principle of the measurement implies that $\Gamma_a$ is kept constant and the s-parameters of the attenuator and of the phase shifter change when $Z_w$ changes. In the following we will discuss the relationship between input and output for each two-port, starting with the package.

A drawing of a typical package is given in Fig. IX-5a. The outer surface of the ceramic is considered as port 1 and the diode metalizations as port 2. The equivalent circuit we use is given in Fig. IX-5b. $C_p$ represents the ceramic wall and $L_p$ is due to the pedestal and wires. In practice it also contains a resistive part because the wires have skin-effect losses. Since it is difficult to separate this resistance from the residual diode resistance it will be included in the latter.

An estimate can be obtained from the impedance of a package where the wires are bonded directly onto the pedestal.

The relation between the package impedance $Z_p$ and the diode impedance $Z_d$ now is:

$$Z_p = \left( j\omega C_p + \frac{1}{j\omega L_p + z_d} \right)^{-1}$$  \hspace{1cm} (IX-1)
The next step is to find an equivalent circuit for the diode mount. A simplified drawing of its physical structure is given in Fig. IX-6. A symmetric obstacle like this in a waveguide can generally be represented by the equivalent circuit of Fig. IX-7 [71]. The reference plane for ports 1 and 2 is the plane through the centre of the obstacle. $Z_s$ can be related to the symmetric higher-order waveguide modes excited around the obstacle and $Z_a$ to the antisymmetric modes. $Z_a$ and $Z_s$ can be calculated separately by placing an electric ($E_{\text{tan}} = 0$) or magnetic ($H_{\text{tan}} = 0$) wall, respectively, through the centre of the obstacle.

![Fig. IX-6. Simplified drawing of waveguide-mounted package.](image)

![Fig. IX-7. Equivalent circuit of a symmetrical obstacle in a waveguide.](image)

It is clear that this equivalent circuit is meaningful only when a characteristic impedance for the waveguide is defined. We use the definition that relates the power flow in the dominant mode to the potential difference between the upper and lower wall in the center of the waveguide and which therefore can be related directly to the
low-frequency definition of impedance for an obstacle placed there. It gives the result:

\[ Z_0 = \frac{2b \lambda_g}{a \lambda_o} \left( \frac{\mu_o}{\varepsilon_o} \right)^{\frac{1}{2}} \]  \hspace{1cm} (IX-2)

where \( \lambda_g \) is the guide wavelength and \( \lambda_o \) the free space wavelength. \( Z_a \) and \( Z_s \) can be measured by terminating port 2 with two different impedances \( Z_{il} \) and \( Z_{i2} \), and measuring the corresponding impedances \( Z_{w1} \) and \( Z_{w2} \) seen looking into port 1. They are related by:

\[ Z_{wk} = \frac{Z_a}{\left( \frac{1}{Z_a} + \frac{Z_s - Z_a}{Z_{ik}} \right)^{-1}} \quad k = 1, 2 \]  \hspace{1cm} (IX-3)

In the bridge measurement we mount in the first case a short-circuit at \( \frac{3}{4} \lambda_g \) behind the diode plane, giving \( Z_{il} = \infty \) theoretically, and in the second case a short at \( \frac{1}{2} \lambda_g \) behind the diode, giving \( Z_{i2} = 0 \). In practice \( Z_{il} \) and \( Z_{i2} \) differ slightly from these values.

Inverting (IX-5) we have

\[ Z_{a,s} = -\frac{B-D}{A-C} + \left( \frac{B-D}{A-C} \right)^2 + \frac{BC-AD}{A-C} \]  \hspace{1cm} (IX-4)

with

\[ A = Z_{w1} - Z_{il}, \quad B = Z_{w1} Z_{il}, \quad C = Z_{w2} - Z_{i2}, \quad D = Z_{w2} Z_{i2} \]  \hspace{1cm} (IX-5)

To determine which sign belongs to \( Z_s \) we first note that

\[ |Z_{i1}| \gg |Z_{w1}| \gg |Z_{i2}| \]

\[ |Z_{w2}| \approx |Z_a| \ll |Z_{w1}| \]  \hspace{1cm} (IX-6)

So that we may write

\[ Z_s \approx Z_{w1} + \left( \frac{Z_{w2}^2}{Z_{w1}} \right)^{\frac{1}{2}} \]

Since we are measuring impedances \( Z_s \) with a small real part we may
assuming that $Z_s$ and consequently in $Z_{wl}$ the imaginary part will be dominant. So let us write:

$$Z_s \approx Z_{wl} + j\left(-\frac{Z_{wl}^2}{2}\right)$$

Now, since $Z_a$ is small, we must have $Z_s \approx 2Z_{wl}$ so that the plus sign must be taken when $\text{Im}Z_{wl} > 0$ and the minus sign in the other case. Applying this to (IX-4) we obtain:

$$Z_s = -\frac{B-D}{A-C} + j\text{sgn} (\text{Im}Z_{wl}) \left\{ \frac{B-D}{A-C}^2 + \frac{AD-BC}{A-C} \right\}^{\frac{1}{2}}$$

(IX-7)

Fig. IX-8. Equivalent circuit of the diode mount.

$C_a = \varepsilon_0 \pi d^2 / 4b$, $n = J_0(2\pi d/\lambda_0)$.

Following Getsinger [72] and Van Iperen [73] we subdivide the mount by considering the outer surface of the package as a third port which is loaded with an impedance $Z_p$ representing the package with its contents. Accordingly the equivalent circuit is divided into two parts, one representing the package and one representing the transformation from the diode plane in the waveguide to the circumference of the package. The latter part consists of a series impedance $Z_{sm}$, a transformer and a negative capacitance $-C_a$ parallel to port 3, see Fig. IX-8. That $Z_{sm}$ is
independent of the exact contents of the package has been concluded by
Getsinger [72] from the expressions of Marcuvitz [74] for a dielectric
post in a waveguide. According to Van Iperen [73] we can also use this
circuit when the package is replaced by a metal dummy, provided we put
$Z_p = 0$ in this case. This gives a method to measure $Z_{sm}$. However, $Z_a$
is not independent of the package contents and should be determined for
each measurement individually, at least in principle.

The measurement described above gives the total $Z_s$ in the parallel
branch from which $Z_p$ then can be found (by reference to Fig. IX-8):

$$Z'_p = \frac{j\mu_0^{\frac{2nd}{o\lambda_0}}(Z_s - Z_{sm})}{Z_p}$$

(IX-8)

$$Z_p = \left(\frac{1}{Z'_p + j\omega C_a}\right)^{-1}$$

So far it has been assumed that the height of the ceramic outer wall
of the package is equal to the waveguide height. From work of
Heijnemans [75] and Versnel [76] it can be concluded that the same
equivalent circuit may be used when the ceramic is lower than the
waveguide height. This is of great practical importance since now
packages of different dimensions can be measured in the same mount.
The only difference is that a fringe capacitance appears parallel to
port 3. For normal package dimensions, however, this capacitance is
so small that it can be neglected.

From now on it will be more practical to speak in terms of reflection
coefficients instead of impedances. From the foregoing the reflection
coefficient $\Gamma_w$ looking into port 1 of the mount is:

$$\Gamma_w = \frac{Z_w - Z_o}{Z_w + Z_o}$$

(IX-9)

When the taper is reflectionless, this reflection coefficient under-
goes only a fixed (but frequency dependent) phase shift upon transfer
to the taper input. The taper we use has a maximum VSWR of 1.07 (at
7 GHz). Neglecting this introduces a too larger error. Therefore the
full set of s-parameters has to be used so we get:
The same reasoning applies to the phase shifter. Here however the VSWR is never greater than 1.03 which is considered small enough to neglect the contribution of \( s_{11} \) and \( s_{22} \). Then \( s_{12} \) remains which in the ideal case only contains a phase shift. In practice it also contains some damping. The reflection coefficient \( \Gamma_p \) looking into the phase shifter from the hybrid side then can be written as:

\[
\Gamma_p = \exp(-2\alpha_p + 2j\varphi_p)\Gamma_t \tag{IX-11}
\]

The variable attenuator has a maximum VSWR of 1.07 which again is too much to be neglected. So we have for the reflection at the hybrid-T:

\[
\Gamma_a = s_{11} + \frac{s_{12}^2}{1 - s_{22}^2}\Gamma_p \tag{IX-12}
\]

where the \( s \)-parameters now refer to the attenuator.

In principle all these parameters are dependent on the setting. But over the limited range which is used in this measurement (typically from 2.5 to 3 dB) it turns out that \( s_{11} \) and \( s_{22} \) as well as the phase of \( s_{12} \) can be considered as constants. \( \Gamma_a \) is kept the same during the measurement but since \( s_{11} \) is constant we can also use:

\[
\Gamma_a \Gamma_t = \frac{\exp(-2\alpha + 2j\varphi)\Gamma}{1 - s_{22}^2}\tag{IX-13}
\]

As said already in Sec. 1 we start with a reference measurement. For this the diode mount is empty and a \( \frac{3}{4} \lambda \) short-circuit is mounted behind it so that the \( Z_S \) of Fig. IX-7 is infinite and \( Z_w = Z_{11} \). The latter result is obtained because the diode plane is used as the reference plane in defining \( Z_w \). The fixed phase shift occurring between port 1 of the mount and the diode plane is included in the \( s \)-parameters of the taper. Now, since \( Z_{11} \) is a known quantity, \( \Gamma_t \) is known too.
Let us denote this value by \( \Gamma_{to} \). The phase shifter and attenuator are at the settings \( \phi_0 \) and \( \alpha_0 \). Then with (IX-11,12,13) we have:

\[
\Gamma' = \frac{\Gamma_{to} \exp\{-2(\alpha_1 + \alpha_0) + 2j(\phi_1 + \phi_0)\}}{1 - s_{22} \exp\{-2\alpha_0 + 2j\phi_0\} \Gamma_{to}}
\]  

(IX-14)

Here \( s_{22} \) refers to the attenuator.

If now an unknown impedance is mounted we have an unknown \( \Gamma_{t1} \) and phase and attenuation settings \( \phi_1 \) and \( \alpha_1 \). An expression of the same form as (IX-14) applies and by combining the two we get:

\[
\Gamma_{t1} = \frac{\Gamma_{to} \exp\{2(\alpha_1 - \alpha_0) - 2j(\phi_1 - \phi_0)\}}{1 - s_{22} \Gamma_{to} (\exp\{2(\alpha_1 - \alpha_0) - 1\}\exp\{-2\alpha_1 + 2j\phi_1\})}
\]  

(IX-15)

When all components are ideal \( \Gamma_{to} \) and the denominator are equal to one and (IX-15) reduces to the expression given by Van Iperen and Tjassens [69].

It is interesting to note that the correction factor in the denominator disappears when the two attenuation settings are equal, that is, when the unknown impedance is purely imaginary. This means that the correction introduced by \( s_{22} \) is proportional to \( \text{Re}Z_w \), not to \( |Z_w| \).

One also notes that, unlike the ideal case, not only the change in phase shift but also its absolute value must be known.

Once \( \Gamma_{t1} \) is known Eqs. (IX-9,10) can be used to calculate \( Z_w \).

IX-4. Calibrations

A number of calibrations is necessary before a measurement can be made. In principle the full set of s-parameters should be determined for every two-port in the measuring branch. Fortunately we need to know only those which appear in Eq. (IX-15). For others, notably \( s_{11} \) and \( s_{22} \) of the phase shifter, it is sufficient to know that they are small.
The s-parameters of the taper as well as $s_{22}$ of the attenuator can be determined in the usual way, terminating the component with known impedances and measuring the standing wave on the other side with a slotted line. In the case of the attenuator a commercial sliding load was used. The taper was terminated on the reduced-height side with short-circuit pieces that are employed in the measurements to terminate the mount, as well as with a sliding load made in reduced-height waveguide.

The slotted-line method does not give the necessary accuracy for the determination of $|s_{12}|$ of the attenuator and $\arg(s_{12})$ of the phase shifter. Here other methods must be employed which will now be described.

The accuracy of the attenuator is specified by the manufacturer as 0.1 dB in the range below 10 dB. From Eq. (IX-15) we note that not the absolute values of attenuation need to be known but the difference between two settings, so probably the accuracy in our measurements is better than the specification, especially since we use only a limited part of the range (mostly that between 2 and 3 dB). Nevertheless it was judged necessary to calibrate the attenuator. This has been done with the help of a power meter, simply measuring the output power as a function of the attenuator setting at constant input power. Of course this puts high demands on the linearity of the power meter. For instance, a non-linearity of 0.1 percent gives an attenuation error of 0.004 dB, which is about the accuracy we want. Somewhat surprisingly this turned out to be possible, using an M.I. Sanders, Model 6460 with a General Microwave Power Head, Model 6420. The output of the Power Meter was read on a digital voltmeter. It should be mentioned that only the range of 0-4 dB was calibrated which means that the output power varies a range of 0.4 to 1. The linearity was first tested by doing the calibration twice at power levels differing by a factor of 2. The result was the same within 0.002 dB. Then it was compared with a waveguide-below-cut-off attenuator at a frequency of 30 MHz. Here the linearity was found to be better than the accuracy of the 30 MHz attenuator which is 0.005 dB. The measured attenuation as a function of reading was represented by a 5th-degree polynomial whose coefficients were determined by the least squares method.
The phase variation of the attenuator is specified by the manufacturer to be less than 3 degrees over the whole range. Most of this is in the range of 0-1 dB. In the part of 2-3 dB where we use it mostly it was measured to be less than 0.05 degrees which is certainly negligible.

Calibration of the squeeze section was done in the bridge substituting the taper by a movable short with a micrometer scale which serves as a reference phase shifter. The short had to be specially made for this purpose since in commercial movable shorts the variation of the waveguide width along its length is so large that intolerable phase deviations result. The measured phase as a function of the micrometer reading on the squeeze section was represented by a fifth degree polynomial.

As already said in the previous section \( Z_{sm} \) can be determined from two measurements on a metal dummy package. One cannot rely on the theoretical value of \( Z_{sm} \) since the bias feed-through produces some field disturbance [69] which can be represented by an impedance in series with \( Z_{sm} \) [77]. The value of \( Z_a \) found in this measurement can not be used for other configurations. Since \( Z_a \) is dependent on the package contents one in principle has to do two measurements for each impedance to be determined. In our experiments it has been found however that \( Z_a \) differs very little between an empty package, an internally short-circuited package and package with diode. Therefore \( Z_a \) is determined on an empty package and this value is used in the diode measurements.

The most tricky calibration perhaps is that of the package. \( C_p \) is found from the measurements on an empty package mentioned above. It varies very little from package to package. \( L_p \), however, can show significant variations because it is influenced by the position of the mounting wires. Van Iperen and Tjassens [69] have demonstrated that \( L_p \) can be calibrated for each package individually using the 1 MHz C-V data. The method is to first determine \( Z_p \) when the diode is biased at a voltage below punch-through. As discussed in Ch. VIII the diode in this case can be represented by the series circuit of a resistor and a capacitor. The value of the latter is measured at 1 MHz. Subtracting from...
\[ Y_p = \frac{1}{Z_p} \] the admittance of \( C_p \) leaves an impedance of which the imaginary part is equal to \( \omega L_p - \frac{1}{\omega C_d} \). Subtracting the contribution of \( C_d \) yields \( L_p \). One has to be sure in this case that the depletion capacitance in the microwave region is the same as at 1 MHz. The authors mentioned above give an example of a \( p^+ - n - n^+ \) Impatt diode where this goes wrong, which they ascribe to the circumstance that the metalization of the n-layer behaves as a Schottky contact. The mechanism is probably the same as the one discussed in Sec. VIII-3.

In our Baritt diodes the contacts are not likely to cause this kind of trouble because they consist of \( p^+ \)-layers metalized with metals that have a low barrier for holes, like gold or platinum. So for low-leakage diodes the method can be trusted.

IX-5. Measuring at elevated temperatures

When higher temperatures are applied the mount, the short-circuit behind it and part of the taper expand which changes their characteristics. In this section we will discuss how these changes can be measured and accounted for.

In principle one could do all the calibrations together with the measurements at the elevated temperature. This however means that parts have to be changed that are hot. Besides after each change one has to wait until the temperature is stable again. Therefore a different approach was taken. All calibration measurements were done once as a function of temperature so that the temperature coefficients of all parameters could be determined. These were then processed afterwards together with the measurement data. This procedure has the advantage that one can do the diode measurements at different temperatures directly after each other without removing the diode or any other part. This not only means a great saving in time but also improves the accuracy, especially that in the relative positions of curves taken at different temperatures.

The quantities that are affected by temperature changes can be listed as follows:
- the width and height of the diode mount change so that the guide wavelength and characteristic impedance change. These changes can be calculated knowing the thermal expansion coefficient of copper;
- the length of waveguide between the diode and the taper cooling changes which gives an extra phase shift. This has to be determined empirically since the temperature gradient along the taper is not exactly known;
- the impedances of the $\frac{1}{2} \lambda g$ and $\frac{5}{4} \lambda g$ short-circuit pieces as well as $Z_{SM}$ change. For the $\frac{5}{4} \lambda g$ piece it has to be calculated, the others can be measured;
- one expects also the package parameters to change. It has been found however that these changes are smaller than the measurement accuracy.

It has been found that all variations can be described as linear functions of temperature within the accuracy of the measurements.

IX-6. Measuring under pulsed bias

If one wants to separate the influences of bias current and temperature on the impedance it is necessary to measure under pulsed bias. In Sec. VIII-4 the importance of the current pulse rise time has already been discussed. The same considerations are valid here but the pulse rise time will be longer than in the I-V measurements because the construction of the bias filter produces a larger parallel capacitance (100 pF) than the mount used for the I-V measurements.

Besides there are other time constants to be considered. To begin with, there is the detection system. Being a heterodyne receiver it has a restricted i.f. bandwidth and this limits its response time. We use a 70 MHz i.f. amplifier with 35 MHz bandwidth, giving a rise time of about 10 nsec.

Another, less obvious, time constant is that of the bridge itself. The whole waveguide circuit can be looked upon as a resonator and the time it needs to settle to a stationary field distribution depends on its quality factor. This means the less damping the two attenuators in the reference and measuring branches give the longer time constant.
A better way to describe this is the following: when the diode current is switched the diode impedance changes suddenly and a wavefront is created that travels from the diode to the hybrid-T, is (partly) reflected there, travels back to the diode where it is reflected again and so on. Depending on the reflection coefficient of the hybrid port and the attenuation setting a number of round trips is necessary. The effect can be observed quite clearly in the r.f. signal from the detector port, see Fig. IX-9. Here the attenuator was set at 2.5 dB and the steady state is reached after two round trips or about 40 nsec.

Fig. IX-9. Transient behaviour of the bridge output signal after a step change of diode bias. Horizontal scale: 10 nsec/div.

Clearly it is not advisable to have the variable attenuator at a too low setting. Other reasons to avoid this are that the diode can oscillate when the damping is too low and that the attenuator's phase shift variation is higher at low settings.

A third time constant is introduced by the PIN-diode switch. This is not such an important one since it modulates the detector signal but does not influence the instant at which a null is obtained. The one we use has a rise time of 50 nsec.

Taking it all together it appears that the bias circuit is the main limitation. It has been found possible to restrict the overall response time to 0.2 μsec. at the highest currents, which keeps the temperature rise of the diode above the heat-sink temperature below 5°C. At lower currents the response time becomes longer but the dissipation is less so the temperature rise does not increase.
X. R.F. NOISE MEASUREMENTS

X-1. Theory

In Ch. VI the noise produced by Baritt diode has been represented by an equivalent noise voltage source and expressions have been derived to calculate the mean-squared noise voltage. To measure it a small-signal amplifier can be built using the diode, Fig. X-1. The diode is connected through a tuning reactance and an impedance transformer to a source and a load. Circulator, source and load are supposed to be matched to the transmission-lines.

When a noise power $P_n$ is fed into the input of the amplifier, at the output a noise power $P_{out}$ is obtained given by:

$$P_{out} = G P_{in} + P_{nd}$$  \((X-1)\)
where $G$ is the gain of the amplifier and $P_{nd}$ is the noise added by the diode, assumed to be uncorrelated with the input noise. Using Fig. X-1 one can readily express $G$ and $P_{nd}$ in terms of the diode impedance $Z_d$ and the equivalent noise voltage:

$$G = 1 - \frac{4RR_o}{(R+R_o)^2+X^2} \quad (X-2)$$

$$P_{nd} = \frac{V_n^2 R_o}{(R+R_o)^2+X^2} \quad (X-3)$$

where $R = \text{Re} (Z_d)$ and $X = X_t + \text{Im} (Z_d)$.

For a resistive attenuator a formula similar to (X-1) can be derived:

$$P_{out} = G_A P_{in} + (1-G_A)kT_A \Delta f \quad (X-4)$$

where $T_A$ is the temperature of the attenuator and $G_A$ its gain, i.e. the inverse of its attenuation. To come to this result it must be assumed that the attenuator contains only components in thermal equilibrium, e.g. no PIN-diodes. Otherwise $T_A$ must be replaced by an effective noise temperature.

![Schematic diagram of noise measuring set-up.](image)

To measure $P_{nd}$ the circuit of Fig. X-2 is used where the amplifier of Fig. X-1 is embedded between two attenuators of which the one at the input side is connected to a calibrated noise source and the other to a detector. All components are matched to the transmission-lines.
It will be handy in the following to make use of the noise measure [78] which for a negative-resistance diode is defined as [29]:

\[ M = \frac{\nu^2}{-4Re(Z_d)kT_0 \Delta f} \]  \hspace{1cm} (X-5)

where \( T_0 \) is a standard temperature, usually 290 K.

Now, when the noise source delivers a noise power \( kT \Delta f \) into a matched load we can write for the cascade of Fig. X-2:

\[ P_{det} = \left[ G_3 G_2 \left( G_1 kT_0 + (1-G_1)kT_1 + M(1-1/G_2)kT_0 \right) \right] \Delta f \]  \hspace{1cm} (X-6)

Suppose now that we move attenuator 1 to a value \( G_1' \). The noise input to the amplifier then is changed and its noise output too but not in the same proportion since the diode noise is not affected. To obtain the same noise power as before at the detector we have to move attenuator 3 to a reading \( G_3' \). By eliminating \( P_{det} \) we now can calculate \( M \):

\[ M = \frac{G_2}{G_2 - 1} \left( \frac{T_3}{G_2 T_0} - \frac{T_1}{T_0} - \frac{G_3 G_1 - G_4 G_3' G_1'}{G_3 - G_3'} \cdot \frac{T_5 - T_1}{T_0} \right) \]  \hspace{1cm} (X-7)

For completeness' sake it should be mentioned that, although this result is derived for a specific amplifier circuit, it is valid also when a general lossless impedance-transforming network is inserted between the diode and the circulator.

A problem can arise when the noise is detected with a heterodyne receiver. This one always detects the noise in two sidebands whose widths are equal to the bandwidth of the i.f. amplifier. Eq. (X-7) is valid in this form only when the intermediate frequency and the i.f. bandwidth are such that both sidebands fall within the flat part of the amplifier's gain curve (Fig. X-3a). Then \( G_2 \) is the peak gain of the amplifier. When one of the sidebands falls outside the amplifier's range (Fig. X-3b), or when the i.f. bandwidth is even larger than the amplifier bandwidth (Fig. X-3c), then in Eq. (X-6) \( G_2 \Delta f \) has to be replaced by an integral and in (X-7) \( G_2 \) has to be replaced by:
\[ G_2^* = \frac{1}{B} \int_B G_2 df \]  

(X-8)

where \( B \) is the frequency range covered by the two sidebands and \( G \) is the frequency-dependent gain of the amplifier. It is assumed here that \( M \) does not vary over the bandwidth considered. It is clear from (X-7) that the influence on the result is not great as long as \( G_2^* \) is large.

---

**Fig. X-3.** Possible bandwidth configurations in the noise measurement.
Another point that should be made concerns the diode resistance. In Eqs. (X-2,3,5) the diode impedance $Z_d$ appears which in previous chapters always has been understood to be the impedance of the Baritt diode proper, excluding the parasitic series resistance $R_s$ due to the substrate and the mounting wires. In the measurement of $M$, however, this resistance is present and reduces the effective negative resistance. Thus in these equations $Z_d$ must be understood to include $R_s$. Also in theoretical calculations of $M$ allowance has to be made for $R_s$ to make a comparison with experiments valid.

X-2. Experiment

To insure stability and avoid signal leakages the whole set-up, except the diode holder, is built from waveguide components. The diode holder is a coaxial oscillator circuit with three movable $\frac{1}{4}\lambda$ slugs. It is modified by placing an additional $\frac{1}{4}\lambda$ impedance transformer close to the diode which lowers the load impedance to a value close to the negative resistance of the diode. The additional tuning provided by the slugs is necessary since a wide range of negative resistances has to be accommodated.

The attenuators are both of the rotary-vane type. From (X-9) it can be seen that especially on the second attenuator high demands are put. Here the same one as used in the impedance bridge is employed.

The noise source is an argon-filled gas discharge tube which is mounted in a waveguide under an angle of 10 degrees with the guide axis. It was calibrated by comparison with a hot load following a method described in [76]. The excess-noise ratio was found to be $15.3 + 0.3$ dB, corresponding to a noise temperature $T_s$ of $10150$ K $\pm 7\%$.

The receiver is the same one as used in the impedance measurements. Since it has an i.f. of 70 MHz and 35 MHz i.f. bandwidth, the extreme points of its sidebands are 175 MHz apart. It is not always possible to tune the diode amplifier to have this bandwidth and reasonable amplification as well. In these cases Eq. (X-8) has to be applied. To assess the gain curve a sweep oscillator in the place of the noise
source and a video detector instead of the receiver are used. The
detector output is displayed on an oscilloscope.

Naturally it has been tried to adapt this measurement for pulsed-bias
operation. This has been found possible although at the cost of reduced
accuracy and increased circuit complexity.

The output noise signal now is pulse modulated and the detected signal
from the i.f. amplifier is measured by a box-car integrator, i.e. a
pulsed gate, synchronized with the bias pulse and followed by an
averager.

The amplifier shows gain only during short pulses and the oscilloscope
display now also consists of a series of pulses the tops of which
follow the gain curve. Since the duty-factor is very low (0.001) these
tops are drowned out on the display by the much higher intensity of
the base line. One therefore has to apply z-axis modulation to suppress
the beam outside the pulses.

Since the detected noise pulses are short their amplitude fluctuates
considerably and a fairly long averaging time is necessary to obtain
a stable output. To keep the averaging time within reasonable limits the
gate width of the box-car integrator cannot be made shorter than 1
μsec. During this time the diode temperature and with it the negative
resistance changes. The amplifier curve which is quite sensitive to
variations of the negative resistance then changes too and this limits
the accuracy.
XI. RESULTS AND CONCLUSIONS

XI-1. Introduction

In the previous chapters theoretical models and experimental techniques have been outlined that can be used to study the behaviour of Barritt-diodes. Rather than giving the results of each individual technique in the corresponding chapter it was preferred to bring them all together in a separate chapter. The reason for this is that the results are interrelated in many ways (e.g. the doping and width obtained from the r.f. impedance measurements are used to evaluate the I-V measurements) and a clear picture of a particular Barritt diode can only be obtained by considering all relevant data in their mutual connection.

In this chapter then the results of I-V, impedance and noise measurements, grouped by diode, are presented and whenever appropriate compared with theoretical calculations. Three diodes have been selected, each representing a different type: a p-n-p diode, an M-n-p diode operating above flat-band at all temperatures considered and another one operating partly above, partly below flat-band.

Impedance data of the first two diodes have been published before [79]. At that time the observed characteristics of the M-n-p diode were not fully understood because it was not recognized that these diodes are operating above flat-band. Thanks to the combined analysis of impedance data below punch-through and I-V characteristics this has now been established beyond doubt, as will be demonstrated in Sec. 3.

XI-2. P-n-p diode series F

The starting material for this series was n-type epitaxial silicon grown on a p+-substrate, oriented in the 1.1.1. direction. In the same batch one n-type substrate was included which made it possible to make single Schottky diodes for diagnostic purposes. The n-layer thickness and concentration were given by the manufacturer as 8.0 μm and 1.2 x 10^{21} m^{-3}. C-V measurements on the single Schottky diodes gave 7.3 μm and 1.1 x 10^{21} m^{-3}.
The p-n-p diodes were made by a boron diffusion from a doped-oxide source as described in Sec. VII-3. The p-n junction resulting from this diffusion was measured to be located 0.4 μm beneath the surface. Contact to the diffused p-layer was made by the platinum silicide process of Sec. VII-2, with the heat treatment at 550 °C.

Fig. XI-1. Impedance below punch-through of p-n-p diode F2.

F=7.03 GHz. T=24 °C.

Fig. XI-1 gives a plot of the impedance of a diode from this series at voltages below and partly above punch-through. As expected the graph shows a straight line with a sharp kink at punch-through, this in spite
of the fact that this diode was rather leaky and no distinct punch-through point could be discerned in the I-V characteristic. From this graph the n-layer width and doping are calculated by the method of Sec. VII-3 to be 7.1 µm and $1.6 \times 10^{21} \text{ m}^{-3}$. The first figure looks surprising at first sight. From the layer thickness measured on the companion single

![Graph](image)

**Fig. XI-2.** Current-voltage characteristics of diode F2.

*Temperature is parameter.*

--- measured.

--- calculated.

Schottky diodes and bearing in mind what has been said in Ch. VII about the reduction of the layer width by the diffusion one would expect a thickness around 6.5 µm. On the other hand the junctions formed are not abrupt but graded so that the depletion layers stretch out some distance
Fig. XI-3. Impedance above punch-through of diode F2.

Temperature is parameter. $f = 7.03$ GHz.

--- measured.

--- calculated, including 1.5 $\Omega$ series resistance.

$N_p = 1.6 \times 10^{21}$ m$^{-3}$, $L_d = 7.1 \times 10^{-6}$ m, $A = 3.0 \times 10^{-8}$ m$^2$. 
into the p-regions. In all experiments it is the total depletion layer width that counts, not the metallurgical width of the n-layer. The doping value is substantially higher than measured on the single Schottky diodes. An explanation for this has not been found. From the punch-through point one can also deduce that the parasitic series resistance is about 1.5 ohms.

Fig. XI-2 shows the I-V characteristic of the same diode at two temperatures. The result of a theoretical calculation along the lines of Ch. IV have been drawn as dotted lines in the figure. To obtain a reasonable match at 25°C the values of layer width and concentration had to be taken at 7.1 μm and 1.6 x 10^21 m^-3, which agree well with the values obtained from the impedance below punch-through. At 125°C there is a slight disagreement. However, the results are so sensitive to the values of doping and width that a decrease of either N₀ by 2% or L_d by 1% is sufficient to remove the discrepancy.

Fig. XI-3 gives the r.f. impedance measured at three temperatures and at a frequency of 7.03 GHz. Also the results of a simulation are shown using the model of Ch. V for operation below flat-band. A series resistance of 1.5 ohm has been added to the calculated resistance. The width and doping values are those deduced from the previous measurements.

Fig. XI-4 finally shows the noise measure of this diode at the same temperatures and frequency, together with calculated values. The same parameter values as in the impedance calculations were used, also adding 1.5 ohm to the resistance. The general form of the measured curves is reproduced quite well in the simulation, as well as the temperature dependence. However, the calculated noise measures are much lower than the measured ones. The calculated values are of the same magnitude as measured by others [80] or calculated by different methods [81,82] for cases that were very similar, so one suspects that our diodes contain an additional noise source. In [5] it is stated that carrier multiplication noise starts to be significant when the peak electric field rises above 150 kV/cm. From our d.c. calculations it turns out that the peak field varies from 170 to 200 kV/cm so it is possible that carrier multiplication is the cause of the excess noise.
In the noise calculations it is found that, when multiplication noise is neglected, thermal noise is the predominant one in these diodes. Shot noise is important only at the lowest currents. At higher currents it is smoothed out by the velocity modulation in the diffusion region.

![Graph](https://via.placeholder.com/150)

Fig. XI-4. Noise measure of diode F2.

Temperature is parameter. F=7.03 GHz.

--- measured.

--- --- calculated, including 1.5 Ω series resistance.

The properties of these diodes can be summarized as follows:

- the peak negative resistance is not strongly dependent on temperature but the peak shifts to higher currents at higher temperatures. This means that the negative resistance of a diode, set at a fixed current, still can be strongly temperature dependent.
the diode resistance is quite strongly dependent on current and
temperature. This has a consequence among others that an oscillator
using these diodes will have a strong temperature dependence of the
oscillation frequency unless special stabilizing measures are taken.
- the noise measures are too high to make these diodes suited for small-
signal amplifiers. On the other hand, they are still lower than those
of Gunn and Impatt diodes and since the small-signal noise measure is
an indication of the oscillator noise one may expect, Baritt diodes
are envisaged to make low-noise oscillators.

It is evident from the foregoing that the impedance features are
reproduced quite well by the theoretical model. For the noise this is
not so. A possible explanation is is offered by the circumstance that
multiplication noise was not included in the model. Analyzing the model
further one finds that the low-field region plays an important role.
By its transit delay it increases the negative resistance (cf. Secs.
II-3 and V-3). By the same mechanism it also increases the electronic
capacitance. The shift of the peak negative resistance to higher
currents at higher temperatures appears to be largely due to the tem-
perature dependence of the low-field mobility. When the latter is kept
at its room temperature value the impedance curves at higher temperatures
are very close to the room temperature curve.
XI-3. M-n-p diodes series G

The starting material for these diodes was a slice from the same batch as the one used for the F-series. A Schottky contact was made to the n-layer using the process of Sec. VII-2 with the sintering temperature of 550°C.

![Graph showing impedance below punch-through of M-n-p diode G19.](image)

Fig. XI-5. Impedance below punch-through of M-n-p diode G19. F=7.03 GHz, T= 23°C.

The impedance below punch-through is shown in Fig. XI-5. From this graph the layer width and concentration were deduced to be 7.0 μm and $1.2 \times 10^{21}$ m$^{-3}$, values that are in good agreement with those of the single Schottky-diodes. The series resistance is about 1.5 ohm.
Fig. XI-6 gives the I-V characteristics. Compared with those of the p-n-p diodes they show a less steep slope and a much greater variation with temperature. With the method of Sec. IV-5 the electric field at the contact is calculated from the I-V data using the doping and width values quoted above. It is found to be positive at all but the lowest currents for the three lowest temperatures which means the diode is operating above flat-band. The relationship between contact field and current is displayed in Fig.XI-7. The slope of these curves gives for the proportionality field $E_s$, defined by Eq.IV-7, a value around 0.9 kV/cm, nearly independent of temperature. Extrapolating the lines to zero field gives the saturation currents which should obey Eq. II-2 (with $V_m = 0$). In Fig. XI-8 $\ln(I_s/T^2)$ is plotted as a function of $1/T$, a so-called Arrhenius plot. It gives a straight line as expected but the values of $A^*$ and $\phi_h$ deduced from it are far from the theoretical ones.
$A^* = 16 \text{ Am}^{-2}\text{K}^{-2}$ and $\phi_h = 0.14 \text{ V}$. In Ch. VII it was shown already that this manufacturing process gives Schottky barriers that are far from ideal. This experience is confirmed here.

**Fig. XI-7.** Current as a function of the electric field at the injecting contact for diode G19. Temperature is parameter.

**Fig. XI-8.** Arrhenius plot for diode G19. $I_s$ is the extrapolated zero-field current of Fig. XI-7.
In Fig. XI-9 the r.f. impedance is given, measured at 7.03 GHz, and compared with calculations. The model of Ch. V for operation above flat-band was used with the parameters deduced from the previous measurements. The agreement is not perfect, especially at 75°C but the general features of the measured curves are reproduced.

![Graph showing r.f. impedance measurements and calculations.]

**Fig. XI-9.** Impedance above punch-through of diode G19.
Temperature is parameter. F=7.03 GHz.

--- measured.

--- calculated, including 1.5 Ω series resistance.

\[ N_D = 1.2 \times 10^{21} \text{ m}^{-3}, \quad L_d = 7.0 \times 10^{-6} \text{ m}, \quad A = 3.3 \times 10^{-8} \text{ m}^2. \]
In Fig. XI-10 the noise measures at the two highest temperatures are given. At the lowest temperature a noise measurement is not possible because there is no net negative resistance.

Calculated noise measures are also shown. The good agreement at 75°C is somewhat fortuitous in view of the impedance data. Nevertheless on the whole the agreement is better than in the previous case. The peak electric field in this diode assumes values between 160 and 170 kV/cm so that less multiplication noise is expected.

![Graph](image)

*Fig. XI-10. Noise measure of diode G19. Temperature is parameter. F=7.03 GHz.*

In the calculations it is found that now the shot noise makes a larger contribution to the total noise and the latter is therefore higher than in the p-n-p diode.

Looking at the general features of this diode we see some striking differences with the p-n-p type. They can be listed as follows: - the peak negative resistance is lower and more temperature dependent. At low temperatures it almost disappears, a fact that has already been noted by Snapp and Weissglas [20].
- The diode reactance varies much less with temperature and current.
- The noise measure on the whole is higher but it drops with increasing temperature.

All these phenomena can be explained by the fact that the d.c. field at the injecting contact $E_c$ is positive. First, $n_c$ increases with decreasing temperature and this reduces the negative resistance. This cannot be the only reason, however, since the decrease of $-R_d$ with $n_c$ is rather slow (cf. Fig. II-9). More important is the fact that at low temperatures $E_c$ is high (cf. Fig. XI-7) so that the low-field region is virtually absent. This not only is detrimental for the negative resistance but it also increases the noise since the shot noise is not smoothed out. At higher temperatures $E_c$ decreases and the low-field region grows in importance, increasing the negative $R$ and decreasing the noise. However, it never becomes as influential as in the p-n-p diode. This explains why the reactance variation with current and temperature is smaller.
XI-4. M-n-p diodes series K

Starting material for these diodes was n-on-p+ epitaxial silicon with a layer width of 6.0 μm and a donor concentration of $1.5 \times 10^{21} \text{ m}^{-3}$, according to the manufacturer. A Schottky contact was made to the n-layer with the platinum silicide process but now with the sintering done at 650°C.

\[ \text{Fig. XI-11. Impedance below punch-through of M-n-p diode K19.} \]

\[ F=7.03 \text{ GHz.} \quad T=24 \text{ °C.} \]

Fig. XI-11 gives the impedance below punch-through. This graph yields a layer width of 6.7 μm and a concentration of $1.6 \times 10^{21} \text{ m}^{-3}$. The series resistance is about 2 ohms.
The I-V characteristics are shown in Fig. XI-12 and the I-Ec relationship derived from them in Fig. XI-13. It turns out that these diodes are operating above flat-band at temperatures of 50°C and lower, and below flat-band at higher temperatures. The Arrhenius plot, Fig. XI-14, yields the values: $A^* = 2.8 \times 10^4 \text{ Am}^{-2} \text{K}^{-2}$ and $\phi_h = 0.24 \text{ V}$. These values are much closer to the theoretical ones than in the case of the G-series which agrees with the findings of Ch. VII. This once more demonstrates that there is a correlation between process parameters of Schottky barriers and their physical properties.

In Fig. XI-13 another interesting phenomenon can be noticed. At the temperature of 10°C the curve shows an oscillatory deviation from the straight line. This is not accidental. It is caused by the quantum-mechanical interference of hole wave functions reflected at the metal-
semiconductor interface and at the potential barrier which lies a small distance inside the semiconductor. A more detailed description of the phenomenon has been given elsewhere [78].

Fig. XI-13. Current as a function of electric field at the injecting contact for diode K19. Temperature is parameter.

Fig. XI-14. Arrhenius plot for diode K19. $I_0$ is the extrapolated zero-field current of Fig. XI-13.
Fig. XI-15. Impedance above punch-through of diode K19.
Temperature is parameter. F=7.03 GHz.
--- measured.
--- --- calculated, including 2 Ω series resistance.

\[ N_p = 1.6 \times 10^{21} \text{ m}^{-3}, \quad l_d = 6.7 \times 10^{-6} \text{ m}, \quad A = 3.0 \times 10^{-8} \text{ m}^2. \]
The difference between the operating regimes above and below flat-band also shows up quite distinctly in the impedance at 7.03 GHz, Fig. XI-15. At room temperature the negative resistance is practically absent and also the reactance variation is very small. Both resistance and reactance have their strongest variation at the lower currents where the diode is still below flat-band (the saturation current at this temperature is 7 mA). The constantness of the reactance is even more pronounced than in the G-diodes. This is because the contact field increases more strongly with current so that soon after passing the flat-band point the carrier velocity is close to saturation throughout the diode.

![Graph](image)

Fig. XI-16. Noise measure of diode K19.

Temperature is parameter. F=7.03 GHz.

--- measured.

--- calculated, including 2 Ω series resistance.

Theoretical curves are also given in Fig. XI-15. At 30°C the model for operation above flat-band is used and at the other temperatures the below-flat-band model. The agreement between theory and experiment is good enough to claim that the observed phenomena are explained by the models used. Note that the above-flat-band model gives large discrepancies when the saturation current is approached. This is because the
curvature of the energy bands was neglected in the theory of the Schottky effect and also because the diffusion region was left out in these calculations which should lead to significant errors when $E_c$ is close to zero.

Fig. XI-16 shows the noise behaviour of this diode. As in the case of the G-diodes it is not possible to measure a noise figure at room temperature. At the higher temperatures the behaviour is a mixture of those of the F and G diodes: the minimum noise measure has about the same values as in the F-diodes but it decreases with increasing temperature as in the G-diodes. The agreement between experiment and theory is good. The peak electric field has about the same range as in the G-diodes: 160-175 kV/cm.

Looking over these results we can say that these diodes are intermediate between the F and G diodes. The negative resistance at the higher temperatures behaves like that of the p-n-p diodes but the variation of the reactance with current and temperature is smaller. Also the noise is of the same magnitude as in the F-diodes but decreases with temperature as in the G-diodes. These effects can be explained by noting that, although the diode operates below flat-band, the contact field is closer to zero so that the low-field region is shorter than in a comparable p-n-p diode.
XI-5. Conclusions

The aim of the work reported in this thesis was to develop theoretical and experimental methods that could provide insight into the behaviour of Baritt diodes. Measurement set-ups were built that allow accurate determination of the microwave impedance and noise, and of the d.c. characteristics. The influences of diode temperature and bias current could be separated by performing all measurements under pulsed-bias conditions.

Along with the experimental work theories were developed with which the d.c. and small-signal a.c. properties can be calculated. Although the a.c. theory contains several approximations, a satisfactory agreement between theory and experiment could be obtained. In the course of the calculations it was found that the results are quite sensitive to the values of the width and doping of the central layer. Therefore much attention has been paid to experimental techniques that allow an accurate determination of these parameters. Especially the analysis of the r.f. impedance below punch-through proved useful for this purpose.

It may be stated that only by combining the results of different techniques applied to one diode a clear insight into its behaviour can be obtained. This has been demonstrated on the examples of one p-n-p and two M-n-p diodes, each representing a different type. It has also been shown that M-n-p diodes can operate in the regime above flat-band much more easily than assumed generally. This is due to the circumstance that, depending on the manufacturing process, the properties of the Schottky barriers can vary widely. Notably the saturation current can be much lower than predicted by theory. Useful negative resistances can be obtained above flat-band when the proportionality field $E_s$ is low as is the case in our G-diodes. Furthermore it appears that for both p-n-p and M-n-p diodes the a.c. properties of the injecting contacts can be derived with reasonable accuracy from their d.c. properties.

Finally, it seems that at peak electric fields above 170 kV/cm multiplication noise starts to make a significant contribution to the total noise.
As to the relative merits of p-n-p and M-n-p Baritt-diodes it can be said that p-n-p diodes seem to give higher and less temperature-dependent negative resistances and lower noise measures. On the other hand M-n-p diodes offer a lower reactance variation. M-n-p diodes are also easier to manufacture. Although this thesis is devoted exclusively to the small-signal characteristics it is interesting to make an observation on the large-signal behaviour: the output powers from all three types of diodes were in the same range, 5-10 mW. In fact, the best results were obtained from the G-diodes which show the smallest small-signal negative resistance. This leads one to suspect that in operation above flat-band the magnitude of the negative resistance decreases slower with signal amplitude than below flat-band, a matter that seems worth further investigation.

It is hoped that this work will provide a better understanding of Barritt diodes and that it will lead to a usable design theory.
REFERENCES


2. Müller, I., Elektronenschwingungen in Hochvakuum. Hochfrequenztechnik u. Elektroakustik 14, 156-167 (1933); ibid. 43, 195 (1934).


29. Haus, H.A., Statz, H. and Pucel, R.A., Noise Measure of Metal-
Semiconductor-Metal Schottky-Barrier Microwave Diodes. Electron.

Resistance Device Utilizing Injection from a Schottky-Barrier.

31. Vlaardingerbroek, M.T. and Van de Roer, T.G., On the theory of


34. Blötekjaer, K., High-Frequency Conductivity, Carrier Waves and
Acoustic Amplification in Drifted Semiconductor Plasmas. Ericsson

35. Jacoboni, C., Canali, C., Ottaviani, G. and Alberigi Quaranta, A.,


37. Canali, C., Jacobini, C., Ottaviani, G. and Alberigi Quaranta, A.,
27, 278-280 (1975).

38. Canali, C., Majni, G., Minder, R. and Ottaviani, G., Electron and
Hole Drift Velocity Measurements in Silicon and their Empirical
Relation to Electric Field and Temperature. I.E.E.E. Trans. ED-22,
1045-1047 (1975).

Ottaviani, G., Correction to above paper, I.E.E.E. Trans. ED-23,
1113 (1976).


40. Caughey, D.M. and Thomas, R.E., Carrier Mobilities in Silicon Empirically

-131-


47. De Cogan, D., to be published.


50. Sellberg, F., Private Communication.


57. Langevin, P., Comptes Rendus 146, 530 (1908).


70. Kwaspen, J.J.M. and van de Roer, T.G., to be published.
SUMMARY

In this thesis the study of the small-signal impedance and noise properties of Baritt diodes is described. Along with it also the d.c. behaviour is taken into consideration. The work contains a theoretical part and an experimental one. The experimental part includes measuring the d.c. I-V and C-V characteristics as well as the microwave impedance and noise measure. The theoretical part is concerned with the development of models that can explain the observed microwave properties. The results of the d.c. measurements are used to obtain the values of diode parameters that play a role in the theory. A program to manufacture Baritt diodes was set up in cooperation with the Semiconductor Laboratory of the Department of Electrical Engineering at Eindhoven University of Technology.

Modeling a Baritt diode is complicated by the circumstance that the electric field rises from a low value at the injecting contact to a high value at the collecting contact. This means that there is a low-field region where diffusion is necessary to transport the carriers and a high-field region where the drift velocity is nearly saturated. In operation one can distinguish two regimes: the non-conducting regime when the voltage is below the punch-through voltage and the current-carrying regime above punch-through (PT). In the latter regime the diode is usually operated. In metal n-p diodes (with the metal as the injecting contact) it can be further divided into two regimes: below and above flatband where flatband (FB) denotes the bias point at which the electric field at the injecting contact is zero. Below FB a potential barrier exists at some distance from the injecting contact, above FB this barrier is at the contact.

The first models published simplified matters considerably by assuming that the carrier drift velocity is saturated throughout the diode. In spite of this simplification these models were able to give a qualitative explanation of many of the characteristics of the diodes. Vlaardingerbroek and the author gave an extended model in which the importance of the low-field region was stressed. A survey of these models is given in chapter II. The last-mentioned model was further
extended and refined by the author and is presented in its final form in chapters V and VI.

Chapter III is devoted to a discussion of the equations that govern carrier transport in a semiconductor. Its aim is to make clear what approximations are included in the equations as they will be used in the following three chapters. Of these chapter IV gives the d.c. analysis. The low-field region and the high-field region are treated separately. In the high-field region an analytical solution could be found owing to the fact that diffusion here is of minor importance. In the low-field region diffusion is predominant and no analytical solution was obtained. Here the equations are solved by a numerical technique. It turns out that for a p-n-p diode with abrupt p-n junctions the solution depends only very weakly on the doping of the p-regions. On the other hand it strongly depends on the values of doping and width of the n-region. This gives a method to calculate these parameters by comparison of measured and calculated I-V characteristics. In M-n-p diodes this is not possible since the parameters of the M-n contact have a great influence on the results. On the other hand, in this case it is possible, when the n-layer width and doping are known, to determine the contact parameters from the I-V measurements.

Chapters V and VI discuss, as already mentioned, the extended a.c. model. Contrarily to chapter IV in these chapters several approximations are introduced. Three regions are distinguished in the diode: the contact region including the injecting contact and the potential barrier, the diffusion region, from the potential barrier to a point where the influence of diffusion has become negligibly small, and the drift region comprising the high-field part of the diode. Approximations are introduced to make analytic solutions possible for the contact and diffusion regions. Chapter V then gives the calculation of the a.c. impedance and in chapter VI the noise is calculated. To do this two noise sources are taken into consideration: shot noise in the injected current and thermal noise throughout the diode. A third noise source, carrier multiplication noise, is not considered here.
The techniques that have been applied to manufacture p-n-p and M-n-p Barritt diodes are outlined in chapter VII. For the M-n contacts platinum silicide is used as the metal. Measurements on single Schottky diodes (M-n-n$^+$) show that the parameters of the manufacturing process have a great influence on the properties of the junctions formed.

The following chapters are devoted to the measurements. Chapter VIII describes diagnostic techniques that are used to obtain information about diode parameters. The use of I-V characteristics for this purpose has already been mentioned. It has been found that from the r.f. impedance below punch-through the n-layer width and doping can be deduced. To avoid self-heating effects one can bias the diode with short current pulses and measure during these pulses. In this chapter it is discussed what demands have to be put on the rise time and duty factor of these pulses.

In chapters IX and X measurement set-ups are described to measure the r.f. impedance and noise. Both have been made suitable for measurement during short bias pulses. Besides, the diode mountings can be heated so that measurements at elevated temperatures are possible. All measurements have been done at a frequency of 7 GHz.

The results of all measurements and their discussion are relegated to the last chapter. Also the comparison with the theoretical model is done here. This has been done because it is necessary to combine the results of various experiments in order to obtain a good picture of a particular diode. Three diodes are described: one of p-n-p type and two of M-n-p type of which one is operating nearly always above FB and the other one partly below, partly above FB. First the n-layer width and doping are determined from the r.f. impedance below PT. For the p-n-p diode these results could be checked by the I-V measurements. Excellent agreement was obtained between the two methods. For the M-n-p diodes the I-V measurements are used to determine the contact parameters. These data are then used in the theoretical models for the r.f. impedance above PT and the noise.
For all three diodes good agreement has been found between theory and experiment as far as the impedance is concerned. In the case of the p-n-p diode the noise data show less good agreement: the measured noise is substantially higher than the calculated one. This is ascribed to the fact that in these diodes the peak electric field rises high enough to produce a significant amount of multiplication noise. An analysis of the data shows furthermore that the low-field region plays a great role in the p-n-p diode and somewhat less in the M-n-p diodes, especially when operating above $FB$. As a consequence the p-n-p diode shows a larger peak negative resistance and a lower noise measure. On the other hand the M-n-p diodes show a smaller variation of the diodes reactance with current and temperature.
SAMENVATTING

De Baritt-diode is een drielaags halfgeleiderstructuur waarbij de buitenste twee lagen als contacten fungeren. In dit proefschrift worden Baritt-diodes van de samenstelling p-n-p of metaal-n-p beschreven waarbij de halfgeleider silicium is. De diode is isolerend beneden een bepaalde spanning, genoemd de doorslag ("punch-through")-spanning. Bij hogere spanningen is stroomgeleiding mogelijk doordat gaten vanuit het positieve kontakt in het n-gebied geïnjecteerd worden. Onder deze omstandigheden kan de impedantie voor mikrogolf frequenties een negatief reéel deel hebben. Als mogelijke toepassingen kunnen genoemd worden: locale oscillator in ontvangers voor satelliet-TV en kleine Doppler-radars voor detektie van bewegende objecten.


De elektrische veldsterkte in Baritt-diodes loopt op van een lage waarde aan het positieve kontakt tot een hoge waarde aan het negatieve kontakt. Dit maakt de modelvorming niet eenvoudig: in het gebied van laag veld geschiedt het transport van de ladingsdragers voornamelijk door diffusie terwijl in het hoog-veldgebied de driftsnelheid onder invloed van het veld tot zijn verzadigingswaarde nadert. Men kan bovendien nog twee regimes onderscheiden: beneden en boven flatband (FB). Beneden FB is het elektrisch veld aan het injecterende (positieve) kontakt negatief en oeënt een remmende werking op de ladingsdragers uit, zodat diffusie voor het transport moet zorgen. Boven FB is het veld aan dit kontakt positief en speelt diffusie een minder grote rol. P-n-p diodes werken altijd beneden FB terwijl M-n-p diodes in beide regimes
kunnen werken. Aan de konsequenties hiervan wordt in dit proefschrift veel aandacht besteed.

De eerste modellen die voor Baritt-diodes gepubliceerd werden gingen uit van de vereenvoudigende aannemer dat de driftsnelheid in de hele diode verzadigd, dus constante, dus konstant is. Ondanks deze vereenvoudiging konden ze toch al aan kwalitatief inzicht in het gedrag van de diodes geven. Door Vlaardingerbroek en de auteur werd gewezen op het belang van het laagveldgebied. Een overzicht van deze modellen wordt gegeven in hoofdstuk II. Het laatstgenoemde model werd door de auteur uitgebreid en verfijnd en wordt in zijn uiteindelijke vorm beschreven in de hoofdstukken V en VI. Hoofdstuk III is gewijd aan een discussie van de vergelijkingen die het transport van ladingstragers in een halfgeleider beschrijven. Uiteindelijk leiden deze tot een tweede-orde, niet-lineaire partiële differentiaalvergelijking voor het elektrische veld, die de basis vormt van de analyse in de volgende drie hoofdstukken.

In hoofdstuk IV wordt de gelijkstroomtheorie beschreven. Het gebied van laag elektrisch veld en het hoog-veldgebied worden apart beschouwd. Voor het laatste gebied kon een analytische oplossing gevonden worden dank zij het feit dat diffusie hier van geringe invloed is. In het laagveldgebied werd een numerieke oplossingsmethode gebruikt. Het blijkt dat voor een p-n-p diode met abrupte p-n overgangen de oplossing weinig afhankelijk is van de dotering van de p-gebieden. Dit levert een middel om door aanpassing van berekende aan gemeten stroom-spannings- karakteristieken de dotering en dikte van de n-laag te bepalen. Bij een M-n-p diode is dit niet mogelijk daar de parameters van het metaal-halfgeleidercontact een te grote invloed hebben. Anderzijds is het wel mogelijk, als dotering en laagdikte bekend zijn, uit de I-V karakteristieken de contactparameters te bepalen.

Hoofdstuk V en VI zijn gewijd aan respectievelijk de wisselstroomimpedantie en de ruis. In diodes die beneden FB werken worden drie gebieden onderscheiden: het kontaktgebied, van het injecterende contact tot het punt waar het elektrisch veld door nul gaat; het diffusiegebied, vanaf dit laatste punt tot een punt waar de invloed van de diffusie verwaarloosbaar klein is geworden en tenslotte het driftgebied dat de rest van
De diode beslaat. Boven FB worden twee gebieden onderscheiden: het
injekterende kontakt en het driftgebied dat dus de hele n-laag beslaat.
Om de ruis te berekenen wordt aangenomen dat er twee ruisbronnen onderschei­den kunnen worden: hagelruis in de geïnjekteerde stroom en ther­
mische ruis verdeeld over de hele diode. Een derde bron van ruis, de
vermenigvuldiging van ladingsdragers bij hoge veldsterkten, wordt niet
in beschouwing genomen.

De toegepaste fabrikagetechnieken voor p-n-p en M-n-p diodes worden be­
schreven in hoofdstuk VII. Het toegepaste metaal is platina-silicidle.
Door metingen aan enkelvoudige Schottky-diodes (M-n-n+ struktuer) is
gevonden dat de parameters van het fabrikageproces grote invloed geb­
ben op de eigenschappen van de gevormde metaal-halfgeleiderovergangen.

In hoofdstuk VIII worden metingen van de stroom-spanningskarakteristiek
nen van de differentiele capaciteit en de hoogfrekwent-impedantie
als functie van de spanning beneden doorslag besproken. Aangetoond wordt
dat uit de laatste de dikte en de dotering van de n-laag afgeleid kunnen
worden. Doordat de dichtheid van de gedissipeerde energie bij stroom­
voerende Baritt-diodes zeer hoog is neemt de temperatuur toe met de
stroomsterkte. Aangezien de temperatuur het gedrag van de diode sterk
beïnvloedt levert dit een complicatie op die vermeden kan worden door
alle metingen uit te voeren gedurende korte stroompulsen. Hier wordt
afgeleid dat hiertoe de pulsdur en kleiner moet zijn dan 1 μsec. en de
herhalingstijd groter dan 1 msec.

In de hoofdstukken IX en X worden meetopstellingen beschreven waarmee
respektievelijk de hoogfrekwent impedantie en ruis kunnen worden ge­
meten. Beide zijn geschikt voor metingen gedurende korte stroompulsen. Ook kan de diode opgewarmd worden zodat de invloed van de temperatuur
bestudeerd kan worden. Alle metingen zijn gedaan bij een frekwentie
van 7 GHz.

De beschijving en discussie van alle meetresultaten zijn samengebracht
in hoofdstuk XI. Ook worden hier de metingen vergeleken met de uit­
komsten van de theoretische modellen. Dit is gedaan omdat alleen een
studie van de resultaten van verschillende metingen aan één diode in
hun samenhang een goed inzicht kan geven in het gedrag van deze diode. Drie diodes worden beschreven: een p-n-p diode en twee van het M-n-p type waarvan er één vrijwel uitsluitend boven FB opereert en de andere gedeeltelijk boven, gedeeltelijk beneden FB. Uit het verloop van de hoogfrequente impedantie beneden doorslag worden eerst de dikte en de dotering van de n-laag bepaald. Voor de p-n-p diode worden deze ook bepaald uit de I-V karakteristiek. Beide metingen geven goede overeenstemming wat een nuttige controle op de nauwkeurigheid van de eerste methode is. Voor de M-n-p diodes worden dan de parameters van het metaal-halfgeleiderkontakt bepaald uit de I-V-metingen. Deze gegevens worden vervolgens gebruikt in de theoretische berekeningen van de impedantie in het stroomvoerende gebied (boven doorslag) en de ruis.

Voor alle drie de diodes wordt goede overeenstemming gevonden tussen de gemeten en de berekende impedanties. Voor de ruis is de overeenstemming minder goed bij de p-n-p diode. Dit wordt toegeschreven aan het feit dat de elektrische veldsterkte in deze diode hoger oploopt dan in de andere twee, zodat vermenigvuldigingsruis hier een grotere rol kan spelen. Een nadere beschouwing van de theoretische modellen leert dat het gebied van laag veld een grote rol speelt in de p-n-p diode en een vrij bescheiden rol in de M-n-p diodes, in het bijzonder boven FB. Positieve effekten van het laagveldgebied zijn een vergroting van de negatieve weerstand en een vermindering van de hagelruis. Een negatief effect is de toename van de stroom- en temperatuurafhankelijkheid van de diodereaktantie.
Levensbericht

29 juni 1937 geboren te Brunssum

1949-1953 MULO A te Brunssum, MULO B te Treebeek
1953-1955 HTS te Heerlen, afd. Chemische Techniek
1955-1958 idem, afd. Elektrotechniek
1958-1959 Nat. Lab. N.V. Philips, onderwerp snelle pulsschakelingen
Afstudeerhoogleraar prof. dr. A.A.Th.M. van Trier.
Onderwerp: propagatie in een golfpijp gedeeltelijk gevuld met anisotroop ferriet.
Studie van ladingstransport in halfgeleiders.
Experimenteel werk aan lawinedoorslag in Indium-antimonide.
1968-heden werkzaam in de Vakgroep Theoretische Elektrotechniek.
STELLINGEN

behorende bij het Proefschrift van Th.G. van de Roer

1
Voor uniform gedoteerde Baritt- en Impatt-diodes levert een meting van de
mikrogolfimpedantie als functie van de spanning beneden door slag een nauw-
keurige methode om de dikte en de dotering van de depletielaag te bepalen.

_Dit proefschrift, hs. VIII en XI._

2
Voor uniform gedoteerde Baritt-diodes met een p-n-p structuur levert de ver-
gelijking van gemeten en berekende I-V-karakteristieken een bruikbare methode
om dikte en dotering van de depletielaag te bepalen.

_Dit proefschrift, hs. IV, VIII en XI._

3
Baritt-diodes met een M-n-p structuur kunnen in het werkgebied boven flat-
band een bruikbare negatieve weerstand vertonen mits de karakteristieken van
het metaal-halfgeleidercontact voldoende afwijken van de ideale.

_Dit proefschrift, hs. XI._

4
De specifieke voordelen van p-n-p en M-n-p Baritt-diodes zijn van zodanige
aard dat het van de toepassing afhangt aan welk type men de voorkeur zal geven.

_Dit proefschrift, hs. XI._

5
Het bepalen van de warmtestroomweerstand van Baritt-diodes door vergelijking
van I-V-karakteristieken opgenomen met gepulseerde, resp. continue, stroom
levert alleen voor M-n-p diodes betrouwbare resultaten.

_S. Ahmad, J. Freyer, Electron. Lett. 12, 527-528 (1976)._

6
De door Sze et al. afgeleide uitdrukking voor de spanningsafhankelijkheid van
de capaciteit van een M-S-M diode bij kleine spanningen is onjuist.

De empirische formule die Canali et al. vinden voor de driftsnelheid van
gaten in sillicium als functie van het elektrisch veld geeft voor de ver-
zadigingsnelheid een geextrapoleerde waarde die vermoedelijk te laag is
en te sterk afhangt van de temperatuur.

C. Canali, G. Majni, R. Minder en G. Ottaviani,

De "thermodynamische paradox" die wordt gevonden bij berekeningen aan
een gedeeltelijk met anisotroop ferriet gevulde golfpijp kan verklaard
worden uit de omstandigheid dat bij de gebruikelijke aanname van per-
fekt geleidende golfpijp-wanden het elektromagnetisch veld aan het
grens- vlak van ferriet en lucht een integreerbare singulariteit vertoont.


De weergave die T.S. Kuhn geeft van de geschiedenis van de fysische
optica is onvolledig en doet vermoeden dat hij op dit punt gepoogd
heeft de feiten aan zijn theorie aan te passen.

T. S. Kuhn, The Structure of Scientific Revolutions,

Het heropenen van de Limburgse kolenmijnen is bij de huidige stand van
de stofbestrijdingstechniek uit het oogpunt van de volksgezondheid
onaanvaardbaar.

Het verdient aanbeveling voor jonge pas afgestudeerde academici een
aparte doctorstitel te creeren. De eisen hiervoor zouden vergelijkbaar
moeten zijn met die voor een Amerikaanse of Britse Ph.D. zodat het
promotiewerk in drie jaar afgerond zou kunnen worden.