

ABSTRACT

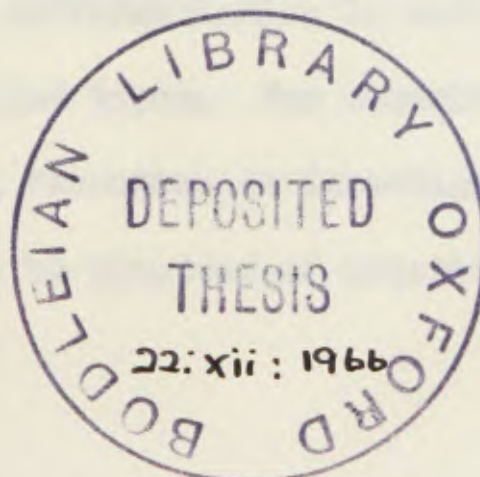
of a thesis entitled
Properties of Semiconductors
at Low Temperatures

by

C. O. Newton.

Oriel College,

September 1966.



ABSTRACT

Consideration has been made of bulk differential negative resistance as a low temperature property of semiconductors. The following topics have been investigated:

- (a) criteria for the existence of a differential negative resistance;
- (b) the application of irreversible thermodynamics to spontaneous instability;
- (c) double carrier injection as a cause of negative resistance in bulk material;
- (d) impact ionisation as a cause of negative resistance of the current controlled type;
- (e) methods of procuring a voltage controlled negative resistance, together with the properties of subsequent travelling domains;
- (f) thermally induced negative resistance.

(a) General criteria for bulk differential negative resistance have been deduced; differentiation is made between current controlled and voltage controlled types. The occurrence of instability against filament and domain formation is investigated (1); domains and filaments are shown to be symptoms of negative resistance rather than causes.

(b) The application of irreversible thermodynamics to non-linear processes and in particular to negative resistance has been examined. It has been shown that maximum (2) and minimum (3) entropy production principles as presently employed are not valid for non-linear processes. An alteration which corrects this has been made and justified from statistical mechanics. While the new stationary entropy production principle provides no solution to the domain or filament situation to first order, it is shown that it does determine the solution to second order. A physical justification has been suggested in terms of the conservation laws.

(c) As a first example of bulk negative resistance, double carrier injection (4) has been studied. As the current flow in a diode is increased, the injected carriers penetrate further into the crystal; if the crystal contains minority carrier traps, it has been noted that a trap saturation front can be pushed through the crystal as the current is raised. Since the carrier density for two carrier flow is higher, negative resistance may be developed in the saturation front region. In a short specimen this may dominate the characteristic over a small range. The two carrier equations have been written in a way which is particularly suitable for computation; in this way parameters for the recombination functions have been found which lead to a bulk negative resistance. Experimentally this negative resistance has been observed in overcompensated zinc-doped germanium; it has been studied as a function of specimen

length, probe position, compensation, temperature and magnetic field. Examination of the Hall effect for double injection has been found to demand a special interpretation, due to carrier stagnation at a blocking electrode, resulting in Hall voltage inversion.

(d) A study of impact ionization as a cause of a current controlled negative resistance has been made. From the application of criteria from section (a), it has been found that present explanations (5), (6) are rather unsatisfactory. The analysis is confined to an examination of what conditions can and what conditions cannot give rise to a negative resistance, employing an ionization rate equation. A distinction is made between spatially localized negative resistance processes leading to filament formation and non-localized ones.

(e) Two methods of obtaining a voltage controlled bulk negative resistance have been treated in the literature, a band structure effect (7) and a carrier energy dependent capture cross-section (8); a third method has been considered, a carrier energy dependent scattering rate leading to distortion of the distribution function, treated in outline for a number of physical processes. It is shown that for all these processes, the existence of space-charge leads to the phenomenon of travelling domains, observed as the Gunn Effect (9). Consideration has been made of the carrier flow equations; it has been found that the discrete domain solution of P. N. Butcher (10) is a limiting case of a set of continuous waveform solutions, all of

higher frequency. Thus the suppression of the discrete domains would lead to higher frequency operation.

(f) Thermal heating has been considered as a cause of negative resistance, for hot filaments and domains in an otherwise isothermal lattice, and also for specimen heating due to film boiling at the boundaries (11). Experimental results for which the above analysis was undertaken are explained under this heading.

References (Abstract)

- (1) Ridley B.K. 1963 Proc. Phys. Soc. 82, 954.
- (2) Ziman J.M. 1956 Canad. J. Phys. 34, 1256.
- (3) de Groot S.R. 1959 Thermodynamics of Irreversible Processes p.196, North Holland Publishing Co., Amsterdam.
- (4) Lampert M.A. 1964 Rep. on Prog. in Phys. p.329, Inst. Phys., Phys. Soc., London.
- (5) Rediker R.H., McWhorter A.L. 1960 Proc. Intern. Conf. Semicond. Phys. p.154, Prague.
- (6) Kurosawa T. 1965 J. Phys. Soc. Jap. 20, 1405.
- (7) Ridley B.K., Watkins T.B. 1961 Proc. Phys. Soc. 78, 293.
- (8) Ridley B.K., Watkins T.B. 1961 Proc. Phys. Soc. 78, 710.
- (9) Gunn J.B. 1964 Plasma Phys. in Solids p.199, Dunod, Paris.
- (10) Butcher P.N. 1965 Phys. Lett. 19, 546.
- (11) Eastman P.C., Detars W.R. 1963 Canad. J. Phys. 41, 161.

PROPERTIES OF SEMICONDUCTORS
AT LOW TEMPERATURES

A Thesis

submitted for the degree of Doctor of Philosophy
in the University of Oxford

by

C. O. Newton.

Oriel College,
September 1966.

Table of Contents

	page
Preface.	1
Chapter 1 Introduction.	1
Chapter 2 Irreversible Thermodynamics.	12
Chapter 3 Experimental Methods.	60
Chapter 4 Double Injection.	71
Chapter 5 Impact Ionisation.	143
Chapter 6 Voltage Controlled Negative Resistance.	176
Chapter 7 Effects Induced by Specimen Heating.	208
Chapter 8 Conclusion.	224
Appendix.	Z1
References.	R1
Acknowledgements.	u

List of Sections

page

Chapter 1 Introduction.

1.1	Aims.	1
1.2	Negative Resistance.	1
1.2.1	Negative Resistance Form.	1
1.2.2	Negative Resistance Criteria.	3
1.2.3	Instability.	5
1.3	Survey of Impurity Levels.	7
1.3.1	General Levels in Germanium.	7
1.3.2	Zinc Levels in Germanium.	9

Chapter 2 Irreversible Thermodynamics.

2.1	Introduction.	12
2.2	Applicability to Negative Resistance.	13
2.3	Condition for the Steady State.	14
2.4	Recirculatory Illustration.	15
2.5	Derivation of the Minimum Entropy Production Principle.	16
2.6	Application of the Minimum Entropy Production Principle.	20
2.7	Derivation of the Maximum Entropy Production Principle.	22
2.8	Application to Electrical Networks.	27
2.9	The Search for a New Entropy Production Principle.	32
2.10	Justification of the New Entropy Production Expression.	34
2.11	The Physical Basis of the Entropy Alteration.	39
2.12	Employment of the Entropy Alteration.	40
2.13	Appraisal of the Entropy Production Calculation.	42
2.14	Applicability of Entropy Production Principles.	44
2.15	The Relationship between Maximum and Minimum Principles by Analogy.	46
2.16	Topological Representation of Entropy Production.	47
2.17	Application to Filament Formation.	50
2.18	Application to Domain Formation.	55
2.19	Aims and Results.	57

Chapter 3 Experimental Methods.

3.1	Crystal Preparation.	60
3.2	Contact Mounting.	61
3.3	Contact Wiring.	62
3.4	General Purpose Cryostat.	63
3.5	D.C. Systems.	64
3.6	Pulsed Systems.	65
3.6.1	Pulsed Systems for Radiation Amplification.	65
3.6.2	Pulsed Systems for Rise and Decay Time Observation.	65
3.7	Radiation Generation - Detection System.	66
3.8	Cryostat for Spectral Analysis of Recombination Radiation.	67
3.9	Magnetic Fields.	67
3.10	Crystal Materials.	68
3.11	Specimen Classification.	69
3.12	The Advantages in the Use of D.C. Techniques.	70

Chapter 4 Double Injection.

4.1	Theory.	71
4.1.1	Introduction to Injection.	71
4.1.2	Qualitative Approach to Double Injection.	74
4.1.3	Negative Resistance through Double Injection.	76
4.1.4	Equations of Carrier Flow.	78
4.1.5	The Poisson Equation.	79
4.1.6	The New Equation.	80
4.1.7	Development of the Equation.	84
4.1.8	The First Boundary Condition.	84
4.1.9	The Carrier Density at the Boundary.	87
4.1.10	Recombination and Trapping.	88
4.1.10.1	Shockley - Read Theory.	88
4.1.10.2	Hole Trap Recombination.	90
4.1.10.3	Fixed Hole Lifetime.	91
4.1.10.4	Saturated Trap Recombination.	92
4.1.10.5	Recombination Functions for Computation.	93
4.1.11	Objections to Planar Flow.	94
4.1.12	Magnetic Field Effects.	97
4.1.13	The Hall Coefficient.	98

4.2	Computation.	104
4.2.1	The Basis for Computation.	104
4.2.2	Methods of Computation.	106
4.2.3	Predictor - Corrector Formula.	107
4.2.4	Programmes Employed.	108
4.2.5	Results for PCBR.	108
4.2.6	Results for PCSTF.	110
4.2.7	Summary of Computation.	112
4.3	Experimental Results.	113
4.3.1	Introduction.	113
4.3.2	Pattern of Results.	113
4.3.3	The Results.	115
4.3.3.1	The Shape of the Current-Voltage Characteristic.	115
4.3.3.2	The Characteristic for Long Specimens.	116
4.3.3.3	The Change in Characteristic with Specimen Length.	118
4.3.3.4	Section Voltages.	119
4.3.3.5	The Local Potential.	120
4.3.3.6	The Hall Voltage as a Function of Current and Position.	122

4.3.3.7	The Hall Voltage as a Function of Compensation.	125
4.3.3.8	The Hall Voltage as a Function of Temperature.	126
4.3.3.9	The Hall Voltage as a Function of Magnetic Field.	128
4.3.3.10	The Effect of a Magnetic Field on the Voltage Distribution.	130
4.3.4	Discussion of Results.	131
4.3.4.1	Aims.	131
4.3.4.2	Measurements.	131
4.3.4.3	The Sample of Information Reproduced.	133
4.3.4.4	The Consistency between the Data and the Hypotheses.	134
4.3.4.4.1	Self-consistency of the Data.	134
4.3.4.4.2	Extraneous Effects.	135
4.3.4.4.3	Consistency of the Hypotheses.	136
4.3.4.5	Inconsistency of Alternative Hypotheses.	136
4.3.4.6	Experimental Results in the Literature.	137
4.3.4.7	Conclusion to the Discussion of Experimental Results.	140
4.4	Conclusions on Double Injection.	142

Chapter 5 Impact Ionization.

5.1	Introduction.	143
5.2	Two Carrier Impact Ionization.	144
5.3	Single Carrier Impact Ionization.	145
5.3.1	Special Features.	145
5.3.2	Domain and Filament Formation.	146
5.3.3	The Ionization Rate Equation.	147
5.3.4	Classification of Physical Influences.	150
5.3.5.1	Case 1. Normal Ionization.	152
5.3.5.2	Case 2. A carrier Density Dependent Ionization Rate.	155
5.3.5.3	Case 3. A Current Density Dependent Mobility.	158
5.3.6	Variation in the Number of Carriers.	162
5.3.7	Critique of an Analysis Provided by T. Kurosawa.	163
5.3.8	Hot Lattice Filaments.	167
5.3.9	Filament Flow in an Isothermal lattice.	169
5.3.10	Application to Semiconductor Physics.	171
5.3.11	Device Potential in Impact Ionization.	173

6.1	Introduction.	176
6.2	Crystal Features Leading to Voltage Controlled Negative Resistance.	177
6.2.1	Band - Structure Effects.	178
6.2.2	Effects due to an Energy Dependent Capture Cross- Section.	179
6.2.3	Distribution Function Effects.	180
6.2.3.1	The Negative Resistance Criterion.	180
6.2.3.2	The Effects of Electric Field and of Scattering.	181
6.2.3.3	Dominant Inelastic Scattering with Threshold.	183
6.2.3.4	Non-dominant Inelastic Scattering with Threshold.	184
6.2.3.5	Perfect Elastic Scattering with Threshold.	186
6.2.3.6	Elastic Scattering with Threshold and Persistence of Velocity.	186
6.2.3.7	Methods of Producing a Mobility Variation.	187
6.2.3.7.1	Anomalies in Normal Scattering.	188
6.2.3.7.2	Inter-valley Exchange.	189
6.2.3.7.3	Acoustic Phonon Production.	190
6.2.3.7.4	Optical Phonon Production.	191
6.2.3.7.5	Plasmon Production.	191

6.2.3.8	Critique of Papers by T. Musha and F. Yoshida.	192
6.2.3.9	Summary.	194
6.3	Domain Formation.	195
6.3.1	Introduction.	195
6.3.2	Equations of Carrier Flow.	196
6.3.3	Extended Use of Butcher's Equal Areas Rule.	198
6.3.4	Small Signal Solutions.	199
6.3.4.1	Equal Areas.	199
6.3.4.2	Ridley's Sinusoidal Solution.	202
6.3.5	Hall Effect Techniques.	203
6.4	Parallels between Current and Voltage Controlled Negative Resistance.	205
6.4.1	Intrinsic Instability.	205
6.4.2	Filament and Domain Structure.	207

Chapter 7

Effects Induced by Specimen Heating.

7.1	Introduction.	208
7.2	Theory.	209
7.3	Experimental Results.	211
7.4	Elimination of Alternative Explanations.	212
7.5	Evidence for Film Boiling.	215
7.5.1	The Power Dissipation Formula.	216
7.5.2	Dependence on Material.	217
7.5.3	Specimen Disintegration.	217
7.5.4	Fluctuations.	218
7.5.5	Pulsed Breakdown.	219
7.5.6	Radiation.	220
7.6	Conclusion.	223

Chapter 8

Conclusion.

224

List of Figures

Figure	Title	Facing page
1.1	Negative Resistance Types.	2
2.1	The Effect of Domains on the i - V Characteristic.	13
2.2	Recirculatory System.	16
2.3	Energy Contours for Two Elastic Strings.	46
2.4	Filament Diagrams.	50
2.5	Filament Stability.	55
2.6	Domain Diagrams.	56
2.7	Domain Stability.	56
3.1	The Furnace.	61
3.2	The Hot Gas Jet System.	62
3.3	D.C. System.	64
3.4	D.C. Circuit.	65
3.5	Pulse Circuit for Radiation Observation.	65
3.6	Pulse Circuit for Oscilloscope.	65
3.7	Generation - Detection System.	66
3.8	Magnetic Field Pulser.	67
3.9	Table of Specimens.	69
3.10	Specimen Sizes.	69
3.11	Contact Sizes.	70

4.1.1	Barrier Structure.	71
4.1.2	Crystal for Double Injection.	75
4.1.3	Radial Diagrams.	95
4.1.4	Hall Voltage Inversion Curves.	103
4.2.1	Variation of Injected Carrier Density across Crystal for PCSR.	109
4.2.2	Hole Densities across Crystal for Different Numbers of Steps.	109
4.2.3	Current - Voltage Characteristic for PCSR.	109
4.2.4	Injection Hall Voltage for PCSR.	110
4.2.5	Variation of Injected Carrier Density across Crystal for PCSTF.	111
4.2.6	Current - Voltage Characteristic for PCSTF.	111
4.2.7	Injection Hall Voltage for PCSTF.	111
4.3.1	Rectifying Characteristics at 20° K, 77° K.	115
4.3.2	Rectifying and Ohmic Characteristics.	116
4.3.3	Rectifying Characteristics at Four Temperatures.	117
4.3.4	Characteristic Variation with Contact Separation.	118
4.3.5	Characteristic Variation with Contact Separation.	119

4.3.6	Sectional Voltages.	119
4.3.7	Sectional Voltages.	120
4.3.8	The Local Potential.	121
4.3.9	Recording of Hall Voltage.	122
4.3.10	Hall Voltages at Two Points.	123
4.3.11	Hall Voltages for Two Directions of Flow.	124
4.3.12	Hall Voltages for Uncompensated Material.	125
4.3.13	Hall Voltages at Three Temperatures.	126
4.3.14	Hall Voltages near 77° K.	127
4.3.15	Hall Voltages at Eight Magnetic Fields.	128
4.3.16	Hall Voltages at Four Magnetic Fields.	129
4.3.17	Hall Voltages at Five Magnetic Fields.	129
4.3.18	Magnetic Field Dependence of Hall Voltage Inversions.	129
4.3.19	Hall Voltages at Three Magnetic Fields at 20° K.	129
4.3.20	Hall Voltages at Three Magnetic Fields at 20° K.	129
4.3.21	Sectional Voltages in Magnetic Fields.	130

5.1	The Impossibility of Negative Resistance for Case 1.	153
5.2	Curve Forms for Case 2.	156
5.3	Contour Plot Showing Negative Resistance Region in Terms of the Ionization Rate Coefficient Variation.	157
5.4	The Origin of Negative Resistance.	166
6.1	Sub-bands in a Semiconductor.	178
6.2	Scattering Functions.	183
6.3	Distribution Function Distortion.	184
6.4	Domain Representation.	196
6.5	The Equal Areas Rule.	199
6.6	Division of Region between Hall Contacts into Elements.	203
7.1	Ohmic Characteristic Showing Hysteresis.	211
7.2	Recording of Hall Voltage Showing Hysteresis.	212
7.3	Temperature Dependence of Breakdown.	212
7.4	Reversibility of Ohmic Characteristic.	213
7.5	Breakdown Statistics for Overcompensated Germanium.	216
7.6	Breakdown Statistics for Different Materials at 77 K.	217
7.7	Input Power for Pulsed Breakdown.	219
7.8	Radiation Signal Passed by Filters.	220

Preface

Under the general title "Properties of semiconductors at low temperatures", the work presented here is concerned exclusively with bulk differential negative resistance, which serves as a subtitle. While some of the ideas presented may have application outside this field, development of each is entirely in terms of its usefulness within the field of negative resistance.

The initial aim was an understanding of the recombination properties of excess carriers in overcompensated zinc-doped germanium. A number of negative resistances have been observed experimentally for which a self-consistent set of explanations was not immediately obvious. Hence a wide-ranging survey of negative resistance was undertaken and no attempt has been made subsequently to confine the discussion to specific properties of germanium. Hence negative resistance possibilities are also treated in materials such as gallium arsenide and indium antimonide. No claim is made to "comprehensiveness" in the field of negative resistance; the work is an investigation of certain topics.

In presentation it has been necessary to restrict the length of theoretical sections quite severely. Hence intermediate stages in mathematical derivations have been omitted, though every attempt is made not to disguise assumptions and approximations in the

process. Minimal reproduction of the literature is also intended; on each topic, a well-established paper is taken as a starting point and, to facilitate reference, the notation used there is continued, subject to limitations of internal self-consistency. It is hoped that in terms of physical principles, this work will present a self-contained whole, but for a detailed understanding, reference to quoted papers is probably essential. The foundation stones may be listed as:

- Chapter 1 Negative resistance --- B. K. Ridley (1963, P1).
- Chapter 2 Irreversible thermodynamics --- S. R. de Groot
(1959,P2); J. M. Ziman (1956, P3).
- Chapter 4 Double injection --- M. A. Lampert (1964,P4).
- Chapter 5 Impact ionisation --- S. H. Koenig, R. D. Brown,
W. Schillinger (1962, P5).
- Chapter 6 Voltage controlled negative resistance ---
B. K. Ridley, T. B. Watkins (1961, P6).
- Chapter 7 Thermal effects --- P. C. Eastman, W. R. Datars
(1963, P7).

The reference system is devised to be fully comprehensive in both directions. Each reference appears with author and date; where the reference has been previously employed in the same chapter,

this is shown by an absence of initials and by the letters l.c. (loc. cit.). The index at the back is in alphabetical order of author's name, all authors being listed for jointly written papers. At the same time, all papers quoted are numbered in sequence as employed, whether newly or previously quoted, and these numbers also appear in the index, to make reference back possible to all relevant points.

Chapter 1

Introduction.

1.1 Aims.

The aim of this work is to investigate bulk differential negative resistance in semiconductors with particular reference to overcompensated zinc-doped germanium. In this chapter an examination is made of the fundamental features of such a negative resistance; also the relevant properties of impurity doped germanium are considered. From this base, a number of topics in negative resistance will be explored. A frequent question will be: "Can this process provide a negative resistance or not?" In each case, an answer will be sought through the application of the negative resistance criteria to be deduced here.

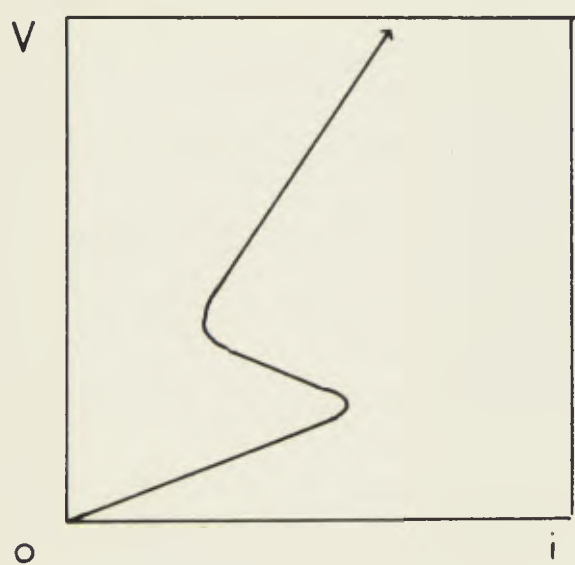
1.2 Negative Resistance.

1.2.1 Negative Resistance Forms.

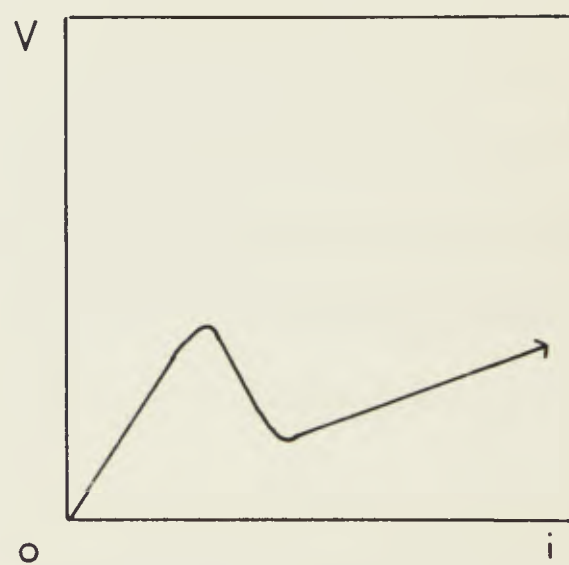
Differential negative resistance is of considerable interest because of its applicability to amplification; if a stable working point can be found, a negative resistance dissipates a negative amount of a.c. power, i.e. it produces amplification; voltage amplification may be obtained directly by employing a simple potential dividing circuit. Thus the discovery of new forms of negative resistance may lead to new types of amplifier and oscillator.

Figure 1.1

Negative Resistance Types



Voltage Controlled



Current Controlled

Various differential negative resistances have been found in solids, primarily associated with junction and contact phenomena. We concern ourselves here only with bulk properties. A starting point for the consideration of the fundamental concepts is provided by a paper of B. K. Ridley (1963, 101).

Differential negative resistances may be divided into two classes, which will be termed voltage controlled and current controlled, following Ridley. These are shown diagrammatically in figure 1.1. The notation N-type and S-type will be avoided because this is ambiguous and normally refers to axes plotted conversely. Throughout this work the current will be employed as the abscissa; this is because the main interest is in domain formation in which the current is the solenoidal function.

The current-voltage characteristic observed may be a considerable modification of the natural one, due to the external circuitry used for the demonstration. For an applied voltage V_0 and external circuit resistance R , the load-line is given by:

$$iR = V_0 - V \quad .$$

If the load-line cuts the true characteristic in one point, the observed characteristic may be expected to be stable and a reproduction of the natural one; if the load-line cuts in several points, spontaneous transitions along load-lines may be anticipated, with possible hysteresis effects on cycling. Consideration of the stable operating points for a bulk negative resistance will be made in

chapter 2. Here we examine the physical criteria for the existence of a negative resistance.

1.2.2 Negative Resistance Criteria.

We wish to find conditions under which negative resistance can take place in solids; the simple case is taken of one carrier flow without diffusion or space charge; electron flow is considered.

Then the current is given by:

$$J = \sigma E$$

In a negative resistance region dJ/dE is negative, by definition.

Taking the total differential results in:

$$\frac{dJ}{dE} = \frac{\partial(\sigma E)}{\partial E} + \frac{\partial(\sigma E)}{\partial J} \frac{dJ}{dE}$$

$$\frac{dJ}{dE} = \frac{\sigma + E \frac{\partial \sigma}{\partial E}}{1 - \frac{J}{\sigma} \frac{\partial \sigma}{\partial J}}$$

This can be negative because of a negative value either for the numerator or the denominator. This procedure is valid provided that the partial derivative of the conductivity used includes all current or voltage dependences.

In examining the expression, we consider first the case of a negative numerator and positive denominator; this requires that $\partial \sigma / \partial E$ be negative and hence that the conductivity should decrease with the electric field. In passing through the negative resistance region, dJ/dE goes through the value zero on two occasions;

thus the negative resistance is voltage controlled. Secondly, consider the case of a negative denominator and a positive numerator; this occurs if $d\sigma/dJ$ has a large enough positive value and hence the conductivity is to increase markedly with the current. In reaching the negative resistance dJ/dE goes through the value infinity and thus the negative resistance is current controlled.

There is also the possibility that $\partial\sigma/\partial E$ may go through the value of plus infinity to a large negative value. This is equivalent to the condition that $\partial E/\partial\sigma$ should go through the value of zero. It has just been shown that in ordinary circumstances for positive resistance, $\partial\sigma/\partial E$ is not a large negative number, so it is clear that if $\partial E/\partial\sigma$ passes through zero, a current controlled negative resistance results. Similarly if $\partial J/\partial\sigma$ passes through zero a voltage controlled negative resistance results.

The simultaneous assumption of negative values by both the numerator and denominator will be ignored. This would represent a portion of characteristic with positive resistance traced in the reverse direction, which does not appear to be useful.

We now consider the converse to the criteria. For a single carrier field-driven current flow, a voltage controlled negative resistance is only possible under the alternative conditions:

$$(a) \quad \sigma + E \frac{\partial\sigma}{\partial E} \leq 0 \quad ;$$

$$(b) \quad \frac{\partial J}{\partial\sigma} \leq 0 \quad .$$

A current controlled negative resistance for these circumstances is only possible for the alternatives:

$$(a) \quad 1 - \frac{J}{\sigma} \frac{\partial \sigma}{\partial J} \leq 0 \quad ;$$

$$(b) \quad \frac{\partial E}{\partial \sigma} \leq 0 \quad .$$

In all cases the zero values have to be taken during the process. This means that for a voltage controlled negative resistance, either the conductivity must be a function of the electric field, or the conductivity must already be multivalued in the current. Similarly for a current controlled negative resistance, either the conductivity must be a function of the current density, or it must already be multivalued in the electric field.

An analysis of this kind for the criteria for negative resistance does not seem to exist in the literature. In practice there seems to be considerable confusion between the two types of negative resistance and the conditions under which each can occur. Our criteria will be found to be vital for the discussion of impact ionization in chapter 5.

1.2.3 Instability.

The possibility of electrical instability in a negative resistance region is discussed by B. K. Ridley and T. B. Watkins (1961, 102). In a current controlled negative resistance region, several values of the voltage are possible for one current. Thus a spontaneous

transition from one voltage value to another is possible. But the fact that it is possible does not mean that it must take place. Such breakdowns are experimentally observed, and several are described in this work. But considerable care is required in the explanation, between the stages of showing that a crystal can exist in two different 'phases', and of deducing that a spontaneous transition between the two must take place. We emphasize this because in some papers it seems to have been ignored, as will be shown. Indeed the whole subject of when and why spontaneous transitions take place seems to have been scantily explored. This is the realm of irreversible thermodynamics which forms the subject of the next chapter; in the rest of this section we consider the different 'phases' themselves.

An example of the possibility of different 'phases' is given by filament formation. Consider the case where the mobility rises with the current density. Then it may be imagined that the flow of electrons may exist both in a low current state, and also in a high current state for the same applied voltage, or in some intermediate stage with a high current filament; by adjusting the cross-section of the filament, the applied voltage can supply any of the intermediate currents. A spontaneous transition into filament flow at a given voltage may thus be expected to increase the current by orders of magnitude, while a transition out of filament flow would do the reverse. We see that such transitions are possible, but are as yet

undetermined. A similar situation exists for domain formation. We will show (section 2.2) that a spontaneous transition into or out of domain structure can lead to a marked change in the working point on the current-voltage characteristic.

Thus we see that filaments and domains, whether growing or decaying, are manifestations of negative resistance rather than causes. If a bulk negative resistance exists it may prove unstable against them, but this is not in itself a necessary result. We shall turn to thermodynamics for principles for determining stable working points.

1.3 Survey of Impurity Levels.

1.3.1 General Levels in Germanium.

Electrically active impurities in germanium are classified as shallow level and deep level impurities. Shallow level impurities, which consist of lithium and elements of columns III and V of the periodic table, introduce levels about 0.01 eV from the band edges. The nature and properties of such levels have been reviewed by W. Kohn (1957, 103). A survey of the deep level impurities has been provided by W. W. Tyler (1959, 104); deep level acceptor impurities with three or more levels are given as copper, cobalt, chromium, gold, silver and platinum; and those with two levels as iron, nickel, manganese, cadmium and zinc; selenium and tellurium act as deep level donors.

Considerable progress has been made in the understanding of the shallow levels in germanium even in degenerate material (reviewed by J. I. Pankove 1965, 105) and for impurity band conduction. The main problem in dealing with carrier interaction with a deep level is the lack of concrete knowledge of the impurity wavefunction (B. K. Ridley, T. B. Watkins 1961, 106). Capture cross-sections may be considered in terms of the giant trap model of M. Lax (1959, 107) but no detailed treatment is available. Deep levels are of interest because of their influence on the trapping and recombination of minority carriers. In particular there is the question as to whether radiative recombination at an impurity site takes place. The evidence for this in germanium and silicon is not fully conclusive. In these elements the band to band recombination process is indirect due to the band structure, so that only one transition in 10^9 is radiative (S. H. Koenig, R. D. Brown 1960, 108). Hence the radiation intensities available are very low. Initial studies of the intrinsic recombination radiation in Ge and Si by J. R. Haynes, M. Lax and W. F. Flood (1959, 109) demonstrated the phonon spectrum but not an impurity effect. Y. E. Pokrovskii, K. I. Svistunova (1964, 110) have reported observation of radiative recombination at zinc levels in germanium. Since then further reports of extrinsic radiative recombination have appeared in the Russian literature (I. V. Karpeva, Y. E. Pokrovskii 1965, 111; Y. L. Ivanov, A. V. Yuhnevich 1965, 112; A. A. Gippius, V. S. Vavilov 1965, 113).

So far intrinsic or near-intrinsic radiation has been considered, resulting from transitions between the conduction band and an acceptor level. An alternative possibility is to consider extrinsic radiation for transitions between the conduction band and a donor level or an acceptor level and the valence band. Radiative recombination for the 0.01 eV transition to the arsenic level has been reported by Koenig and Brown (l.c. 1960, 114). To observe transitions from a deep level acceptor site, a method of excitation is required. This may be provided by the use of a recombination centre; if electron-hole pairs are injected into a region where the main recombination process operates through a deep level site, then a high rate for transitions of the sort to be investigated is produced. The carrier injection may be either by illumination or by electrical injection at a rectifying contact. In this work carriers have been injected electrically into overcompensated zinc-doped germanium.

1.3.2 Zinc Levels in Germanium.

Zinc is an element in group II and is hence to be expected to have two acceptor levels as an impurity in germanium. These have been found to be situated at 0.03 eV and 0.09 eV above the valence band (W. W. Tyler, H. H. Woodbury 1956, 115).

Three states of compensation are hence available for examination:

(1) uncompensated by shallow donor impurities, giving a p-type crystal with behaviour dominated by the lower zinc level; Hall coefficient and photoconductivity studies give the ionization energy of this lower level;

(2) partially compensated, in which donors have been introduced to fill the lower level entirely, giving p-type behaviour dominated by the upper level, whose energy may be found by Hall coefficient or photoconductivity measurements;

(3) overcompensated, in which donors have been introduced to fill both acceptor levels, giving n-type behaviour, the Hall coefficient and photoconductivity being dominated by the surplus donors.

The photoconductive sensitivity in these cases has been measured by Tyler and Woodbury (l.c. 1956, 116); for (1), (2) the carrier lifetime is short and the photoconductive sensitivity correspondingly low, but for (3) the lifetime is longer and the sensitivity higher. This is due to hole trapping at the zinc sites, with the photoelectron free to contribute to the conductivity. The hole may then be excited thermally back into the valence band to recombine normally, or the trap may be able to capture an electron and act as a recombination centre itself. The recombination probability at a zinc site has been studied by A. I. Morozov and S. G. Kalashnikov (1961, 117); the conclusion drawn there is that recombination at a zinc site is improbable compared with the probability of thermal ejection and of recombination by other processes.

The aim has thus been set as the detection of radiative transitions between the upper zinc level and the valence band; this provides an energy of 0.09 eV corresponding to 15μ wavelength. It is certainly true that a recombination centre such as nickel or copper would be expected to provide a better chance of observing a radiative transition; the choice of zinc rested entirely on the availability of a suitable overcompensated specimen.

Chapter 2

Irreversible Thermodynamics.

2.1 Introduction.

The general theory of irreversible thermodynamics is of fairly recent origin, being based on Onsager's Theorem of 1931. Just as classical thermodynamics --- thermostatistics --- deals with equilibrium conditions, so irreversible thermodynamics is concerned with steady flows under given constraints.

In the standard formulation, e.g. that of S. R. de Groot (1959, 201), the interaction effects of gradients of temperature, concentration, pressure, electric fields and such are discussed. In this way, certain principles applicable to general steady flow conditions are deduced. As applied to electric field gradients in ohmic specimens, the thermoelectric properties are successfully explained. Application to bulk negative resistances may be expected to give rise to curious features; the only work on this seems to be due to B. K. Ridley (1963, 202), as a direct result of experimental results similar to those reported here. Our interpretation of the principles of irreversible thermodynamics leads to conclusions considerably different from his, but these seem to be better supported by recent experimental results.

1. The first part of the paper is devoted to the study of the properties of the function $f(x)$ defined by the equation

$$f(x) = \int_0^x \frac{1}{1+t^2} dt$$

It is shown that the function $f(x)$ is increasing and concave down on the interval $(-\infty, \infty)$.

2. In the second part of the paper, we consider the function $g(x)$ defined by the equation

$$g(x) = \int_0^x \frac{t}{1+t^2} dt$$

It is shown that the function $g(x)$ is an odd function and that it is concave up on the interval $(-\infty, \infty)$.

3. In the third part of the paper, we consider the function $h(x)$ defined by the equation

$$h(x) = \int_0^x \frac{t^2}{1+t^2} dt$$

It is shown that the function $h(x)$ is an even function and that it is concave down on the interval $(-\infty, \infty)$.

4. In the fourth part of the paper, we consider the function $k(x)$ defined by the equation

$$k(x) = \int_0^x \frac{t^3}{1+t^2} dt$$

It is shown that the function $k(x)$ is an odd function and that it is concave up on the interval $(-\infty, \infty)$.

5. In the fifth part of the paper, we consider the function $l(x)$ defined by the equation

$$l(x) = \int_0^x \frac{t^4}{1+t^2} dt$$

It is shown that the function $l(x)$ is an even function and that it is concave down on the interval $(-\infty, \infty)$.

6. In the sixth part of the paper, we consider the function $m(x)$ defined by the equation

$$m(x) = \int_0^x \frac{t^5}{1+t^2} dt$$

It is shown that the function $m(x)$ is an odd function and that it is concave up on the interval $(-\infty, \infty)$.

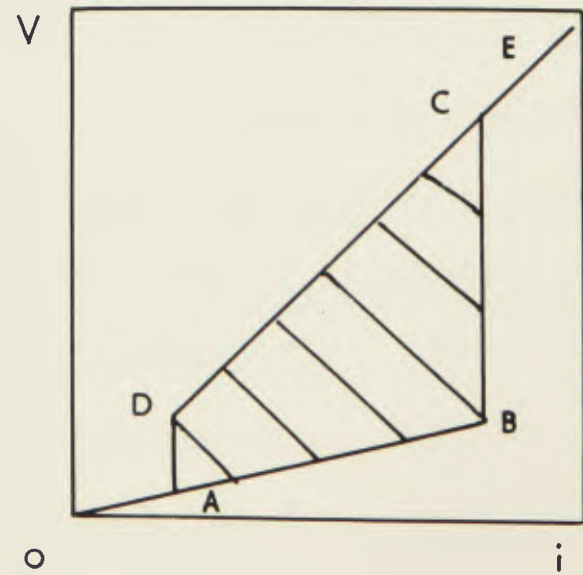
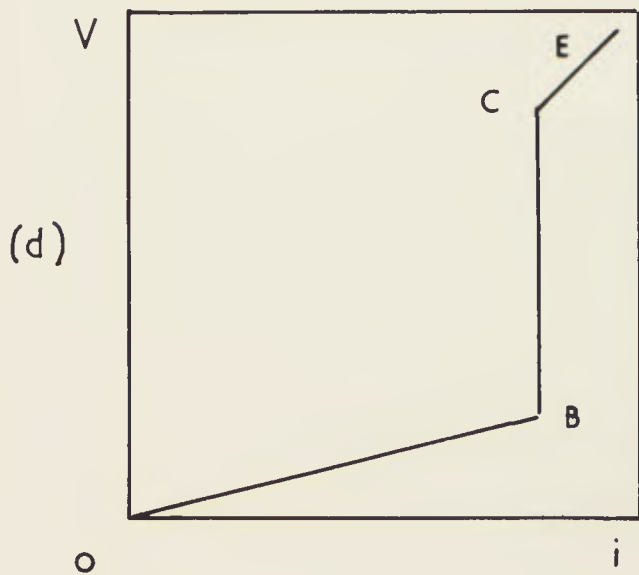
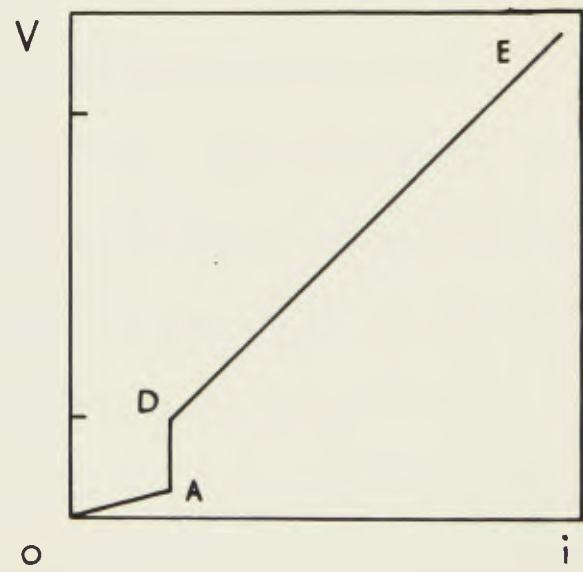
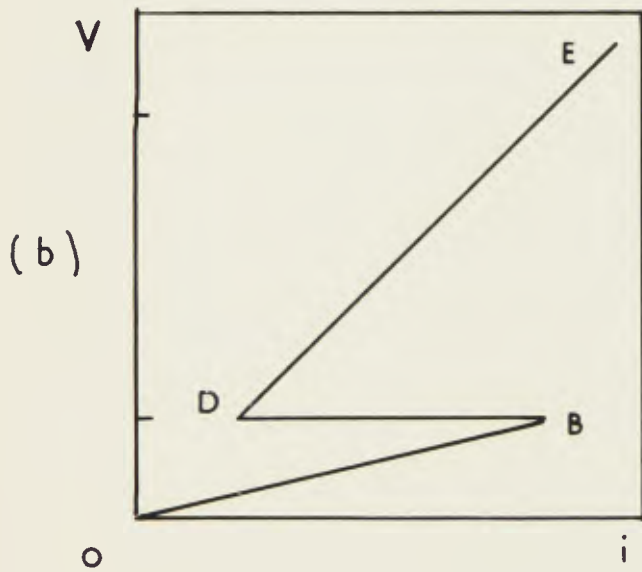
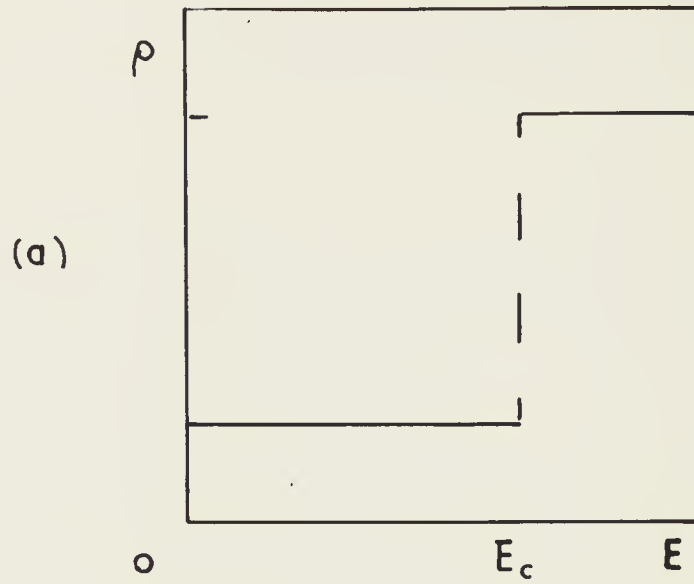
7. In the seventh part of the paper, we consider the function $n(x)$ defined by the equation

$$n(x) = \int_0^x \frac{t^6}{1+t^2} dt$$

It is shown that the function $n(x)$ is an even function and that it is concave down on the interval $(-\infty, \infty)$.

Figure 2.1

The Effect of Domains on the i - V Characteristic



2.2 Applicability to Negative Resistance.

The importance of a correct formulation of irreversible thermodynamics for negative resistance is best illustrated by the following example. Consider a hypothetical crystal in which there is a sudden rise in the resistivity at a certain electric field E_c ; this may be imagined to be due to some field-dependent carrier capture cross-section for an impurity site, or to a field-dependent carrier mobility. Thus the bulk resistivity plotted against the electric field shows a jump at E_c which will be shown as a discontinuity (figure 2.1a). Assume first that the crystal remains in a homogeneous state; then the current-voltage characteristic will be of the form of figure 2.1b. Suppose however that the crystal breaks into domains; then the current may stay at its lowest value, while the extra voltage is taken up in the form of a high resistance domain; this will grow until it envelopes the whole crystal, giving the curve of figure 2.1c. Alternatively domains may form at the highest point, giving the characteristic of figure 2.1d. In fact the only basic condition is that there must be current continuity, and for different domain sizes, it is clear that any position is possible within a trapezium ABCD (figure 2.1e).

The question arises as to which path the crystal actually takes as the voltage is raised; there are three governing features:

- (1) the external circuitry;
- (2) internal conditions on domain formation;
- (3) irreversible thermodynamics.

External circuitry will limit any spontaneous movement to a load line:

$$i = (V_o - V)/R_o$$

where zeros apply to external circuit parameters, as Ridley (l.c. 1963, 203) explains, but it cannot do more. What is desired is the equilibrium point on a given load line. We set (2) on one side for the time being and look to thermodynamics.

2.3 Condition for the Steady State.

The formulation of irreversible thermodynamics is still incomplete and paradox is immediately encountered. P. Ehrenfest in his *Encyclopaedia of Mathematical Physics* (1911, 204, p.82) first asked the vital question: "Which state variable, and under what auxiliary conditions, has its maximum when the stationary state is reached?"

Quoting de Groot (l.c. 1959, 205, p.196): "We see that the answer to this question is given in (the) thermodynamics of irreversible processes; the characteristic state variable is the entropy production which has a minimum value in the stationary state, compatible with the auxiliary conditions which express that certain external forces are kept at fixed values." Thus the answer is to be minimum entropy production.

Quoting J. M. Ziman (1960, 206, p.235): "In the steady state the currents are such that the production of entropy has the largest value consistent with its subsequent conservation." The answer is

then to be maximum entropy production.

So "you pays your money and you takes your choice."

In this situation Ridley has chosen minimum entropy production (l.c. 1963, 207) and it is intended here to show that in his case this is unsatisfactory. It is hoped to give some insight into the nature of the paradox but a general formulation from the principles of quantum statistical mechanics will not be attempted.

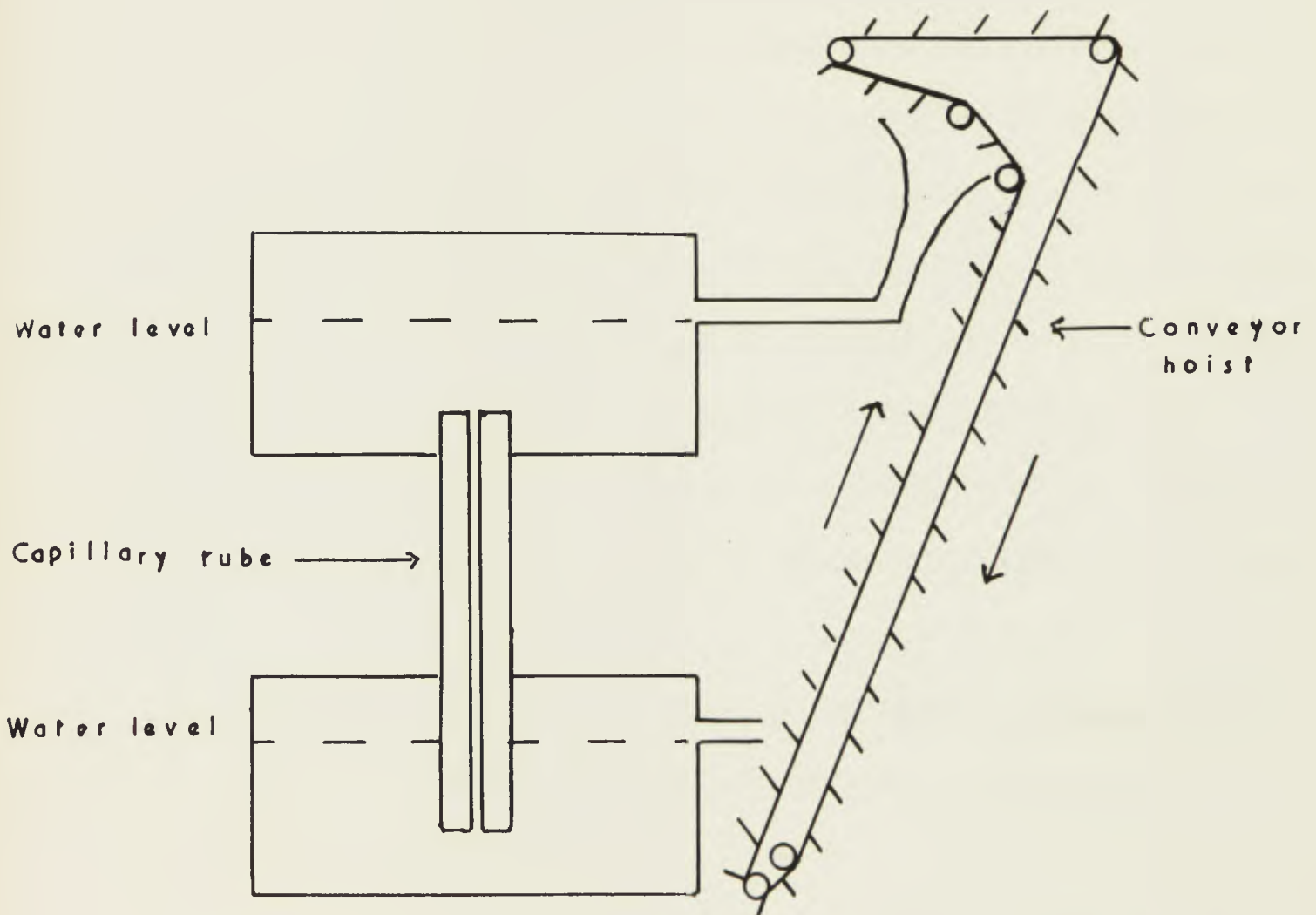
Ridley's justification for his choice is as follows (l.c. 1963, 208): "In the steady state the overall entropy accumulation rate is zero, as are the overall accumulation rates of energy and of charge, and in each case the production is balanced by the divergence of the flow. At equilibrium both the production and the flow of entropy are zero. Since it is to be expected that the steady state will be as near the equilibrium state as the various constants will allow, we expect the entropy production rate to be as small as possible. This plausible argument is at the basis of the principle of least entropy production." But is plausibility enough?

2.4 Recirculatory Illustration.

That the principles of minimum and maximum entropy production are open to abuse, is demonstrated by the following example. Consider a recirculatory system at constant temperature; a constant head of water in one tank is connected by a capillary tube to a second tank which has an overflow pipe (figure 2.2).

Figure 2.2

Recirculatory System



The water is being continuously replaced at its constant head height by a fast continuous conveyor hoist, whose capacity is well in excess of the flow capacity of the tube constriction. We now ask how fast the water circulates round the system. A simple-minded application of minimum or maximum entropy production would give respective flow rates of zero or infinity. In fact the microscopic processes inside the constriction govern the flow giving rise to a macroscopic resistance.

This case is closely analogous to that of the special crystal introduced in section 2.2; but for this crystal the internal processes are such that the resistance may take any of a range of values, due to domain formation. Thus we are thrown back on the thermodynamic question "Minimum or maximum entropy production (if either)?" These will now be examined in turn.

2.5 Derivation of the Minimum Entropy Production Principle.

In deriving the principle of minimum entropy production for the present purpose, particular care is to be taken over the conditions under which it is valid. The methods and notation are those of de Groot (l.c. 1959, 209, p.6, p.196).

The rate of entropy production is written down for the system as the sum of a number of terms:

$$\dot{\Delta S} = \sum_i J_i X_i$$

where the J_i represent fluxes and the X_i forces. This is based on

a definition of the entropy value as being a quadratic function of the deviations of the state parameters from their equilibrium value. We leave further discussion of this until section 2.10, accepting meanwhile that the above expression for the entropy production rate is to be limited to conditions close to equilibrium.

Phenomenological relations are chosen for the system between the fluxes and the forces so that:

$$J_i = \sum_j L_{ij} X_j \quad ,$$

where the L_{ij} may or may not be constant. Choice of the J_i and X_i is not unique, but all choices lead to the same relationship; for particular J_i and X_i , the L_{ij} represent the transport coefficients of macroscopic physics.

Onsager's theorem now states that for linear phenomenological laws in the absence of magnetic fields:

$$L_{ij} = L_{ji} \quad .$$

This is obtained through the application of the principle of microscopic reversibility to the time average for energy fluctuations. Two assumptions are made; it is assumed firstly that the mechanisms and rates of decay of a microscopic fluctuation are exactly the same as for a macroscopic one; so that macroscopic coefficients may be inserted into the microscopic reversibility relations; this is true for a Maxwell-Boltzmann transport equation and is well justified by the results; and secondly that the characteristic molecular

collision time is much shorter than the decay time of a macroscopic fluctuation; kinetic transport theory requires this. Thus the Onsager relations are justifiable in terms of Boltzmann transport theory in the neighbourhood of equilibrium, the criterion being the validity of the quadratic entropy approximation.

We are now in a position to carry out our aim, which is to compare the stationary state of a system with the state of minimum entropy production. We consider a system with n independent parameters (such as pressure, temperature) and k forces ($k < n$) $X_1 \dots X_k$ held constant by external constraint (such as a temperature gradient); then the steady state is given, by definition, when all fluxes (e.g. the particle flow) are zero, whose complementary forces are not fixed by constraint; thus:

$$J_{k+1} = J_{k+2} = \dots = J_n = 0 \quad .$$

This is to be compared with the state of minimum entropy production.

The entropy production rate is:

$$\dot{S} = \sum_i J_i X_i = \sum_{ij} L_{ij} X_i X_j \quad .$$

This is stationary when:

$$\frac{\partial}{\partial X_i} (\dot{S}) = 0 \quad , \quad i = k+1, k+2, \dots n.$$

Differentiating:

$$\sum_j (L_{ij} + L_{ji} + X_i \frac{\partial}{\partial X_i} (L_{ji})) X_j = 0 \quad , \quad i = k+1, k+2, \dots n.$$

Using Onsager's relations,

$$\sum_j 2L_{ij}X_j + \sum_j X_iX_j \frac{\partial}{\partial X_i} (L_{ji}) = 0, \quad i = k+1, k+2, \dots n.$$

Hence:

$$J_i = \sum_j L_{ij}X_j = -\frac{1}{2} \sum_j X_iX_j \frac{\partial}{\partial X_i} (L_{ji}), \quad i = k+1, k+2, \dots n.$$

If the phenomenological relations are linear then the state of minimum entropy production gives:

$$J_i = 0,$$

which is the steady state. If the phenomenological relations are not linear, then it becomes difficult to determine what the Onsager relationship is. This question is discussed by D. T. J. Hurle, J. B. Mullin and E. R. Pike (1965, 210); straightforward application gives the result as above:

$$J_i = -\frac{1}{2} \sum_j X_iX_j \frac{\partial}{\partial X_i} (L_{ji}),$$

from which it is apparent that the steady state is not that of minimum entropy production. It is the departure from linearity of the phenomenological laws which invalidates the minimum entropy production principle rather than the departure from equilibrium. In the next section it will be shown that with linear laws, least entropy production can remain a reasonable steady state criterion even quite far from equilibrium.

2.6 Application of the Minimum Entropy Production Principle.

In standard texts on irreversible thermodynamics, consideration is generally confined entirely to systems with linear phenomenological equations. This is clearly stated, but it is also stressed that the reason for this is that transport properties of physical systems fall into this category. The confusion to be found in the literature seems to be due to the assumption that this should apply to all systems.

Particular stress is placed here on the distinction between the effects of non-linearity in the phenomenological equations and of closeness to equilibrium, with respect to the validity of the principle of minimum entropy production. Neglect of these conditions is the starting point for misunderstanding; often the entropy production rate is obtained directly from the physical system and set to be a minimum without further ado. An interesting instance of this is provided in the calculations and comments of M. J. Klein (1954, 211; 1955, 212; 1960, 213). He considers two definite examples.

Firstly two identical containers of an ideal gas separated by a constriction are held at different temperatures by thermal contact with heat baths. Individual energy states are defined and the rates of change of the occupation probabilities of the energy states in the two containers due to the heat baths and to gas flow between baths are calculated. In the steady state the rates of occupation change are zero. The system entropy is also written down and the minimum

production rate calculated. It then becomes apparent that this minimum production rate only gives the steady state if certain assumptions are made about the conditions. It should be noted that the effective phenomenological equations used are very far from linear, and the main effect of the approximations is to linearize these rather than, as is assumed to be necessary, to bring the system nearer to equilibrium.

In the second example, that of a spin system subjected to a radiation field, this feature is still more apparent; here again the phenomenological equations are very non-linear and the minimum entropy production principle only approximates to the steady state condition when certain assumptions are made about the conditions. However Klein made numerical calculations for a particular set of conditions well away from equilibrium, and it was found that the least entropy production principle gave a close approximation to the steady state.

Quoting Klein (l.c. 1960, 214, p.203): "Although the results are limited to the simple case treated, they are interesting in so far as they suggest the possibility of quite another range of conditions (apart from close to equilibrium) in which the principle of minimum entropy production may be valid." In fact the effect of the change in conditions is to linearise the phenomenological equations, thus giving computational confirmation to the hypothesis put forward here that the validity of the principle of minimum entropy production relies

mainly on the linearity of the phenomenological equations.

We have considered Klein's examples as more than a passing reference because they are reckoned to be of more than marginal importance. For in C. Kittel's textbook "Elements of Statistical Physics", (1958, 215, p.166) the second example is quoted in its original form. This indicates the need for a fundamental appraisal of the whole of irreversible thermodynamics.

The results may now be applied to electric field effects in crystals; as long as all phenomenological equations are linear, the minimum entropy production principle may be expected to give valid results. Results for thermoelectricity and for galvanomagnetic and thermomagnetic effects confirm this. But application to specifically non-ohmic effects and in particular to differential negative resistance, which is the basis of Ridley's work (l.c. 1963, 216), is not justifiable and hence is likely to lead to invalid conclusions.

2.7 Derivation of the Maximum Entropy Production Principle.

With this negative conclusion we turn to an investigation of the principle of maximum entropy production; here the formulation and the notation to be used are due to J. M. Ziman (1956, 217; 1960, 218, p.257).

Boltzmann's equation for electrical conduction is introduced as:

$$\dot{f}_k = (\dot{f}_k)_{\text{diffusion}} + (\dot{f}_k)_{\text{fields}} + (\dot{f}_k)_{\text{scatter}} = 0 \quad .$$

For simplicity, magnetic fields and diffusion effects will be set aside. Thus:

$$(\dot{f}_k)_{\text{field}} + (\dot{f}_k)_{\text{scatter}} = 0 .$$

The effect of the electric field E may be written, if the change from equilibrium is small, as:

$$(\dot{f}_k)_{\text{field}} = -eE \cdot v_k \frac{\partial f_k^0}{\partial E_k} ,$$

where v_k is the group velocity and the zero suffix denotes the equilibrium value. The electron scatter rate depends on the cross-section for the process and on the occupation probabilities, with Fermi-Dirac statistics. Thus:

$$(\dot{f}_k)_{\text{scatter}} = -\int \{f_k(1 - f_{k'})Q(k, k') - f_{k'}(1 - f_k)Q(k', k)\} dk' .$$

Transition reversibility ensures that $Q(k, k') = Q(k', k)$.

For elastic scattering, $f_k^0 = f_{k'}^0$.

Defining $f_k = f_k^0 - \phi \frac{\partial f_k^0}{\partial E_k}$,

$$P(k, k') = f_k^0 (1 - f_{k'}^0) Q(k, k') ,$$

we obtain:

$$-kTeE \cdot v_k \frac{\partial f_k^0}{\partial E_k} = \int \{\phi_k - \phi_{k'}\} P(k, k') dk' .$$

This equation may be represented as:

$$\chi(k) = P \Phi .$$

Defining:

$$\langle \Phi, \Psi \rangle = \int \Phi(k) \Psi(k) dk ,$$

Ziman shows that if $\bar{\Phi}$ is a solution of

$$\chi = P \bar{\Phi}$$

and ψ of

$$\langle \psi, \chi \rangle = \langle \psi, P \bar{\Phi} \rangle,$$

then of all possible ψ , $\bar{\Phi}$ is the one which makes $\langle \psi, P \bar{\Phi} \rangle$ a maximum. To employ this variational technique to solve the equation

$$\chi = P \bar{\Phi},$$

a set of solutions to

$$\langle \psi, \chi \rangle = \langle \psi, P \bar{\Phi} \rangle$$

is obtained, and a linear combination of these with variable coefficients is made; then the optimum solution ψ' is the one where the coefficients are such that

$$\langle \psi', P \psi' \rangle$$

has a maximum value. This is valid subject to the conditions that

$$\langle \psi, P \psi \rangle > 0 \text{ for all } \psi \text{ and that } \langle \psi, P \psi' \rangle = \langle \psi', P \psi \rangle.$$

Ziman has shown that this technique has a physical significance. The basic equation is:

$$(\dot{f}_k)_{\text{field}} = - (\dot{f}_k)_{\text{scatter}}.$$

Variational techniques may be applied as long as one side of the equation is independent of the variable. Looking back we note that the distribution function f_k has been eliminated from the field term; so employing any function ψ of the distribution function, we may say that for any ψ such that:

$$\int \psi(f_k) \cdot (\dot{f}_k)_{\text{field}} dk = - \int \psi(f_k) \cdot (\dot{f}_k)_{\text{scatter}} dk ,$$

the required solution for f_k is the one for which:

$$- \int \psi(f_k) \cdot (\dot{f}_k)_{\text{scatter}} dk$$

has a maximum value. This assumes that the variational conditions are fulfilled.

Now choose:

$$\psi(f_k) = \phi_k = - (f_k - f_k^0) \left(\frac{\partial f_k^0}{\partial E_k} \right)^{-1} .$$

The expression to be maximised is

$$- \int \phi_k (\dot{f}_k)_{\text{scatter}} dk .$$

The variational conditions are satisfied in this case provided that the transition probabilities are positive and reversible, and that the scattering is elastic.

The entropy production rates will now be calculated. For an assembly of fermions the entropy production rate is (D. ter Haar, 1954, 219; Ziman l.c. 1960, 220, p.281):

$$T\dot{S} \approx - \int \phi \dot{f}_k dk .$$

The Joule heating entropy is:

$$\begin{aligned} T\dot{S} &= E \cdot J \\ &= E \cdot e \int v_k (f_k - f_k^0) dk \\ &= - \int \phi v_k \cdot eE \frac{\partial f_k^0}{\partial E_k} dk \\ &= - \int \phi (\dot{f}_k)_{\text{field}} dk . \end{aligned}$$

Hence the equations

$$\int \phi(\dot{f}_k)_{\text{field}} dk = - \int \phi(\dot{f}_k)_{\text{scatter}} dk$$

$$- TS_{\text{field}} = TS_{\text{scatter}}$$

are identical, and the expression to be maximised is equivalent to the entropy production. This is ingeniously interpreted by Ziman as follows; the irreversible scattering processes produce entropy at a certain rate, while the electric field sweeps this away into the reservoirs where it appears as Joule heating; in the steady state these two rates balance.

Ziman generalizes the principle, defining the entropy produced by scattering as intrinsic and that due to fields as extrinsic; his statement is (l.c. 1956, 221, p.1263): "Consider all distributions of currents such that the intrinsic entropy production equals the extrinsic entropy production for a given set of forces. Then of all current distributions satisfying this condition, the steady state distribution makes the intrinsic entropy production a maximum."

In the development of the principle, the choice for ψ of ϕ was arbitrary; such a choice enabled the entropy productions to be worked out easily; but more important the variational conditions may be shown to be satisfied for elastic collisions. We have thus shown that the maximum entropy production principle is consistent with elastic collision processes in a fermi gas close to equilibrium;

in general however it is no more than a postulate for irreversible thermodynamics. We would suggest that the above is a more logical treatment of the application of the variational principle than that given by Ziman (l.c. 1960, 222); Ziman makes no reference to the need for one side of the Boltzmann equation to be constant for variational principles to apply and claims that the maximum entropy production principle is being proved. In fact it is just being verified for a rather limited range of conditions; it is the linearized Boltzmann equation which is being employed and this can be shown to lead to transport equations with linear coefficients (C. F. A. Beaumont 1956, 223). Since we are again limited to linear laws, we may wonder whether any progress at all has been made in this section apart from the paradoxical substitution of maximum for minimum in the entropy production principle for the steady state.

2.8 Application to Electrical Networks.

We have thus seen that confusion between principles of minimum and maximum entropy production is alarmingly possible. Ziman points this out (l.c. 1956, 224, p.1264) and adds: "The two theorems seem to be entirely independent of each other." Yet a classic example of their application to two aspects of one problem may be found on adjacent pages of the textbook by J. Jeans (1925, 225, p.322-3), sections 357 and 358.

"Section 357. Theorem. When a steady current flows through

a network of conductors in which no discontinuities of potential occur, (and which therefore contains no batteries), the currents are distributed in such a way that the rate of generation of heat in the network is a minimum, subject only to the conditions imposed by Kirchhoff's First Law." A proof follows based on virtual displacement. This is minimum entropy production.

"Section 358. Theorem. When a system of steady currents flows through a network of conductors of resistances $R_1 \dots R_n$ containing batteries of emfs $E_1 \dots E_n$, the currents $x_1 \dots x_n$ are distributed in such a way that the function

$$\sum R x^2 - 2 \sum E x$$

is a minimum, subject to the conditions imposed by Kirchhoff's First Law." Again a virtual displacement proof follows. Ziman (l.c. 1956, 226) shows that this is a case of maximum entropy production.

So as $E \rightarrow 0$, the same theorem is a demonstration of the principles of minimum and of maximum entropy production. We shall examine the thermodynamics of resistive networks in some detail for two reasons; firstly because a network provides a situation in which the paradox can be examined in its barest essentials, and secondly because it is precisely with negative resistance networks that we are concerned.

The basic feature of maximization is the variable. In this case we consider a network of resistances and emfs maintained at constant temperature by a large heat capacity. The network currents provide

the variables. The imposition of Kirchhoff's laws defines the operating point except under special circumstances; then a maximisation process for the network is meaningful if its application, together with that of one of Kirchhoff's Laws provides the other. We choose to apply in turn the two entropy production principles to a network, together with Kirchhoff's First Law.

A network is defined conventionally as a set of n closed loops. For application of minimum entropy production, no loops contain emfs, these being outside the part of the network being considered. The loop currents are defined as i_r and resistances common to two loops as R_{ij} , the current through such an element being $i_i - i_j$. This definition of loop currents ensures that Kirchhoff's First Law is valid, however many loops converge at one node. Then by the minimum entropy production principle, $\sum R i^2$ is a stationary (minimum) point in the steady state, with respect to variation of the loop currents. Differentiating with respect to i_r ,

$$\frac{\partial}{\partial i_r} \left[\sum_{i,j \neq r} R_{ij} (i_i - i_j)^2 + \sum_{loop} R_{rs} (i_r - i_s)^2 \right] = 0 \quad . \quad (2.1)$$

Assuming ohmic conditions,

$$2 \sum_{loop} R_{rs} (i_r - i_s) = 0 \quad .$$

This is Kirchhoff's Second Law. For confirmation that the stationary point is a minimum, we refer to the variational technique of Jeans (1.c. 1925, 227).

The principle of maximum entropy production is now applied to a similarly defined network, in this case emfs in the loop circuits being permitted and denoted E_{ij} . The principle states that $\sum Ri^2$ has a stationary (maximum) value at the steady state with respect to variation of the loop currents, subject to the condition that the entropy production equals the entropy flow, expressed as

$$\sum_{\text{loop}} Ri^2 = \sum_{\text{loop}} EI \quad .$$

Incorporating these conditions by the method of Lagrange's undetermined multipliers, the stationary point is given by:

$$\frac{\partial}{\partial i_r} \left[\sum_{\text{all}} Ri^2 - \sum_a \lambda_a (\sum (Ri^2 - EI)) \right] = 0 \quad , \quad (2.2)$$

$$\lambda_a = \frac{\partial}{\partial i_a} (\sum Ri^2) / \frac{\partial}{\partial i_a} \left(\sum_{\text{loop } a} Ri^2 - EI \right) \quad .$$

If at this stage all resistances are assumed ohmic, differentiation gives:

$$\lambda_a = \frac{2 \sum_{\text{loop}} R_{rs} (i_r - i_s)}{2 \sum_{\text{loop}} R_{rs} (i_r - i_s) - \sum_{\text{loop}} E_{rs}} = 2 \quad .$$

Hence the condition becomes that $2 \sum EI - \sum Ri^2$ is to be a maximum and differentiation gives:

$$\sum_{\text{loop}} E_{rs} - \sum_{\text{loop}} R_{rs} (i_r - i_s) = 0$$

showing the process to be consistent with Kirchhoff's Second Law.

Again confirmation that the stationary point is a maximum may be found by the variational technique of Jeans (l.c. 1925, 228). There is no contradiction between maximum and minimum principles since in the limit of zero electric field, the function to be maximized is just the negative of the one to be minimized.

So far all resistances have been assumed ohmic. Our aim in this chapter is to examine the thermodynamics of non-linear and in particular negative resistances. The principle of minimum entropy production is not expected to be valid for non-ohmic resistances, following from sections 2.5 and 2.6. Differentiation of equation (2.1) with the resistances as functions of the currents does not produce Kirchhoff's Second Law, but gives a relationship incorporating the resistance differentials. Since Kirchhoff's Laws define the steady state, minimum entropy production is to be discarded as a principle for applying to non-ohmic networks.

We now come to examine Eizen's postulate that the maximum entropy production principle should apply to the steady state of the universe. The incorporation of non-ohmic resistances in the equations (2.2) again results in an equation involving the resistance differentials which is not compatible with Kirchhoff's Second law. There are four alternative conclusions:

(1) A non-ohmic resistance implies the introduction of an effective source or sink of current flow. Thus maximum entropy production holds but Kirchhoff's Laws do not.

(2) Maximum entropy production holds; the calculation has simply not taken into account all entropy flows, such as heat.

(3) The steady state is defined by some stationary principle, but this is not maximum entropy production as defined here in the case of non-ohmic resistances.

(4) No stationary principle exists.

Conclusions (1), (2) can be demolished without difficulty. Whereas it is possible to have active non-ohmic resistors which form current sources or sinks, it is equally possible to have passive ones, for example a crystal with a field dependent mobility --- such forms our test case. Conclusion (2) can be shown to be invalid physically and logically. Physically it is fair to postulate the network being in contact with an enormous heat capacity so that entropy flow due to the removal of the Joule heating can be ignored. Logically there is no difference between the ohmic case in which the principle holds and the non-ohmic case in which it does not, in terms of entropy flow. As scientists we set conclusion (4) on one side until all hope of satisfying conclusion (3) has been abandoned.

2.9 The Search for a New Entropy Production Principle.

Belief in the existence and usefulness of a stationary principle for the running of the universe is influenced by the success of other such principles. These are the minimum potential energy principle, the maximum entropy principle, and Fermat's principle of least action.

To see how to proceed, we set up as close an analogy between entropy production and potential energy as seems possible. We consider a set of elastic strings of different elastic constants but all of negligible natural length. These are initially assumed to obey Hooke's Law; they represent resistors, the length representing the current and the tension the applied potential. The ends of the strings are connected together in a definite way to represent the resistance network. The string system is suspended from one point and supplied with weights which represent the emfs at different points of the network; hence it hangs in one vertical line. The correspondence between the static and dynamic systems is now virtually exact; it will be investigated further in section 2.17.

The equilibrium situation may be examined in terms of virtual work. Equilibrium is defined by two conditions; firstly that there is no net force at any point and secondly that the potential energy field is conservative, leading to the condition that the net force round any loop is zero. These are analogous to Kirchhoff's laws. We may now apply the principle of stationary potential energy; this states that the energy stored is equal to the lost potential energy of the weights. The principles deduced are that for part of the string net where there are no weights, $\sum \frac{1}{2} \lambda x_r^2$ where λ is the elastic constant and x the length, is to have a stationary point, while for a part of the net which includes weights, $\sum W_r x_r - \sum \frac{1}{2} \lambda x_r^2$ is to have a stationary point. These represent applications of the principle of virtual work. Now the vital point

is that the principle of minimum potential energy and the principle of virtual work may still be applied when the strings are not perfectly elastic. All that is necessary is that the total energy stored in a string should be written as $\int_0^x \lambda z dx$.

It becomes at once apparent how the maximum entropy production principle can be generalised to deal with non-linear resistance networks. All that seems to be required is that the entropy production should be written as $\int_0^i 2Ri di$, which gives the accepted result Ri^2 when the resistances are ohmic. This substitution in the earlier equations 2.1, 2.2 prevents the appearance of resistance derivatives, so that Kirchhoff's Laws are consistent with this new form of the entropy production principles. It enables the local slope resistance to be employed rather than the origin based resistance, as is obviously required by Kirchhoff's Laws.

2.10 Justification of the New Entropy Production Expression.

The conclusion derived in the last section, that it is useful to define the entropy production rate in a resistor as $\int_0^i 2Ri di$ rather than as Ri^2 is quite radical. It is clear that we cannot proceed to substitute this into a general principle of maximum entropy production without first examining its implications for the basic structure of entropy theory. In view of the usefulness of the present structure in physics, any lack of self-consistency in the modified structure would suggest that the aim of defining the steady state in terms of

a stationary point in the entropy production rate would have to be abandoned.

It is not intended to work back step by step from the new production rate to see what changes would be useful. Instead we postulate a change to the foundation of entropy theory and show where this leads. Discussion of this foundation was deferred in the derivation of the minimum entropy production principle (section 2.5) to simplify the presentation there. The foundation in question is the quadratic entropy approximation. We discuss this, following the methods and notation of P. Mazur (1960, 229).

A microcanonical ensemble is considered with N particles described by position and momentum vectors $(r^N, p^N) = (r_1, r_2, \dots, r_N; p_1, p_2, \dots, p_N)$, the energy being in the range $E, E + dE$. The system is described by n macroscopic variables $A_1, \dots, A_n$, (where $n \ll N$), as $A = A(r^N, p^N)$; the probability of finding the system in a state in the range $A, A + dA$ is defined as:

$$f(A)dA = \mathcal{N}(A)/\mathcal{N} \quad ,$$

where $\mathcal{N}$ is the volume of phase space in the energy shell $E, E + dE$ and $\mathcal{N}(A)$ the volume in the subrange $A, A + dA$. We assume a Gaussian distribution about mean values A_i^0 and write $\phi_i = A_i - A_i^0$. Then

$$f(\phi) = \sqrt{(\text{Det}(\varepsilon_{ij})/(2\pi k)^n)} \cdot \exp(-(1/2k) \sum_{i,j=1}^n \varepsilon_{ij} \phi_i \phi_j) \quad .$$

This is now compared with the entropy value. Boltzmann's entropy postulate is:

$$S(\phi) = k \log \mathcal{N}(\phi) \quad ,$$

so that: $\Delta S = S(\phi) - S(0) = k \log(\mathcal{N}(\phi)/\mathcal{N}) = k \log f(\phi) + c$.

or: $f(\phi) = f(0) \exp(\Delta S/k)$.

Hence: $\Delta S = - \frac{1}{2} \sum_{j=1}^n E_{ij} \phi_i \phi_j$.

This may be regarded as a form of Taylor series expansion up to order ϕ^2 for $S(\phi)$ around its maximum value $S(0)$. Any appreciable departure from equilibrium may be looked at as requiring the addition of higher order terms in the Taylor expansion, or alternatively as destroying the validity of employing a Gaussian distribution.

The Gaussian or normal distribution thus forms the foundation to entropy production theory and it is precisely this which we wish to question. The derivation of this will now be sketched in; notation for the one dimensional case is drawn from H. Jeffreys (1961, 230) and for the extension to many dimensions from R. von Mises (1964, 231).

In one dimension the distribution is based on the binomial theorem. If the chance of success in an event H is x, of failure $1 - x$, then the probability of $l + m$ trials giving l successes and m failures is:

$$P(l, m | H) = (l + m)! / (l! m!) \cdot x^l (1 - x)^m$$
 .

This represents a term in the binomial expansion of $(x + (1 - x))^{l+m}$. Expansion by the use of Stirling's theorem gives (Jeffreys, l.c. 1961, 232, p.62):

$$P(l, m | H) = 1 / \sqrt{(2\pi(1 + m)x(1 - x))} \cdot \exp(-(1 - (1 + m)x)^2 / (2(1 + m)x(1 - x)))$$
 .

This may be written as:

$$p(x) = h/\sqrt{\pi} \cdot \exp(-h^2(x-a)^2) ,$$

where $x = a$ is the position of the maximum of the density function and h the measure of precision. Extension to multidimensional space provides the product of the probabilities of each individual dimension together with cross-terms, giving a sum within the exponential (von Mises l.c. 1964, 233, p.424); the result is:

$$f(\sigma) = f(0) \exp(-1/2k \cdot \sum_{i,j=1}^n \delta_{ij} \sigma_i \sigma_j) .$$

It is now suggested that a more general distribution function is given

$$\text{by: } f_1(\sigma) = f_1(0) \exp(-1/2k \cdot \sum_{i,j=1}^n \int_0^{\sigma_j} 2\delta_{ij} \sigma_i d\sigma_j) .$$

In one dimension this is equivalent to:

$$p_1(x) = p_1(a) \exp(-\int_a^x 2h^2(x-a)dx) .$$

It is clear that this form of distribution is related to a stochastic process. The probability of a given set of events is provided by the probability of building it up from equilibrium when the individual event probabilities are not independent of the configuration at any stage; in the one dimensional case, this is when x depends on l, n . Thus we are led to consider the probability of one set of events as compared with a neighbouring set.

$$P(l, n | H) = f_1 \exp(-\frac{(1 - (1+n)x)^2}{2(1+n)x(1-x)})$$

$$P(1 - \delta l, n + \delta l | H) = f_2 \exp(-\frac{(1 - \delta l - (1+n)x)^2}{2(1+n)x(1-x)}) .$$

Then we may say that the relative probability of $P(1 - \delta l, m + \delta l | H)$ to $P(1, m | H)$ is given by their ratio:

$$f_3 \exp\left(-\frac{\delta l(1 - (1 + m)x)}{(1 + m)x(1 - x)}\right) .$$

So a probability may be related to the maximum probability as:

$$P(1, m | H) = f_4 \cdot P(\max) \cdot \exp\left(-\int_{l(\max)}^l \frac{(1 - (1 + m)x)}{(1 + m)x(1 - x)} dl\right) ,$$

or more simply:

$$P(x) = f_5 \cdot P(\max) \cdot \exp\left(-\int_a^x 2h^2(x - a) da\right) .$$

Generalization to many dimensions gives the desired result:

$$f(\phi) = f_1(0) \exp\left(-1/k \cdot \sum_{j=1}^n \int_0^{\phi_j} g_{1j} d\phi_j\right) .$$

The normalization factors f_n are defined with the aid of integration over all space. For entropy applications, all calculations are made with respect to the probability maximum, which is taken to represent the equilibrium point. A problem which arises in considering the distribution in terms of a stochastic process is whether the result is independent of the path. The expression arrived at is symmetric in all terms; assuming the g_{ij} to be single-valued, this integral assigns a unique value to each probability, so that the difficulty seems to resolve itself.

2.11 The Physical Basis of the Entropy Alteration.

The formula for the entropy has thus been varied from

$$S = - \frac{1}{2} \sum_j \epsilon_{ij} \delta_i \delta_j$$

to

$$S = - \frac{1}{2} \sum_j \int \epsilon_{ij} \delta_i \delta_j$$

in order to cope with situations where the probability of a given system state is not independent of the probability of a state with neighbouring parameters. We point out that the δ 's represent the values of macroscopic parameters for ultra-small regions of the crystal. Physically a given state develops by a temporal sequence from a neighbouring one, instead of just happening at random, so we may hope that an entropy computation which employs a stochastic process from equilibrium will be more satisfactory than one which uses random probabilities; the two should then come to the same value when non-random influences are removed. It is our claim that the altered entropy formula meets these conditions.

The non-random influence may be said to be local order. Imagining a set of δ 's to represent the departure from equilibrium of the temperature for a large number of small regions in a crystal, it is clear that a state where neighbouring regions have similar temperature values is more probable than a random array of the same values. In normal circumstances this ordering is of such a short range that it can be neglected in the entropy sum; such is the case for linear processes. However for cooperative phenomena, the effect

of order on the entropy may be expected to be considerable, and to lead to non-ohmic processes. In this thesis, we shall be concerned with two types of ordering in negative resistance, in real space and in momentum space (sections 5.3.9, 6.3).

2.12 Employment of the Entropy Alteration.

We first show the orthodox treatment of de Groot (l.c. 1959, 234). The quadratic entropy expression is given from section 2.10

$$\text{as:} \quad S = -\frac{1}{2} \sum_{ij} \epsilon_{ij} \dot{x}_i \dot{x}_j \quad .$$

The fluxes are defined as:

$$J_i = \dot{x}_i \quad .$$

and the forces as:

$$X_i = \frac{\partial}{\partial x_i} (\Delta S) \quad .$$

$$\text{Then:} \quad \frac{d}{dt} (\Delta S) = \sum_i \frac{\partial}{\partial x_i} (\Delta S) \frac{dx_i}{dt} \quad ,$$

$$S = \sum_i J_i X_i \quad .$$

This provides the starting point for the discussion of minimum entropy production in section 2.5; it is precisely because this has been found to be inadequate for non-linear processes that a more generally valid treatment has been attempted.

The new entropy expression is given by:

$$S = -\frac{1}{2} \sum_{ij} \int_0^{\dot{x}_j} \epsilon_{ij} \dot{x}_i d\dot{x}_j \quad .$$

Differentiation produces an entropy production rate:

$$\dot{\Delta S} = - \frac{1}{T} \sum_j \int_0^{J_j} \epsilon_{ij} \sigma_i d\sigma_j \quad ;$$

$$T\dot{\Delta S} = \sum_i \int_0^{J_i} 2\chi_i dJ_i \quad .$$

It is necessary here to assume that the entropy production rate due to the rate of change of the forces σ_i is small compared with the entropy production rate due to the currents which they produce. In the limit of the steady state this becomes exactly true, the forces σ_i becoming constraints such as applied emf's which are assumed not to run down. This is the situation for which we are attempting to obtain a stationary principle. It is clear that the new entropy principle is also applicable far from the steady state for non-linear phenomena, but for these cases the entropy production is given by:

$$T\dot{\Delta S} = \sum_i \int_0^{J_i} 2\chi_i dJ_i + f(\dot{E}, E, \dot{\sigma}_i, \sigma_i) \quad .$$

Reverting to quasi-steady state conditions, we define the phenomenological laws as:

$$J_i = \sum_j L_{ij} \chi_j \quad ,$$

or inverting the matrix to work in terms of flow resistances,

$$\chi_i = \sum_j R_{ij} J_j \quad .$$

Hence:
$$T\dot{\Delta S} = \sum_i \int_0^{J_i} 2R_{ij} J_j dJ_i \quad .$$

Here the cooperative features have been transferred to the coefficient R_{1j} . In an electrical network, cross-effects between different components can be ignored. This results in the entropy production being given for the network as the sum of the production rates in individual components, these being:

$$1/T \cdot \sum_0^1 2R_{1j} dI \quad .$$

Looking back to section 2.9, it is to be seen that this is the form required for a stationary principle to apply to a non-ohmic network, and also that suggested by analogy from the equilibrium of a net of elastic strings.

2.13 Appraisal of the Entropy Production Calculation.

The derivation of the entropy production rate as $1/T \cdot \sum_0^1 2R_{1j} dI$ is not completely satisfactory. Firstly there is the trouble over the differentiation, where $f(\dot{g}, g, \dot{x}_i, x_i)$ is neglected; this is explained away on physical grounds and seems to limit the use of the entropy expression to slowly varying conditions. Secondly there are problems over the summation notation. The aim has been to provide a derivation which is physically valid, notationally self-consistent and which reduces to de Groot's form (l.c. 1959, 235) for linear laws. This has been made more difficult by inconsistencies in de Groot's notation; his analysis on pages 6-7 is:

$$\Delta S = - \frac{1}{2} \sum_{i,k} \epsilon_{ik} a_i a_k ,$$

$$\chi_i = - \sum_k \epsilon_{ik} a_k = \frac{\partial}{\partial a_i} (\Delta S) ,$$

$$J_i = a_i ,$$

$$\dot{\Delta S} = - \sum_{i,k} \epsilon_{ik} \dot{a}_i a_k = \sum_i J_i \chi_i .$$

This analysis is valid if the meaning of a summation is treated with some degree of ambiguity. Licence of this sort is required in dealing with the relationship between homomorphic ($\epsilon_{ii} a_i^2$) and heteromorphic ($\epsilon_{ij} a_i a_j$) terms, as is used by de Groot on p.21 (l.c. 1959, 236). A more detailed investigation of these sums is required than is given here; similarly the search for a completely new method of derivation of the entropy production rate, instead of the rather conventional approach adopted here is outside the scope of this work. An approach from the standpoint of quantum statistical mechanics should be possible, following on from the work of I. Prigogine, G. Severne (1963, 237) and H. Baur, J. R. Jordan, P. C. Jordan, J. E. Mayer (1965, 238).

Extension of the entropy derivation gives for a reversible process:

$$T\Delta S = \Delta U - \Delta \sum_i \int_0^{n_i} \mu_i dM_i .$$

This shows that departure from the normal TdS equation is caused by some of the chemical potentials not being truly intensive quantities. Under these circumstances it is imperative to use the local slope resistance, rather than the origin based slope resistance for the

computation of quantities such as the entropy. The preceding sections have provided nothing but an elaborate justification for employing a slope resistance.

2.14 Applicability of Entropy Production Principles.

The aim of the preceding five sections has been to obtain a stationary principle for the description of the steady state for irreversible processes. No originality is claimed for the underlying choice of the entropy production rate for this function; we quote Ziman (l.c. 1956, 239): "It appears to be a well known law of physics, yet for some reason it is not stated anywhere." We have shown that considerable changes in the method of computation of the entropy production are required for the principle to be applicable to non-ohmic networks. Changes have been made employing the following points:

- (1) The expressions for the entropy and for its production rate are symmetrical in the variables.
- (2) The expressions reduce in the case of linear phenomenological equations to the well-founded expressions of linear irreversible thermodynamics.
- (3) The expressions are valid for non-linear isothermal single-component flows, as in the case of electricity in a network.

Irreversible thermodynamics is largely concerned with heat flows and cross-flows which have not so far been considered. The new expressions may be applied to these, with the following questions being

raised:

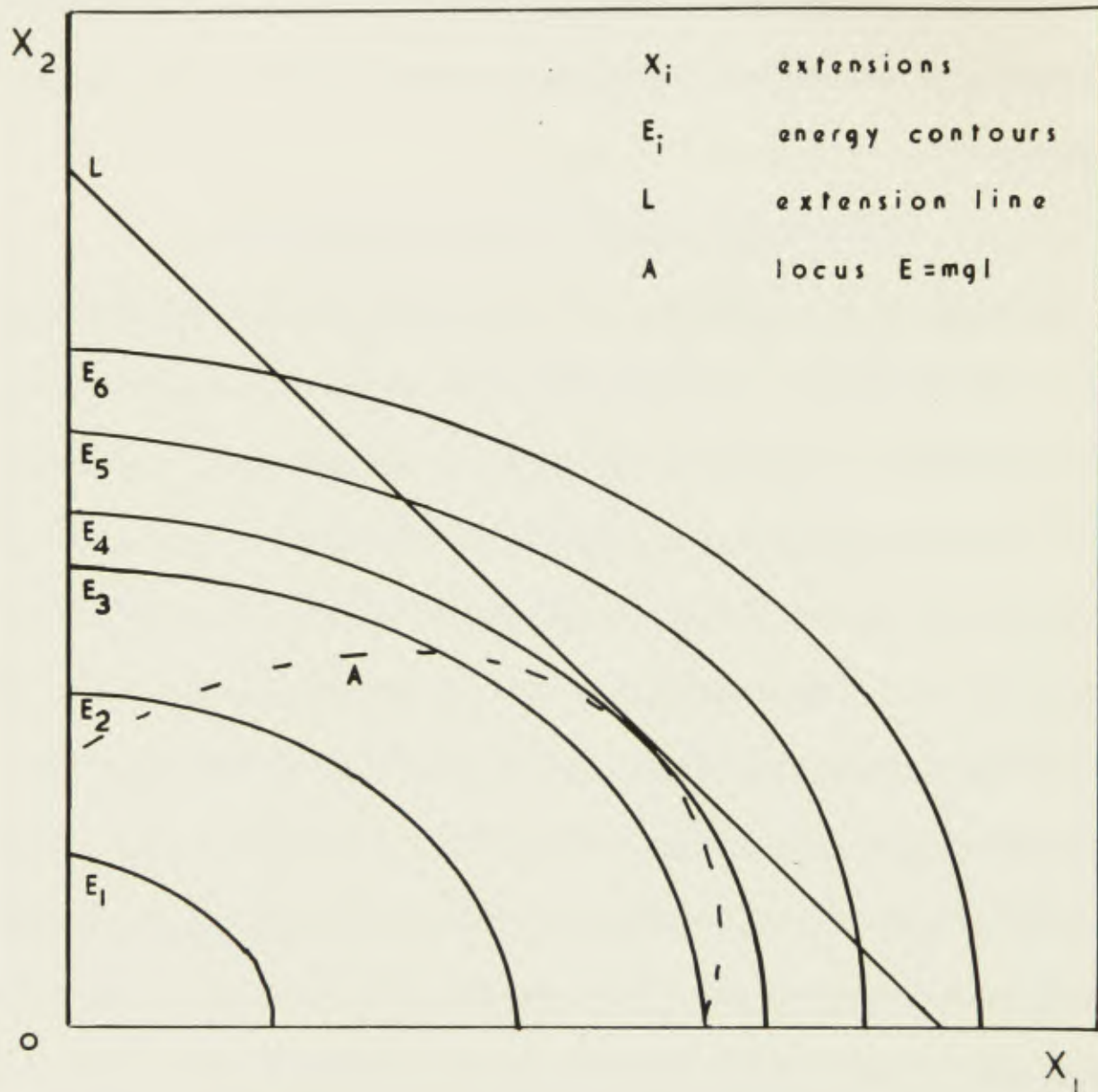
- (1) Are deductions from the expressions self-consistent?
- (2) Are deductions consistent with the main body of the theory of physics?
- (3) Are deductions experimentally confirmed?

We examine these in outline. The symmetry of the expressions may be expected to result in uniquely defined cross-flow parameters, so that self-consistency should be no problem. Linear theory is already embraced by the new expressions. So far little theoretical interest has been shown in the analysis of non-linear flows. Equations for heat transfer in non-ohmic conductors have been derived by C. A. Domenicali (1953, 240; 1954, 241) and these have been extended by D. T. J. Hurle, J. B. Mullin and E. R. Pike (l.c. 1965, 242; 1965, 243). Such work is based on the orthodox entropy production expressions so that it would not be surprising if the results were not consistent. Indeed an investigation of the differences or similarities in the expressions derived would be of considerable interest. However the possible existence of rival theories in this matter does not in itself throw any light on their respective merits. For this, experimental results on non-linear cross-flows, such as the thermoelectric properties of non-ohmic resistances, are required and these do not seem to exist. This kind of subject may be expected to come into vogue in view of the increasing interest in bulk negative resistance. Hence we feel justified in leaving the general theory of cross-flows and concentrating



Figure 2.3

Energy Contours for Two Elastic Strings



Minimum Principle --- the equilibrium point on line L for fixed x_1+x_2 is where it touches a constant energy ellipse; then the stored energy is a minimum ($E_4 < E_5, E_6$).

Maximum Principle --- the equilibrium point for the case where the weight hangs freely, when the stored energy and the potential energy are equal, is given by the tangential situation on curve A; other allowed points are inside so the stored energy is a maximum ($E_4 > E_3, E_2$).

on the application of the stationary entropy production principle to isothermal non-ohmic networks, content that the principle has some fundamental basis and that further analysis, though useful and indeed relevant, would not serve to determine its validity.

2.15 The Relationship between Maximum and Minimum Principles by Analogy.

A primary feature of this work on irreversible thermodynamics is the solution of the paradox of the apparent simultaneous validity of principles of minimum and maximum entropy production. Both were found to be valid for linear phenomenological laws and we have shown that with a revised definition of entropy production they may be extended to non-linear processes. To show quite how this arises we return to the elastic string analogy. The simple case is considered of two strings of negligible natural length joined in series to support a given weight. The energy stored in the two strings is plotted by contour for different extensions of each x_1, x_2 . For elastic strings of different elastic constants, the result would be a set of central ellipses. The equilibrium situation depends on how the system is constrained. One possibility is to specify that the weight be held at a fixed height, so that $x_1 + x_2 = \text{constant}$; then the strings adjust their lengths so that the energy stored is a minimum. This is shown in figure 2.3 as the point at which the constant energy contours touch the constraint line. This gives the principle of minimum stored energy. Alternatively it may be specified that the weight

is to hang freely. Then the strings adjust so that the energy stored in the strings is a maximum, subject to the condition that the energy stored in the strings is equal to the potential energy of the weight. This is shown in figure 2.3 as the point where the line and contour meet tangentially. Hence maximum and minimum principles represent different methods of using the same diagram. A similar line of approach may be applied to the entropy production principles. The problem here is that it is very much easier to visualize an entropy contour diagram than an entropy production rate one; yet no entity corresponding to the entropy exists in the string analogy, as the energy corresponds to the entropy production. Thus while it is valuable to know that maximum and minimum properties can exist side by side, it is useful to construct an entropy picture to see how they apply.

2.16 Topological Representation of Entropy Production.

We now proceed to construct a suitable topology for entropy production representation. This will be seen to have many features in common with the topology from which the production rate expression was justified from statistical principles in section 2.10.

An isolated subsystem of the universe is chosen, to be called 'the system', (e.g. a fixed volume containing gas). This system is split up into a very large number N of subsystems (volumes); in each subsystem, each macroscopic state variable is defined (temperature, pressure, density, etc.), n in all. Thus the state of the whole

system is defined by Mn parameters relating to subsystems, where Mn is extremely large. We now assume that an expression of the type derived earlier exists for computing the total entropy of the system from a knowledge of the values of the parameters. An $Mn + 1$ dimensional space is employed to plot the total entropy against all possible values of all Mn parameters. In fact we are only interested in a fairly small compact region of this volume, where neighbouring subsystem parameters have similar values, elsewhere the entropy values being ridiculous.

We now set the system off from a given macroscopic state --- this is the reason for choosing macroscopic parameter values; that is from a given point on the surface. How does it evolve? By the principle of maximum entropy production, the system remains on the surface to satisfy the conservation principle (by which production equals flow) and takes the path along which the entropy increases fastest; this is not necessarily the steepest curve, but the one which maximizes $\text{grad}_k S \cdot dk/dt$, k being the direction vector. Equilibrium is reached when the system attains a point where all values of $\text{grad}_k S \cdot dk/dt$ are less than zero. This may not be the maximum entropy of the surface but it represents the maximum attainable value.

Suppose now that the system is no longer isolated, but that a certain number of the parameters (in practice some of the boundary ones) are artificially maintained at fixed values by external agency. On the $Mn + 1$ dimensional surface, the locus represented by the fixing

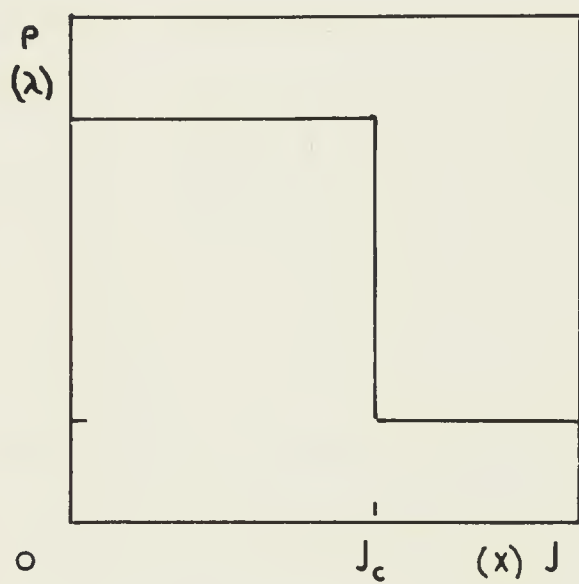
of certain coordinates is a subsurface. The system is set off from some point on this subsurface --- how does it evolve? There are two apparent answers. Firstly it may move along the subsurface to the maximum attainable point on the subsurface R . This is in the direction k given by the maximum of $(\text{grad}_k S \cdot dk/dt)_R$. Secondly it may evolve in the direction given by the component of k parallel to the subsurface where k is the direction of maximum production in the absence of the constraint.

The following argument shows that the second is the case. Consider a system whose constraint is that one subsystem is kept at constant temperature by contact with a large isolated heat reservoir at an elevated temperature. We compare the evolution of the whole system with that of reservoir plus constrained system. By postulate, the whole system evolves freely at maximum entropy production. The flow of heat at constant temperature out of the reservoir into the constrained system means that the reservoir is absorbing entropy which is being produced by the system. In fact the system is evolving as if it were free, but all velocities normal to the subsurface are effectively zero. In practice the reservoir is not infinite, so that some progress is made in this direction (i.e. the reservoir temperature falls marginally) but the entropy contours are so steep that this can be neglected. Thus the system is constrained to move on the subsurface. At every instant the system evolves in the direction of maximum entropy production, but because of the difference in the scale

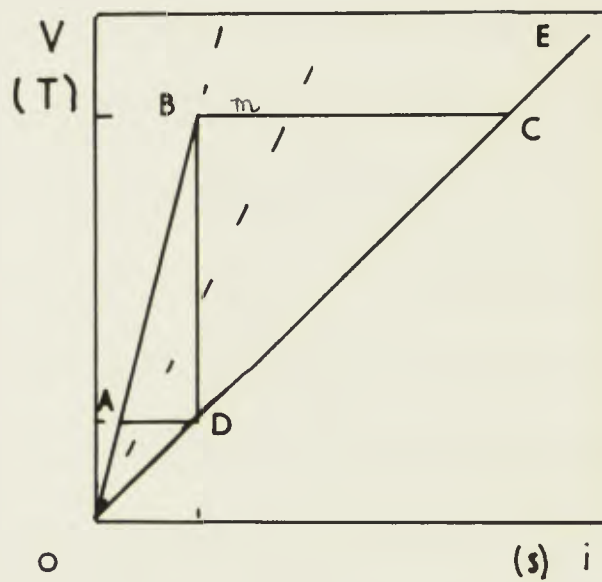
Date	Particulars	Amount

Figure 2.4

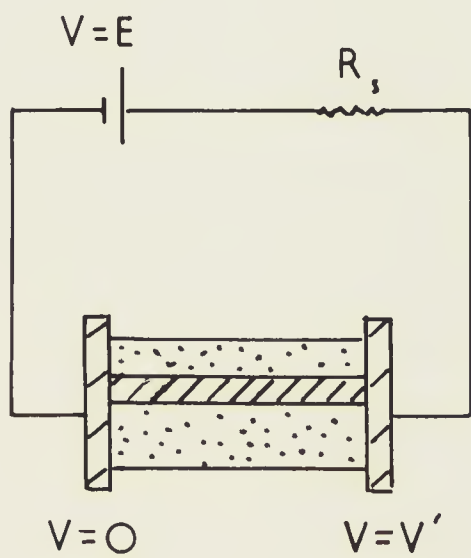
Filament Diagrams



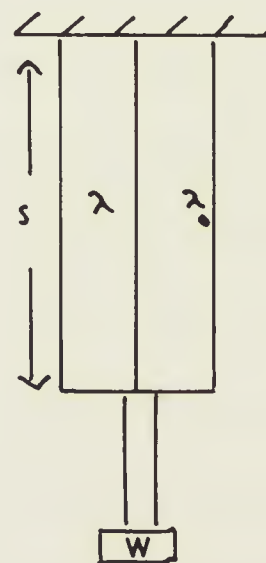
(a)



(b)



(c)



(d)

in the constrained direction, it seems to be moving in the 'plane' normal to this. Thus in computing the direction of motion in this 'plane', it is the total entropy production rate which is to be maximised, not just the production rate in the constrained 'plane'. This suggests that under the action of constraints the system may even move in a direction which decreases the system entropy as seen in its 'plane' because this maximizes the total rate. The separation of matter due to Knudsen gas flow and thermal diffusion are examples of this (I. Prigogine 1955, 244, p.84). This analysis gives both logical and experimental justification for the choice of the second answer. It is a point which is felt to have very considerable importance in practical application.

2.17 Application to Filament Formation.

Consideration is first made of filaments because these are simpler than domains. Diagrams comparable with those at the beginning of this chapter are shown in figures 2.4a, b. For negative resistance with filament formation, the resistivity has to be a sharp function of the current density, as was shown in section 1.2. Figure 2.4a shows a step variation of the resistivity at a critical current density. The resulting i - V curve for the whole crystal is shown in figure 2.4b, the equilibrium paths being OA and CB. Any point within the area ABCD is available for a suitable choice of filament structure. The dotted line is the locus of conditions such that the ratio of high to low

current density is $m : 1 - m$. We point out that even if the resistivity change is less spectacular than that of figure 2.4a, filament formation with a finite allowed area (figure 2.4b) is still available. Irreversible thermodynamics is now to be applied to a determination of the steady state points for the circuit of figure 2.4c. We return to an elastic string analogy; we consider a string whose local modulus of elasticity λ as a function of local extension x is given by figure 2.4a. Then the total tension T of the string as a function of the total length s , is shown by figure 2.4b. The system being examined, figure 2.4d, comprises this string, effective modulus L , a normal elastic string, modulus λ_0 , and a weight W . Natural lengths are again assumed to be negligible. The criterion for equilibrium is that the potential energy has a stationary value with respect to extension:

$$Ws - \frac{1}{2} \lambda_0 s^2 - \int_0^s L x dx = 0 \quad ,$$

or that $W = (L + \lambda_0)s$.

This requires simply that the total tension should match the weight. But this is valid for all points in the allowed area ABCD. Hence the position of stationary potential energy does not determine the extension of the string. Any combination of weight and total length allowed by the diagram area is a state of neutral equilibrium. We draw four conclusions:

(1) This situation does not rely on a special form for the modulus curve, but only on its general shape.

(2) There is no problem involved in matching the tensions in the regions where the modulus is varying.

(3) The principle of stationary potential energy is not useful in determining behaviour to first order.

(4) The system is extremely sensitive to second order effects such as the variation in the gravitational potential.

In returning to the electrical circuit we point out that an elastic net requires inversion before it becomes equivalent to an electrical network. Thus strings in series become resistors in parallel and strings in parallel, resistors in series. This arises because it is necessary to make the tension along a string constant, or the voltage across a set of resistances the same; strings represent equipotentials while resistors represent streamlines. Hence filament formation is compared to series extension of the elastic string and the source internal resistance is represented by the string λ_0 in parallel.

The steady state for the circuit of figure 2.4c is to be determined by the entropy production principle. This is postulated to be that the rate of entropy production is to be stationary subject to the condition that the entropy production equals the entropy flow. Our aim in the preceding sections has been to show that this requires the stationary points with respect to the current of:

$$2\mathcal{E} - R_0 i^2 - \int_0^l 2Ri di \quad .$$

This is just twice the equivalent expression for the elastic string case, and leads to the condition:

$$R = (R_s + R)l,$$

which is Kirchhoff's Law. This again is valid for all points in the area ABCD so the argument is exactly as before, with similar conclusions.

(1) The situation does not rely on a special form for the resistivity curve as a function of the current, but only on its general shape. The only requirement is that negative resistance shall be possible by the criterion of section 1.3.

(2) There is no problem involved in matching filament current densities in the regions where the resistivity is varying.

(3) The principle of maximum entropy production is not useful in determining behaviour to first order.

(4) The system is extremely sensitive to second order effects such as temperature variations.

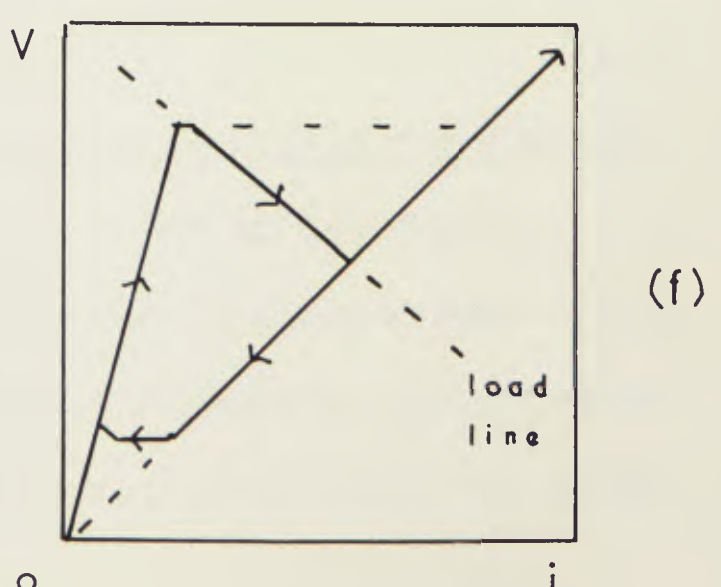
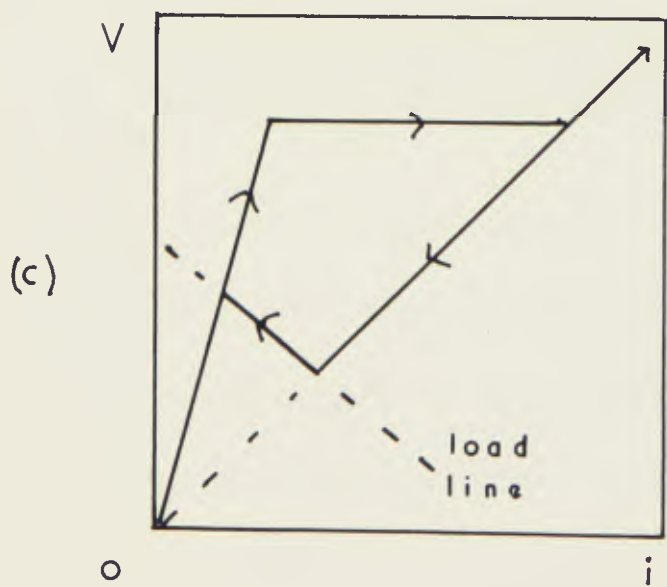
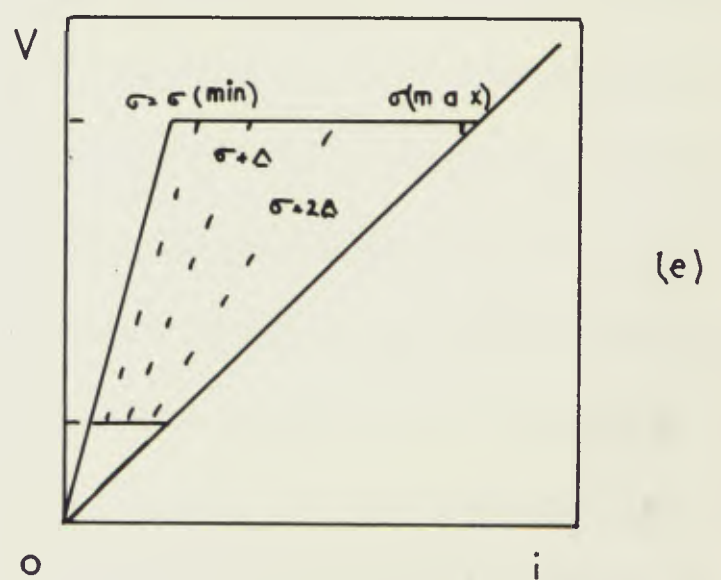
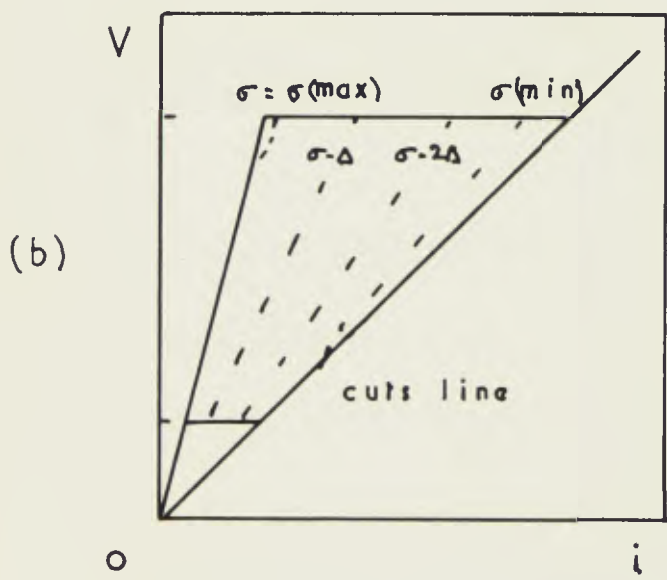
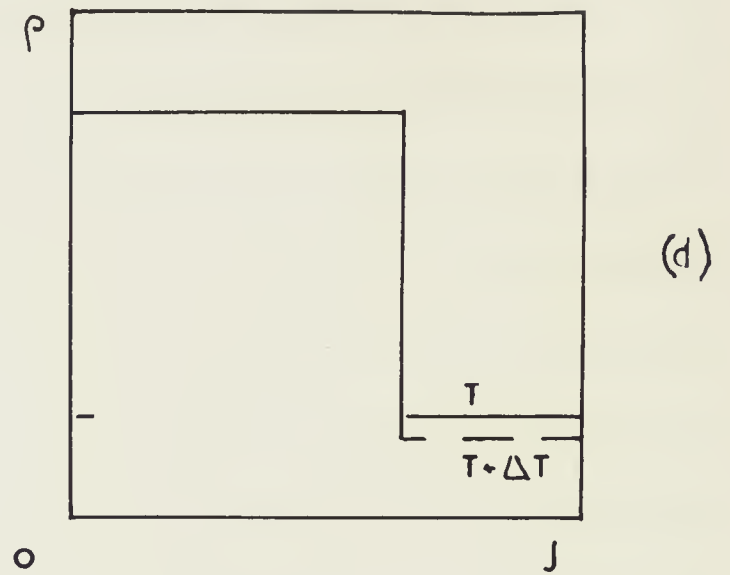
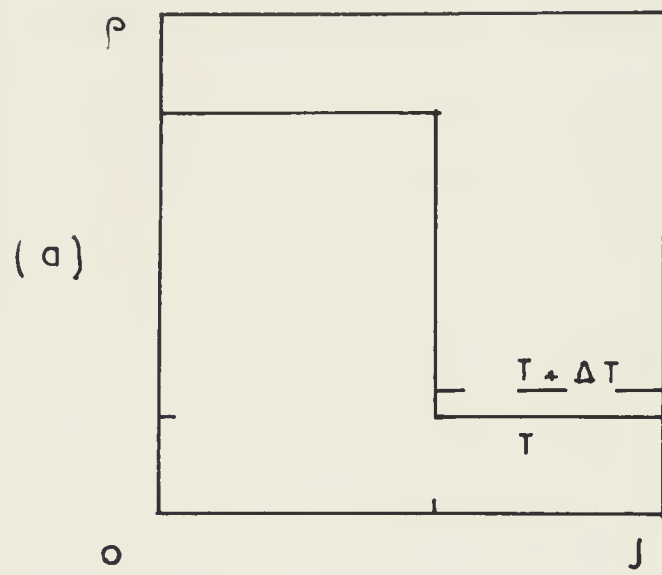
It thus seems that no progress at all has been made on the problem of determining the steady state configuration by the application of entropy production principles. To first order this is true; we are to expect that a system which obeys Kirchhoff's Laws is in technical equilibrium just as a system in which all forces are balanced is in technical equilibrium. What we are interested in is which path the crystal takes across the area ABCD when it is cycled up or down. For this we have to examine second order effects.

The nature of any spontaneous fluctuations in the current flow properties of such a crystal is of considerable significance. This is especially true in terms of the initial formation of the filament. It is reasonably certain that there is a finite minimum size of filament which can exist indefinitely against spontaneous fluctuations and diffusion effects from the filament interface region. The formation of the filament is thus likely to be inhibited for a considerable distance along the line OAB, and it may be necessary to reach B. Similarly the formation of a higher resistance filament may be inhibited along the path ECD. In this way the system resembles a superheated or a supercooled liquid in a phase diagram; the situation is rather more complex however because of the interdependence of current and applied voltage. The result is a hysteresis loop, rather than a true negative resistance.

Once a two phase system has been established, it is important to determine how the system functions. It has been shown that the entropy production at all points is the same in purely electrical terms; there are, however, bound to be other effects, of which temperature is likely to be the most important. These we denote as second order effects; they raise the apparent degeneracy of the entropy production in the region ABCD and give it a set of constant value contours. Then we may expect the steady state position to be the maximum accessible point on a given load line; a spontaneous transition along a load line is thus quite possible. Some idea of the path can

Figure 2.5

Filament Stability



be deduced for a crystal where the second order effect is temperature dominated. Temperature is likely to be important because it has a considerable effect on the resistivity and because heat conduction from high current filaments cannot be perfect. Two cases may be distinguished; the effect of a rise in local temperature may increase the resistivity (a mobility effect) or reduce it (a carrier density effect). Figures 2.5a, b, c, show the effect of increased resistivity and figures 2.5d, e, f, that of reduced resistivity. Figures b, c, show entropy contours and figures e, f, the idealized path which the crystal might be seen to take on cycling. This treatment is rather qualitative; its results will be discussed in terms of impact ionization in chapter 5.

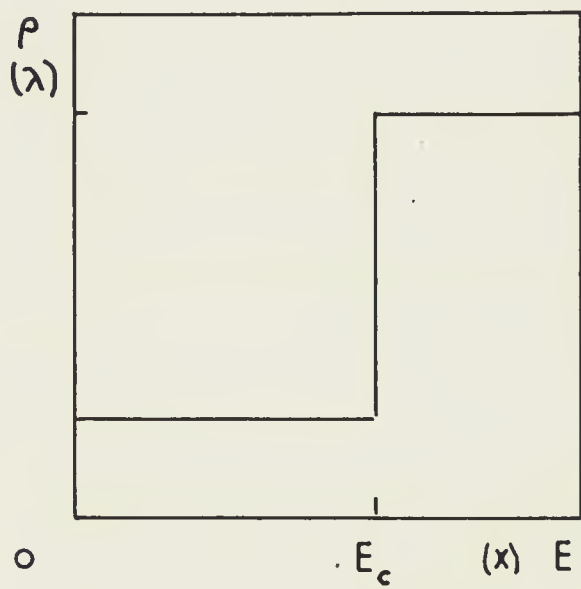
In principle the methods of non-linear irreversible thermodynamics discussed in this chapter are applicable to non-isothermal systems. The analysis involves temperature and electrical cross-flow terms giving the thermoelectric properties as a byproduct; vector notation is required since transverse pressure and temperature gradients are set up. Such an analysis will not be attempted here.

2.18 Application to Domain Formation.

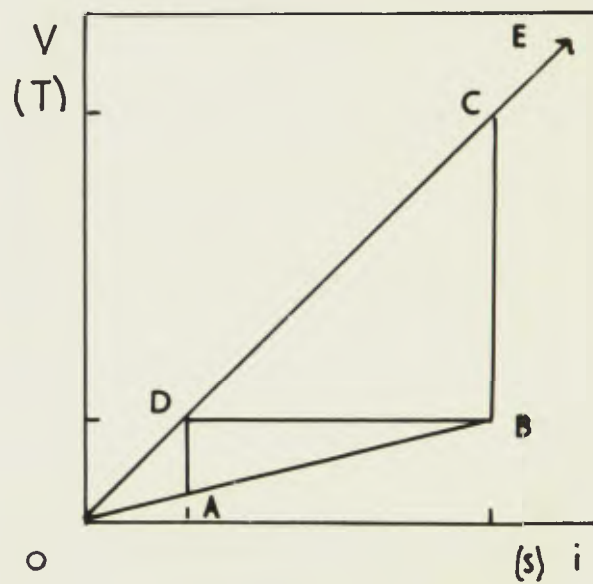
Domain formation may be treated in a form similar to that used for filament formation. Again we consider the elastic analogy. Equivalent diagrams to those of figure 2.4 are given in figure 2.6. First we consider a very large number of parallel strings, and then move

Figure 2.6

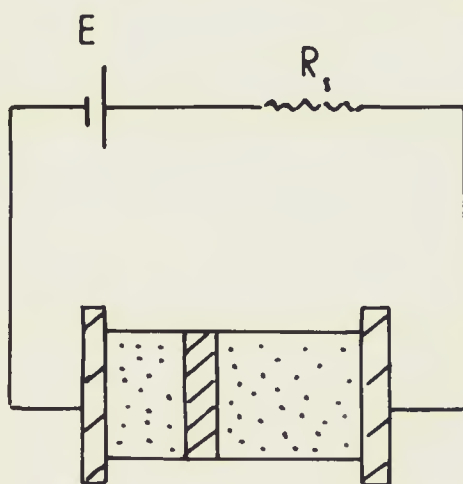
Domain Diagrams



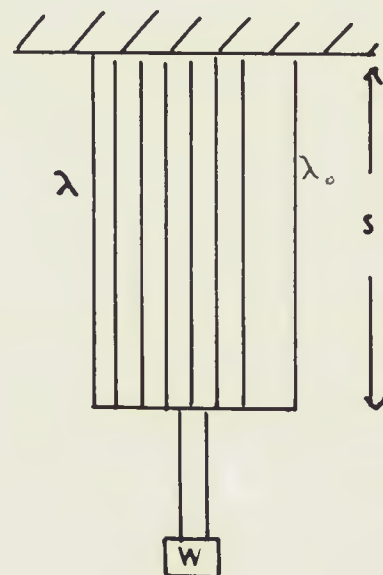
(a)



(b)



(c)

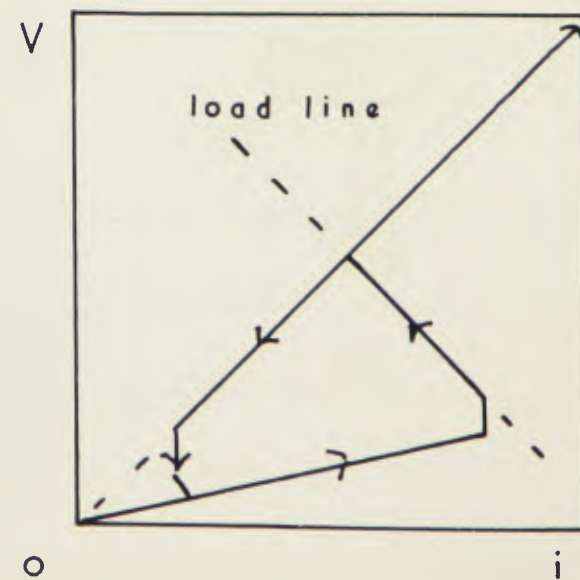
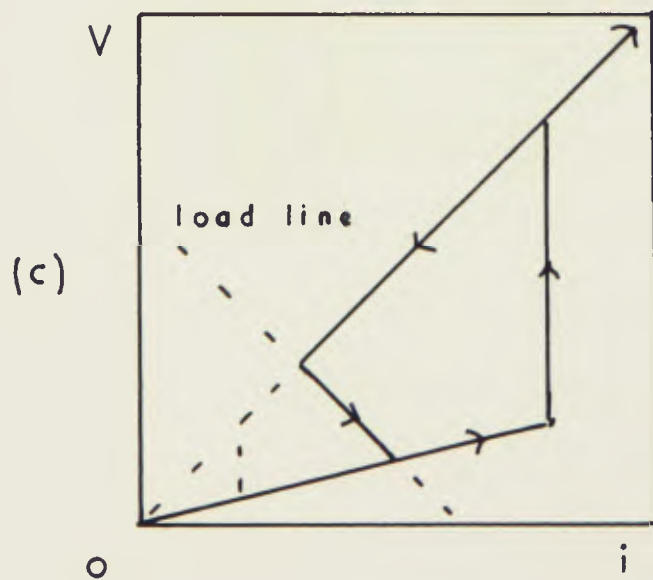
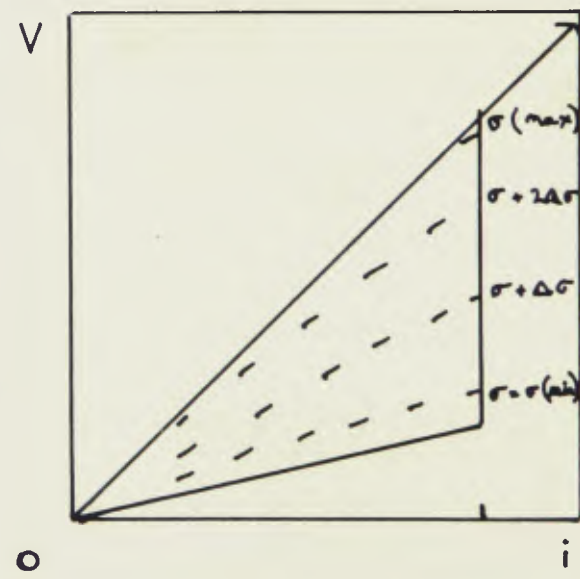
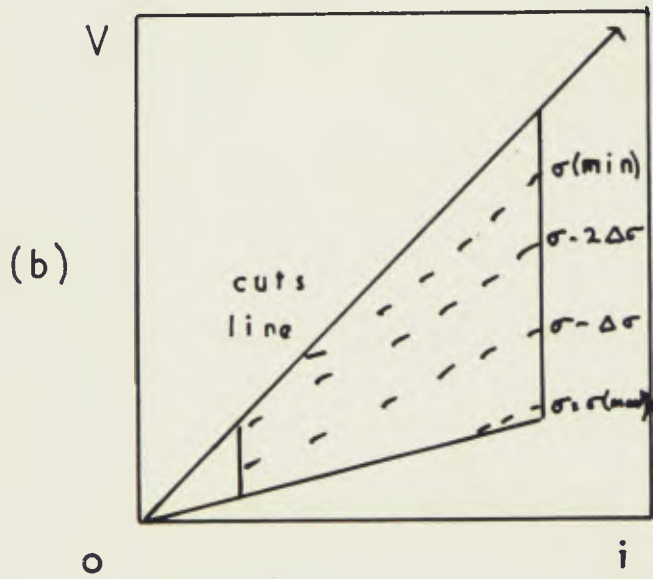
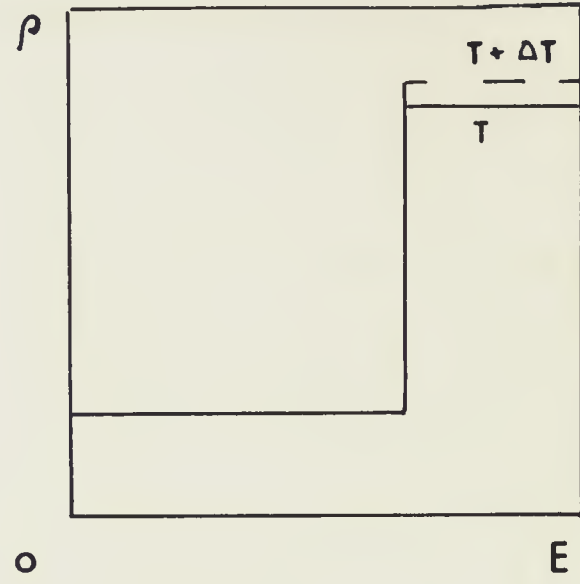
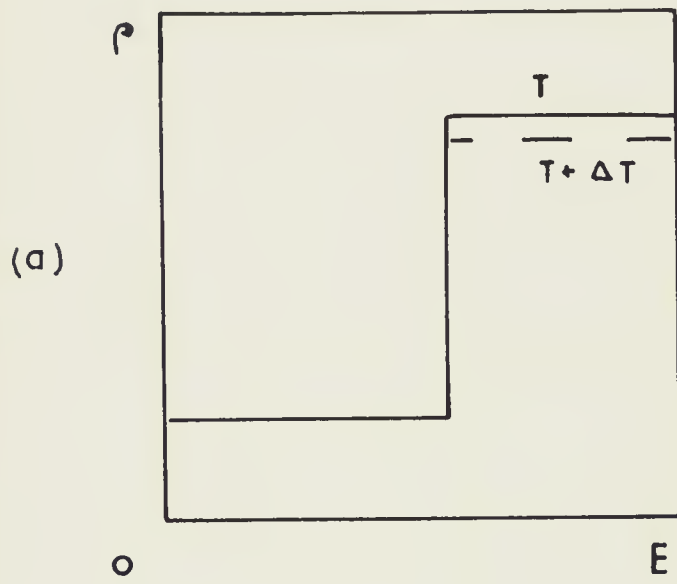


(d)

TABLE I	
Year	Amount
1950	100.00
1951	105.00
1952	110.00
1953	115.00
1954	120.00
1955	125.00
1956	130.00
1957	135.00
1958	140.00
1959	145.00
1960	150.00
1961	155.00
1962	160.00
1963	165.00
1964	170.00
1965	175.00
1966	180.00
1967	185.00
1968	190.00
1969	195.00
1970	200.00
1971	205.00
1972	210.00
1973	215.00
1974	220.00
1975	225.00
1976	230.00
1977	235.00
1978	240.00
1979	245.00
1980	250.00
1981	255.00
1982	260.00
1983	265.00
1984	270.00
1985	275.00
1986	280.00
1987	285.00
1988	290.00
1989	295.00
1990	300.00
1991	305.00
1992	310.00
1993	315.00
1994	320.00
1995	325.00
1996	330.00
1997	335.00
1998	340.00
1999	345.00
2000	350.00
2001	355.00
2002	360.00
2003	365.00
2004	370.00
2005	375.00
2006	380.00
2007	385.00
2008	390.00
2009	395.00
2010	400.00
2011	405.00
2012	410.00
2013	415.00
2014	420.00
2015	425.00
2016	430.00
2017	435.00
2018	440.00
2019	445.00
2020	450.00
2021	455.00
2022	460.00
2023	465.00
2024	470.00
2025	475.00
2026	480.00
2027	485.00
2028	490.00
2029	495.00
2030	500.00
2031	505.00
2032	510.00
2033	515.00
2034	520.00
2035	525.00
2036	530.00
2037	535.00
2038	540.00
2039	545.00
2040	550.00
2041	555.00
2042	560.00
2043	565.00
2044	570.00
2045	575.00
2046	580.00
2047	585.00
2048	590.00
2049	595.00
2050	600.00
2051	605.00
2052	610.00
2053	615.00
2054	620.00
2055	625.00
2056	630.00
2057	635.00
2058	640.00
2059	645.00
2060	650.00
2061	655.00
2062	660.00
2063	665.00
2064	670.00
2065	675.00
2066	680.00
2067	685.00
2068	690.00
2069	695.00
2070	700.00
2071	705.00
2072	710.00
2073	715.00
2074	720.00
2075	725.00
2076	730.00
2077	735.00
2078	740.00
2079	745.00
2080	750.00
2081	755.00
2082	760.00
2083	765.00
2084	770.00
2085	775.00
2086	780.00
2087	785.00
2088	790.00
2089	795.00
2090	800.00
2091	805.00
2092	810.00
2093	815.00
2094	820.00
2095	825.00
2096	830.00
2097	835.00
2098	840.00
2099	845.00
2100	850.00

Figure 2.7

Domain Stability



over into the limit of a single block of this substance with the peculiar local property shown in diagram a. Any point within the region ABCD appears to satisfy the condition for neutral equilibrium, so the conclusions are as before and second order effects are seen to dominate the path over a cycle, with fluctuations important in the initiation of the domain. Qualitative diagrams relating to temperature dependence are given in figure 2.7; a, d, show the resistivity dependence on temperature, b, e, the entropy contours and c, f, the path over a cycle. As in the case of filament formation a more complete analysis will not be attempted.

A serious complication arises in the case of the domain structure of a negative resistance crystal over oscillation. A mechanical oscillation by which the filament cross-sectional area oscillates about its steady state position is possible, but the analysis in terms of entropy production suggests that the oscillation should have no gain; similarly a stationary domain may oscillate without gain about its steady state length. A filament or domain disturbance may also be propagated through an elastic system, resulting in reflection and a standing wave pattern. In the domain structure of a negative resistance crystal however the flow properties exert themselves for the first time. By Poisson's Law the domain walls where the electric field varies contain space charges. Since all charge is moving down the crystal at a drift velocity, this suggests that the domain itself should move down the crystal at the drift velocity in one direction

only. The mechanics of such domain movement will be considered in section 6.3.

2.19 Aims and Results.

The aims of this chapter may be listed as follows:

- (1) To find the relation between the principles of maximum and minimum entropy production.
- (2) To obtain a stationary principle for the determination of the steady state for non-linear processes and negative resistance in particular.
- (3) To apply such a principle to find the stable operating points for a crystal in the presence of filaments or domains.

The questions involved are far from trivial and have received some attention in the literature without any apparent progress being made. Thus the fact that we can claim to have achieved these aims in this chapter is felt to be a considerable accomplishment. Several subsidiary results have been obtained as by-products:

- (a) Light has been thrown on the examples of M. J. Klein of section 2.6; minimum entropy production is reckoned to be valid provided it is calculated by a different formula.
- (b) It has been shown that Ziman's treatment of maximum entropy production leaves something to be desired (section 2.7) but that the maximum principle is valid provided the method of computation is changed.

(c) A justification of the new entropy production rate formula for the steady state stationary principle in terms of basic statistical thermodynamics has been produced in sections 2.10 - 2.13.

(d) The effect of constraints on a dynamic system has been demonstrated in terms of the stationary principle.

(e) We have tried to show that the principle of maximum entropy production is not primarily a principle of metaphysics ("the universe runs downhill at the fastest possible rate") but a consequence of the conservation laws of physics, just as the principle of minimum potential energy is a consequence of the requirement that all forces should be balanced in equilibrium.

The objection may be raised that the aims in this chapter have been too ambitious and that the results are too speculative for an experimental investigation of negative resistance. In practice our aim in this chapter, as elsewhere, has been to pursue any line of thought to its logical conclusion whether or not the results are surprising or even unpalatable. Even if the logic is imperfect, it is felt that the exercise has been scientifically valuable.

Finally it is interesting to note the almost universal failure of scientists to look for a scientific philosophy behind irreversible thermodynamics. Almost the only exception is found in the work of E. I. Blount; his paper (1963, 245), especially section VI, both substantiates this assertion and provides a commentary on this chapter. Had it been noted before the analysis had been completed,

we might have accepted his point of view that "not too much significance should be attributed to the maximum entropy production principle" and abandoned the search for a stationary principle.

Chapter 3

Experimental Methods.

3.1 Crystal Preparation.

A single crystal ingot of overcompensated zinc-doped germanium has formed the base material for the experimental work, supplied by Hallard as a reject in the production of fractionally undercompensated zinc-doped germanium.

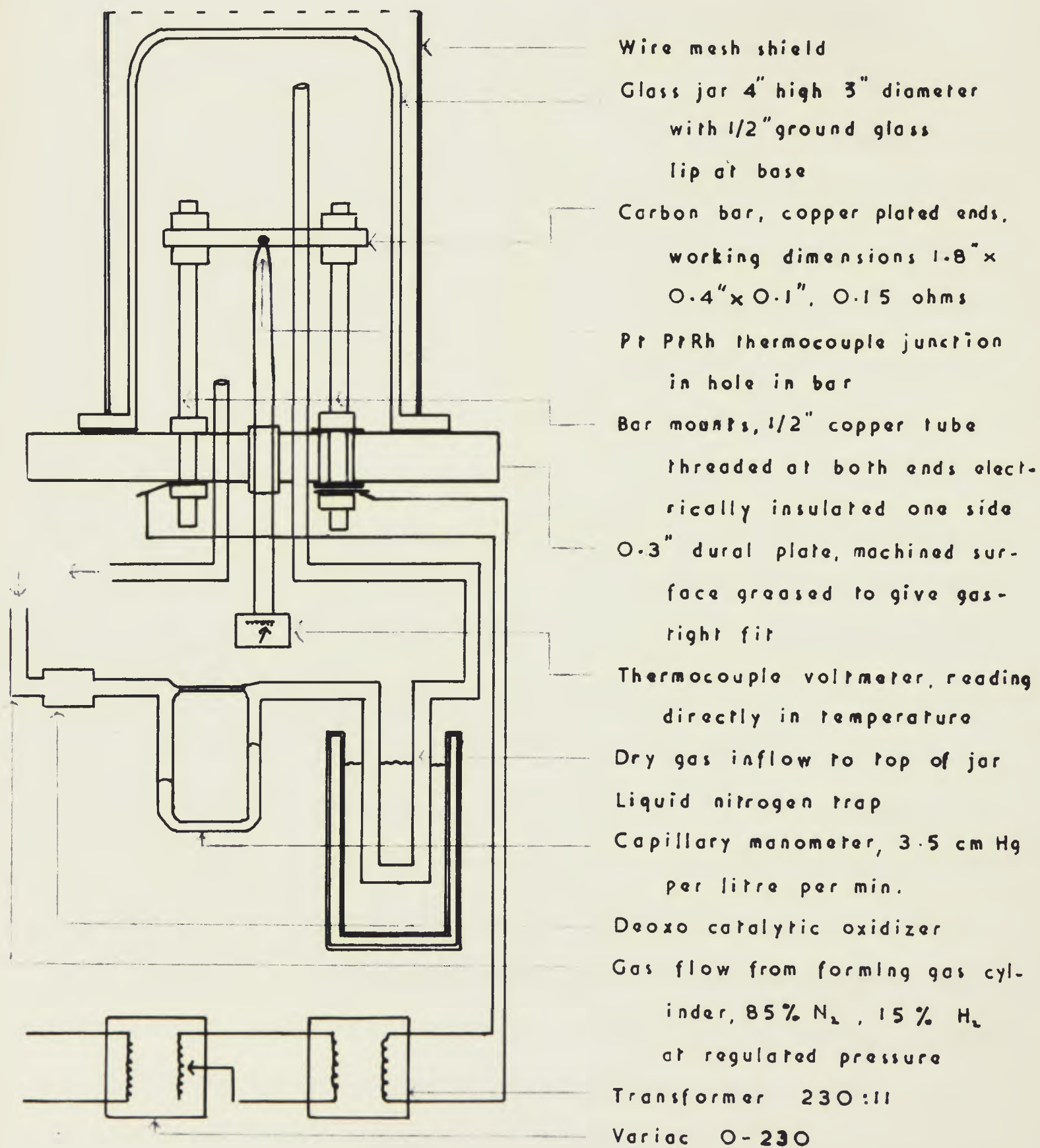
The use of extension side arms for potential and Hall contacts has been widely advocated and employed (E. H. Putley 1960, 301, p.39). This practice has not been followed at Oxford since 1958, rectangular specimens being cut with a diamond impregnated saw (A. D. Petford 1953, 302, p.20); it has been felt that the current flow is disturbed as much by a side arm as by an applied contact. In the present case, extremely small, accurately aligned contacts have been required, down to a lateral dimension of 0.4 mm. The side arms reported are considerably larger.

Specimens for which results are reported have been polished on four grades of carborundum paper and etched in CP4. This is sufficient to remove all traces of surface damage as observed under a microscope. Polishing to a higher degree of perfection before etching was thus unnecessary; without etching, damage may still remain under an apparently good surface smear. Experimental results taken for specimens over the range of surface preparation, from maximum to none at



Figure 3.1

The Furnace



Working conditions

temperature of carbon rod

400 C

gas flow

1.2 lit./min.

voltage applied

90 volts

all, suggested that surface effects were not important. This agrees with reported surface currents --- see section 7.4.

3.2 Contact Mounting.

Rectifying and ohmic contacts have been employed. Rectifying contacts have been made from $1/8$ " diameter Johnson and Matthey 0.5% gallium indium wire. All reported results have been obtained with ohmic contacts of tin-lead solder; sulphur-doped indium solder was tried but some rectification was observed at low temperature with such contacts.

The contacts were cut from the bulk inside of the wire and alloyed in a furnace designed by G. D. Peskett. The furnace is shown diagrammatically in figure 3.1. Complete cleanliness was aimed at in all phases and aspects of the operation. Alloying was carried out in an atmosphere of forming gas, a non-explosive reducing mixture, 85% nitrogen, 15% hydrogen, catalytically deoxygenized and cleaned by a liquid nitrogen trap. The specimen was heated by resting partly or wholly on an electrically heated carbon block. All contacts were alloyed from a position on the upper surface of the crystal. Early ohmic contacts were alloyed without flux. It was found that a small application of Baker's fluid did not affect the ohmic characteristic but considerably improved contact wetting. In order that contact position could be adjusted to give maximum accuracy for the Hall effect, flux was not applied until contacts had been alloyed once.

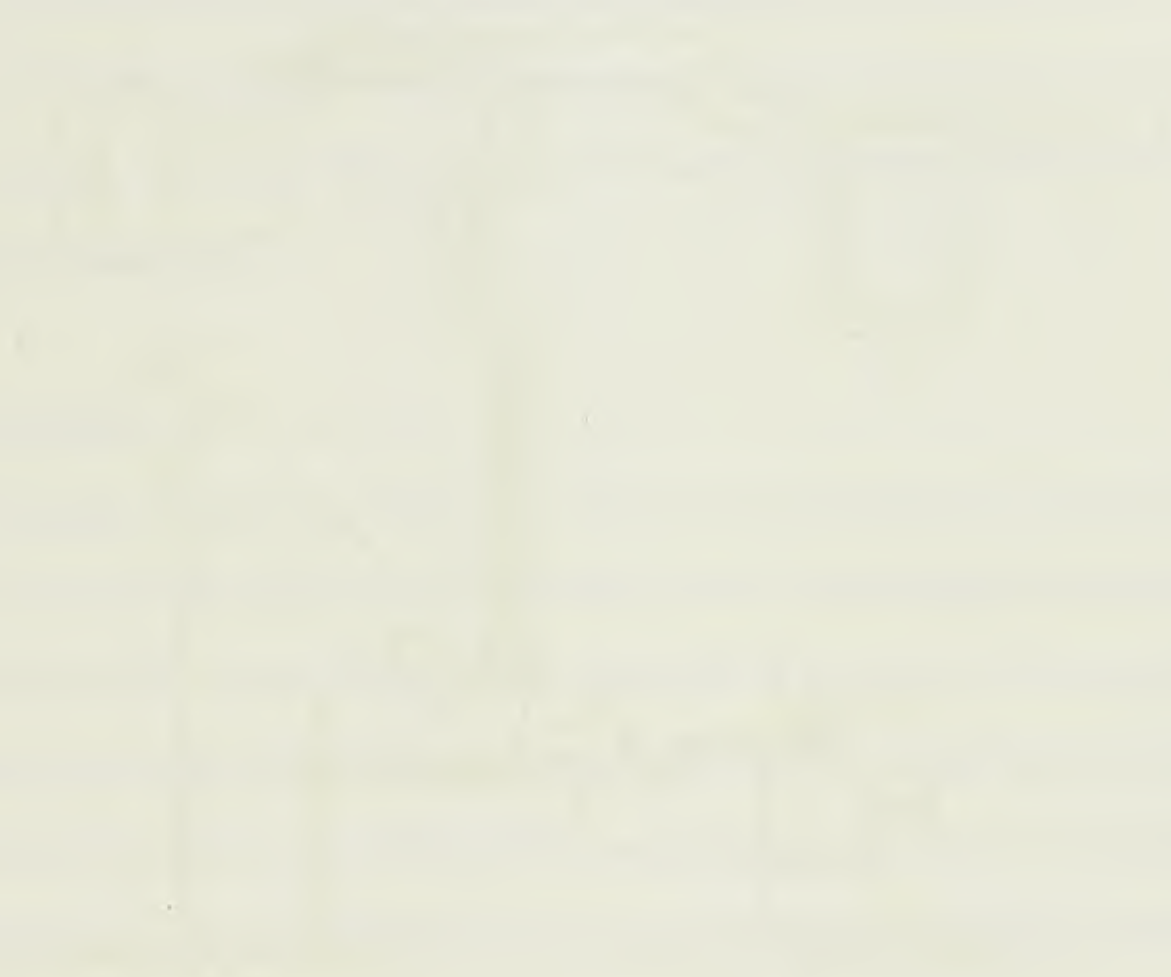
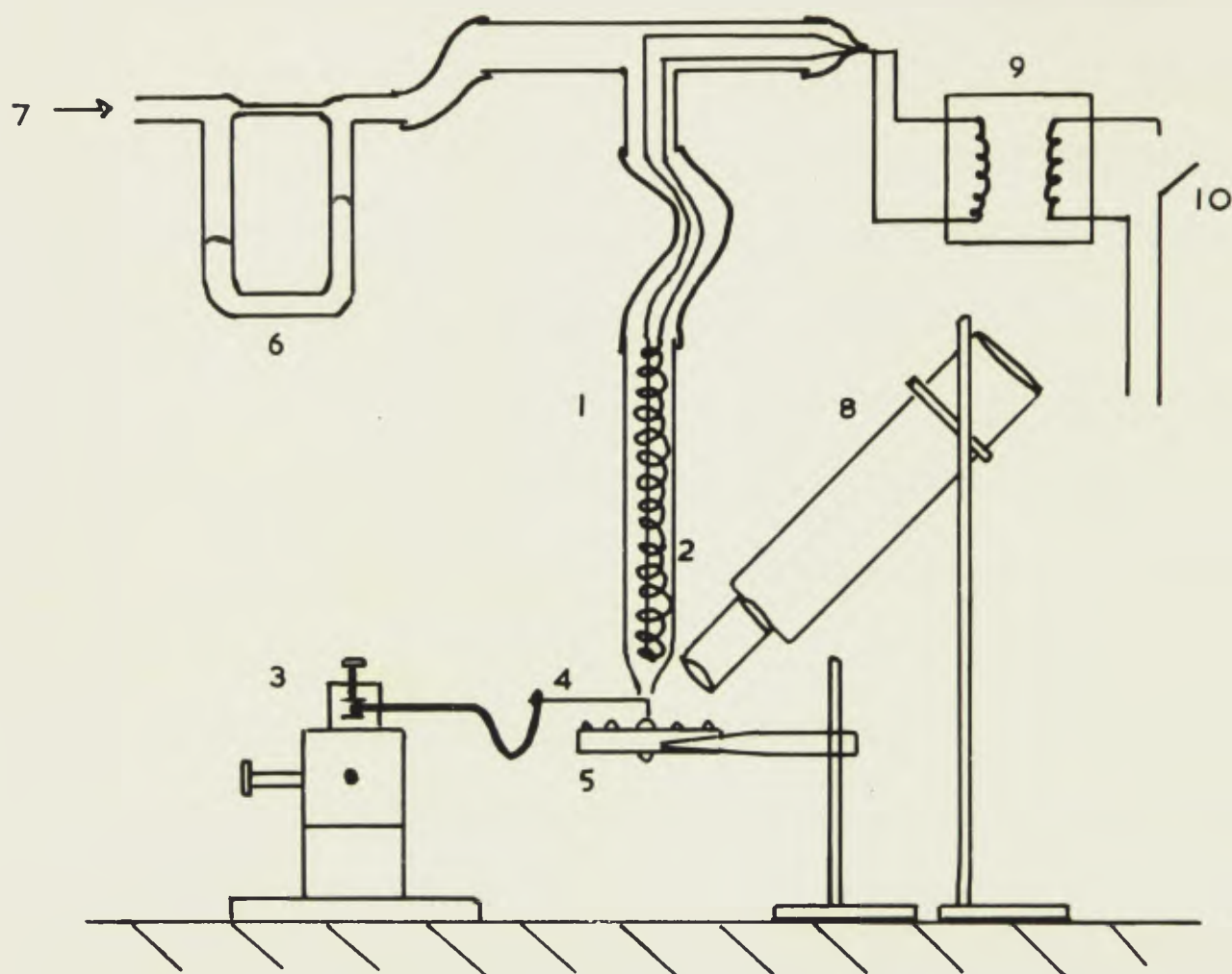


Figure 3.2

The Hot Gas Jet System



- | | | | |
|---|--|----|---|
| 1 | 6" pyrex tube, 0.4" diameter, 3/4 mm. nozzle. | 5 | Crystal specimen with contact dots, held by fixed tweezers. |
| 2 | 20 g. eureka wire on mica former, 2.5 ohms. | 6 | Capillary manometer 3.5 cm. Hg per litre per min. |
| 3 | Three way micromanipulator for positioning contact wire. | 7 | Forming gas flow from cylinder at 0.6 lit./min. |
| 4 | 40 g. copper wire for specimen lead, temporarily attached to micromanipulator by solder. | 8 | Microscope for observation of contact melting. |
| | | 9 | Transformer 230: 6.75 volts. |
| | | 10 | Switch. |

The temperature was monitored on a direct reading voltmeter from a platinum - platinum rhodium thermocouple placed in a hole in the carbon bar; the temperature was held at 400°C for 1 minute for alloying, and reduced over a period of three minutes for recrystallization from the bulk material.

3.3 Contact Wiring.

The original method of mounting wires in the contacts consisted of heating a tinned wire close to its end with a soldering iron until heat conducted down to the end caused a joint to be made. The contacts made in this way were not found to be satisfactory. They were noisy, the resistance was high, rectification properties were poor and fracture on cooling was not infrequent. Properties improved after running hot, which effected a forming process; this was potentially destructive.

As an alternative method, a hot gas jet instrument (W. B. Archey 1964, 303) was designed and built by G. D. Peskett; it is shown diagrammatically in figure 3.2. In this method an electrically heated stream of forming gas is directed onto the contact; the contact is melted and a tinned copper wire inserted by means of a micromanipulator. Such contacts are of great mechanical strength and have consistent properties.

Multi-contact crystal specimens were affixed to nine-pin valve stubs; two contact specimens were put into i-V characteristic test rigs.

3.4 General Purpose Cryostat.

There existed initially a cryostat suitable for measuring Hall effect properties of crystals down to liquid helium temperature; it has been described by J. P. Russell (1962, 304, p.21) and by B. V. Rollin, J. R. Mills, J. P. Russell (1960, 305). This apparatus has been completely redesigned and rebuilt, employing the same principle of coaxial tubes. The apparatus is employed in conjunction with a new water cooled 20 kgauss magnet supplied by Newport Instruments, Type F, 2" gap, 5" pole faces, 0.8 ohms resistance, 16 gallons/minute, maximum current 150 amp. The new system consists of a detachable metal dewar mounted for use with the magnet, together with all required pumping lines and measuring facilities. The dewar may be filled in situ or employed as a transfer vessel. The working space is defined by a 3 cm. german silver tube; pumping and measuring facilities are so designed that this tube is detachable. Thus in effect a whole range of apparatuses is available, each one consisting of a 3 cm. german silver tube incorporating particular facilities.

Tubes currently available are:

- (1) A high insulation Hall coefficient tube for use with exchange gas, similar to the original design, (Russell l.c. 1962, 306); the specimen is contained in a small brass can which forms a constant-volume space for helium gas thermometry; the can is in thermal contact with the main tube through helium exchange gas which may be pumped away; the can is equipped with a heating coil so that

Figure 3.3

D.C. System



the temperature of the specimen may be raised from that of the coolant up to room-temperature. Nine leads are available for Hall measurements on the specimen; by the use of high quality glass-ceramic seals the wiring insulation is better than 10^{10} ohms. This apparatus is suitable only for low currents (less than 10 milliamperes) since all wiring is of 36 gauge eureka.

(2) A lower insulation Hall coefficient tube for use with direct immersion; this consists merely of a valve-base supported at the foot of a tube and wired by 36 gauge copper leads so that high currents may be employed; it is hence unsuitable for use with liquid helium. Glass seals are employed at the top so that the coolant may be pumped to give a temperature variation, measured from the vapour pressure.

(3) A fast pulse tube containing coaxial lines.

(4) A pulse rise time tube containing matched leads.

An attempt has been made to standardize on common-geometry nine-pin valve stubs as specimen holders. Tubes 1, 2, a Hall voltage trial rig and a helium 3 cryostat take these, so that specimens once mounted may be transferred between them.

3.5 D. C. Systems.

The system employed for steady currents with tube 2 (section 3.4) is illustrated in figure 3.3. The recorder is a Moseley X-Y Autograf Model 135 used throughout on the calibrated ranges. The switchbox was designed to make virtually all potential differences required for plotting immediately available. The electrical circuit is shown in fig. 3.4.

Figure 3.4

D.C. Circuit

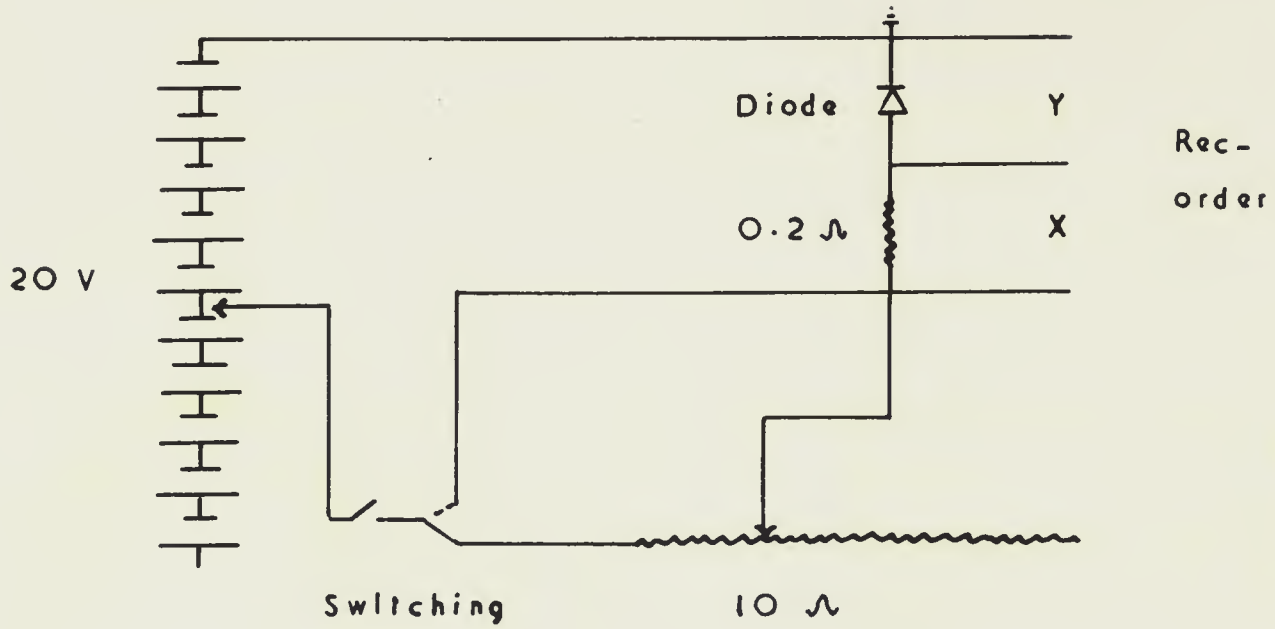


Figure 3.5

Pulse Circuit for Radiation Observation

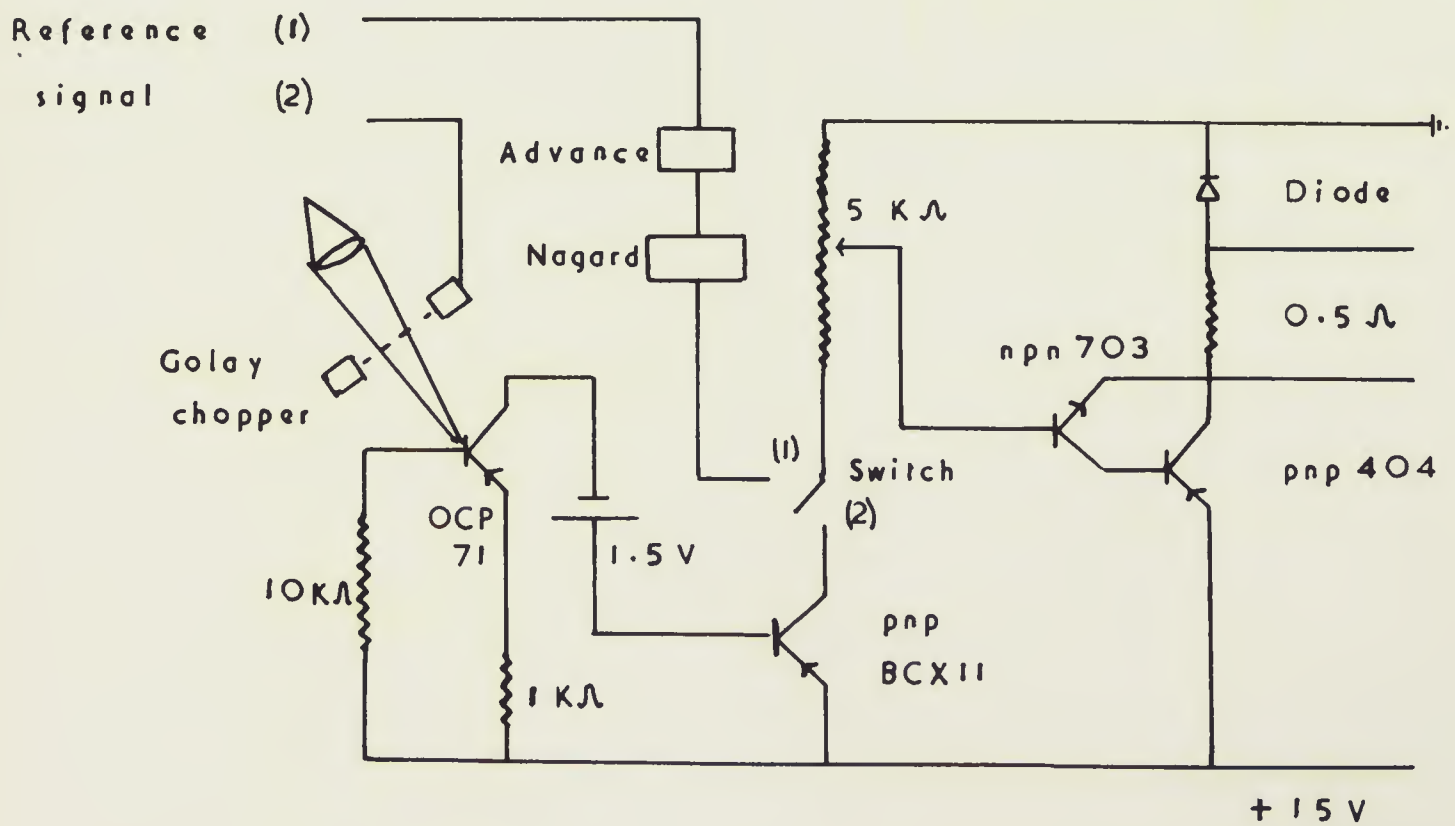
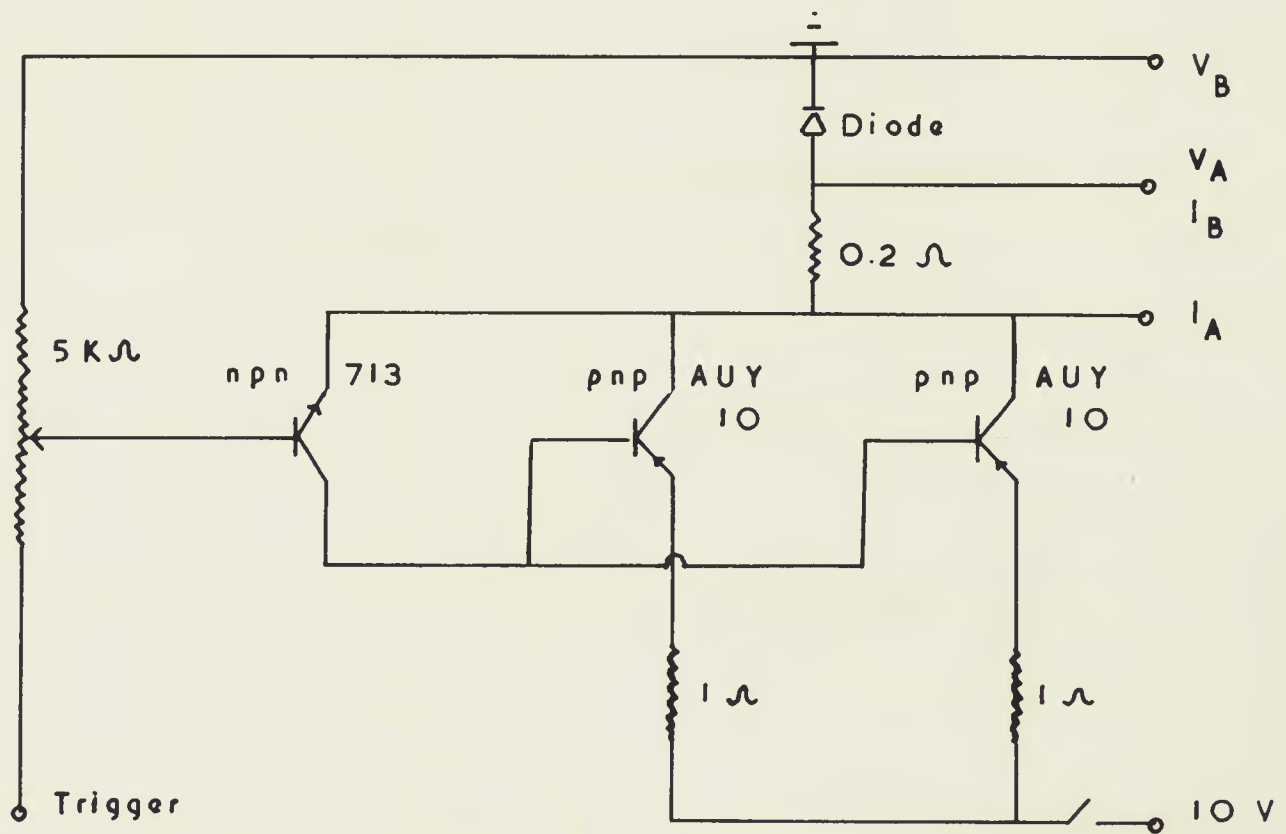


Figure 3.6

Pulse Circuit for Oscilloscope



3.6 Pulsed Systems.

Two kinds of pulsed system were designed:

- (1) for amplification of emitted recombination radiation;
- (2) for observation of rise and decay characteristics.

3.6.1 Pulsed systems for radiation amplification.

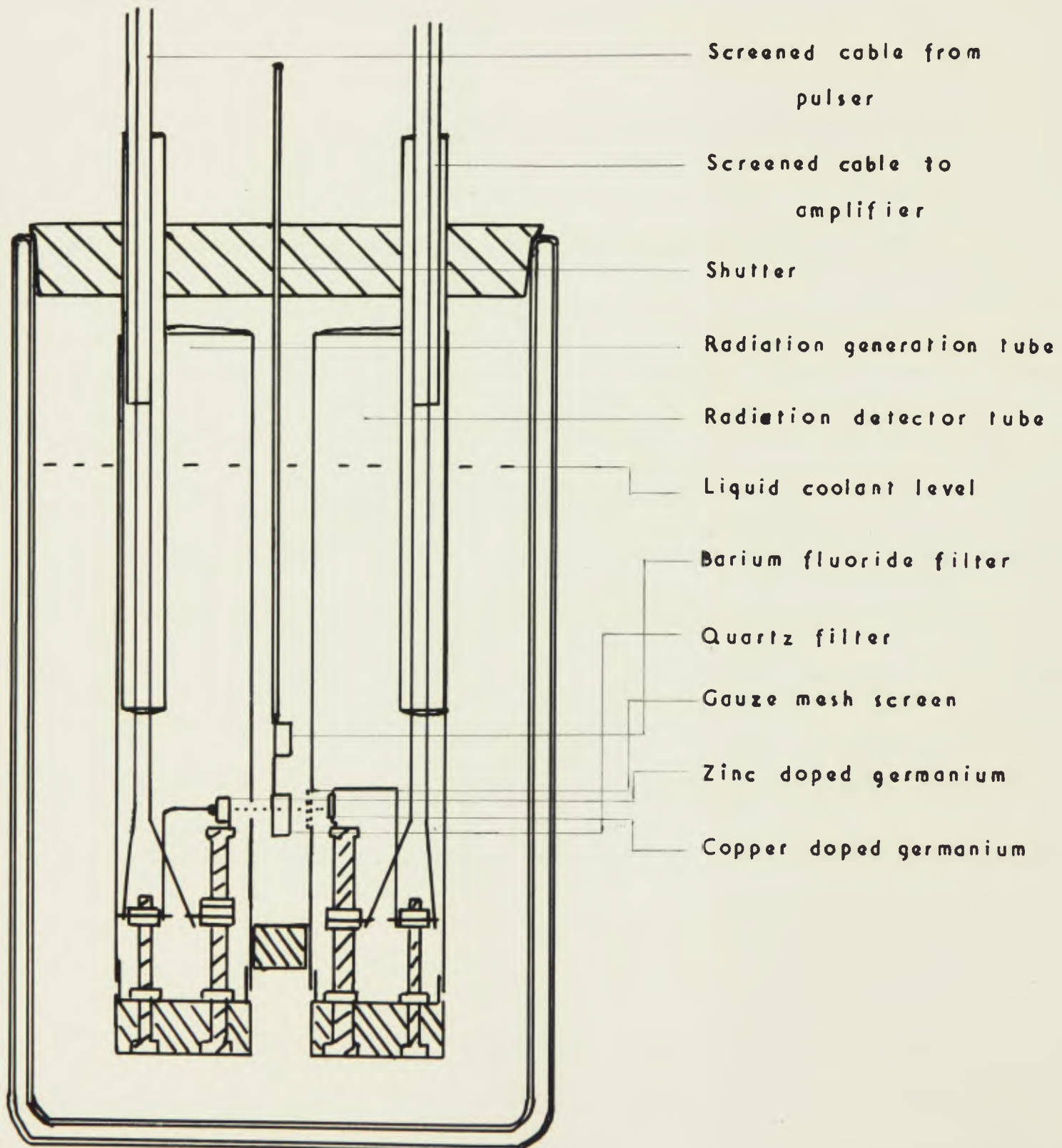
The detection of small low energy radiation signals is usually effected through the use of chopped amplification. When the source crystal is maintained at a constant low temperature, it is not in general possible to employ an external mechanical radiation chopper because of the ballast signal provided by the room temperature chopper blades. Thus it is essential to chop the electrical current supplied to the crystal. For use with a copper-doped germanium detector specimen, the circuit of figure 3.5 was devised, with the switch in position (1). When a Golay cell is employed for detection, it is essential to use the signal from the Golay chopper as the reference source, since the amplifier will not tolerate any other waveform. A circuit employing this as the trigger was devised and is shown in figure 3.5 with the switch in position (2).

3.6.2 Pulsed systems for rise and decay time observation.

The circuitry of figure 3.6 was devised for observing pulsed characteristics on a Tektronix 502 oscilloscope. For a fast rise time Mullard AUY10s limited to 0.75 μ s. were used rather than Newmarket 404s.

Figure 3.7

Radiation Generation - Detection System



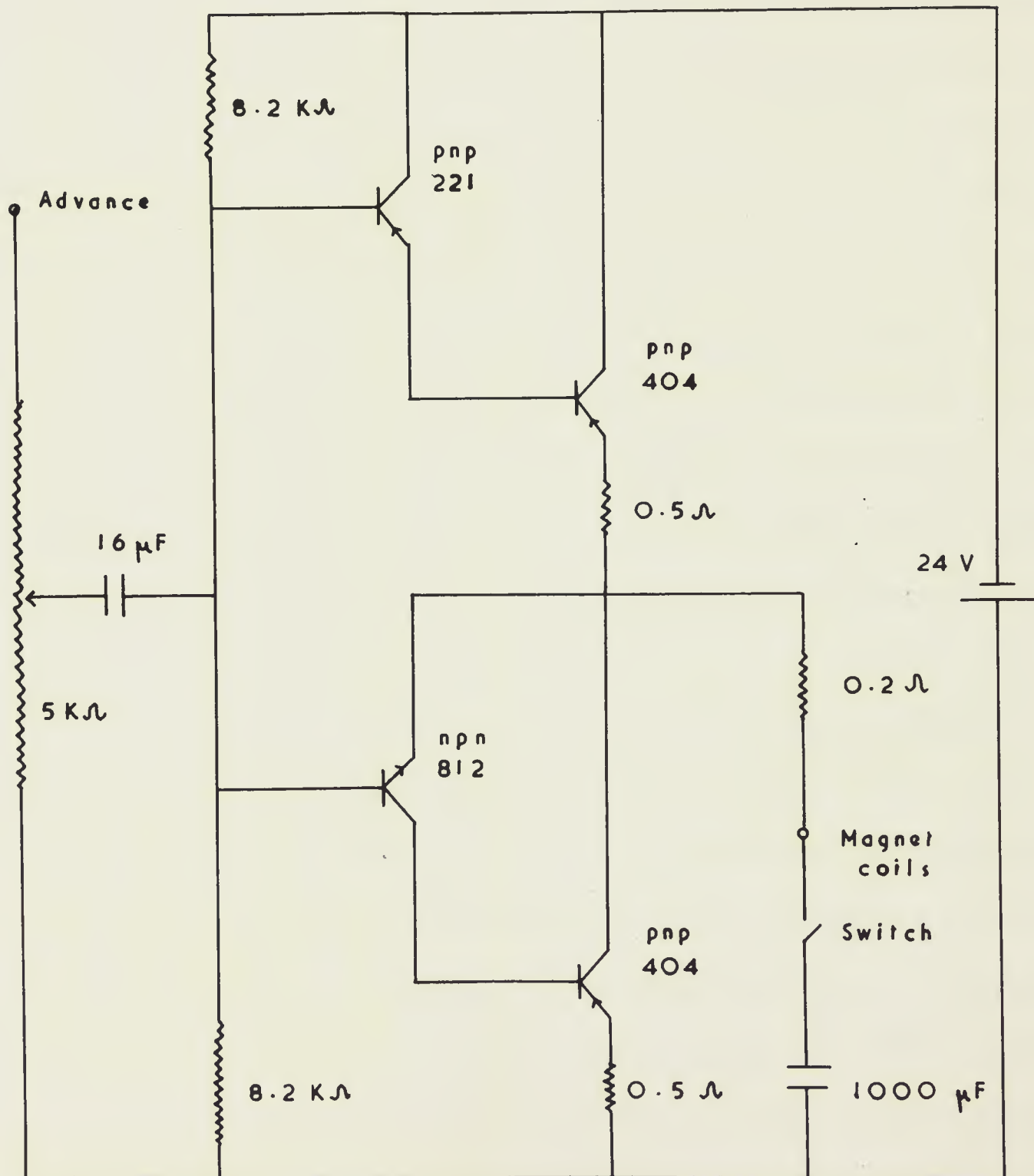
3.7 Radiation Generation - Detection System.

The aim in the construction of a generation-detection system was to have the two crystal specimens as close together as possible so that the absence of a signal could not be put down to misalignment of the optics; a rough spectral check was to be obtained through the use of filters; the crystals were to be immersed in hydrogen to give positive cooling. The radiation energy to be looked for was 0.09 eV. (section 1.3) so a copper-doped germanium crystal was chosen, specimen GC18 (reference B. V. Rollin, J. P. Russell 1963, 307).

Anticipating pick-up troubles at the high current levels to be used, complete shielding was specified, the diode and detector being placed in separate german silver tubes as close together as possible with a minimum size pair of windows for radiation transmission. Initial pick-up troubles were overcome by rigorous attention to the earthing circuit, including covering the detector entrance window with fine copper gauze; in-phase pick-up was reduced below the noise level and 90° capacitative pick-up was only just visible. The filter system consisted of a four-position sliding shutter, the alternatives being open, shut, a quartz filter and a barium fluoride filter. The quartz filter provides a cut-off at 2μ and passes intrinsic radiation. The barium fluoride filter absorbs radiation of a longer wavelength than 13μ . The generation-detection system is shown in figure 3.7.

Figure 3.8

Magnetic Field Pulser



3.8 Cryostat for Spectral Analysis of Recombination Radiation.

A cryostat was constructed for the spectral analysis of long wavelength radiation. For this purpose, in view of the large power inputs in use, it was felt imperative that the crystal diode should be immersed in the liquid refrigerant. The provision of a window to retain liquid helium against a vacuum yet transmit the long-wavelength radiation has presented serious problems. As a material, a slice of pure germanium of high resistivity in a covar holder has been used; the snag has been a tendency to fracture under the simultaneous application of pressure and strain due to cooling. A partial solution has been found to be to arrange for the atmospheric pressure thrust and the differential contraction strain to act in perpendicular directions, using an epoxy resin glue. A tendency to leak after a number of temperature cycles has been noted.

3.9 Magnetic Fields.

The steady magnetic field for Hall voltage measurement has already been described (section 3.4). It later became important to examine the Hall coefficient of a set of slowly varying "domains". For this purpose an oscillating magnetic field was devised, produced by two 11.2 cm. diameter coils with 41 turns of 26 gauge copper wire mounted in Helmholtz fashion, the pulsing circuit being as shown in figure 3.8. The field swing obtained was 10 gauss at 4 amp. peak to peak at 1000 c/s.

3.10 Crystal Materials.

Crystals involved in this work have been:

- (1) Overcompensated zinc-doped germanium, 10^{15} free carriers.
- (2) Uncompensated n-type germanium, 10^{14} free carriers.
- (3) Uncompensated n-type germanium, 10^{15} free carriers.
- (4) Uncompensated n-type germanium, 10^{16} free carriers.
- (5) Copper-doped p-type germanium.
- (6) N-type indium antimonide.

The exact composition of material 1, the zinc-doped germanium overcompensated with arsenic, is not known, but is believed to be similar to a p-type crystal of the same series, given by the manufacturers as:

$$\begin{array}{lll} \text{Zn} & 3.3 & \cdot 10^{16} \quad , \\ \text{Sb} & 6.5 & \cdot 10^{16} \quad , \\ \text{P} & 1.8 & \cdot 10^{15} \quad . \end{array}$$

At 77°K values of the carrier density per c.c. and of the carrier mobility in cm./volt sec. were obtained as:

$$\begin{array}{lll} n & 2.81 & \cdot 10^{15} \quad , \\ \mu & 7.81 & \cdot 10^3 \quad . \end{array}$$

Reference to mobility graphs computed for the dependence of scattering mechanisms on temperature and impurity densities (J. M. Rowell 1961, 308, p.38) suggests that the compensation ratio of material (1) is lower than that of the p-type specimen. We suggest:

$$\begin{array}{lll} \text{Zn} & 5.0 & \cdot 10^{15} \quad , \\ \text{As} & 1.281 & \cdot 10^{16} \quad , \\ n & 2.81 & \cdot 10^{15} \quad . \end{array}$$

Table 1

continued

Number of plants		Number of plants		Number of plants		Total
1950	1951	1952	1953	1954	1955	
1	1	1	1	1	1	5
2	2	2	2	2	2	10
3	3	3	3	3	3	15
4	4	4	4	4	4	20
5	5	5	5	5	5	25
6	6	6	6	6	6	30
7	7	7	7	7	7	35
8	8	8	8	8	8	40
9	9	9	9	9	9	45
10	10	10	10	10	10	50
11	11	11	11	11	11	55
12	12	12	12	12	12	60
13	13	13	13	13	13	65
14	14	14	14	14	14	70
15	15	15	15	15	15	75
16	16	16	16	16	16	80
17	17	17	17	17	17	85
18	18	18	18	18	18	90
19	19	19	19	19	19	95
20	20	20	20	20	20	100

Figure 3.9

Table of Specimens

Material	Number of Specimens	Diodes	Ohmic Pairs	Multi-contact Specimens		
				Name	Rectifying	Ohmic
As	14	25	8	As11G1 α	1	5
				As12H1 α	1	8
				As13H2 α	1	8
				As12I2 α	-	6
				As12H3 α	2	6
1UC	1	-	-	1UC1H1 α	-	6
				1UC1H2 α	1	8
2UC	6	-	5	2UC6H1 α	-	6
3UC	1	-	-	3UC1H1 α	-	6
InSb	1	-	2	InSb1H3 α	-	6

the 100% concentration

concentration

concentration

concentration

concentration

concentration

concentration



concentration

concentration

concentration

concentration

concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%
concentration	100%	100%	100%	100%

concentration

concentration

concentration

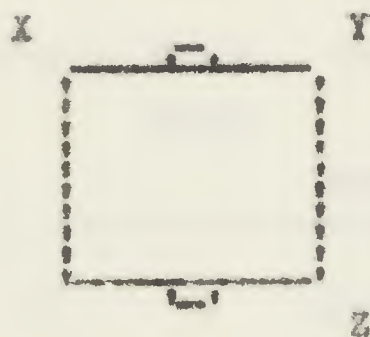
concentration

concentration

concentration

Figure 3.10

Specimen Sizes



Specimen	$E_1 E_2$	XY	YZ
As11G1 ₊	1.335	0.185	0.150
As12H1 ₊	1.050	0.130	0.135
As13H1 ₊	1.016	0.120	0.150
As12H3 ₊	1.027	0.078	0.095
1UC1H1 ₊	0.73	0.09	0.09
2UC6H1 ₊	0.80	0.1	0.1
3UC1H1 ₊	0.75	0.1	0.1
InSb1H3 ₊	0.91	0.2	0.06

(cm.)

The 77° K results for other specimens have been deduced from Hall measurements as:

material (2)	carrier density $3.1 \cdot 10^{14}$	mobility $2.6 \cdot 10^4$
(3)	$2.7 \cdot 10^{15}$	$1.6 \cdot 10^4$
(4)	$4.0 \cdot 10^{16}$	$4.4 \cdot 10^3$

Details of specimen (5) are obtained from the results of J. P. Russell, B. V. Rollin (l.c. 1962, 309) as $n_{cu} 2.3 \cdot 10^{16}$, $n_d 1.4 \cdot 10^{14}$.

The 77° K carrier density of specimen (6) was estimated as $6 \cdot 10^{13}$.

All quoted densities are per c.c. and mobilities in cm./volt sec.

3.11 Specimen Classification.

Specimens have been classified by five features: the crystal material, the section of the material, the method of mounting, the junction contacts and the contact wiring. Thus As12H3₊ refers to arsenic compensated material, the twelfth piece of this used, in the eighth type of mount, the piece being used for its third set of contacts wired for the first time.

A list of specimens employed is shown in the table of figure 3.9. This includes only those for which useful information was obtained. The ten multi-contact specimens have provided the most significant results; these are listed by name. Their dimensions are given in the tables of figures 3.10 and 3.11. The nomenclature of the specimen contacts is also shown in the diagram at the head of figure 3.11. R and R' represent injecting contacts, all others being non-injecting.

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

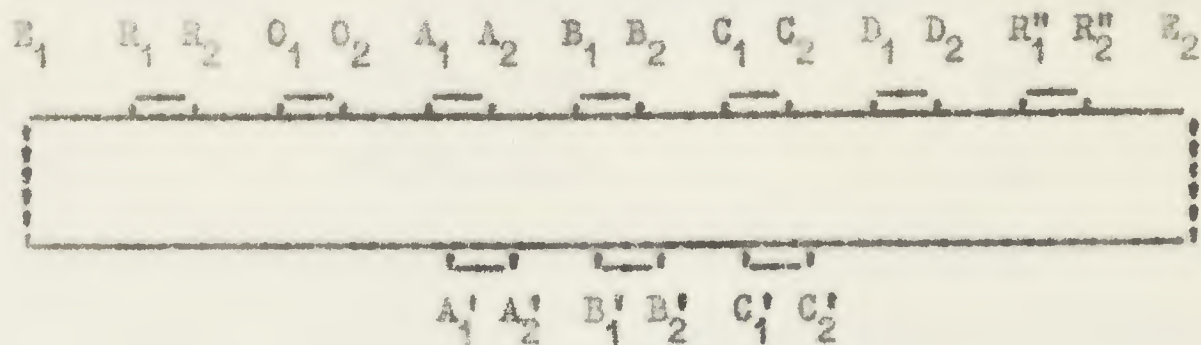
11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

11.2. 11.2. 11.2

Figure 3.11
Contact Sizes



Contacts	As11G1 α	As12H1 α	As13H1 α	As12H3 α
R ₁	2.390	7.950	13.562	13.824
R ₂	2.335	7.850	13.464	13.761
O ₁		8.105	13.422	13.688
O ₂		7.980	13.364	13.634
A ₁	2.255	7.740	13.290	13.500
A ₂	2.195	7.640	13.238	13.452
B ₁	1.830	7.520	13.052	13.250
B ₂	1.770	7.415	12.998	13.201
C ₁	1.425	7.330	12.811	
C ₂	1.355	7.265	12.762	
D ₁	1.270	7.190	12.697	13.001
D ₂	1.225	7.115	12.640	12.960
R ₁ ^{''}				12.905
R ₂ ^{''}				12.854
A ₁ A ₁ [']		-0.040	0.011	0.018
A ₁ A ₂ [']		0.100	0.051	0.052
B ₁ B ₁ [']	0.015	-0.055	0.012	0.004
B ₁ B ₂ [']	0.065	0.160	0.060	0.039
C ₁ C ₁ [']		-0.035	0.011	
C ₁ C ₂ [']		0.165	0.048	(cm.)

In the experimental section, the background data for each graph is given for the following features: specimen classification, current contacts employed, voltage contacts employed, temperature, magnetic field strength, and reference number of recorder trace. These items are inserted, as relevant, into the graph area.

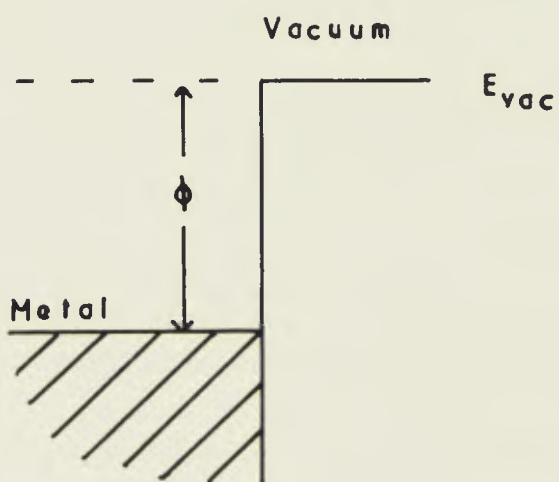
3.12 The Advantages in the Use of D. C. Techniques.

In this work, d.c. measurements have been made wherever feasible, because of the advantages offered by the use of an X-Y recorder. With a recorder in a circuit including rheostats, the unit of information is a complete characteristic, while for visual or photographic recording of pulse heights, the unit is a single point. Assuming 40 useful points per curve, this gives an information capture ratio of 40. Secondly the observation of low time constant negative resistances involving spontaneous transitions, is only possible with continuous recording instruments. Where the use of pulses is essential, as is the case for the extension of the double injection measurements reported here, it seems inescapable that sampling techniques should be devised so that direct recording can be retained.

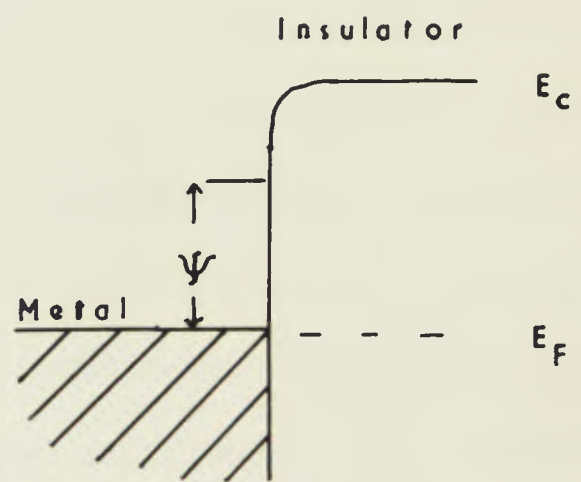
Figure 4.1.1

Barrier Structure

From M. A. Lampert 1964, 401



Thermionic
Emission



Carrier
Injection

Chapter 4

Double Injection.

4.1 Theory.

4.1.1 Introduction to Injection.

The field of current injection in insulators and semiconductors has been reviewed by M. A. Lampert (1964, 401), the summary of the present position being based on his article. The aim of this chapter is basically to provide extensions to theory and experiment at some of the places where he points out that these are lacking. He writes: "(In double injection) a complexity and diversity of behaviour is encountered whose limits at the time of writing have simply not been explored" p.349, and again: "Recent studies of Holcnak et. al. have opened a veritable Pandora's box of double injection phenomena based on the role of deep lying centres in GaAs, Si and Ge" p.361. Experimental and computational techniques are to be developed for these circumstances.

Interest in injection dates from the emergence of the energy band theory of the electronic structure of solids. In 1940, N. F. Mott and R. W. Gurney (1940, 402, p.168) pointed out that analogous to the thermionic emission of electrons from a cathode, a suitable contact should inject carriers into an insulator. This is illustrated in figure 4.1.1. There are two added complications in the solid state --- firstly the action of localized defect states which may be

expected to trap injected carriers, and secondly the dual possibility of hole injection.

In the thermionic diode, the electron current is space-charge-limited; the relationship between the applied voltage and the density of injected carriers is governed by Poisson's law, by which the current-voltage characteristic is defined, the result being Child's law. In the single carrier injection diode, space-charge is similarly the limiting factor, but because of scattering, the transport equations are different, and space-charge immobilized by carrier traps provides a further hazard. By making one contact to an insulator electron-injecting and the second hole-injecting, double injection is obtained; the current need no longer be space-charge-limited since the injected carriers may neutralize each other in the crystal. In the thermionic tube, electrons and positive ions cannot provide effective neutralization because their velocities depend on the potential they have traversed, so that electrons move slowly at high density near the cathode, where positive ions move swiftly at low density, and conversely near the anode. In the solid state, electrons and holes, due to scattering, have velocities dependent on the local field value. The limitation on the current is then due to loss of current carriers by recombination. Thus with double injection in material of high lifetime, a very considerably higher double-injection current may be anticipated, compared with single injection.

Science has been fast to make use of some of the principles of

carrier injection. Injection currents may be used as a probe for the study of localized defect states in insulators (Lampert l.c. 1964, 403, p.332). Injection may be used for the study of plasma instabilities (Lampert l.c. 1964, 404, p.333). A number of solid state devices rely on injection, including types of rectifier, transistor and laser. At the same time ways of avoiding minority carrier injection have been sought for the improvement of photoconductors and particle detectors.

For these reasons a considerable volume of technical and empirical knowledge has been built up about carrier injection; however theoretical and experimental analysis, especially of two carrier processes, has lagged behind. There seem to be several reasons for this. Firstly carrier injection is basically a surface phenomenon, even when the carriers are injected optically. The physics of surfaces and interfaces presents enormous problems; even the technological methods of rectifier manufacture are more trial and error than empirical, let alone scientific (H. K. Henisch, 1957, 405, pp.95, 168). Secondly the mathematical equations of double injection, where they can be formulated for bulk material, are generally so complicated as to be impossible to handle without unjustifiable simplification. Any simplification is difficult to make because so many parameters, especially those of recombination, are not known even approximately, yet these parameters prove critical in terms of the results.

These weaknesses will be apparent in the treatment which follows.

Assumptions have been made about the nature of the contacts and of the recombination laws which may seem arbitrary and unprecedented; such assumptions have not previously been needed, because the mathematical expressions have proved too cumbersome to solve. But our new computational technique has proved so effective that in order to demonstrate it, numerical values are needed. These have been chosen with an eye not to ease of computation --- this is not really necessary --- but to plausibility; it is felt that the technique remains valid, even if the boundary assumptions are dubious.

4.1.2 Qualitative Approach to Double Injection.

The physics of single injection, comprehensively dealt with by Lampert (l.c. 1964, 4.06), does not concern us here; we concentrate on double injection. Throughout we will deal with hole injection into n-type material; application to electron injection in p-type material requires the interchange of corresponding symbols.

Double injection presents three main features: trapping, recombination and contact constraints. Trapping will be set aside for the time-being. Contact constraints are obviously of vital importance within a diffusion length of the contacts, but Lampert claims that for long crystals, a simplified theory neglecting diffusion still gives a useful description of current flow properties and of the overall current-voltage characteristic. Without contact constraints, the question arises as to what limits the double injection

[The body of the page contains extremely faint, illegible text, likely bleed-through from the reverse side of the document. The text is organized into several paragraphs, but the characters and words are too light to be transcribed accurately.]

Figure 4.1.2

Crystal for Double Injection

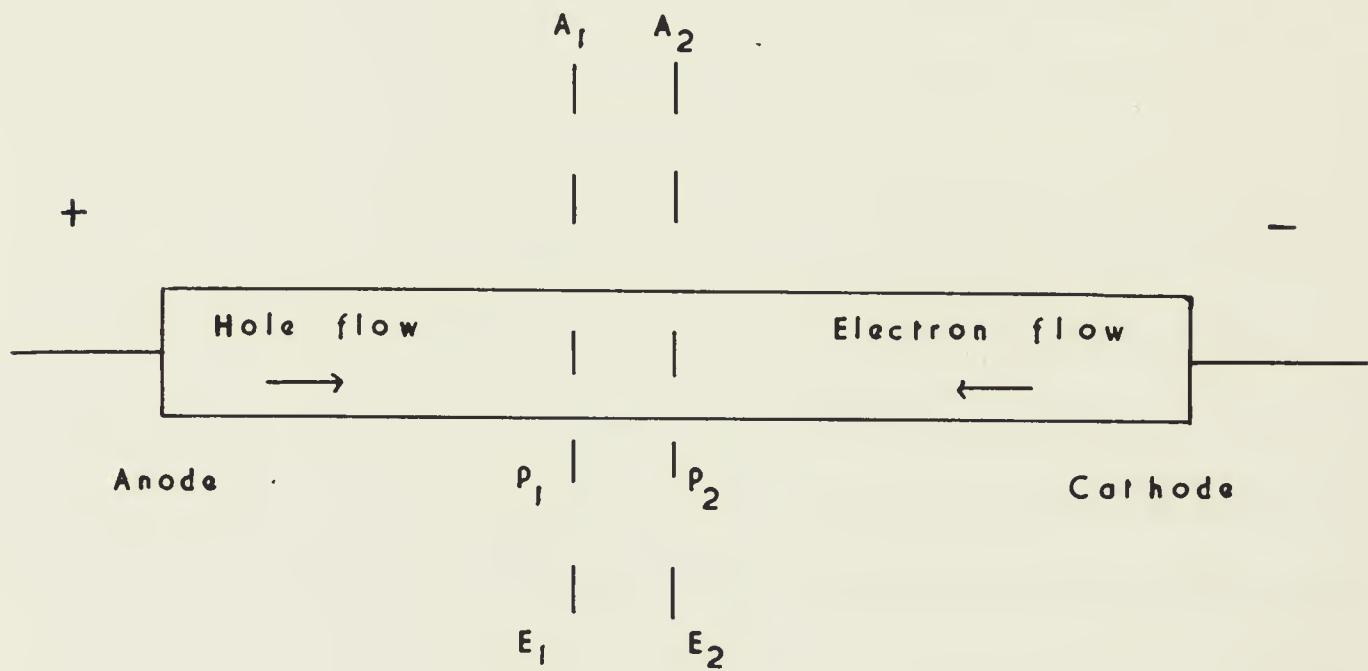


Diagram to illustrate the problem caused by
recombination under conditions of charge neutrality.

current. The answer need not be 'space-charge', since the holes and electrons can neutralize each other. It must then be 'recombination'. Problems immediately arise. Is the hole density greater nearer the anode than the electron density? No, since neutralization is required everywhere, the densities must be equal throughout the crystal. Yet if holes are flowing in from the anode side only and electrons from the cathode side, how can the numbers remain equal in spite of steady recombination? This problem is so fundamental to an understanding of double injection that it will be elaborated here. Consider a section of the crystal far removed from anode and cathode, denoted by A_1A_2 in figure 4.1.2. Holes flow in from the left due to the electric field E_1 , their density being p_1 ; recombination takes place as the holes flow towards the cathode; thus at A_2 either the hole density p_2 is lower than p_1 or the field E_2 is lower so that holes are being concentrated by the field. But in either case the effect on the electrons is exactly the reverse. Thus an electric field cannot maintain neutralization in the face of recombination; carrier movement must be controlled by a combination of electric field and diffusion. Yet in a diffusion solution it is important to know the numbers of carriers at the boundaries A_1 and A_2 . Thus the solution depends on the constraints.

There is a further important point. A contact which is injecting for carriers of one sign of charge is blocking for carriers of the opposite sign. Therefore the electrons will tend to accumulate

or 'stagnate' at the hole injecting contact, building up their density to the point where back diffusion is important, and similarly for the holes. In this way the density of carriers, neutralized as specified, is higher at the electrodes than in the centre of the crystal, and the electric field correspondingly higher in the centre to maintain a constant current.

A complete mathematical treatment begins with the carrier continuity equation, which is a second order differential equation; solution for the i - V characteristic requires two integrations, the contact constraints being involved as the two constants. Lampert (l.c. 1964, 407, p.350) essentially ignores contact problems by considering a region of crystal in which the carrier density varies little over the crystal and is given an averaged value. Our new technique has been devised to handle the original equation. Before describing it we will consider the influence of traps, and their relevance to negative resistance.

4.1.3 Negative Resistance through Double Injection.

Negative resistances in double injection studies have been observed in a number of ways by N. Holonyak et al. (1962, 408; 1962, 409), K. L. Ashley and A. C. Milnes (1964, 410), I. Melngailis and R. H. Rediker (1962, 411), M. C. Steele, K. Ando, M. A. Lampert (1962, 412; 1963, 413), K. Komatsubaka, H. Kurono (1962, 414), M. Kikuchi, Y. Abo (1962, 415), V. I. Stafcev, S. A. Kramareva

(1959, 416; 1960, 417) and others.

A bulk differential negative resistance due to double injection may be explained as follows. Consider a crystal with a fixed number of recombination centres whose cross-section for hole capture is much higher than that for electron capture. At low injection levels the holes have a short free lifetime, being captured near the anode; the current is essentially an electron current. As the current level rises, the injected holes will tend to saturate the recombination centres and diffuse further into the crystal, to give a double injection current. Thus the more holes that are injected, the easier it is for them to traverse the crystal. Furthermore the scattering cross-section for a singly-charged acceptor impurity is lower, so that the mobility may be expected to be enhanced. If these factors are strong enough and the crystal length is favourable, the voltage required to maintain the higher current may actually be lower, giving a negative resistance.

Even when a negative resistance is not apparent in the overall voltage-current characteristic, it may be observed over the portion of the crystal where the saturation wave-front is located. Increased length provides ballast resistance which will dilute or quench the phenomenon; thus the point of appearance and size of the negative resistance in the overall i - V curve is a complex function of all crystal parameters.

The description given here is equivalent to Stafcev's "Modulation

of the diffusion length" (l.c. 1959, 418). Clearly just before the transition the crystal will be very sensitive to external influences on the number of carriers or the mobility, such as illumination or magnetic fields (Kramareva and Stafeev l.c. 1960, 419; E. I. Karakushan, V. I. Stafeev 1961, 420; 1962, 421).

4.1.4 Equations of Carrier Flow.

In this section the basic equations of carrier flow are discussed; the notation is from R. A. Smith (1959, 422, p.235). For electron and hole flows:

$$J_e = ne\mu_e E + eD_e \nabla n ,$$

$$J_h = pe\mu_h E - eD_h \nabla p ,$$

$$J = J_e + J_h .$$

Working in one dimension, assuming planar flow, and incorporating the Einstein relationship $D_e = \mu_e kT$,

$$J_e = ne\mu_e E + \mu_e kT n' ,$$

$$J_h = pe\mu_h E - \mu_h kT p' .$$

Poisson's equation is:

$$dR/dx = e(\Delta p - \Delta n)/\epsilon\epsilon_0 ,$$

where Δp , Δn are the changes in carrier concentration from equilibrium. Carrier continuity equations are required, taking into account generation, trapping and recombination. These are:

$$J'_e = G - R + e \Delta n / \Delta t \quad ,$$

$$J'_h = R - G - e \Delta p / \Delta t \quad .$$

An approach to a solution of these equations as they stand has been given by J. B. Gunn (1958, 423). Under any circumstances some form of simplification is required to obtain any useful results from them. The simplification to be employed in this case is the neglect of the Poisson equation, which will now be justified.

4.1.5 The Poisson Equation.

The Poisson equation may be written as:

$$\frac{e}{\epsilon \epsilon_0} (\Delta p - \Delta n) = \nabla E = \nabla \cdot (J/\sigma) = - \frac{J}{\sigma} \cdot \nabla \sigma \quad .$$

Thus for the conductivity of the crystal to change along its length, a space charge is required. For the situation being considered,

$$J \approx 1 \text{ amp./sq. mm.},$$

$$n \approx 10^{15} \text{ carriers per c.c.},$$

$$\sigma \approx 1 \text{ mho per cm.},$$

$$l \approx 1 \text{ cm. (crystal length).}$$

Then: $\Delta p - \Delta n \approx 10^9 \text{ carriers per c.c.}$

So the extra carrier density due to an infinite decrease in conductivity is negligible compared with the number of carriers present in equilibrium. We may say that conditions of neutrality are satisfied in low resistivity semiconductors.

There is in fact a considerable art to manipulating the diffusion

equations; for if we put $\Delta n = \Delta p$, then automatically no conductivity variation can take place, according to the first equation; yet we do intend to neglect the space-charge. The rather strange treatment of ambipolar mobility in Smith (l.c. 1959, 424, pp.253-4) is an example of this problem. The treatment of the equations to be presented here will show that Smith's procedure for neglecting $\Delta p - \Delta n$ in some terms of his equations but not in others is equivalent to saying that the space-charge is negligible, but not zero.

4.1.6 The New Equation.

We intend to consider steady state flow, with no pair generation.

Thus the equations become:

$$J_e = ne\mu_e E + \mu_e kTn' \quad (4.1)$$

$$J_h = pe\mu_h E - \mu_h kTp' \quad (4.2)$$

$$J = J_e + J_h \quad (4.3)$$

$$J_e' = R \quad (4.4)$$

$$J_h' = -R \quad (4.5) \quad .$$

Equation 4.4 is integrated as:

$$J_e = \int R dx + \text{constant} \quad .$$

An interpretation of this equation is now made. The hole injecting anode is at $x=0$, the cathode at $x=L$. Consider just the electron flow. The electron current which flows away at the anode is put as J_{ea} ;

this is the constant term. The rest of the electron current is contributed by electrons which recombine in the bulk of the crystal; this is given by the integral. So:

$$J_e(x) = \int_0^x R dx + J_{ea} \quad (4.4')$$

From (4.1), (4.2):

$$J = Ee(n\mu_e + p\mu_h) + kT(\mu_e n' - \mu_h p') \quad .$$

Eliminating E , J_e ,

$$J = \left\{ 1 + \frac{p\mu_h}{n\mu_e} \right\} (J_{ea} + \int_0^x R dx - \mu_e kT n') + kT(\mu_e n' - \mu_h p') \quad (4.6)$$

$$J = \left\{ 1 + \frac{p\mu_h}{n\mu_e} \right\} (J_{ea} + \int_0^x R dx) - \frac{\mu_h kT}{n} (pn' + np') \quad (4.7)$$

$$\frac{d}{dx} (np) = \frac{n}{\mu_h kT} \left\{ \left(1 + \frac{p\mu_h}{n\mu_e} \right) (J_{ea} + \int_0^x R dx) - J \right\} \quad (4.7')$$

Direct differentiation of equation (4.6) leads back to the ambipolar equation of Smith (l.c. 1959, 425, p.254), confirming that the space-charge assumptions made here are equivalent to those made in his derivation. Thus we have succeeded in providing a first integration to the equation:

$$D \frac{d^2 p}{dx^2} - \mu E \frac{dp}{dx} - \frac{p}{\tau_p} = 0 \quad ,$$

with $D = \frac{n+p}{n/D_h + p/D_e} \quad , \quad \mu = \frac{n - p}{n/\mu_h + p/\mu_e} \quad .$

In fact the first step to a solution, namely the first integration, has been accomplished at such an early stage, that it has almost gone

unnoticed. This technique does not seem to have been reported previously. It might be termed back-door integration.

We will now return to equation (4.7). Since space-charge is negligible, the relationship of n to p is governed by the impurity concentration and the trapping level. For a simple case, consider n -type material, with n_a donors all ionized, with no traps. Then for an injection level giving a hole concentration f :

$$p = f,$$

$$n = n_a + f,$$

$$\frac{d}{dx}(f) = \frac{n_a + f}{n_a + 2f} \frac{1}{\mu_h kT} \left(\left\{ 1 + \frac{f \mu_h}{(n_a + f) \mu_o} \right\} (J_{ea} + \int_0^x R dx) - J \right). (4.8)$$

If traps are present, the relation between n and p is more complicated but basically it is fully determined. The recombination function R is also some fixed function of the carrier densities, while the boundary electrode current J_{ea} is fixed in some way, as will be considered later.

The tremendous power of equation (4.7) or (4.8) may now be observed. The equation gives the carrier gradient at the point x assuming that the carrier density is known there. This is just what is required for a Runge-Kutta step by step integration. Note that the recombination integral requires values for the carrier densities only as far as integration has proceeded. Thus we are able to use a single Runge-Kutta process which is technically suitable only for first order differential equations, on this second order equation. We would point out

that this method of dealing with recombination integrals by starting integration from one end of the crystal is of far wider applicability than just to this case. It could make computation simpler for recombination in laser diodes, and in other calculations where recombination is important. Yet this technique of applying single equation Runge-Kutta methods to second order transport equations does not seem to be in current use.

The advantages of the above treatment become apparent when previous attempts at solving the double injection problem are examined. The classical approach is employed by Gunn (l.c. 1958, 426), who obtains the differential form of equation (7), which he abandons as unmanageable. N. R. Howard and G. W. Johnson (1965, 427) fall into the trap of confusing $\Delta p - \Delta n$ and dE/dx ; in spite of deleting important terms and using a very simple recombination law to facilitate integration, the boundary conditions are found to require the solution of four simultaneous implicit equations in hyperbolic functions. While suggesting use of a computer, they make no attempt to proceed further. Russian work has followed a different line; here the approach has been by way of calculation of the change in the effective diffusion length (Stafeev l.c. 1959, 428). The ultimate is reached by V. P. Sondaevskii, V. I. Stafeev (1964, 429), who provide an expression for a single current-voltage point which virtually covers a printed page.

4.1.7 Development of the Equation.

So far we have attempted to show that double injection is a current flow process requiring a different approach from ohmic flow and that under certain conditions negative resistance is a distinct possibility. It has been a theme of this work that a solution of the double injection problem depends directly on a knowledge of the boundary conditions --- in fact that contact constraints determine the current flow, in contrast to Lampert's thesis (l.c. 1964, 430). The second theme, in agreement with Lampert, is that the exact recombination function is of vital importance if even a rough picture of the current-voltage relationship is to be given. In the next sections we discuss boundary constraints and recombination at some length before tackling one or two objections to the current flow concepts being put forward. This will conclude the conceptual consideration of carrier flow under the influence of an electric field alone.

4.1.8 The First Boundary Condition.

The solution of the new integral equation (4.7^h) by a Runge-Kutta process relies on two factors --- knowledge of the inverse injection ratio J_{ea}/J and a value for the local electron concentration at the hole injecting anode. These represent the two constants of integration. We set out to find an estimate for the injection ratio.

The current at the injecting anode consists partly of holes injected and partly of electrons removed. Assume a p^+-n junction

with a step concentration gradient. The injection ratio may be expected to depend on the carrier concentrations just on the p^+ side, defined as p_a^+ , n_a^+ and just on the n side, defined as p_o , n_o ; also on the total current J , the electric fields, and possibly on the mobilities. At a low current the forward i - V characteristic for the junction is exponential and is reasonably well understood (Smith l.c. 1959, 431, p.269). Carrier accumulation may take place and have an effect on the injection ratio, as described by J. B. Arthur, A. F. Gibson, J. B. Gunn (1956, 432). These treatments cease to hold at the high currents to be considered here. The i - V characteristic becomes linear, being dominated by the resistance of the bulk material of the diode; the voltage across the junction itself seems likely to grow until it is nearly equal to the band gap, at which value as many carriers as are required can be supplied. For example, for a germanium diode at 1 amp./sq. mm., the voltage across the junction may be 0.7 volts and the total voltage 10 volts. The problem arises over the size of the injection ratio. On this point the literature appears to be silent. Since calculation cannot proceed without some sort of estimate, a hypothesis is provided as follows:

$$\frac{J_{ha}}{J} = \frac{p_a^+ - p_o}{p_a^+ - n_a^+ - p_o + n_o},$$

$$\frac{J_{ea}}{J} = \frac{n_o - n_a^+}{p_a^+ - n_a^+ - p_o + n_o}.$$

Here the energy bands are treated as virtually level, and the flow

rates of carriers across are proportional to their local concentrations. By Boltzmann's equation:

$$p_a^+ n_a^+ = p_o n_o \quad .$$

But $p_a^+ \gg n_o, p_o \quad .$

Hence $J_{ea}/J \approx n_o/p_a^+ \quad .$

Similarly at the other electrode,

$$J_{hc}/J \approx p_L/n_o^+ \quad .$$

Mobilities have been ignored because these do not seem to be directly relevant, especially since the electric field is heavily distorted at the junction.

The simple expression above has a number of advantages. Firstly it depends only on the number of electrons at the electrode, without reference to local fields or density gradients. Secondly it is particularly suitable for dealing with carrier accumulation, because as the carriers accumulate, so the probability of their crossing the junction increases, and the injection ratio falls. Thus, raising the current results in a reduced injection ratio, which at high currents seems to be found experimentally. Problems of infinite barrier recombination velocities, zero electrode fields and other dubious boundary conditions, which have been employed elsewhere (W. M. Webster 1954, 433; J. Lindmeyer, A. Slobodskoy 1963, 434; see K. M. van Vliet 1966, 435), are eliminated. It is emphasized that for the time being the above expression, which is physically meaningful, does allow

calculation to proceed. We have thus fixed one boundary constraint.

4.1.9 The Carrier Density at the Boundary.

A Runge-Kutta integration process requires a value for an initial point, in this case the carrier density at the anode. The fact that this is not a known boundary constraint does not mean that it is undetermined. Consider equation (4.8) written as follows:

$$\frac{d}{dx} (p) = N(p) \left[F(p) \left\{ J_{ea} + \int_0^x R dx \right\} - J \right] ,$$

$$\frac{J_{ea}}{J} = \frac{n_a + p(x=0)}{K_a} .$$

Choose an arbitrary value for $p(x=0)$; then step by step integration provides a value for $p(x=L)$. However the integration could equally well have proceeded by starting at the cathode, where the hole current is to be given by:

$$\frac{J_{hc}}{J} = \frac{p(L)}{K_c} .$$

In fact the correct value of $p(0)$ is such that the current proportions at the cathode J_{hc} , J_{ec} match the carrier concentration $p(L)$ to the cathode boundary injection ratio.

Thus the two constants of integration are provided by the conditions at the two boundaries; matching the carrier concentrations and flows given in the basic differential equation (4.8) is performed by successive approximation on the parameter $p(0)$ until both conditions are simultaneously satisfied.

4.1.10 Recombination and Trapping.

In space-charge neutralized double injection, recombination is a critical feature. It has been pointed out that recombination limits the current flow, section 4.1.2 and that recombination at trap sites governs the phenomenon of negative resistance, section 4.1.3. Since the new integral equation has been explicitly derived to deal with these two features, it is no surprise that a recombination function forms the kernel of the computation. It is thus unfortunate that little is known quantitatively about recombination functions, and about recombination sites.

We will consider for the most part a single type of imperfection level in the forbidden band, acting as both a trap and a recombination site. Four methods of dealing with such a level will be considered, roughly in order of decreasing rigour.

4.1.10.1 Shockley-Read Theory.

The classical formulation for recombination is that of W. Shockley, W. T. Read (1952, 436). The notation employed here is from R. H. Bube (1960, 437, p.305). The density of recombination centres is N_r , their depth below the conduction band E_r , the band gap is E_g , the probability of electron occupation f_r , and the electron and hole capture cross-sections C_n , C_p ; then the electron and hole capture rates U_{en} , U_{op} may be written as:

$$U_{cn} = n(1 - f_r)C_n - f_r C_n N_c \exp(-E_r/kT) ,$$

$$U_{cp} = pf_r C_p - (1 - f_r)C_p N_v \exp(-(E_g - E_r)/kT) .$$

The two terms relate to capture and thermal ejection respectively. We put:

$$n' = N_c \exp(-E_r/kT) ,$$

$$p' = N_v \exp(-(E_g - E_r)/kT) ,$$

$$n'p' = n_i^2 = N_c N_v \exp(-E_g/kT) .$$

Thus n' and p' are the densities of free electrons and holes with the Fermi level at the recombination centre position.

When electrons and holes are being generated at a constant rate U , and in the steady state, we have: $U_{cn} = U_{cp} = U$. Thus f_r may be eliminated to give:

$$U = \frac{C_n C_p (pn - p'n')}{C_n (n + n') + C_p (p + p')} .$$

Because of differential trapping, the relation between n_0 and p_0 is not predetermined. Put:

$$p = p_0 + \delta p ,$$

$$n = n_0 + \delta n ,$$

$$U = \frac{p_0 \delta n + n_0 \delta p + \delta p \delta n}{(n_0 + \delta n + n')/C_p + (p_0 + \delta p + p')/C_n} .$$

For high injection levels, three regimes may be considered:

(a) Take $C_p \gg C_n$, $n_0 \gg p_0$, and a central recombination level.

$$\text{Then } U \approx C_n (n_0 + \delta n) .$$

(b) Take $C_n \gg C_p$, $n_0 \gg p_0$, and a central recombination level.

$$\text{Then } U \approx C_p \delta p \quad .$$

(c) Take $C_n = C_p = C$, $n_0 \gg p_0$, $\delta n = \delta p = \delta$.

$$\text{Then } U = C \delta \frac{n_0 + \delta}{n_0 + 2\delta} \quad .$$

The results for (a), (b) are to be expected; if the carrier capture rates are very unequal, the recombination rate is determined entirely by the more volatile species. For (c), neutrality demands that if the capture rates are equal, so should be the added carrier densities.

4.1.10.2 Hole-trap recombination.

The previous section has been concerned mainly with the neutral centre, which may capture either a hole or an electron. In this section, we consider a negatively charged recombination centre, which can only function by first capturing a hole. The notation used is similar, but f_r' represents the hole occupation probability.

$$U_{cp} = pC_p(1 - f_r') - p'f_r'C_p \quad ,$$

$$U_{cn} = nC_n f_r' \quad ,$$

$$f_r' = \frac{pC_p}{nC_n + (p + p')C_p} \quad ,$$

$$U = \frac{C_p C_n p n}{nC_n + (p + p')C_p} \quad . \quad (4.9)$$

Consider n-type material with an injected carrier concentration δ ; the number of free holes is smaller than that of free electrons due

to trapping.

$$n = n_0 + \delta ,$$

$$p = \delta - N_r f_r' = \delta - \frac{p C N_r}{n C_n + (p + p') C_p} ,$$

$$p = n - n_0 - \frac{p N_r}{n C_n / C_p + p + p'} . \quad (4.10)$$

In a double injection computation for equation (4.7), the recombination to be employed is (4.9) and the relation between the carrier densities n, p is (4.10). It is immediately apparent that equation (4.10), an implicit quadratic relationship between n and p , is not an easy expression to handle.

4.1.10.3 Fixed hole lifetimes.

In this section we examine the case when the hole is assumed to have a lifetime τ in a trap independent of the electron carrier density; thermal ejection is neglected.

Then the recombination rate is:

$$p C_p (1 - f_r') = N_r f_r' / \tau ,$$

$$f_r' = \frac{p \tau C_p}{N_r + p \tau C_p} .$$

Let the number of free holes be p ; then the number of free electrons is:

$$n = n_0 + p + \frac{p \tau C_p N_r}{N_r + p \tau C_p} ,$$

$$= n_0 + p + \Delta .$$

So:

$$\frac{d}{dx} (np) = \left\{ n_0 + 2p + \frac{p\tau C_p N_r^2}{(N_r + p C_p)^2} (2 + p\tau C_p / N_r) \right\} \frac{dp}{dx} ,$$

$$= \{ n_0 + 2p + \Delta' \} \frac{dp}{dx} .$$

Applying this to the double injection equation (4.7),

$$\frac{dp}{dx} = \frac{n_0 + p + \Delta}{n_0 + 2p + \Delta'} \frac{1}{\mu_h kT} \left\{ \left(1 + \frac{p\mu_h}{(n_0 + p + \Delta)\mu_e} \right) (J_{sa} + \int_0^x R dx) - J \right\}$$

where Δ , Δ' are defined in the equations above, and

$$R = \frac{e N_r p C_p}{N_r + p\tau C_p} .$$

With these equations computation is straightforward.

4.1.10.4 Saturated trap recombination.

In this fourth section we assume that the hole capture probability is so great that no holes can remain free in a region where the traps are not fully saturated. The trap saturation front is thus quite sharp. On the cathode side $n=n_a$, $p=0$, $R=0$. On the anode side $n=n_a+n_t$, $p=n-n_a-n_t$ where n_t is the trap density. The recombination rate through the traps will be a function of n_t and n , while some subsidiary recombination mechanism to deal with high carrier densities may be added. So:

$$R = (S n_t + T p) n .$$

Returning to equation (4.7),

$$\frac{d}{dx} (np) = \frac{n}{\mu_h kT} \left\{ \left(1 + \frac{p\mu_h}{n\mu_e} \right) (J_{sa} + \int_0^x R dx) - J \right\} ,$$

it is apparent that on the cathode side of the saturation front F , n and p are constant, and:

$$J = J_{sa} + \int_0^F R dx \quad . \quad (4.11)$$

The integration boundary conditions have thus been affected. The anode condition remains, but the second condition has moved from the cathode to the saturation front. The position of F must satisfy equation (4.11) so that the total current J can be accounted for by recombination and leakage. But at F , we require $p=0$ and $dp/dx=0$. The situation is thus fully defined; equation (4.7) is solved for arbitrary $n(0)$ and F is defined as the point where $p=0$. Then the correct value of $n(0)$ is the one for which $dp/dx=0$ where $x=F$ and $p=0$.

4.1.10.5 Recombination functions for computation.

In choosing a recombination function for computation, two lines of approach have been employed. Firstly it has been felt desirable to demonstrate the use of the new integral equation for Shockley-Read recombination; secondly the aim has been to show that the equation can produce a negative resistance characteristic for recombination at trap sites; since calculations in the literature have only indicated negative resistance thresholds (Ashley and Milnes l.c. 1964, 438) rather than following the curve round its course, this is thought to be a considerable undertaking.

For ordinary recombination, we return to the Shockley-Read formulation of section 4.1.10.1. Thus we use equation (4.8):

$$\frac{d}{dx}(f) = \frac{n_a + f}{n_a + 2f} \frac{1}{r_h kT} \left(\left\{ 1 + \frac{r_{ph}}{(n_a + f)r_e} \right\} (J_{ea} + \int_0^{\infty} R dx) - J \right),$$

with $R = eCf \frac{n_a + f}{n_a + 2f}$.

In this case the constants to be chosen will be based on experimental values.

In the case of negative resistance, the most extreme of the cases considered is chosen, section 4.1.10.4. The constants will be chosen more with an eye to improving the chances of negative resistance than to plausibility.

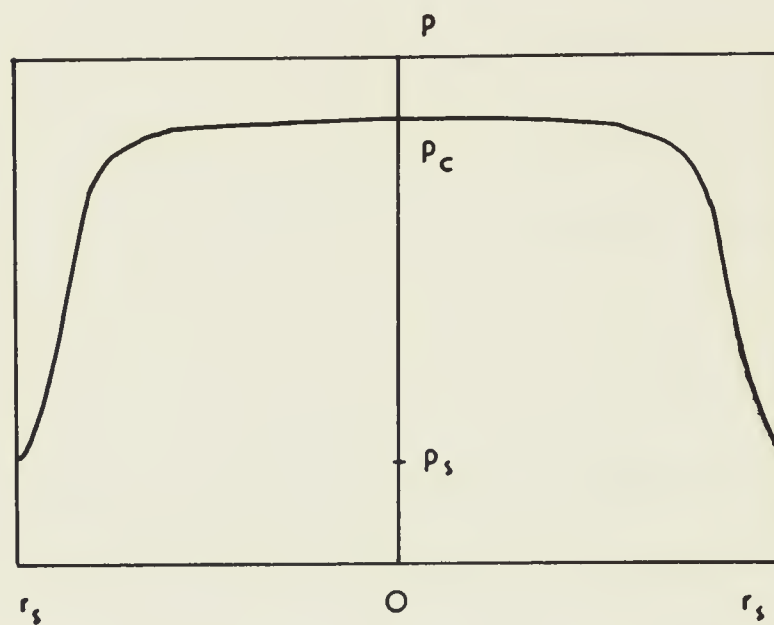
4.1.11 Objections to Planar Flow.

In the treatment given, we have assumed planar flow, by considering a one dimensional analysis. Two features which may invalidate this will be investigated in this section, namely surface effects and funnelled flow.

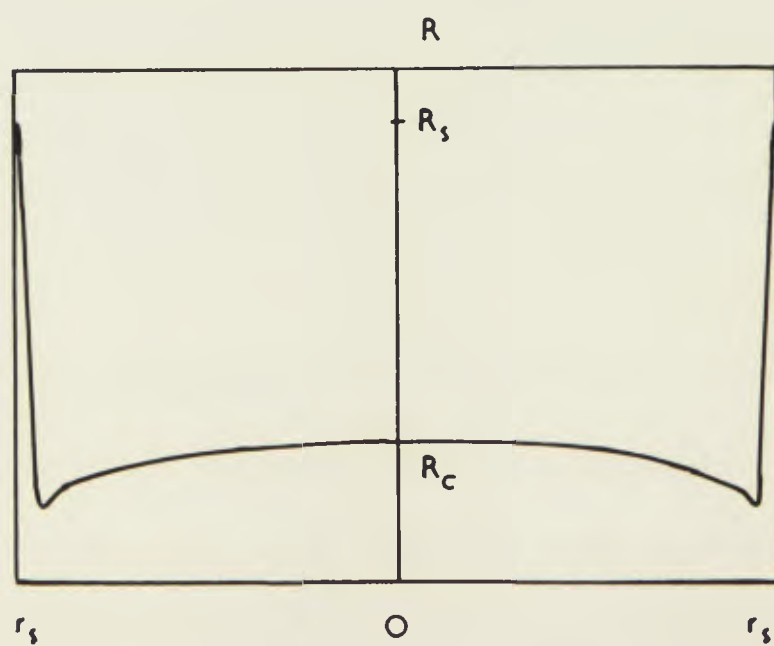
The rate of minority carrier recombination at surfaces is known to be larger than in bulk material (R. Suhl, W. Shockley 1949, 439), due to defects and impurities, which form extra recombinations levels in the forbidden band. Due to these and to other surface layers, a surface current may also flow; this surface current is negligible for low resistance specimens, as will be shown in section 7.4.

Figure 4.1.3

Radial Diagrams



(a) Carrier Density



(b) Rate of Recombination

However the enhanced surface recombination is an important factor.

High surface recombination causes a component of current flow from the bulk of the material towards the surfaces. This flow is diffusion driven, requiring a higher concentration of non-equilibrium carriers at the centre of the specimen. The effect is complex. Consider a long cylindrical specimen. The holes for recombination in an off axis small volume are provided partly by carriers moving parallel to the axis and partly by those flowing radially towards the surface, while the recombination rate depends on the local non-equilibrium concentration. Under these circumstances the equipotentials themselves become distorted. The carrier flow and continuity equations are:

$$j_h = \mu_h e p E - D_h e \nabla p \quad ,$$

$$\nabla j_h = R(p) \quad ,$$

$$j_e = \mu_e e n E + D_e e \nabla n \quad ,$$

$$\nabla j_e = -R^*(n) \quad .$$

The radial picture may be shown as in figure 4.1.3a. Provided that surface recombination is only a small proportion of total recombination, i.e. that the main contribution to the integral under the curve of figure 4.1.3b does not come from the wings, surface recombination will not have an appreciable effect on the i-V characteristic. The surface effect is limited to a depth of the order of the minority carrier diffusion length; if the diffusion length is

10^{-3} cm. and the dimension 10^{-1} cm., this is satisfied. This is to be compared with the field driven minority carrier penetration distance down the specimen of up to 1 cm. Thus we require $L(E=0) \ll d$, $d < l = L(E)$, where $L(E)$ is the diffusion length in a field E and d , l are the specimen width and length. These conditions have been satisfied in the experiments to be described, and are assumed in the computations.

The possibility of funnelled flow has received scanty attention to date. Funnelled flow is most clearly demonstrated as the spark discharge at atmospheric pressure; another example is a gas plasma concentrated by magnetic fields. In solid state physics, heat has been considered as a flow concentrating factor (R. E. Burgess 1960, 440, p.818). Thus at very high field strengths a path of hot material with extra intrinsic carriers forms a lower resistance route which absorbs the main current flow, thus leaving the rest of the crystal in a low current density state. This will be considered further in chapter 5.

Trap sites present a possible cause of funnelled flow. In a semiconductor crystal containing deep level acceptor impurities such as germanium doped with zinc, overcompensation means that each impurity atom has appropriated two electrons. Thus in an impurity dominated scattering regime, the extra electric charge on the impurities will have the effect of increasing the scattering. During hole injection, these impurities trap holes, which may be lost either

by recombination or ejection. While a hole is trapped, the electric charge is reduced and hence the scattering is reduced for both holes and electrons. For zinc-doped germanium, a higher electron mobility for heavy illumination observed by W. W. Tyler, H. H. Woodbury (1956, 441) has been interpreted as the effect of hole trapping on the ionized impurity scattering. Further, if the trap is an important factor in the rate of recombination, the local saturation of traps will lead to a longer hole lifetime in the region of the funnel.

Thus trap saturation may increase the local conductivity either by reducing the scattering or the recombination. As shown above, surface effects also favour this situation. For an appreciable effect, the scattering has to be dominated by acceptor impurity scattering, or the recombination by the trap sites. Under favourable circumstances, the specimen resistance change may be expected to be violent, comparable to the spark discharge. This could explain the change by a factor of 1000 observed for iron-doped germanium by W. W. Tyler (1954, 442). Where circumstances are less favourable, diffusion counteracts funnelling; in the case to be considered, the mobility has been taken as independent of trap occupation, and the hole density planar.

4.1.12 Magnetic Field Effects.

So far we have considered the double injection current-voltage relationship, and have shown that as the minority carriers penetrate

down the crystal, non-linear changes in the local potential may be expected; experimental evidence of this is to be presented. A further method for the investigation of double injection is provided by the Hall effect, with the application of a magnetic field.

We have already pointed out that diffusion plays an important role in balancing the double injection current and that this effect is especially important near the anode; here the electrons are to be expected to stagnate, making the current flow due mainly to holes. We may thus anticipate a corresponding effect on the Hall voltage. The effective electric field mobilities μ_h , μ_e are modified by diffusion to $\mu_h^*(x)$, $\mu_e^*(x)$, which are position and current dependent. A complete reappraisal of the formulae for the Hall coefficients is thus essential, and this will now be tackled. The effect of true magnetoresistance on carrier flow in injection in germanium is small compared with other effects and will be neglected.

4.1.13 The Hall Coefficient.

The effects produced on injected carriers by a magnetic field are listed by W. C. Dunlap (1957, 443, p.173) as follows:

- (1) Photoelectromagnetic effect (PEM).
- (2) Suhl effect.
- (3) Landauer-Swanson effect.
- (4) Magnetic rectification.

Illumination and hence (1) does not concern us here; attention will

be concentrated on (2), (3). Recent interest in magnetic field studies for injected carriers seems to be small. Perhaps a reason for this is the mathematical complexity of the exact picture, as described by R. Landauer and J. Swanson (1953, 444).

The Suhl effect (Suhl, Shockley l.c. 1949, 445) concerns the magnetic deflection of carriers towards a surface where increased recombination takes place. A basic problem exists here, in the two carrier transport equations, which is glossed over in standard texts e.g. Smith (l.c. 1959, 446), W. Shockley (1956, 447); the analysis presented comes from Landauer and Swanson (l.c. 1953, 448). The deflecting force exerted by a magnetic field on a carrier is proportional to the carrier velocity and hence depends on its mobility, while the force exerted by the electric field is not dependent on the mobility. Hence when electrons and holes have different mobilities, a Hall field cannot produce zero transverse currents for both holes and electrons. There are two possible escape routes:

- (1) Take $J_y = 0$, but with J_{ey} , $J_{hy} \neq 0$.
- (2) Assume that the imbalance is corrected by diffusion, so that $J_{ey} = J_{hy} = 0$.

Answer (1) means that electrons and holes are both moving to the same side of the crystal in equal numbers, thus preserving neutrality. It is thus closely connected with surface recombination; for if surface recombination is not large enough to support this transverse flow, a carrier accumulation will set up an opposing diffusion

process. In the limit of negligible surface recombination, answer (2) applies. The mathematical equations for current flow are:

$$J_h = e\mu_h pE - eD_h \nabla p + \mu_h pH \cdot \nabla \mathbf{e} ,$$

$$J_e = e\mu_e nE + eD_e \nabla n + \mu_e nH \cdot \nabla \mathbf{e} .$$

Considering transverse flow, with current flow along the x-axis, Hall field along the y-axis and magnetic field in the z direction, the equations become:

$$J_{hy} + eD_h \frac{dp}{dy} = e\mu_h pE_y + \mu_h^2 pE_x H_z \mathbf{e} ,$$

$$J_{ey} - eD_e \frac{dn}{dy} = e\mu_e nE_y - \mu_e^2 nE_x H_z \mathbf{e} ,$$

$$J_{hy} + J_{ey} = 0 ,$$

$$\frac{dn}{dy} = \frac{dp}{dy} .$$

These equations are similar to those used by R. H. Fowler (1936, 449, p.428). Then we may consider either of the two cases by neglecting the diffusion terms or the field driven terms.

$$(1) \quad R = \frac{E_y}{\sigma H E_x} = \frac{p\mu_h^2 - n\mu_e^2}{\sigma(p\mu_h + n\mu_e)} ;$$

$$(2) \quad R = \frac{p\mu_h - n\mu_e}{\sigma(p + n)} .$$

A partial solution for the complete equation is discussed by P. C. Banbury, H. K. Henisch and A. Many (1953, 450) but is too complex for the expression of the Hall coefficient. It is interesting to note that for $\mu_e = \mu_h$, the two expressions become identical; in fact

the problem itself ceases to arise.

So far in this section we have considered carrier flow where the only longitudinal driving force is the electric field. Our particular concern is for the modification caused by diffusion, due to carrier injection and accumulation. For this we reintroduce the effective mobilities μ_e^* , μ_h^* of section 4.1.12, defined as:

$$J_h = ep\mu_h E_x - \mu_h kT p' = pe\mu_h^* E_x ,$$

$$J_e = en\mu_e E_x + \mu_e kT n' = ne\mu_e^* E_x .$$

Only longitudinal mobilities are to be replaced by asterisked quantities; thus the transverse diffusion terms are unchanged, while one occurrence of μ in the Hall effect terms, the one which balances the transverse field, is also unaffected by the downstream diffusion. For the two cases we have:

$$(1) \quad R = \frac{p\mu_h^* - n\mu_e^*}{(p\mu_h^* + n\mu_e^*)(p\mu_h^* + n\mu_e^*)} ;$$

$$(2) \quad R = \frac{p\mu_h^* - n\mu_e^*}{(p\mu_h^* + n\mu_e^*)(p + n)} .$$

From the definition of μ_h^* , μ_e^* we have:

$$(1) \quad R = \frac{\mu_h J_h - \mu_e J_e}{e(p\mu_h^* + n\mu_e^*)J} ;$$

$$(2) \quad R = \frac{J_h - J_e}{e(p + n)J} .$$

Again both expressions become the same for equal mobilities. For computational purposes, this assumption will be made.

The observable Hall voltage is given by:

$$V_H = c' \frac{HJ}{st} \int_{-t/2}^{t/2} R_y dy ,$$

where s , t are the specimen dimensions and c' a dimensional constant.

We assume that planar flow is a good approximation, so that R_y is independent of y , giving:

$$(1) \quad V_{H(1)} = c' \frac{H}{es} \frac{\mu_h J_h - \mu_e J_e}{(\mu_h p + \mu_e n)} ;$$

$$(2) \quad V_{H(2)} = c' \frac{H}{es} \frac{J_h - J_e}{(p + n)} .$$

Since J_h and J_e are functions of position along the crystal, so also will be the Hall voltage. At a particular point in the middle of the crystal, the current at low levels is entirely an electron current and the Hall voltage is given by:

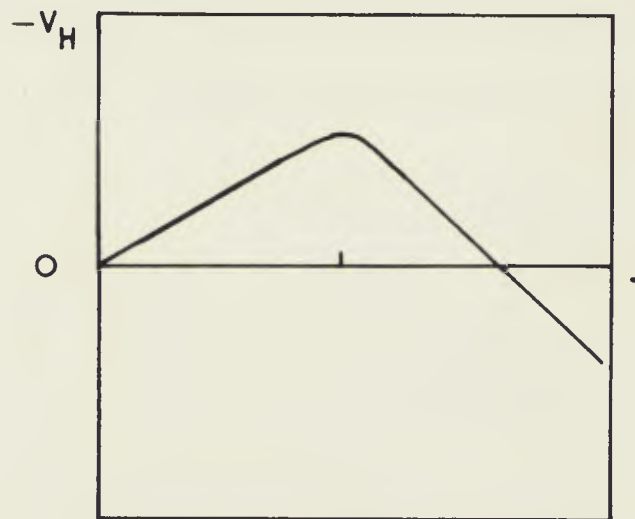
$$V_{H(\text{low level})} = c' \frac{H}{es} \left(- \frac{J}{n} \right) .$$

As the current is increased, the Hall voltage is reduced by two factors; the increase in the carrier density caused by double injection affects the denominator, and the change from an electron to a hole dominated current the numerator. If carrier exclusion is strong enough, inversion of the Hall voltage is to be expected when the numerator passes through zero. This gives an anticipated Hall curve as shown in figure 4.1.4a.

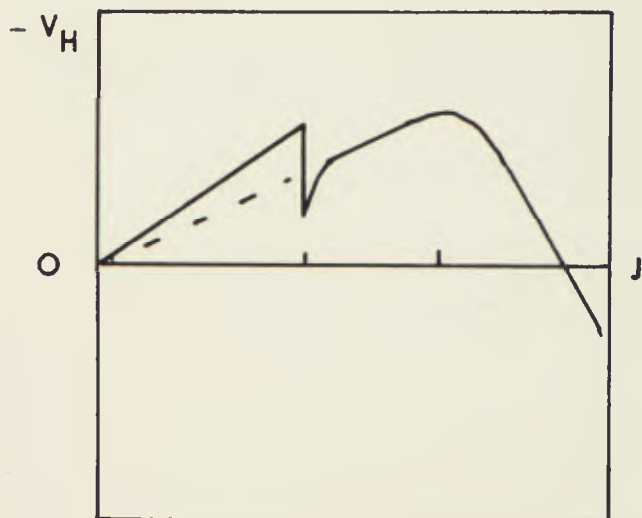
In a crystal containing hole traps, a further phenomenon is possible. In the limit of strong trap recombination, we have shown

Figure 4.1.4

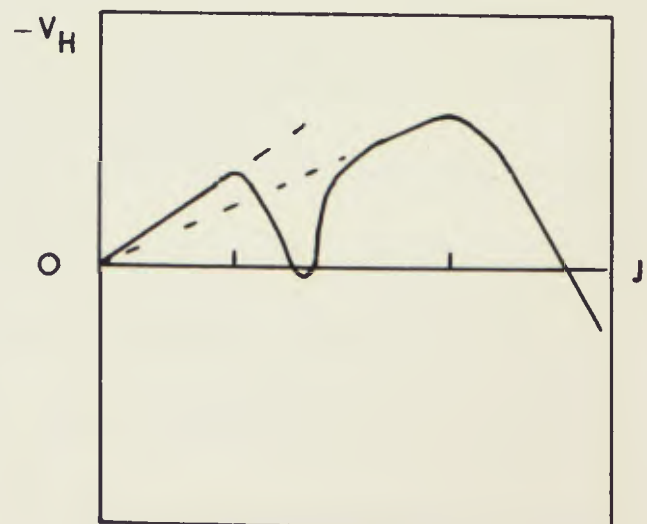
Hall Voltage Inversion Curves



(a)



(b)



(c)

in the section on recombination, that a sharp trap saturation front may exist. This front is driven down the crystal as the current is raised; the trap front passes through a given point at a lower current than the Hall voltage inversion, since the trap front indicates a low hole current proportion, while inversion represents approximately half hole flow. A completely sharp saturation front might give a Hall voltage picture as indicated in figure 4.1.4b; the reason for the spike is local back diffusion of electrons. A broader saturation front is shown in figure 4.1.4c. An increased magnetic field might be expected to broaden the front, to reduce its penetration at a given current and to delay the inversion point, due to increased surface recombination.

Further discussion will be left until the experimental data on injection has been presented. Meanwhile computational results will be described.

4.2 Computation.

4.2.1 The Basis for Computation.

The preceding sections have been devoted to the derivation of carrier flow equations, the analysis of recombination and of boundary conditions, and the consideration of the Hall voltage. There seem to be four effects possible, given the right conditions: a local negative resistivity, an overall negative resistance, a trap saturation front and a Hall voltage inversion. The purpose of the computation is firstly to demonstrate the power of the analytical technique and secondly to define more clearly the conditions required for these effects.

The basic equation (equation 4.7', section 4.1.6) gives the change of carrier concentration with position as:

$$\frac{d}{dx} (np) = \frac{n}{\mu_h kT} \left\{ \left(1 + \frac{p\mu_h}{n\mu_e} \right) (J_{ea} + \int_0^x R dx) - J \right\} .$$

The following assumptions are made:

(a) $\mu_e = \mu_h$. This is for simplicity in calculation; also it evades the problem of the Hall coefficient discussed in section 4.1.13. In practice it seems to be reasonably justifiable for germanium in the temperature range 20° K --- 77° K.

(b) As discussed in section 4.1.8, the injection levels are taken as:

$$\frac{J_{ea}}{J} = \frac{n_o}{p_a^+ + n_o - p_o} ,$$

$$\frac{J_{ho}}{J} = \frac{p_L}{n_o^+ + p_L - n_L} ,$$

(c) (1) Shockley - Read recombination. For this case:

$$p = f ,$$

$$n = n_a + f ,$$

$$R = eCf \frac{n_a + f}{n_a + 2f} ,$$

from section 4.1.10.1; by substitution and rearrangement:

$$\frac{d}{dx}(f) = \frac{1}{\mu kT} \left\{ J \left(\frac{n_a + f(0)}{p_a^+ + n_a} - \frac{n_a + f}{n_a + 2f} \right) + eC \int_0^{\infty} \frac{f(n_a + f)}{(n_a + 2f)} dx \right\} .$$

(2) Trap saturation recombination (section 4.1.10.4).

$$x > F : n = n_a ,$$

$$p = 0 ,$$

$$R = 0 .$$

$$x < F : n = n_a + n_t + f ,$$

$$p = f ,$$

$$R = S \cdot n_t \cdot n + T \cdot p \cdot n .$$

$$\begin{aligned} \frac{d}{dx}(f) = \frac{1}{\mu kT} \left\{ J \left(\frac{n_a + n_t + f(0)}{p_a^+ + n_a + n_t} - \frac{n_a + n_t + f}{n_a + n_t + 2f} \right) + e \int_0^F (S n_t + T f) \cdot (n_a + n_t + f) dx \right\} . \end{aligned}$$

where at $x = F$, $p = 0$ $\frac{dp}{dx} = 0$, unless $F > L$.

4.2.2 Methods of Computation.

The problem has been reduced to a computation of the form:

$$\frac{dy}{dx} = g(x,y) \quad .$$

As pointed out in section 4.1.6, the two integrations of the second order equation are carried out together; the limits of the first integration are entirely to one side of those of the second, which is defined by the equation above; hence the first constitutes a quasi-constant in the main step-by-step integration process.

Wide varieties of step-by-step integration processes for dealing with such a differential equation are available (L. Fox 1962, 451). Basically there are two categories; those which use information from previous steps and those which do not. The disadvantage of using previous information is the difficulty in obtaining a first point or in changing the step length; the advantages lie in fewer computations and in a calculable probable rounding error. Only the briefest justification for the method chosen will be given. Since the calculation is at an experimental stage, it has been felt that there is no point in using one of the more elaborate methods (e.g. that of Adams and Bashford involving six previous points and three successive fresh slope calculations for each step). Because evaluation of the differential requires considerable computation, due to the complication of the integral, a method using information from previous steps has been reckoned to be advantageous, ruling out Runge-Kutta techniques.

So a predictor-corrector method has been chosen, as outlined by D. D. McCracken and W. S. Dorn pp.330-341 (1964, 452).

4.2.3 The Predictor - Corrector Formula.

The solution process is defined for a step of length h from x_m to x_{m+1} for the equation $dy/dx = g(x,y)$;

$$1^{st} \text{ value} \quad y_{m+1}^{(0)} = y_m + 2h g(x_m, y_m) ,$$

$$i^{th} \text{ value} \quad y_{m+1}^{(i)} = y_m + h \{ g(x_m, y_m) + g(x_{m+1}, y_{m+1}^{(i-1)}) \} / 2 ,$$

$$\text{final value} \quad y_{m+1}^{(f)} = (4y_{m+1}^{(r)} + y_{m+1}^{(0)}) / 5 .$$

The process is halted when $|y_{m+1}^{(r)} - y_{m+1}^{(r-1)}| < \epsilon$, a specified limit;

the mean truncation error for the process is $E_T = (y_{m+1}^{(0)} - y_{m+1}^{(r)}) / 5$.

This represents convergence to the wrong point; the process chooses a new point such that the slope of the chord equals the average of the slopes of the two tangents. Clearly it is not useful to pursue convergence to this arbitrary point further than to the extent that this point represents the true curve value.

The predictor-corrector method requires an initial point. This is provided in this computation by a single Euler point:

$$y_{1E}^{(0)} = y_0 + h g(x_0, y_0) ,$$

$$y_{1E}^{(i)} = y_0 + h \{ g(x_0, y_0) + g(x_1, y_{1E}^{(i-1)}) \} / 2 .$$

In order to cope with the lower accuracy of this step, $h(E) = h/32$ is

employed. This involves doubling the interval five times before the ultimate y_1 is obtained for step-by-step progress.

4.2.4 Programmes Employed.

Two Algol programmes have been employed, written for a KDF9 machine; PCSR employs Shockley-Read recombination statistics; PCSTF deals with a trap saturation front. PCSTF, which represents an extended version of PCSR, is shown in the appendix together with the store list.

The most difficult feature of the computation is the change in the boundary condition when the saturation trap front reaches the far end of the crystal as the current is raised. This takes place when trip takes the value 2 and is accomplished without any anomalous behaviour. The only operational problem is a tendency for the carrier density to "overflow" if the upper bound at the first point is chosen to be too large.

4.2.5 Results for PCSR.

The parameters employed have been:

specimen natural electron density	$2 \cdot 10^{15}/\text{c.c.},$
electrode carrier densities	$10^{18}/\text{c.c.},$
carrier mobilities in the specimen	$3 \cdot 10^4 \text{ cm./volt/sec.},$
temperature	$77^\circ \text{ K.},$
minority carrier lifetime	$5 \cdot 10^{-5} \text{ sec.},$
specimen dimensions	$0.5 \cdot 0.1 \cdot 0.1 \text{ cm.},$
magnetic field	200 gauss.

Figure 4.2.1

Variation of Injected Carrier Density
across Crystal for PCSR

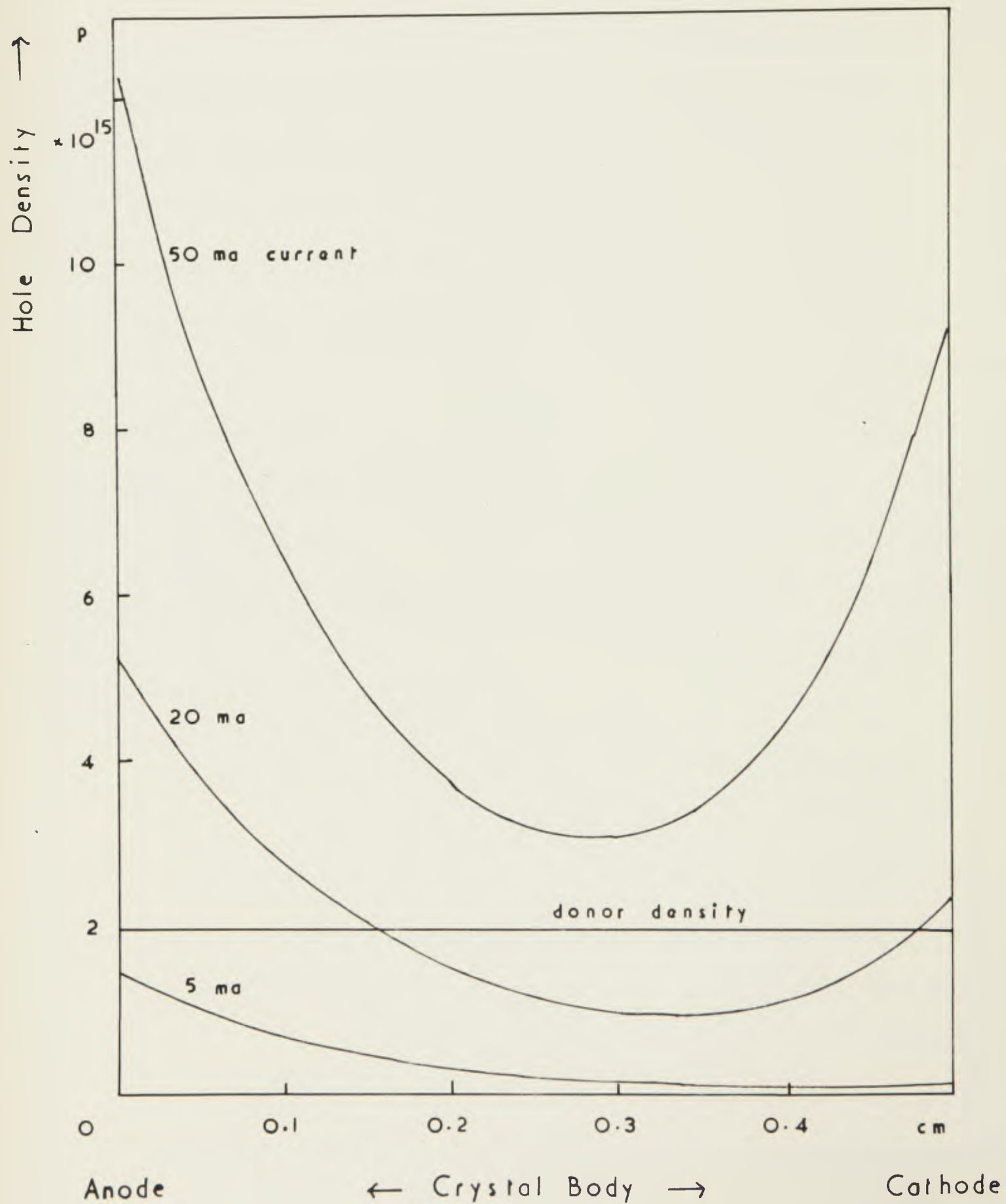


Figure 4.2.2

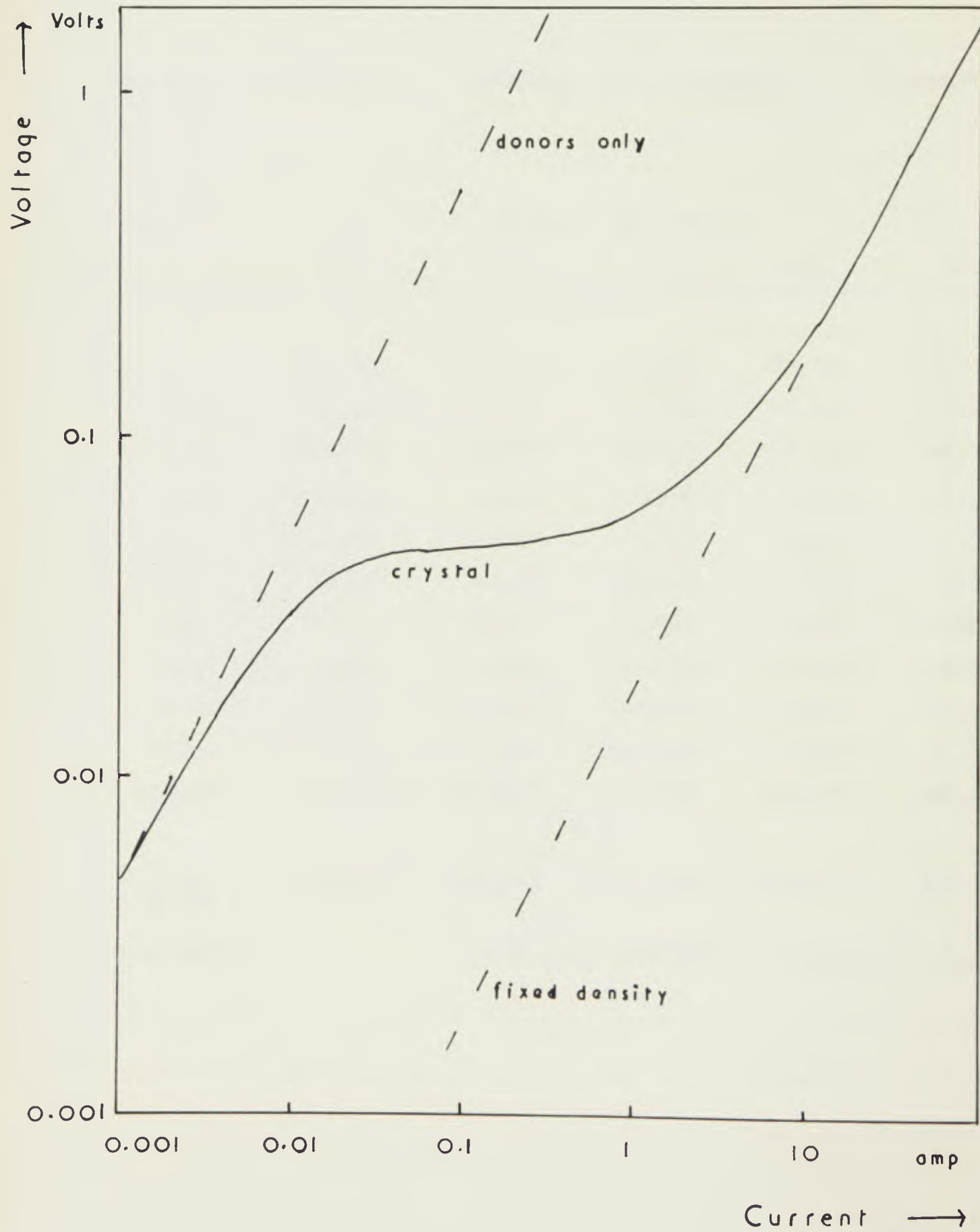
Hole Densities across Crystal for Different Numbers of Steps

Densities $10^{15}/\text{c.c.}$ Current 100 milliamp. Distances cm.

Position	Number of Steps				
	20	40	80	160	320
Anode	23.675	23.410	23.239	23.151	23.106
0.05	16.948	16.777	16.684	16.636	16.612
0.1	12.362	12.288	12.247	12.225	12.214
0.15	9.373	9.352	9.338	9.330	9.325
0.2	7.568	7.573	7.572	7.571	7.570
0.25	6.727	6.747	6.753	6.754	6.755
0.3	6.810	6.844	6.856	6.861	6.863
0.35	7.926	7.985	8.009	8.019	8.024
0.4	10.287	10.396	10.444	10.467	10.477
0.45	14.189	14.394	14.490	14.536	14.588
Cathode	20.070	20.447	20.627	20.716	20.759
Voltage (v.)	46.396	46.867	47.132	47.273	47.345
V. change		1.02 %	0.57 %	0.30 %	0.15 %

Figure 4.2.3

Current - Voltage Characteristic for PCSR



The minority carrier density across the specimen is shown in figure 4.2.1 for three current values. Notable features are that the minimum density is between the electrodes as predicted in section 4.1.2, and that this minimum moves towards the centre of the specimen as the current is increased; this is again to be expected; at low currents the proportion of holes reaching the far end is small so that a normal diffusion length solution is in order, while at high currents and injected carrier densities, the donor electron density becomes negligible, giving symmetrical electron and hole flows for identical mobilities and boundary conditions.

Figure 4.2.2 indicates the computational accuracy of the step-by-step process. It shows the minority carrier density at 100 milli-amp. for 20, 40, 80, 160, 320 step intervals across the specimen. For the constants chosen, there is clearly no cause for concern as to whether a true solution is being obtained; other results for PCSR are taken for 160 steps, with ϵ as 0.00006. Densities were reiterated to be correct to 10 places of decimals so that the error in matching the far boundary condition is less than 1 in 10^4 .

Figure 4.2.3 shows the current-voltage characteristic for the specimen; a pronounced flattening is apparent without a negative resistance. The low field linear region is due to the negligible penetration of minority carriers, giving donor electron behaviour, while the high current steepening is due to a limiting carrier density being reached when leakage at the contacts reduces the

The first part of the paper is devoted to a general discussion of the problem of the origin of life. It is shown that the problem is not only a scientific one, but also a philosophical one. The scientific aspect of the problem is concerned with the question of how life arose from non-life. The philosophical aspect is concerned with the question of whether life is a necessary part of the universe or whether it is a mere accident.

The second part of the paper is devoted to a discussion of the various theories of the origin of life. These theories are divided into two main classes: the theory of spontaneous generation and the theory of biogenesis. The theory of spontaneous generation is the older of the two and is based on the idea that life can arise from non-life. The theory of biogenesis is the newer of the two and is based on the idea that life can only arise from pre-existing life.

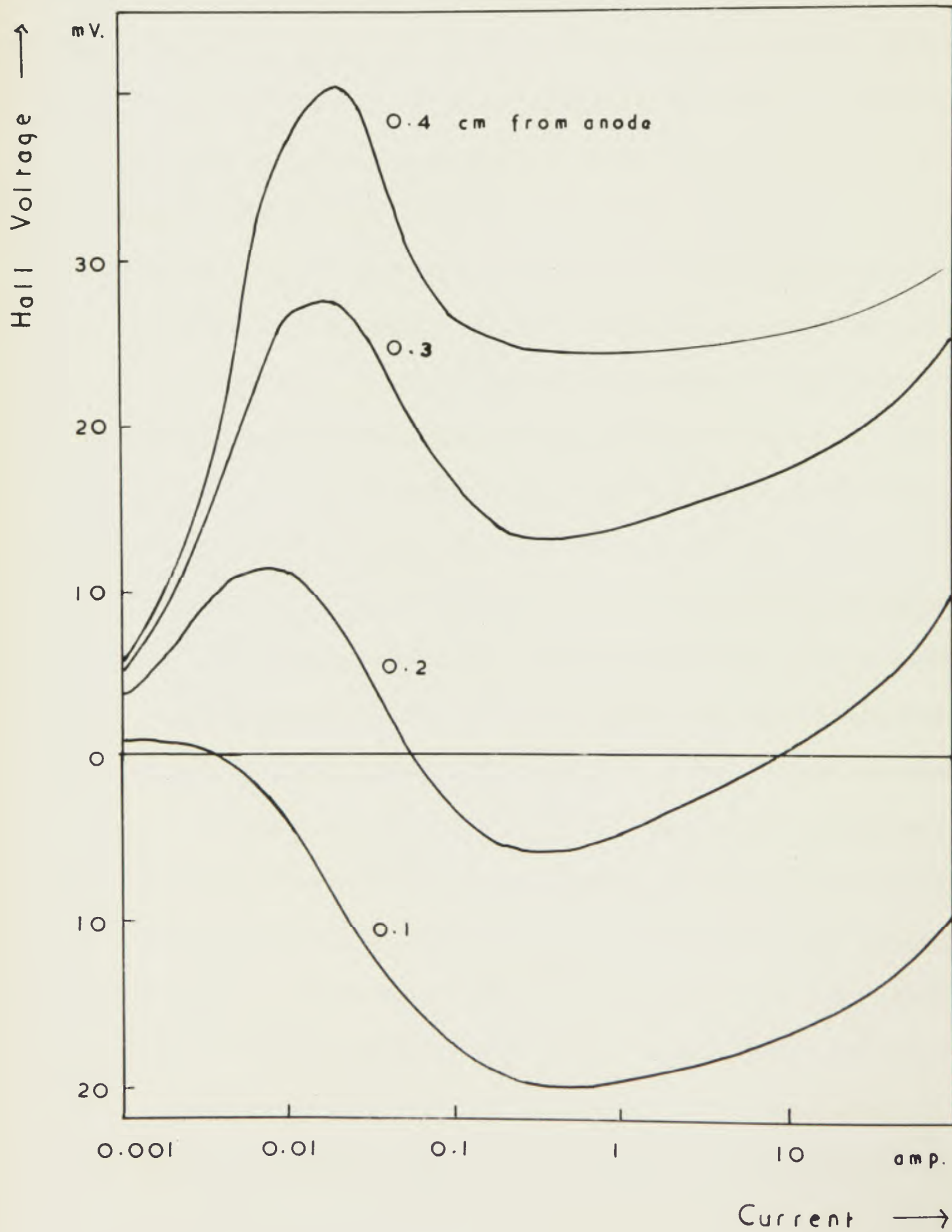
The third part of the paper is devoted to a discussion of the evidence for and against the theory of spontaneous generation. It is shown that there is a great deal of evidence in favor of the theory of spontaneous generation, but that there is also a great deal of evidence against it. The evidence in favor of the theory of spontaneous generation is based on the fact that life has been found to arise from non-life in a number of cases. The evidence against the theory of spontaneous generation is based on the fact that life has never been found to arise from non-life in a single case.

The fourth part of the paper is devoted to a discussion of the evidence for and against the theory of biogenesis. It is shown that there is a great deal of evidence in favor of the theory of biogenesis, but that there is also a great deal of evidence against it. The evidence in favor of the theory of biogenesis is based on the fact that life has never been found to arise from non-life in a single case. The evidence against the theory of biogenesis is based on the fact that life has been found to arise from non-life in a number of cases.

The fifth part of the paper is devoted to a discussion of the philosophical aspects of the problem of the origin of life. It is shown that the problem is not only a scientific one, but also a philosophical one. The philosophical aspect of the problem is concerned with the question of whether life is a necessary part of the universe or whether it is a mere accident.

Figure 4.2.4

Injection Hall Voltage for PCSR



injection ratios. It is interesting to note that for a small region of the specimen over a small current range a local negative resistance did in fact occur. Because of the recombination and boundary conditions, it seems extremely unlikely that any choice of parameters would lead to an overall negative resistance.

Figure 4.2.4 shows the Hall voltage assuming planar flow, and hence neglecting surface effects. An electron dominated Hall coefficient is noted everywhere at first, but inversion spreads from the injecting contact down the crystal as the current is raised. The high current reverse tendency is again due to the increasing leakage at the contacts. In sections 4.2 and 4.3 all graphs show electron dominated Hall voltages as positive.

4.2.6 Results for PCSTF.

The parameters involved in this case have been:

specimen natural electron density	$2 \cdot 10^{15}/\text{c.c.},$
hole trap density	$10^{16}/\text{c.c.},$
electrode carrier densities	$10^{18}/\text{c.c.},$
carrier mobilities in the specimen	$10^4 \text{ cm.}/\text{volt}/\text{sec.},$
temperature	$77^\circ \text{ K},$
S	$4.375 \cdot 10^{-13} \quad ,$
T	$1.875 \cdot 10^{-14} \quad ,$
specimen dimensions	$0.5 \cdot 0.1 \cdot 0.1 \text{ cm.},$
magnetic field	$200 \text{ gauss},$

Figure 4.2.5

Variation of Injected Carrier Density

across Crystal for PCSTF

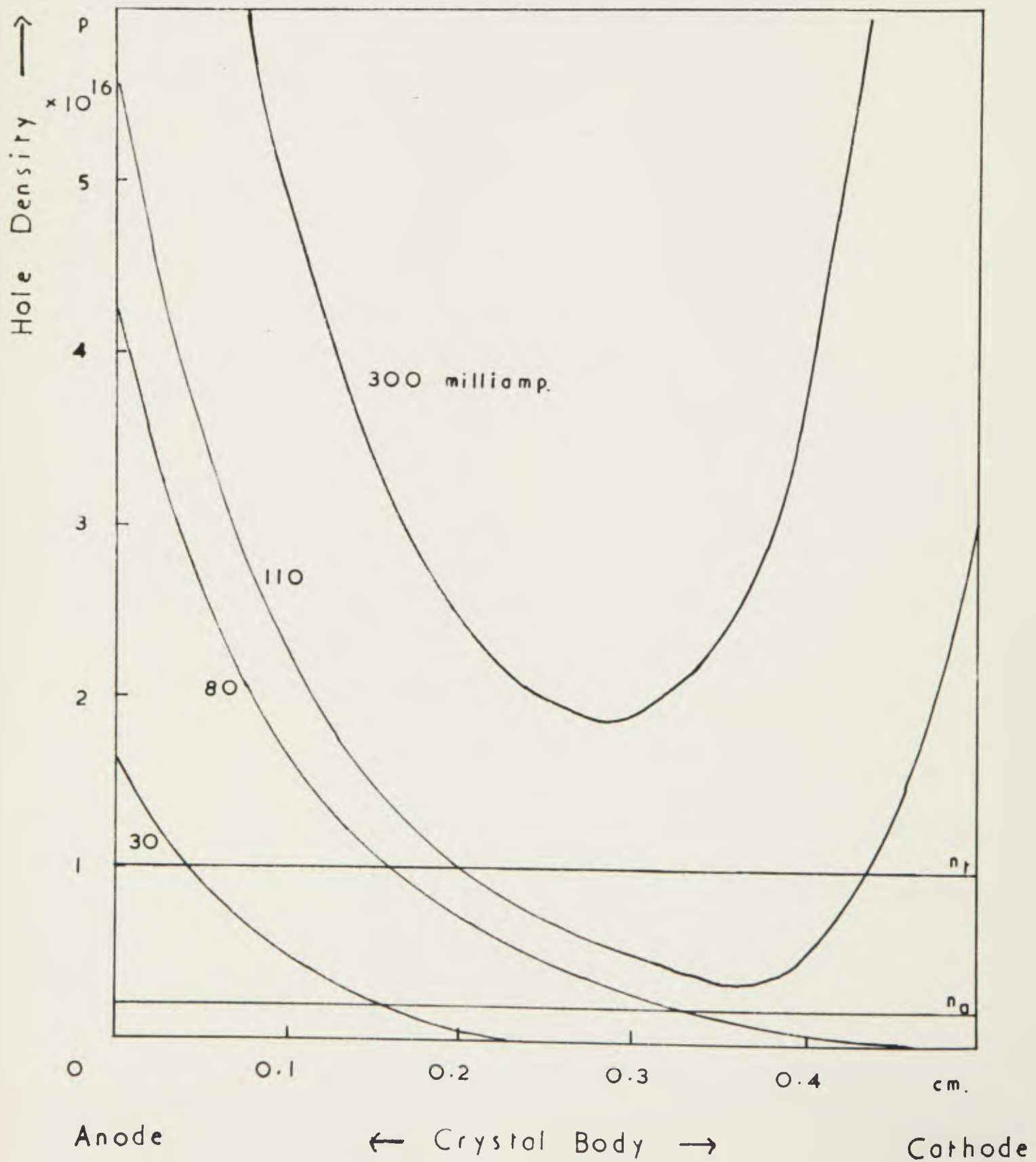
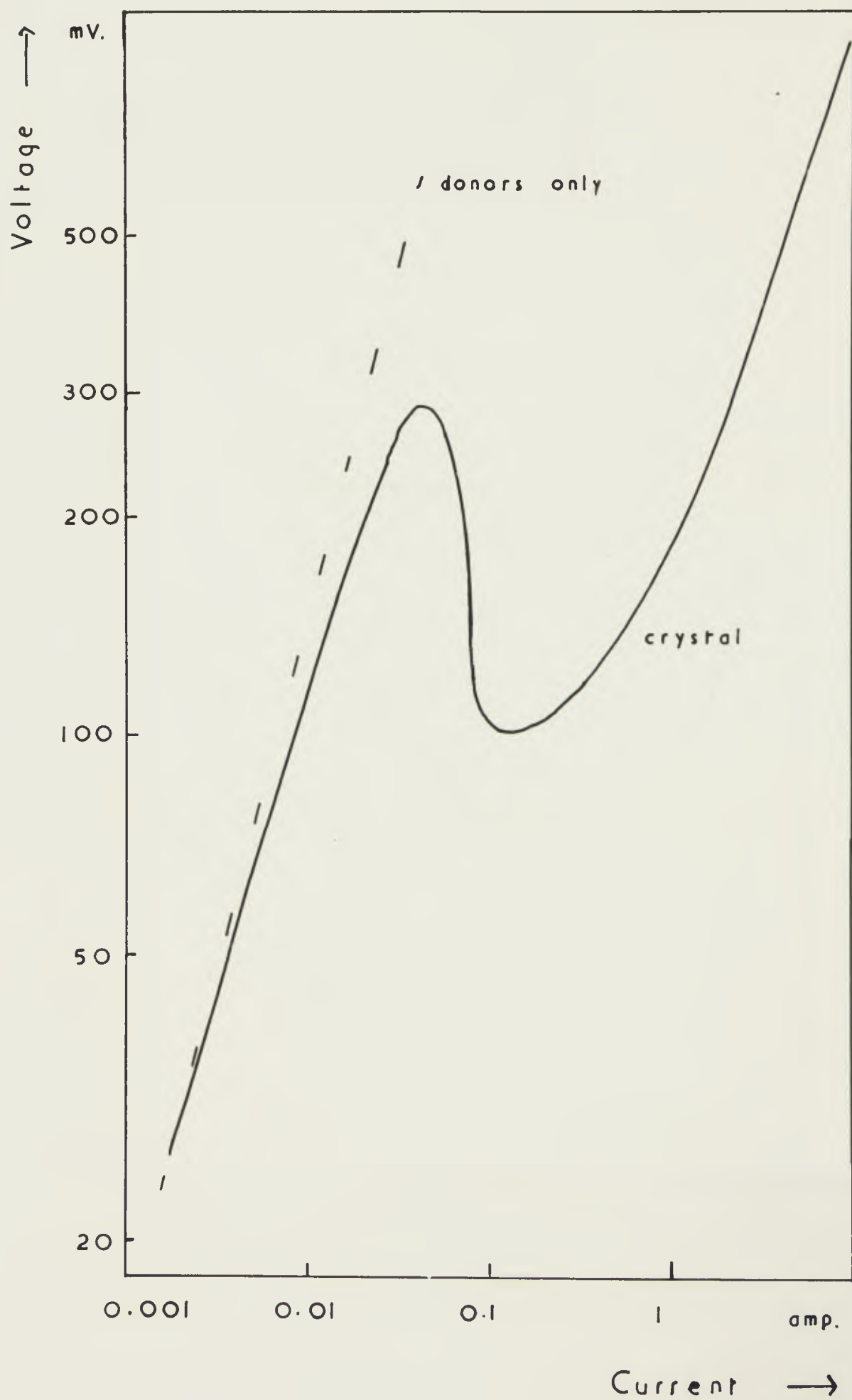




Figure 4.2.6

Current - Voltage Characteristic for PCSTF



1881. 1882. 1883. 1884. 1885. 1886. 1887. 1888. 1889. 1890.

1891. 1892. 1893. 1894. 1895. 1896. 1897. 1898. 1899. 1900.

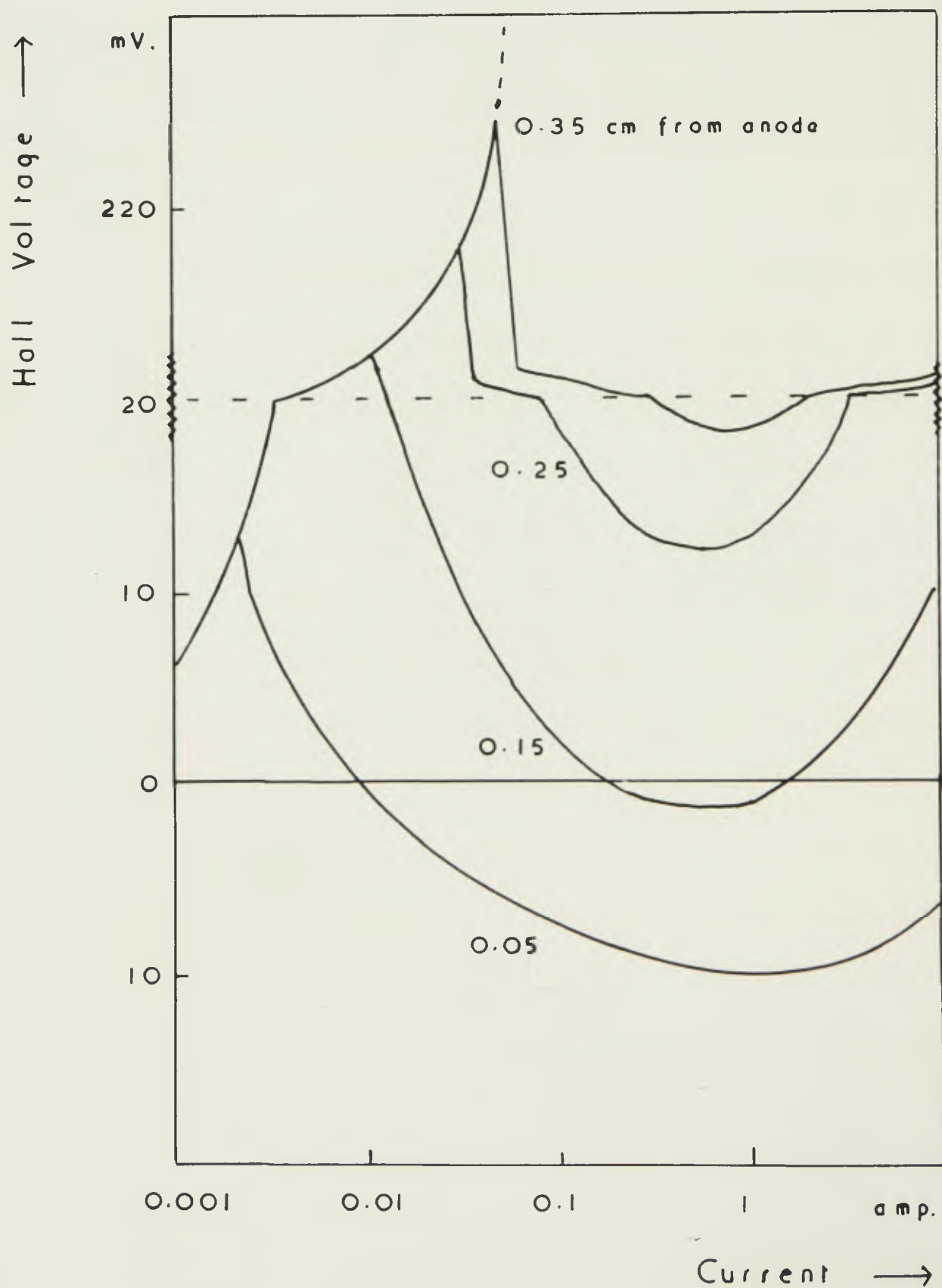
1901. 1902. 1903. 1904. 1905. 1906. 1907. 1908. 1909. 1910.
1911. 1912. 1913. 1914. 1915. 1916. 1917. 1918. 1919. 1920.
1921. 1922. 1923. 1924. 1925. 1926. 1927. 1928. 1929. 1930.
1931. 1932. 1933. 1934. 1935. 1936. 1937. 1938. 1939. 1940.

1941. 1942. 1943. 1944. 1945. 1946. 1947. 1948. 1949. 1950.
1951. 1952. 1953. 1954. 1955. 1956. 1957. 1958. 1959. 1960.
1961. 1962. 1963. 1964. 1965. 1966. 1967. 1968. 1969. 1970.
1971. 1972. 1973. 1974. 1975. 1976. 1977. 1978. 1979. 1980.
1981. 1982. 1983. 1984. 1985. 1986. 1987. 1988. 1989. 1990.
1991. 1992. 1993. 1994. 1995. 1996. 1997. 1998. 1999. 2000.
2001. 2002. 2003. 2004. 2005. 2006. 2007. 2008. 2009. 2010.

2011. 2012. 2013. 2014. 2015. 2016. 2017. 2018. 2019. 2020.
2021. 2022. 2023. 2024. 2025. 2026. 2027. 2028. 2029. 2030.
2031. 2032. 2033. 2034. 2035. 2036. 2037. 2038. 2039. 2040.
2041. 2042. 2043. 2044. 2045. 2046. 2047. 2048. 2049. 2050.

Figure 4.2.7

Injection Hall Voltage for PCSTF



number of steps	80 ,
value of ϕ	0.008 .

Figure 4.2.5 shows the minority carrier density plotted across the specimen at four currents; at 30 , 80 milliamp. the hole saturation front is in the body of the specimen while by 110 milliamp. it has reached the far end. Again in the two higher current cases, the carrier density rises at the far boundary.

Figure 4.2.6 shows the current-voltage characteristic. As in PCSR a linear low current region is found together with a steepening high current region. In between, a negative resistance is to be noted. The computation follows this through the minimum without any difficulty.

Figure 4.2.7 shows the Hall voltage; the change in scale at 20 mV. should be noted. The curve is basically similar to that for PCSR; its exaggerated shape is due to the much sharper change in the number of carriers in the region where the trap front is moving along the crystal.

The programme thus acts as a "crystal simulator" for trap saturation front penetration, with the limiting carrier density controlled either by junction leakage or by recombination. For the demonstration, junction leakage has been used.

4.2.7 Summary of Computation.

No attempt will be made to compare characteristics for different parameters. Nor is it claimed that the parameters employed have any special significance in terms of the experimental results which follow. The aim has been to show what can be done with this kind of approach, employing the new carrier flow equation, and what sort of experimental results may be anticipated. Following the path of the characteristic through a negative resistance is felt to be a triumph which justifies this section in view of the discussion of section 4.1.6 on page 83.

4.3 Experimental Results.

4.3.1 Introduction.

The experimental results form the primary feature of the work on double injection; thus the success of the fore-going theoretical investigation is measured in terms of how well the results are explained. This point will be expanded upon in the discussion of the correlations after all results relevant to this chapter have been presented.

Experimental methods and specimen nomenclature have been described in chapter 3; only specifically relevant points are added.

4.3.2 Pattern of Results.

Experimental information has been obtained for 31 injecting diodes and for 24 ohmic contact pairs, as classified in chapter 3. Our concern here is to demonstrate the effects of injection. For this purpose, ten types of measurement have been made, for both injective and ohmic flow. The measurements and the respective aims are:

- (1) Measurement: the current-voltage characteristic of a short specimen.

Aim: to show the nature of the anomalies for which an explanation is attempted;

- (2) the i - V characteristics of a long specimen;

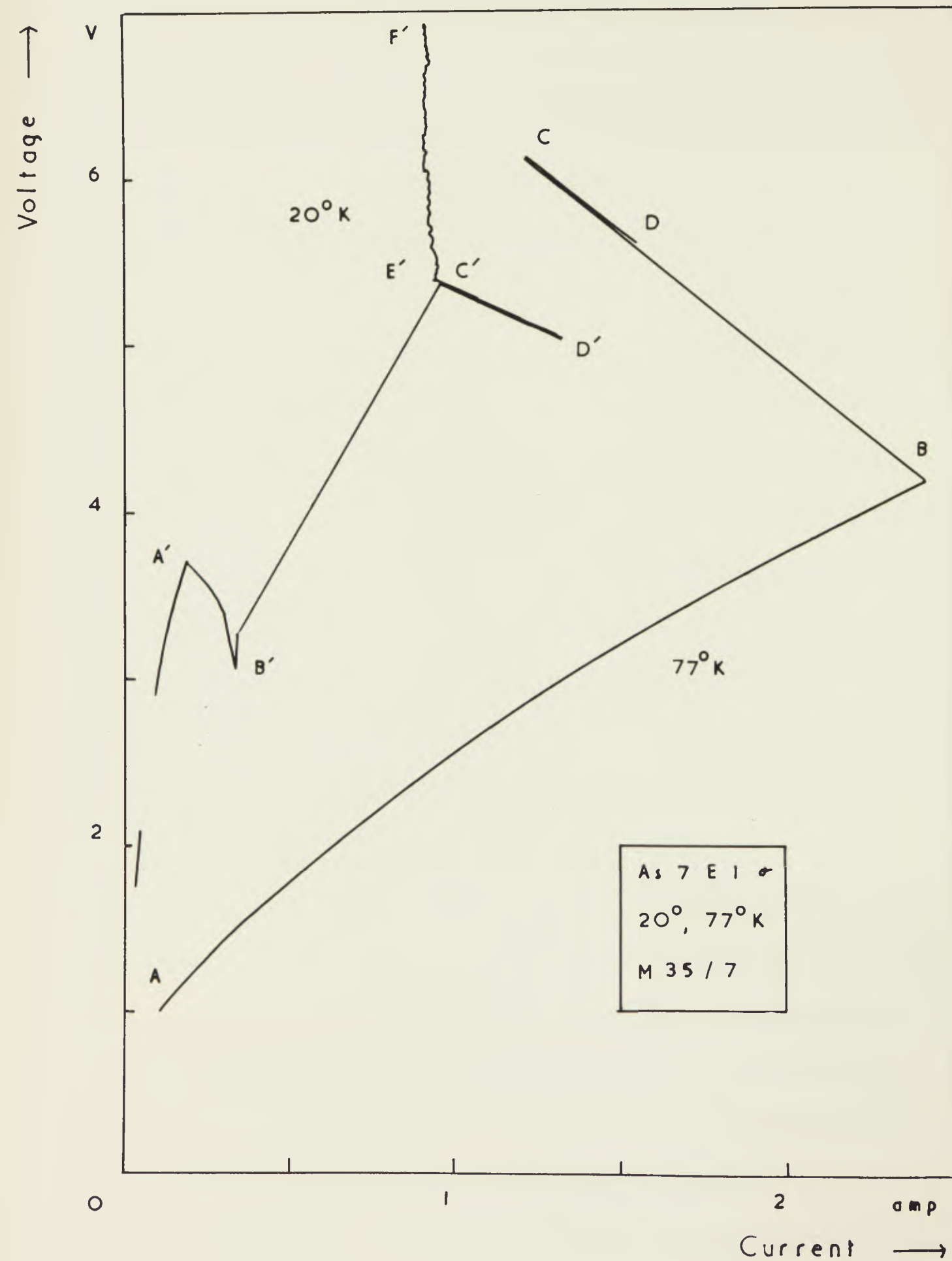
to show the specific effects of injection compared with ohmic

flow;

- (3) the change in characteristic with specimen length;
to show that a non-linearity turns into a negative resistance
as the length is reduced;
- (4) the voltage distribution along the crystal;
to show the spread of higher conductivity from the injecting
contact;
- (5) the local potential;
to demonstrate the presence and movement of a hole saturation
front;
- (6) the Hall voltage as a function of position;
to demonstrate two kinds of Hall coefficient inversion, one
associated with the saturation front, the second with
electron stagnation;
- (7) the Hall voltage as a function of compensation;
to demonstrate the dependence of inversion on the impurity
level;
- (8) the Hall voltage as a function of temperature;
to show the temperature dependence of the carrier flow parameters;
- (9) the Hall voltage as a function of magnetic field;
to show the effect on the inversions of increased surface
recombination (Suhl effect);
- (10) the effect of a magnetic field on the voltage distribution;
to confirm the results of measurement (9).



Rectifying Characteristics at 20°K, 77°K



4.3.3.1 The shape of the current-voltage characteristic.

The nature of the i - V characteristic for injection was found to vary with specimen length, as will be described in section 4.3.3.3. Long thin specimens of higher resistance, which may be fitted with arrays of contacts are more suitable for analysis than short specimens. Short specimens however give a clearer qualitative characteristic.

Figure 4.3.1 shows the current-voltage characteristic of a short specimen, diode As7E1, both at 77°K and at 20°K. All five negative resistances are visible; they are labelled as follows:

negative resistance no. 1	curve portion	A'B'
no. 2		C'D'
no. 3		D'E' or EC
no. 4		E'F'
no. 5		CD

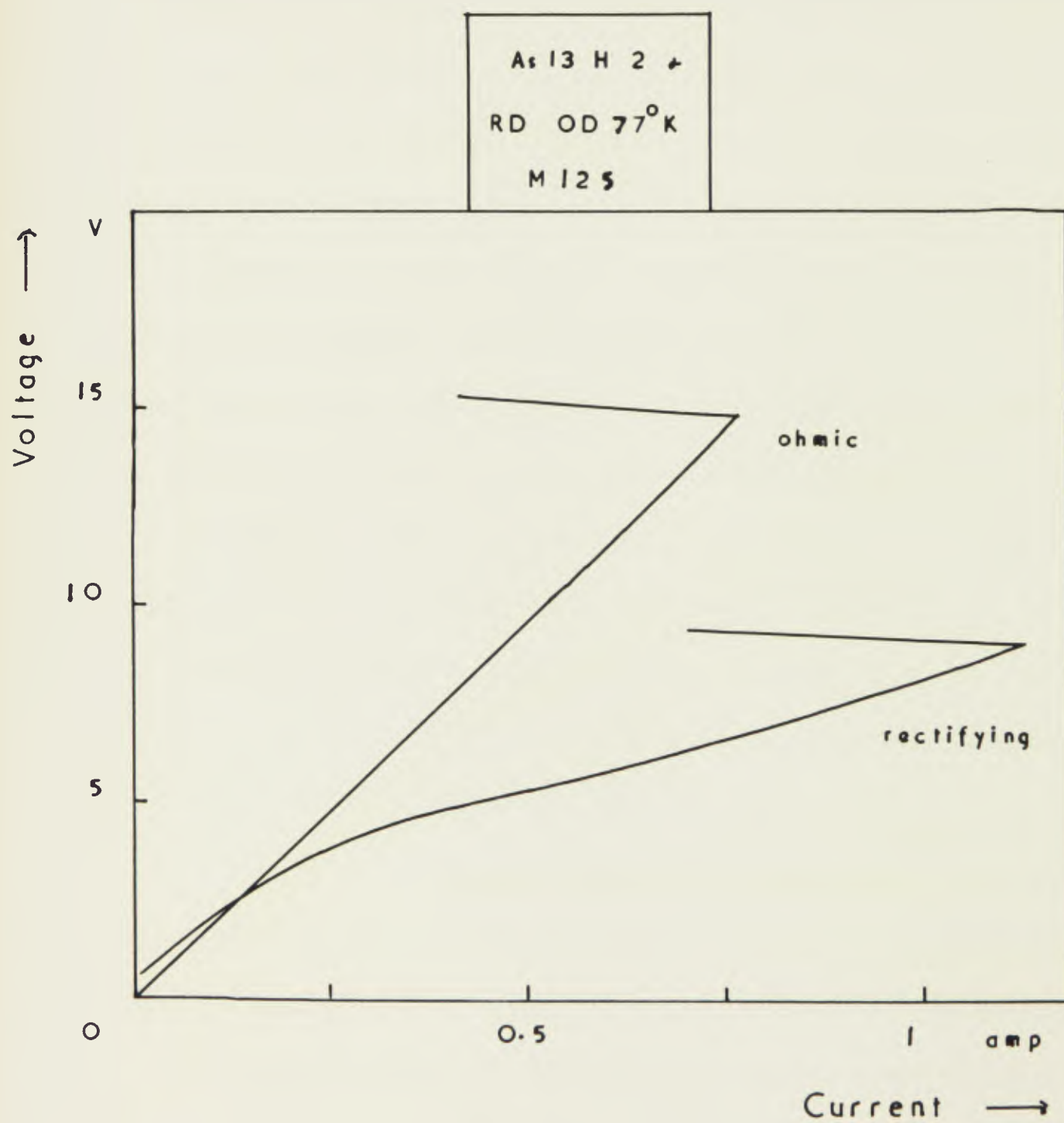
This figure must constitute one of the most anomalous current-voltage characteristic pictures ever seen, especially since the maximum applied field is 20 volts/cm., far below the hot electron threshold described by Paige (l.c. 1964, 453). In this thesis, the aim is to investigate and explain each negative resistance; for this chapter we confine our interest to the regions OAB, OA'E'C', and attribute the non-linearities within this range to the effects of injection, for reasons discussed in the next section.

The range for this injection study is terminated at B, C' by

[The text in this section is extremely faint and illegible. It appears to be a series of paragraphs or a list, but the specific content cannot be discerned.]

Figure 4.3.2

Rectifying and Ohmic Characteristics



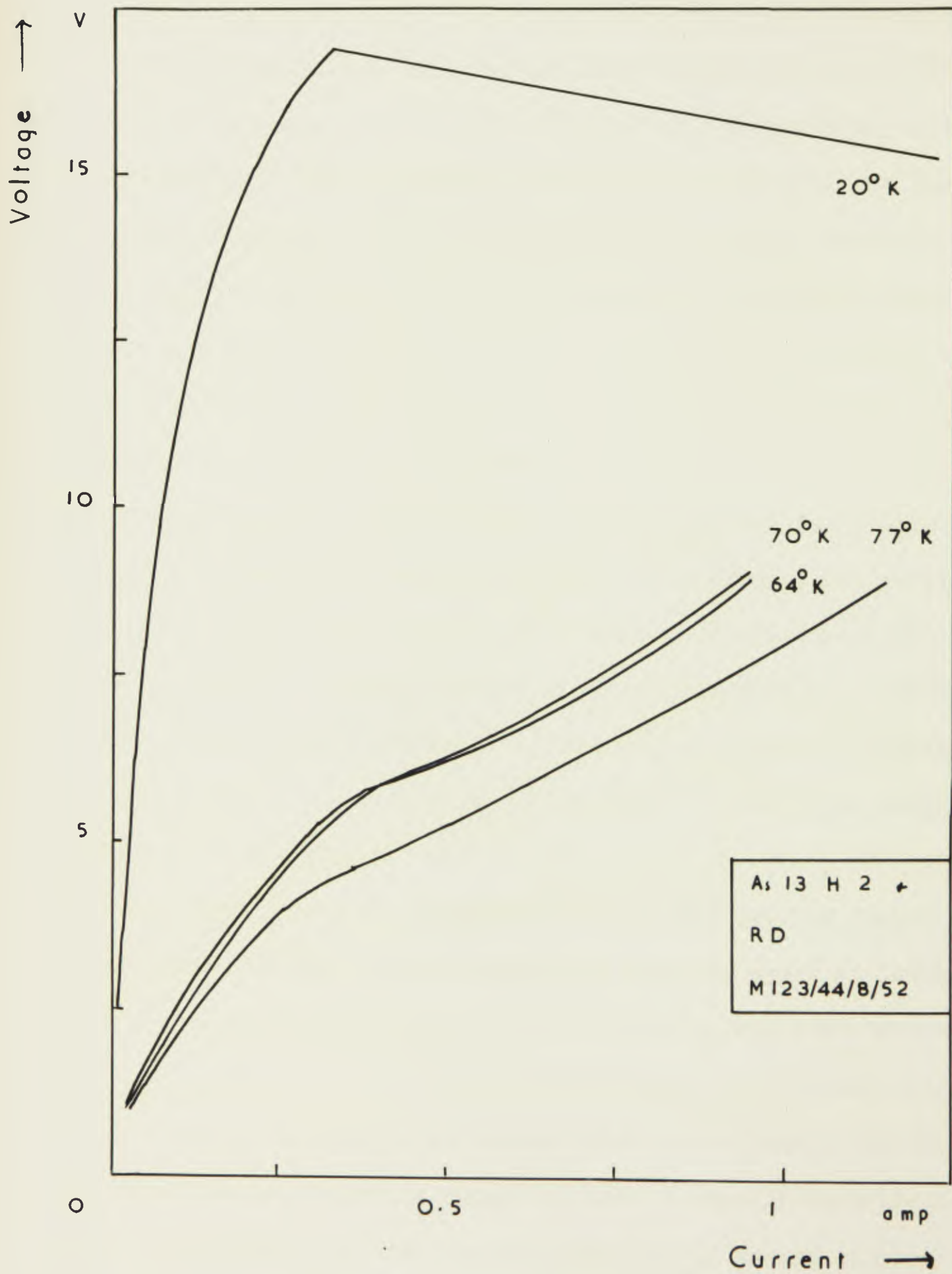
a negative resistance which is not dependent on the injection property, a similar phenomenon being found for ohmic flow. It was noted however that the transition point may be raised by using pulsed currents. Such pulsed conditions make it possible to examine growth and decay transients; these were found to be complex, apparently incorporating two or three diffusion waves of excess carrier density. No attempt will be made to explore these phenomena here. This is not because interpretation presents any special problems but because it is felt that it merits a project of its own which would be a continuation of the present work rather than an integral part of it. Certainly such an analysis would rely on the ideas put forward here more than it would confirm them. There was no evidence that the post-transient values of the pulse heights were dependent on the duty cycle before the terminating negative resistance was reached. Hence d.c. readings are employed because of the advantages in accuracy and information grasp noted in section 3.12.

4.3.3.2 The characteristic for long specimens.

Reasons for considering the non-linearities OAB, OA'B'C' as due to injection will now be advanced. For analytical purposes we turn to the consideration of results taken with the long thin higher resistance specimens 11, 12, 13. Figure 4.3.2 shows the total characteristic for specimen As₁3H₂+, both for injection and ohmic flow at 77° K. Figure 4.3.3 shows the injection characteristic at

Figure 4.3.3

Rectifying Characteristics at Four Temperatures



four temperatures.

We note:

(1) The ohmic contact characteristic is linear, whereas the injection characteristic is not. This is true in general for all specimens above hydrogen temperatures; deviation from linearity for one or two ohmic specimens was attributed to poor contact production. At hydrogen temperature and below, impact ionization is possible --- this is discussed in chapter 5.

(2) Figure 4.3.2 shows that relatively little current flows for an applied voltage of less than 0.7 volts. This feature was common to all diodes but was absent in the voltage curve over portions of the crystal not including the junction as shown by figures 4.3.6, 4.3.7. This confirms the analysis of section 4.1.8 where it was deduced that a potential drop equal to the band gap is required at the junction to supply an appreciable injection current but that at this figure any number of carriers is available. This corresponds well with the germanium band gap of 0.7 eV. It means that the exponential region is soon left behind because the potential drop across the crystal is large compared with that at the junction. Thus the premises on which the computation for the injection ratio is based are fair, even if the formula derived is not.

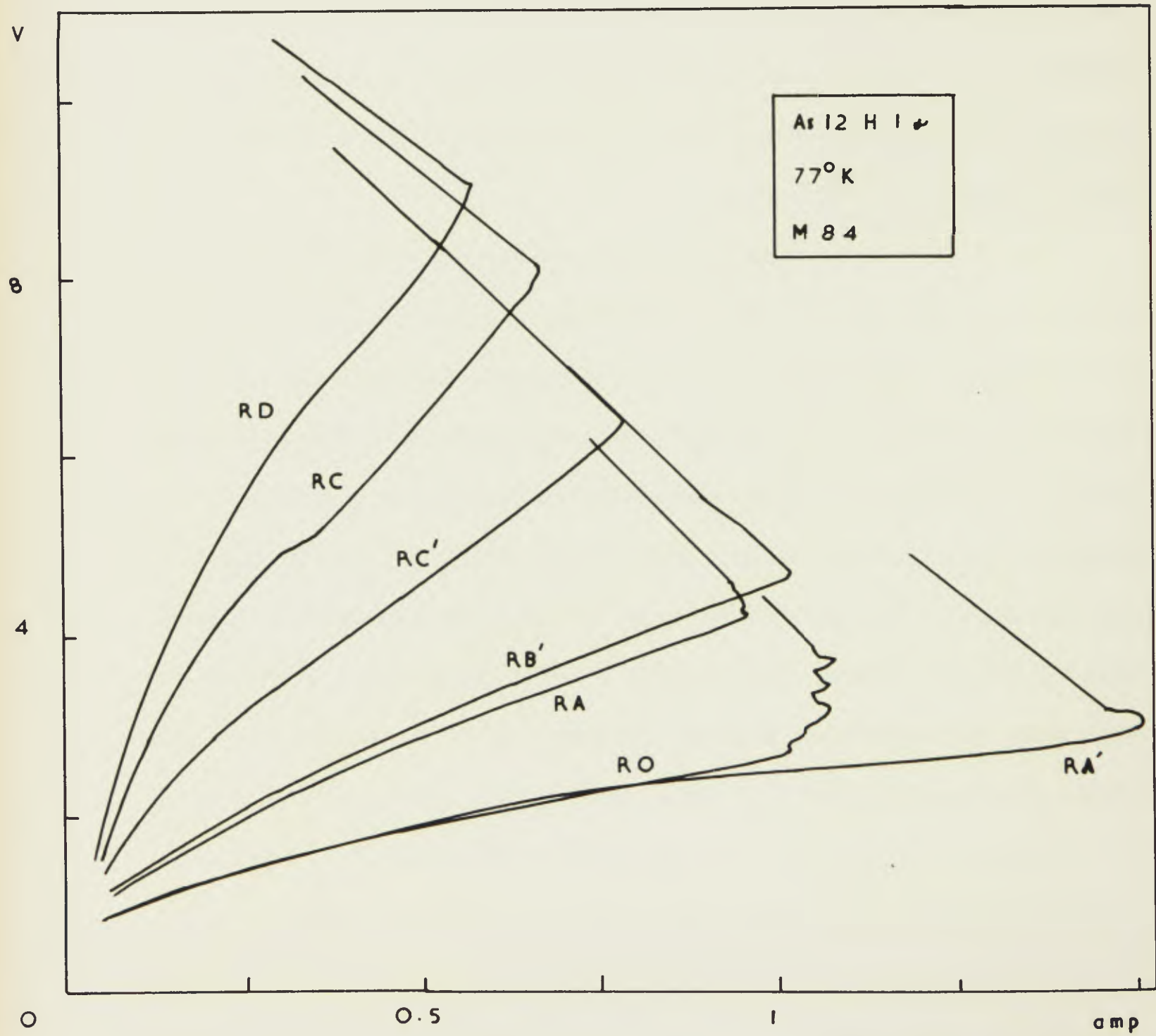
(3) The injection and ohmic curves have the same slope for low applied voltages (figure 4.3.2); at higher currents, the characteristic is again linear with a lower slope. This is attributed to the

[The following text is extremely faint and illegible due to the quality of the scan. It appears to be a list or a series of entries, possibly related to a botanical or geographical study.]

Figure 4.3.4

Characteristic Variation with Contact Separation

Voltage



Current →

change-over from single to double carrier flow. At low currents recombination takes place close to the injecting anode, giving conditions similar to ohmic flow; as the holes penetrate further, increased carrier densities provide a lower resistance. Once holes start to reach the far end, a new regime takes over with a slower resistance variation. Since the line does not pass through the origin, even the new linear range represents falling resistance due to a swelling carrier density.

(4) A double injection negative resistance characteristic is not observed in the 77° K curve of figure 4.3.3, but a discontinuity is observable at 0.5 amp., which is even more apparent at 64° K. This region is interpreted as being where some special distortion of the carrier density is being pushed down the crystal as the current is raised. Following the theory section, we suggest that the distortion represents the front behind which a set of slow hole recombination centres is saturated. Trapping and recombination are considered to be similar in principle to those for the negative resistance computation of section 4.2.6.

4.3.3.3 The change in characteristic with specimen length.

The i - V characteristic variation as a function of diode length is shown in figure 4.3.4 at 77° K and in figure 4.3.5 at 20° K. Specimen length has a clear effect on the negative resistance observed at hydrogen temperature. As the length decreases, so the non-linearity



Figure 4.3.5

Characteristic Variation with Contact Separation

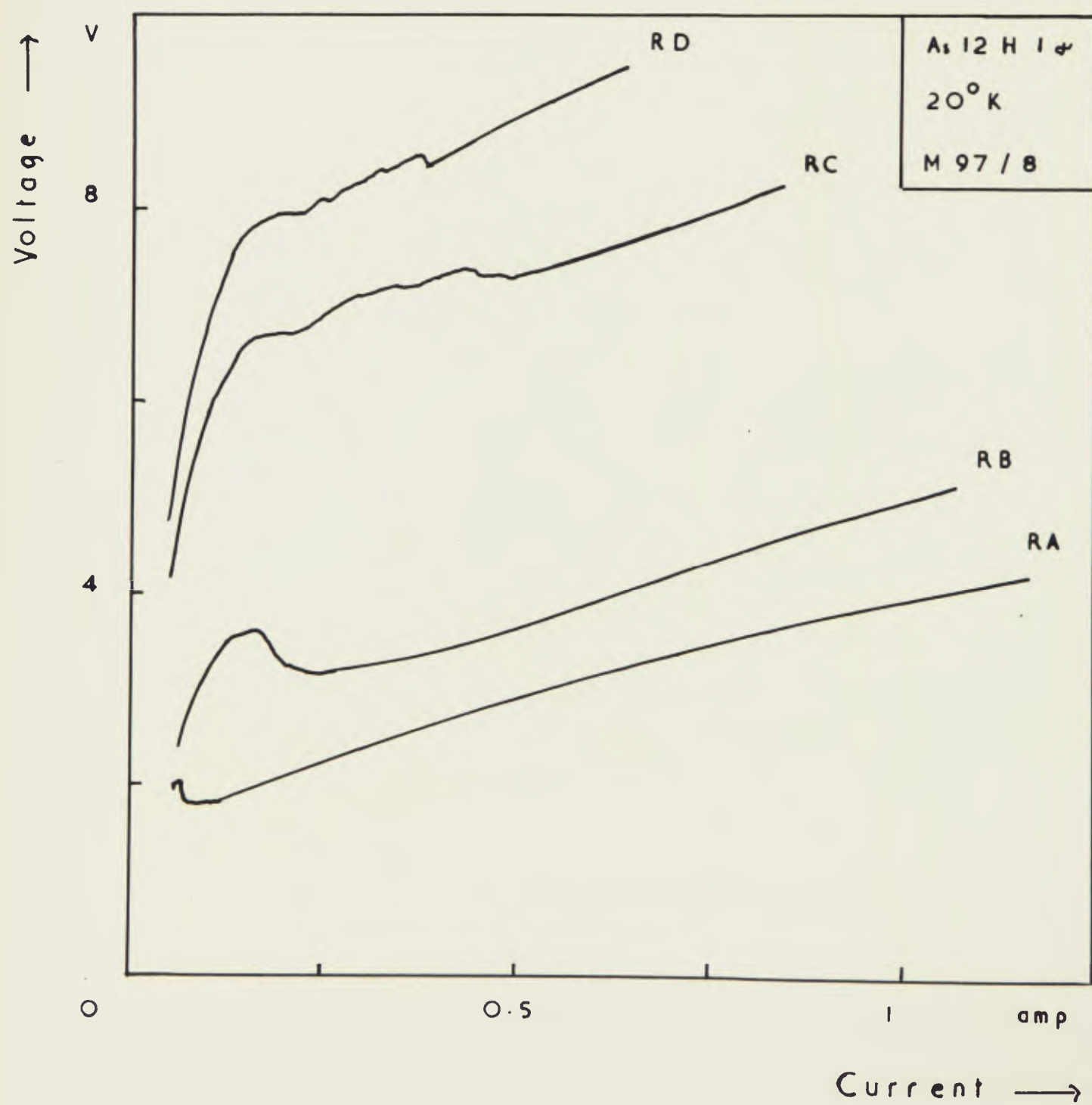
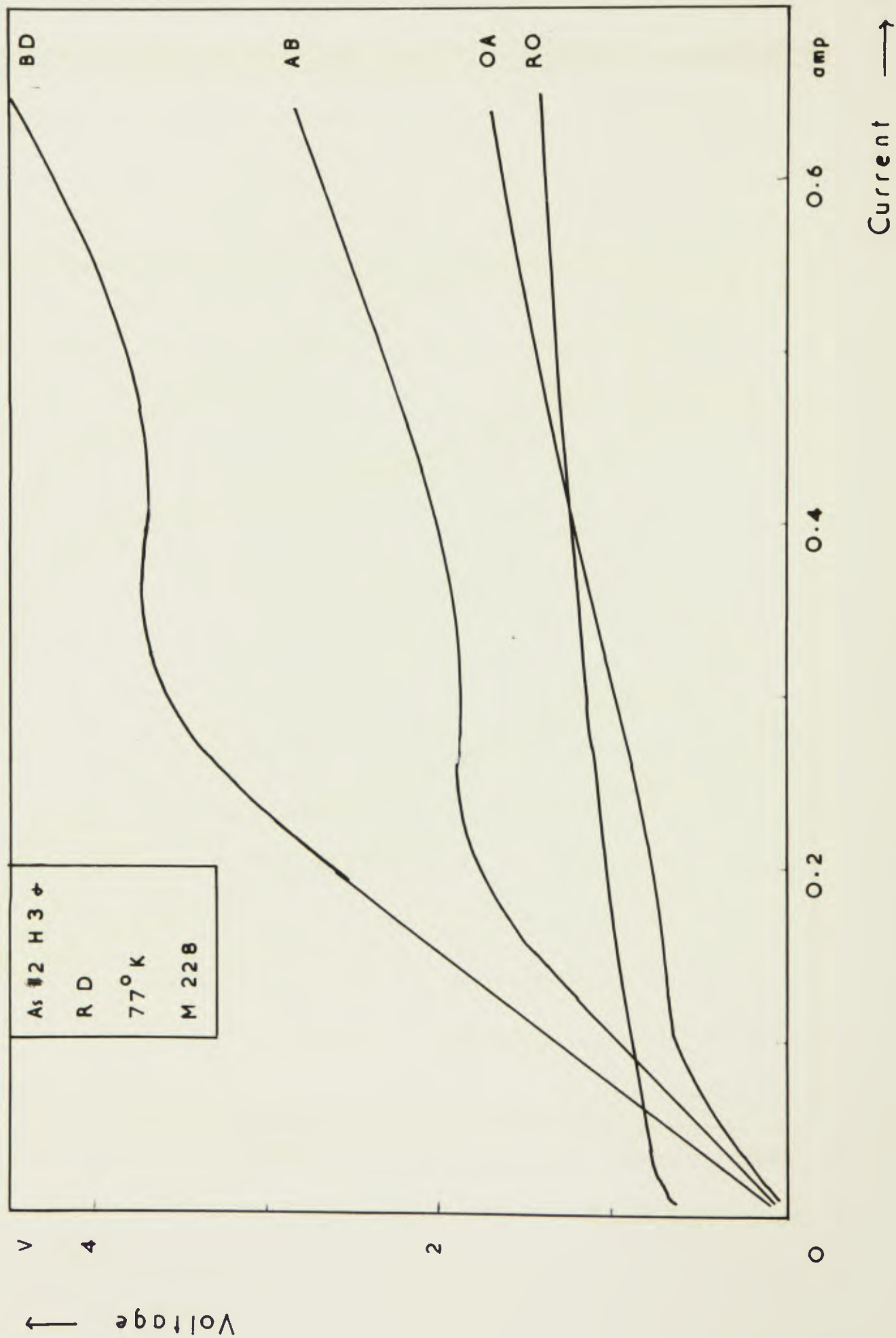


Figure 4.3.6 Sectional Voltages



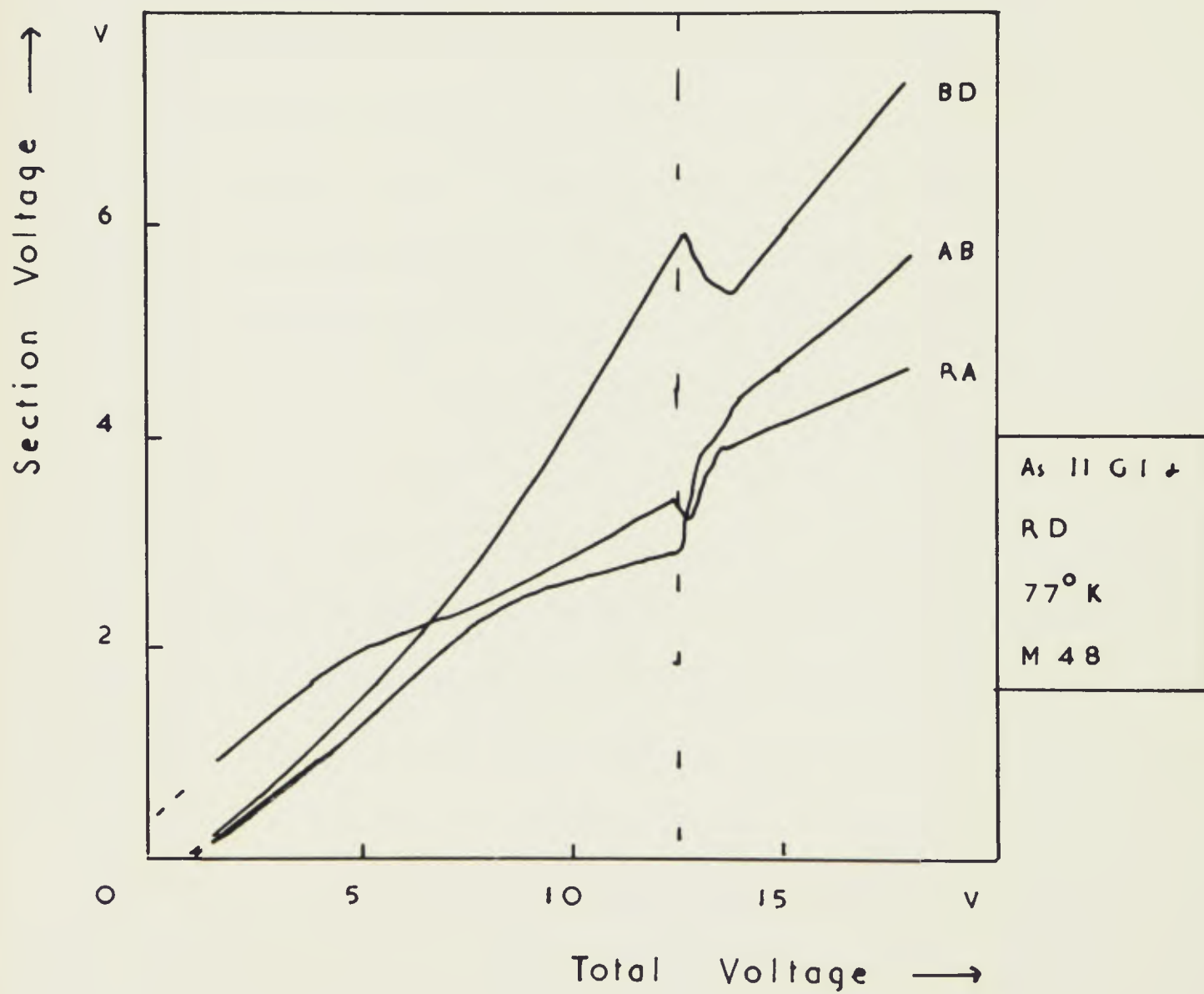
tends towards a negative resistance. The negative resistance was yet sharper for the shorter specimen shown by figure 4.3.1. It is deduced that the negative resistance region is due to the motion of a narrow zone, while the rest of the crystal provides a ballast of positive resistance. Overall negative resistance is observed for fairly short specimens, and for still shorter ones a spontaneous transition, as in 4.3.1 when the negative resistance is steeper than the load line. The negative resistance of 4.3.1 shows little hysteresis, and may be traversed in either direction. This indicates that filament formation is not significant, following from the consideration of filament behaviour in chapter 2. We associate the negative resistance zone with the region of the saturation front.

4.3.3.4 Section voltages.

The potential drop between adjacent contacts down the length of the specimen was measured to show the voltage distribution as a function of the current. This is shown for 77°K in figure 4.3.6. It demonstrates that for low currents all portions of the crystal are ohmic except R0 which incorporates the rectifying junction. Above a junction voltage of 0.7 eV., section R0 shows a fairly linear characteristic; below 0.7 eV. the exponential dependence of current on the junction voltage is apparent. This supports the conclusions drawn from the total characteristic in section 4.3.3.2. The figure shows that the non-linearity delays its appearance to higher currents as the distance

Figure 4.3.7

Sectional Voltages



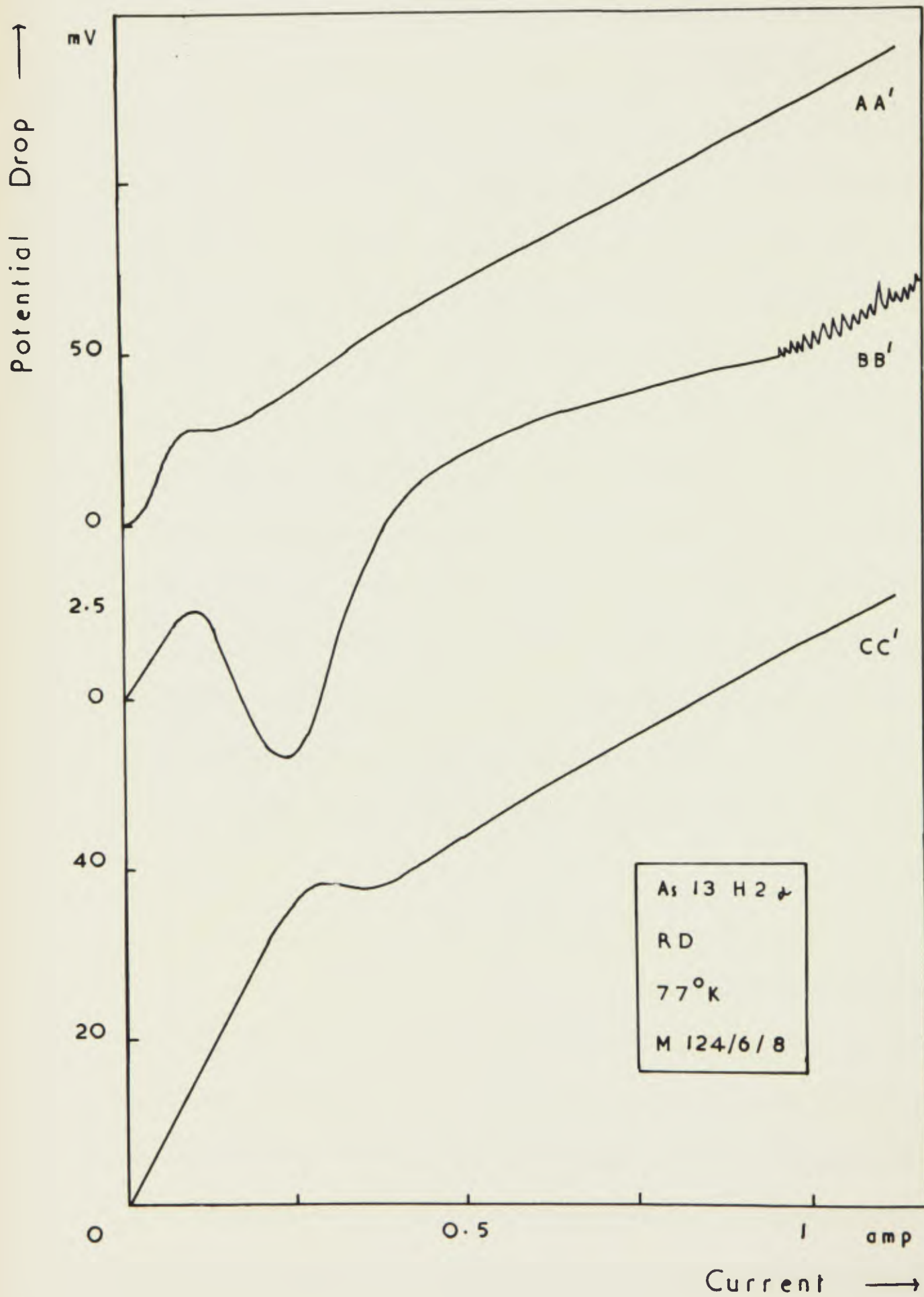
from the injecting contact increases. This is a logical consequence of the increased penetration of holes at higher currents; double injection flow predominates over longer crystal ranges, giving a higher local conductivity. At 0.6 amp. the double injection range has spread through the whole length of the crystal to the cathode. In the region of the change-over from single to double carrier flow, a negative resistance is almost visible for some sections; comparing the sections with the characteristic for the whole crystal which is their voltage sum, the relationship between specimen length and negative resistance is clear. For contact pairs still closer together a more marked effect may be anticipated, which the next section confirms. Figure 4.3.7 shows the sectional voltages for a different specimen plotted against the total applied voltage; only the portion up to 12.5 volts applied is relevant to this chapter. Again the concentration of voltage at the junction for low voltages and its transfer to other regions as the voltage is raised, are both demonstrated.

4.3.3.5 The local potential.

The long specimens are fitted with paired contacts for the examination of the Hall voltage --- with one set (specimen As11G1+), two sets (As12H3+ and 1UC1H2+) or three sets (As12H1+ and As13H2+). Since it is not possible to place these exactly on equipotentials, a small out of balance potential difference exists across the pair

Figure 4.3.8

The Local Potential



in the absence of a magnetic field, due to the longitudinal electric field. This potential difference may be expected to be some measure of the local resistivity, and is termed the local potential.

Both ohmic and injective flows were studied for the same contact pairs by switch control. Results are considered for $\text{As}_{1.5}\text{Hf}_{2.5}$. Electron flow in the crystal gave a linear response to the potential at each contact set. This suggests that there are no basic faults in the contacts themselves. The injection curves at 77°K for the contact pairs are shown in figure 4.3.8. The following features are notable:

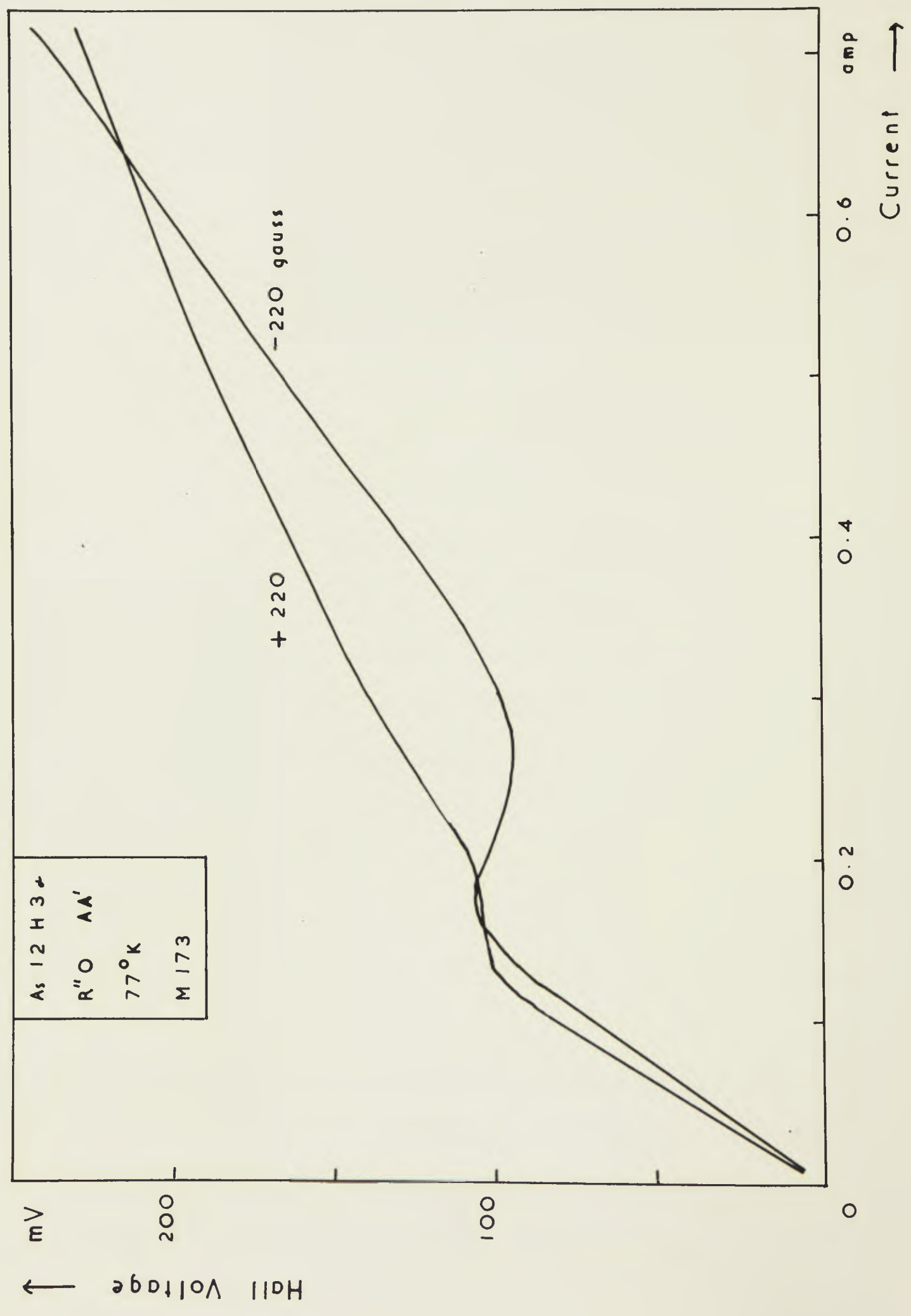
(1) The shape is similar for each crystal locality A, B, C; the anomaly at B is more marked because the Hall contacts are more nearly opposite there.

(2) The anomaly appears at a higher current for a larger distance from the injecting contact. It is thus fair to state that some form of carrier density distortion is being driven down the crystal as the current is raised. This confirms the deductions from the i - V curves.

(3) The total curve of figure 4.3.2 for the characteristic is a combined version of the three at A, B, C; this shows that a negative resistance zone may exist in the material without it being explicitly shown in the whole characteristic, confirming the deduction made in section 4.5.3.5 on the length variation.

(4) A post-anomaly linear region is visible, corresponding to an increased carrier density. The conductivity is seen to be increasing

Figure 4.3.9 Recording of Hall Voltage



towards a value of three times the single carrier one. For constant and equal electron and hole mobilities, this indicates a free carrier population of three times the equilibrium electron density.

(5) For 8B' a small reverse potential, ~ 2 mV., is given at 0.5 sec. This is very difficult to explain phenomenologically. It was mainly to explain this point that the double-ended injection specimen As12H3+ was constructed. Results from this suggested firstly that a local dip in one direction was accompanied by a corresponding peak in the other and secondly that the more nearly opposite the contact pairs are, the less consistent is the behaviour of the local potential for injection conditions. These points are interpreted as evidence of changes in the equipotential surfaces in the crystal as the current is raised; distortion of the equipotentials in the region of the saturation front seems probable, while the double carrier equipotentials are not necessarily the same as the single carrier ones which preceded them, due to the possibilities of differential surface recombination and funnelled flow. No irregularities at all were found for ohmic flow.

We deduce just that the local potentials indicate the movement of a hole saturation front down the crystal.

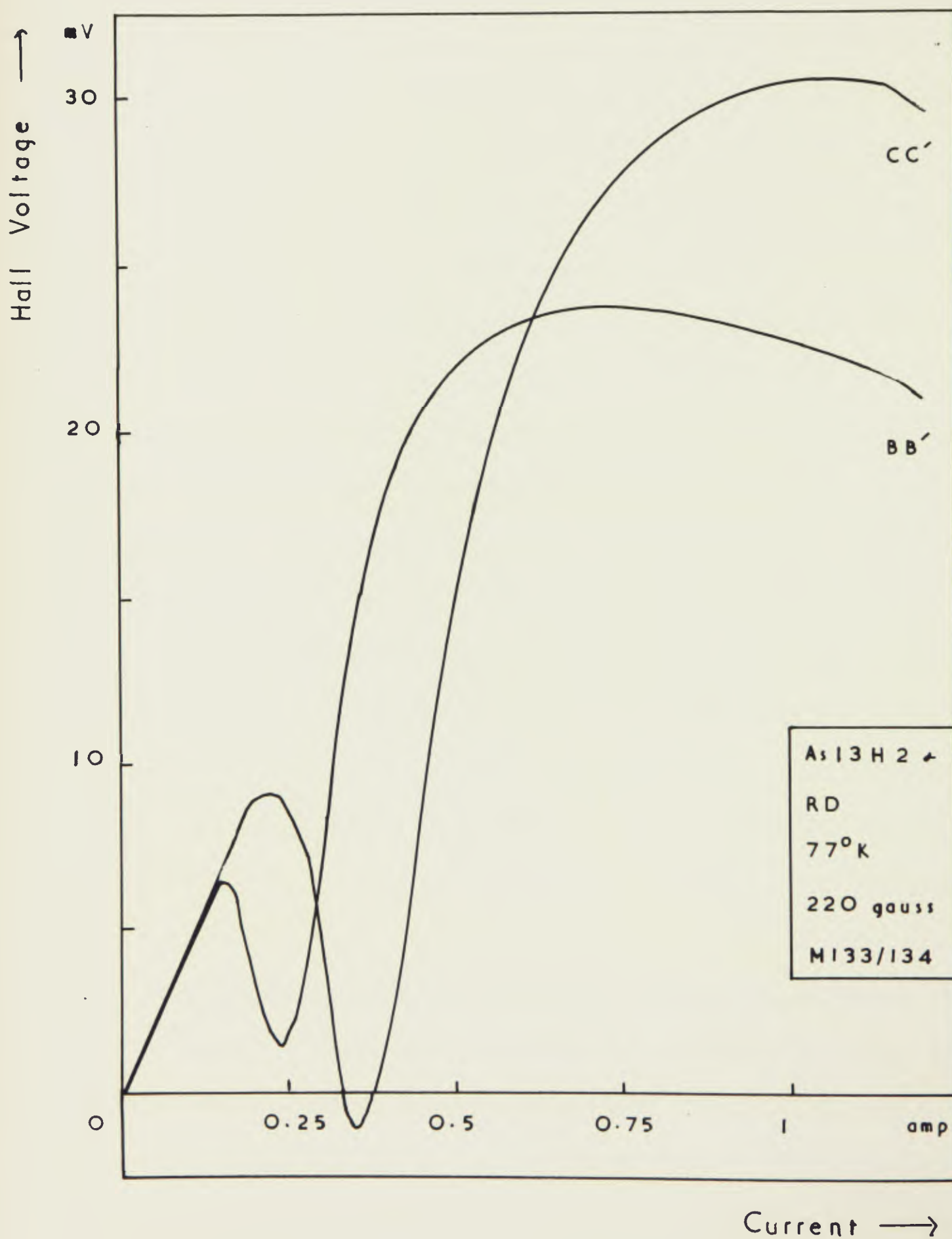
4.3.3.6 The Hall voltage as a function of current and position.

The most impressive and important results in this chapter are those on the Hall effect. This is because their complex structure demands an explanation which is based on an understanding of the



Figure 4.3.10

Hall Voltages at Two Points



fundamental flow processes in the crystal. A reproduction of an example of original data is shown in figure 4.3.9. The two curves represent the potential differences between the Hall contacts, measured in the two opposing magnetic field directions. All subsequent curves show half the difference as the Hall voltage.

It is immediately apparent from figure 4.3.9 that the Hall coefficient is far from linear, under injection conditions; for ohmic flow past the same contacts a linear graph was obtained. Hence we deduce that the distortion of the shape is due to the presence of holes.

Figure 4.3.10 shows how the distortion in the Hall voltage curve varies with position along the crystal for diode $As_{13}M_{2+}$. BB' is a contact pair closer to the injecting contact than CC' . Figure 4.3.11 shows the Hall voltage curve for diode $As_{12}M_{3+}$; here there are two contact pairs AA' , BB' and injection may be effected at either end, so that a saturation front may be driven down either way.

We deduce:

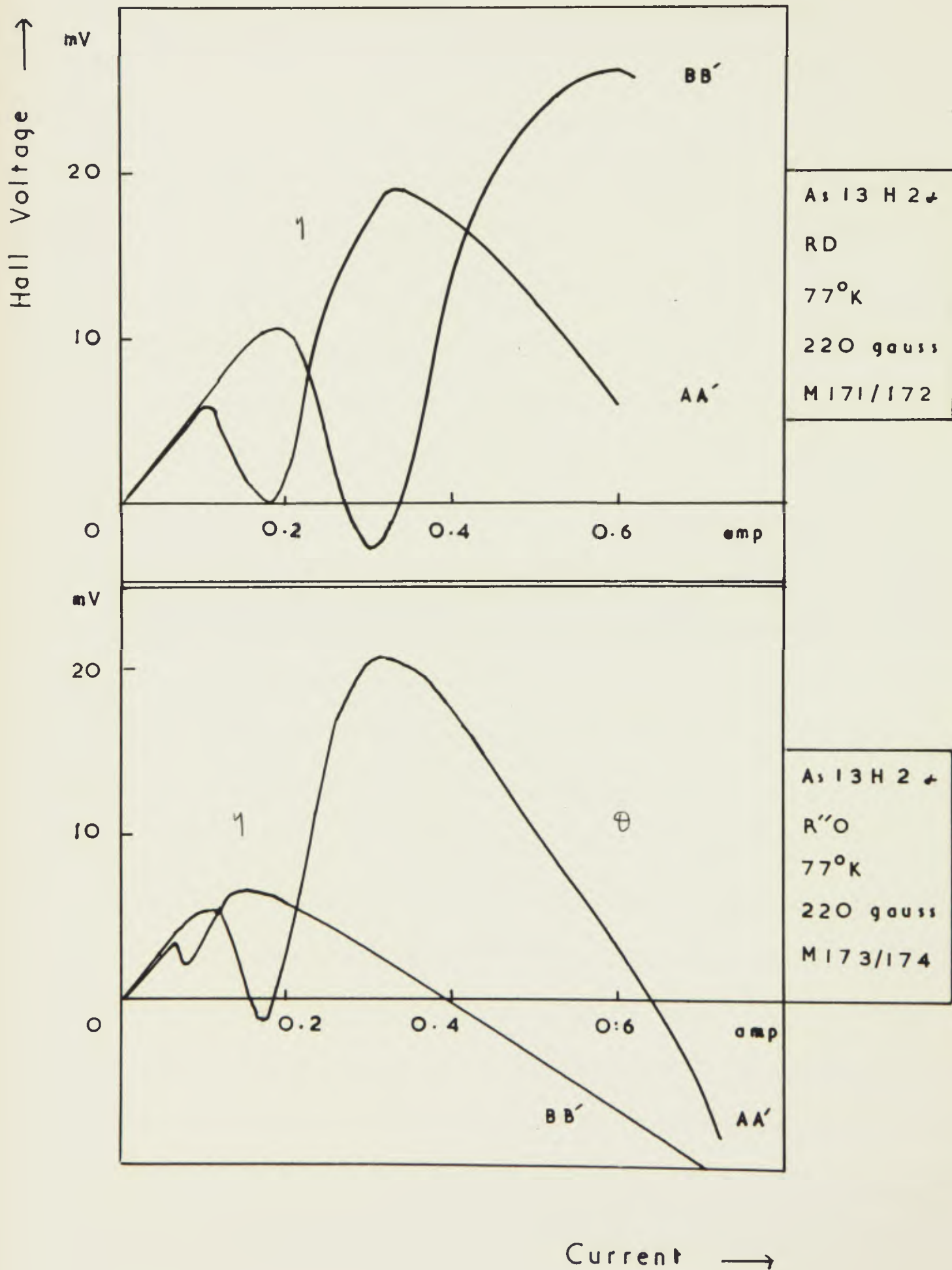
(1) The distortion moves down the crystal from the injecting contact as the current is raised. For $As_{13}M_{1+}$ this might be coincidental. But the results for $As_{12}M_{3+}$ provide stronger evidence since the distortion passes first the pair nearer the injecting contact for each direction of hole flow.

(2) At currents below the point at which distortion begins, the current is simply an electron one; the linear slope in this region



Figure 4.3.11

Hall Voltages for Two Directions of Flow



is exactly that obtained with ohmic flow in this material.

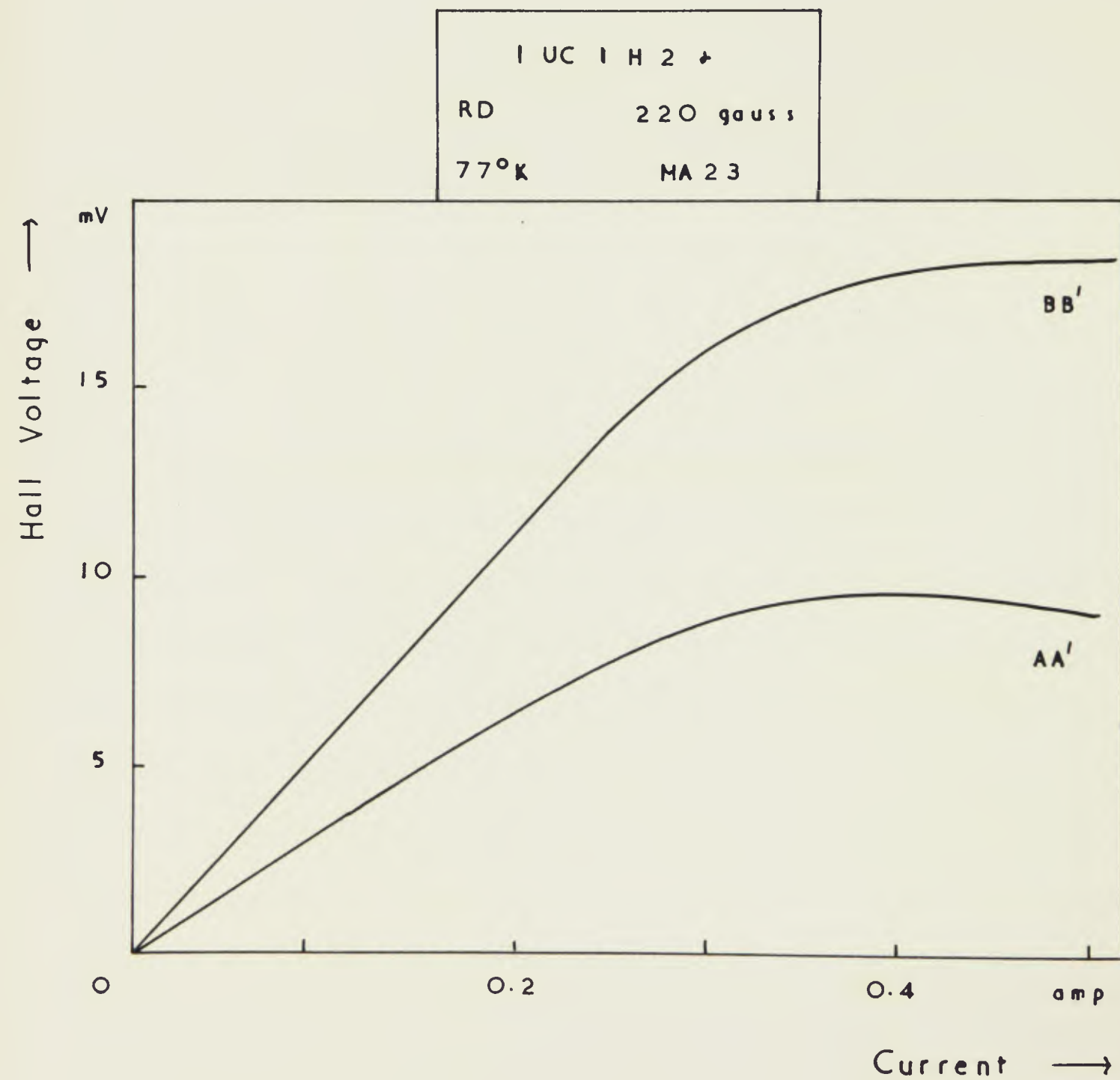
(3) There are two rather different effects, which both tend to produce a change from n-type to p-type behaviour. The first inversion is a passing effect, region γ , which is not necessarily completed. This is exactly correlated with the negative resistance shown by the local potential. The second inversion, region θ , is not visible on all curves before breakdown. Evidence taken from pulsed conditions, under which breakdown is delayed, confirms that the cross-over may be obtained in these cases. The first effect is interpreted as being due to the passing of a hole-trap saturation front, the shape being due to a strong local distortion of the equipotentials by the magnetic field acting on the holes. The second is interpreted as being due to the progressive change from electron flow to hole flow, predicted in the theory section, caused by electron stagnation at the injecting contact and consequent back diffusion.

(4) Following the first inversion effect, the Hall voltage rises to a value which in some instances exceeds the equivalent ohmic value. Yet any increase in the carrier density due to trapped or free holes (necessarily fully neutralized) would reduce the Hall voltage. The rise is considered to be a passing equipotential distortion effect which matches the inversion narrowing.

(5) The characteristics for current flowing in opposite directions for AsI_2H_3 were not identical, in spite of the contact geometry being similar with respect to the crystal. This is shown by the fact

Figure 4.3.12

Hall Voltages for Uncompensated Material



that the current at which the distortion reaches AA' is the same for both directions of hole flow though AA' is not at the centre of the crystal. Also the terminating breakdown currents are seen to be different. Possible explanations are:

- (a) Inhomogeneities in the bulk or surface properties.
- (b) Different recombination rates in the contact localities.
- (c) Different injection ratios.

Of these, (a) is unlikely in the absence of any correlating evidence. (b) and (c) are in practice equivalent; we point out that these are still consistent with the model of the injection ratio of section 4.1.

4.3.3.7 The Hall voltage as a function of compensation.

Since results similar to those of the foregoing sections did not seem to have been reported before, it became important to see how dependent they are on the compensation of the material. Because a range of materials of known compensation ratios was not available, evidence could only be procured for a crystal of approximately the same carrier density, reckoned to be uncompensated. Results for this crystal, mounted as 1UC1H2+, are shown in figure 4.3.12; two sets of Hall contacts are employed, AA' being closer to the rectifying contact. We deduce:

- (1) There is evidence that some sort of disturbance of the Hall voltage is taking place at high currents for AA' and for BB'. The effect is small but steady.

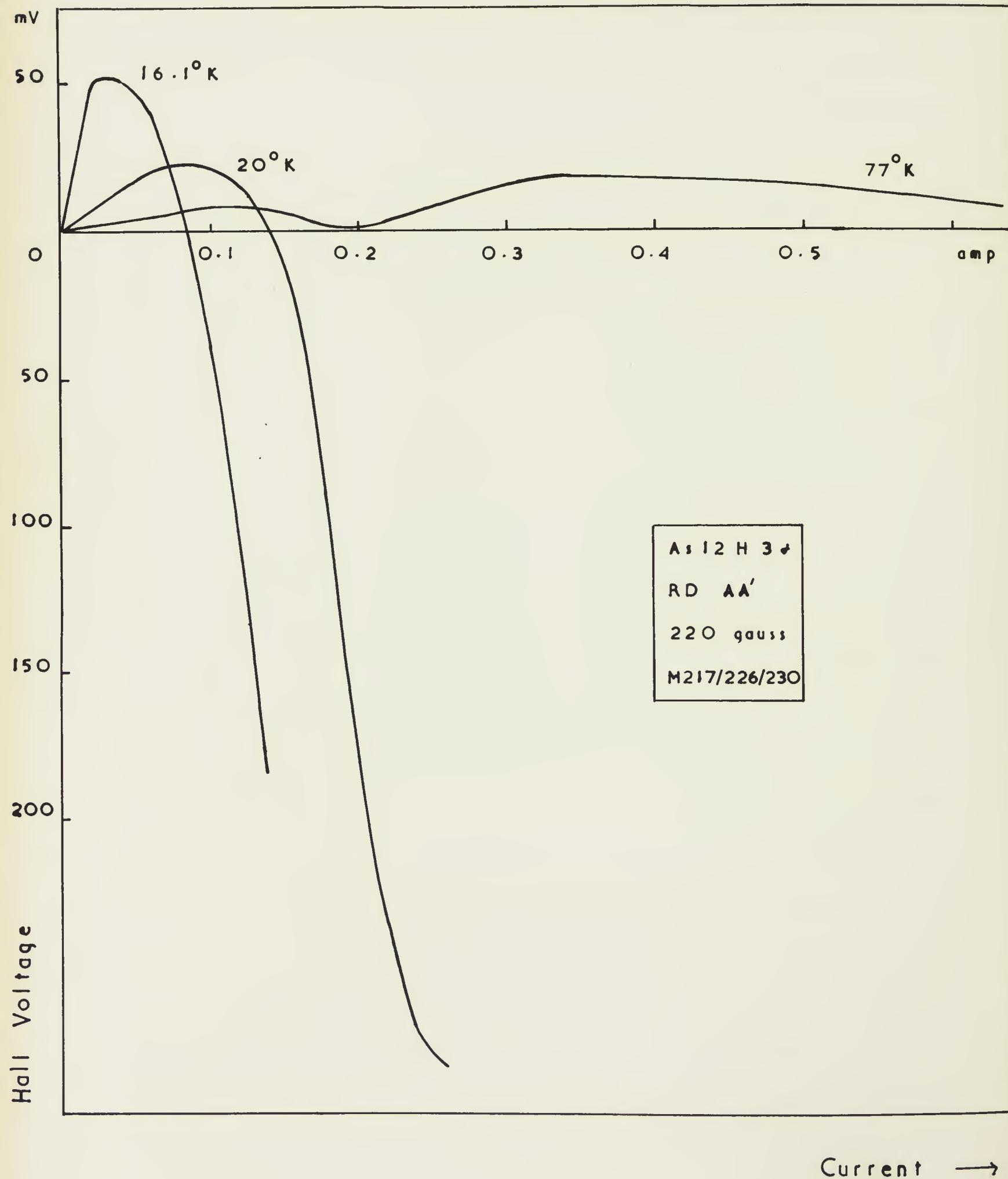
1. The first part of the paper is devoted to a general discussion of the problem.

2. The second part is devoted to a detailed study of the case of a single particle.



Figure 4.3.13

Hall Voltages at Three Temperatures



(2) The effect is far less significant than in the compensated material and it is clear that considerably higher current densities are required in uncompensated material to produce positive evidence for inversion.

The relative difficulty in demonstrating an injection effect in the bulk of the material may be due either to the rectifying contact or to the recombination properties of the material itself. It was found to be more difficult to make strongly rectifying junctions on uncompensated material, though the production of poor ones on compensated material is quite possible. If the anode has a low injection ratio, then the electron leakage current may be too high for hole penetration to become noticeable. Strong recombination in the contact locality produces the same result. It is significant that the best of six junctions on uncompensated material produced less hole penetration than the worst of those on overcompensated material.

These results thus fit in with those on compensated material without providing very much positive evidence. All further results in this chapter will be for the overcompensated zinc-doped material.

4.3.3.8 The Hall voltage as a function of temperature.

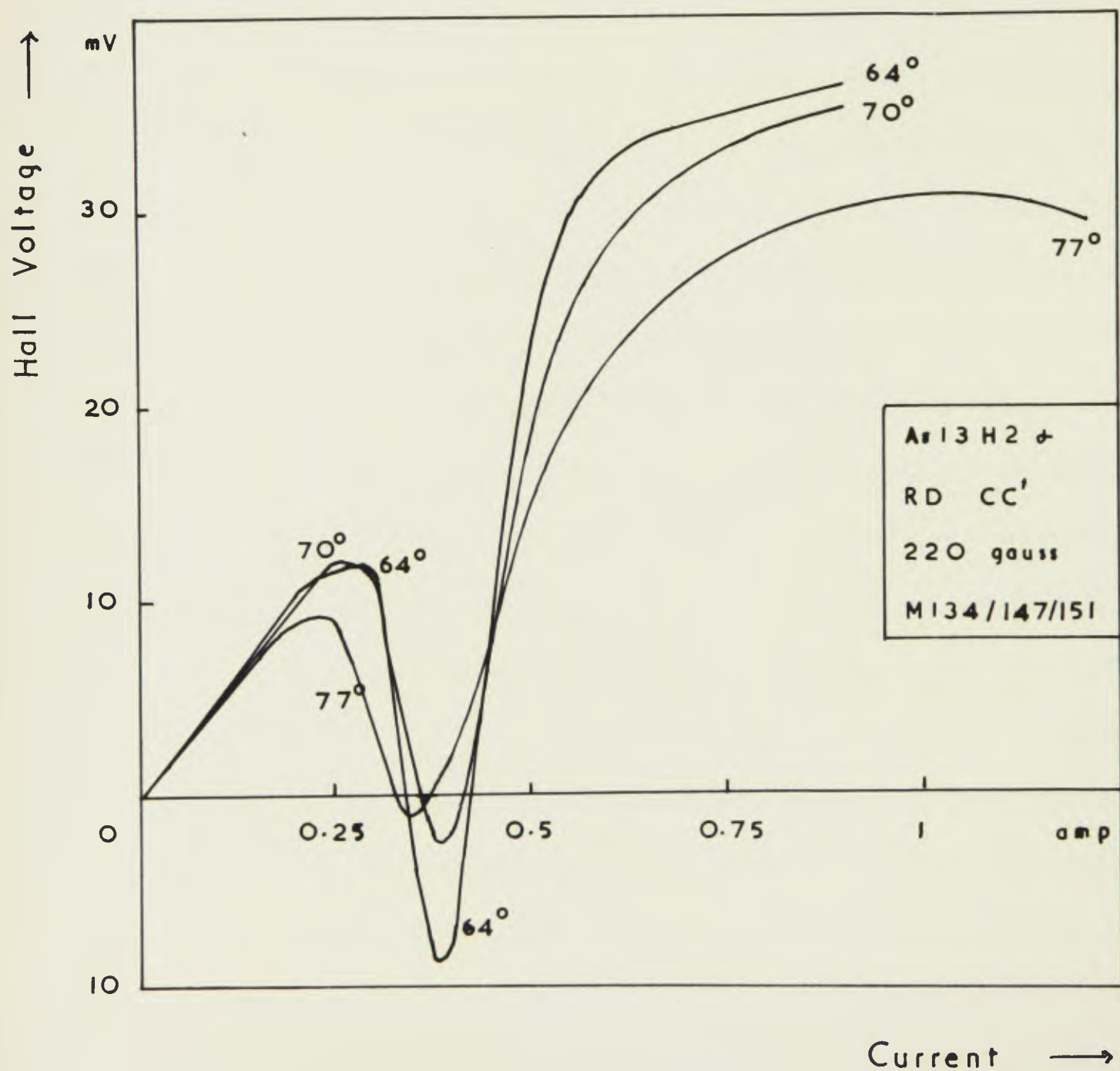
Figure 4.3.13 shows the Hall voltage for a number of temperatures for specimen As12H3+ and figure 4.3.14 for specimen As13H2+. We note:

(1) Between 77°K and 64°K, there is little change in the characteristic pattern. The general tendency for most specimens and contact



Figure 4.3.14

Hall Voltages near 77°K



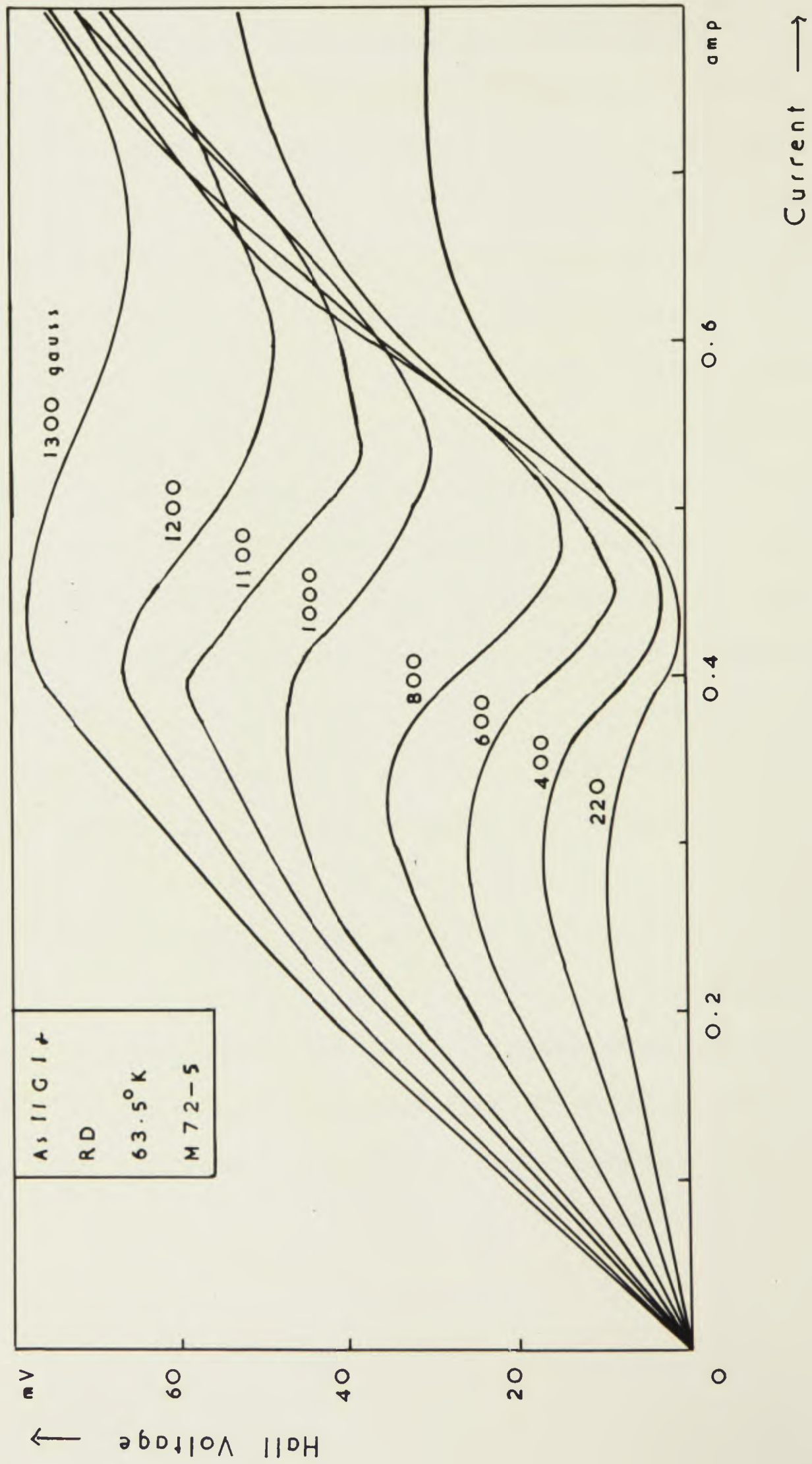
pairs was for the first inversion to be more intense at the lower temperature. No consistent movement to higher or lower current values was found. The greater intensity of the inversion at 64°K is interpreted as a sign that the saturation front is more sharply defined at a lower temperature, while the constancy of the currents may be due to the lack of variation in this temperature range of the number of ionized donors, the carrier mobilities and the recombination cross-sections.

(2) Between 64°K and 20°K the main change is a very large increase in the intensity of the first inversion. This may represent a sharper saturation front which corresponds with the evidence from the increased tendency for a negative resistance at 20°K noted in section 4.3.3.1.

(3) Between 64°K and 20°K the first inversion moves to a higher current. In view of the lack of evidence, no explanation will be attempted.

(4) No results are available on the temperature variation of the second inversion, which was found to take place. If it is due to a change from electron to hole flow, and assuming the mobility ratio moves in favour of the holes as the temperature is lowered, then it may be expected that the second inversion will take place at a lower current.

Figure 4.3.15 Hall Voltages at Eight Magnetic Fields



4.3.3.9 The Hall voltage as a function of magnetic field.

Figure 4.3.15 shows the Hall voltage curve for $\text{As}_{11}\text{Ga}_{19}$ at six magnetic fields. It demonstrates the field dependence of the first inversion. We note:

(1) The first inversion is just complete in the lowest magnetic field; as the magnetic field is raised, the tendency to cross-over is reduced. The interpretation given is that the inversion is due to distortion of the equipotentials as the trap saturation front is pushed through the contacts. Raising the magnetic field increases surface recombination of deflected minority carriers and reduces the sharpness of the front. Equipotential distortion is thus reduced together with its effect on the Hall voltage characteristic.

(2) The point of closest approach to inversion moves to a higher current as the magnetic field is raised. Our interpretation is that the enhanced surface recombination at higher magnetic fields reduces the number of holes available, so that a higher total current is required to force the saturation front a given distance down the crystal.

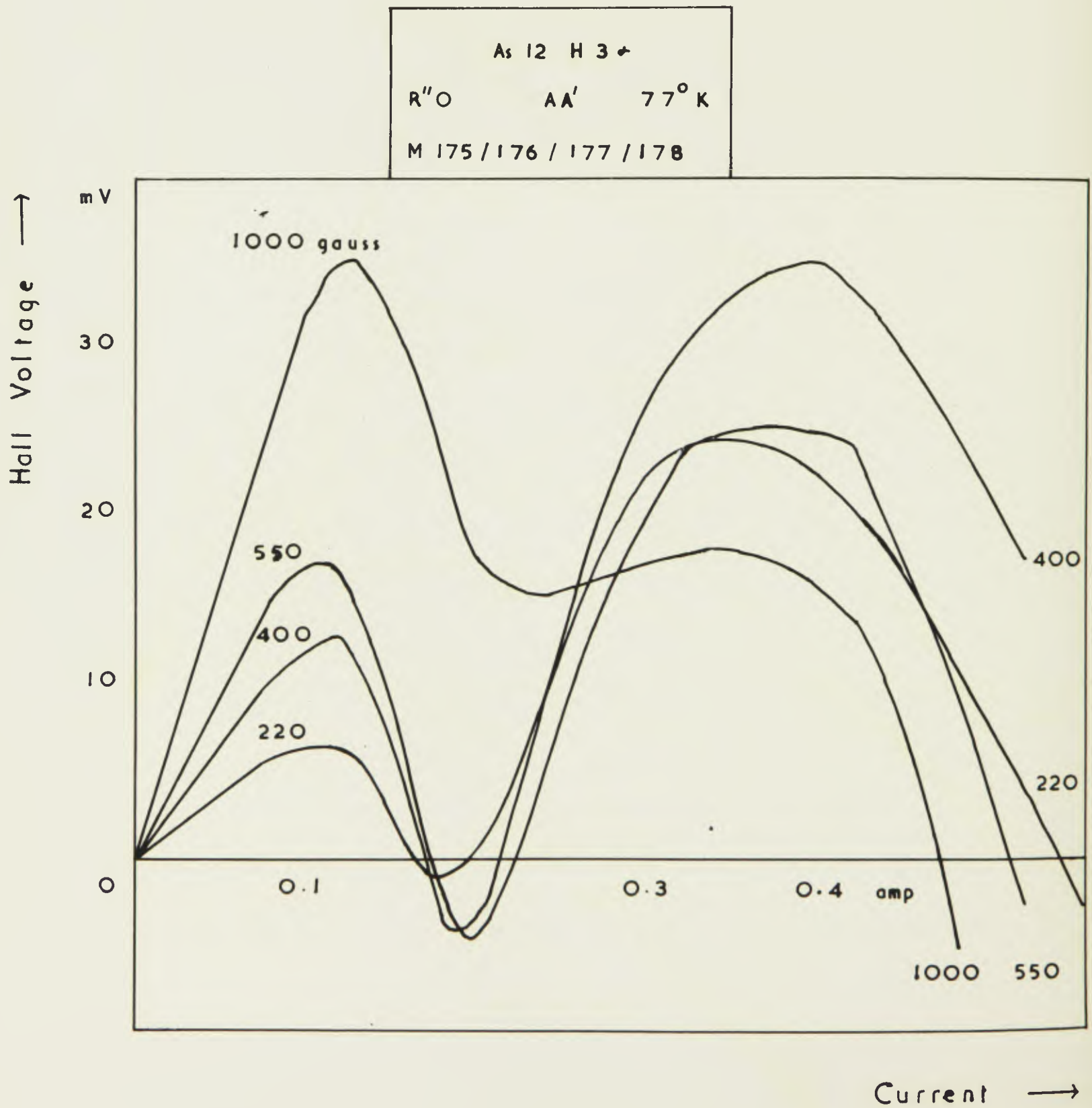
(3) These effects are small at the lowest fields, showing a limiting position to have been reached. According to the discussion of section 4.1.13, this might indicate a change from case (1) to case (2).

(4) The second cross-over is not visible. It was however observed for this specimen under pulsed conditions at a higher current.



Figure 4.3.16

Hall Voltages at Four Magnetic Fields

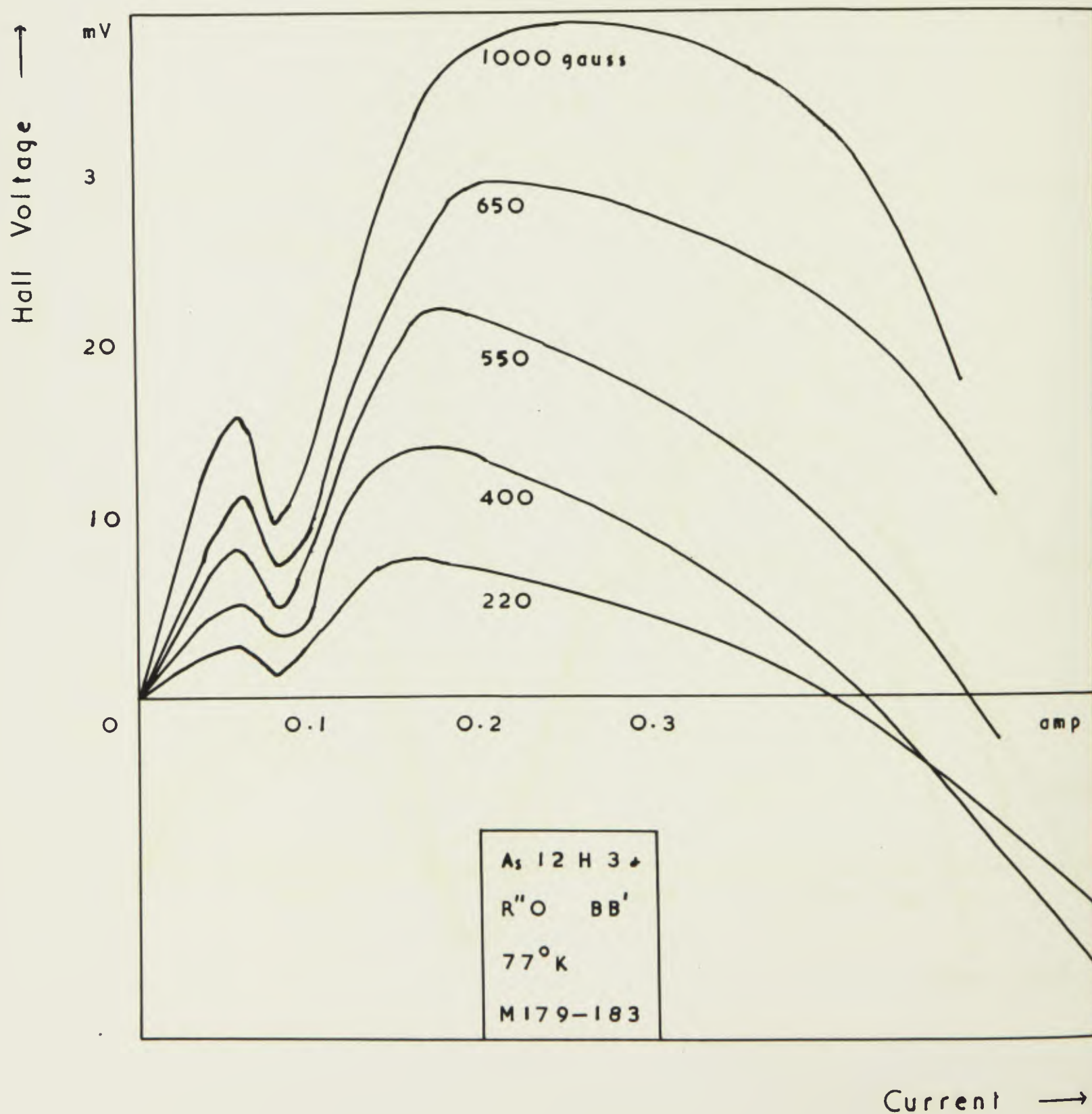


THE [illegible] OF [illegible]
[illegible] [illegible] [illegible] [illegible] [illegible]



Figure 4.3.17

Hall Voltages at Five Magnetic Fields





1000

1000 1000

Magnetic Field

Figure 4.3.18

Magnetic Field Dependence
of Hall Voltage Inversions

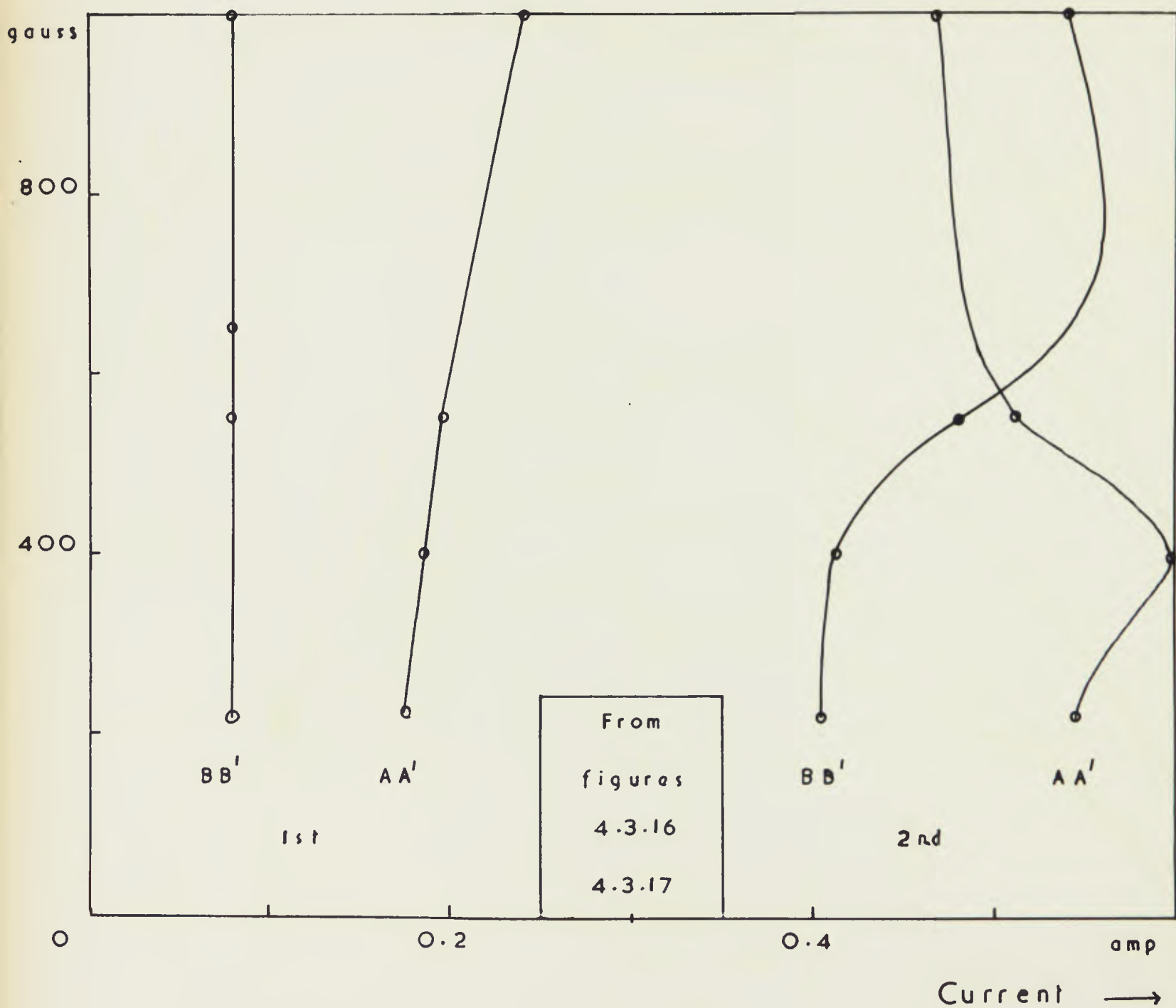




Figure 4.3.19

Hall Voltages at Three Magnetic Fields at 20°K

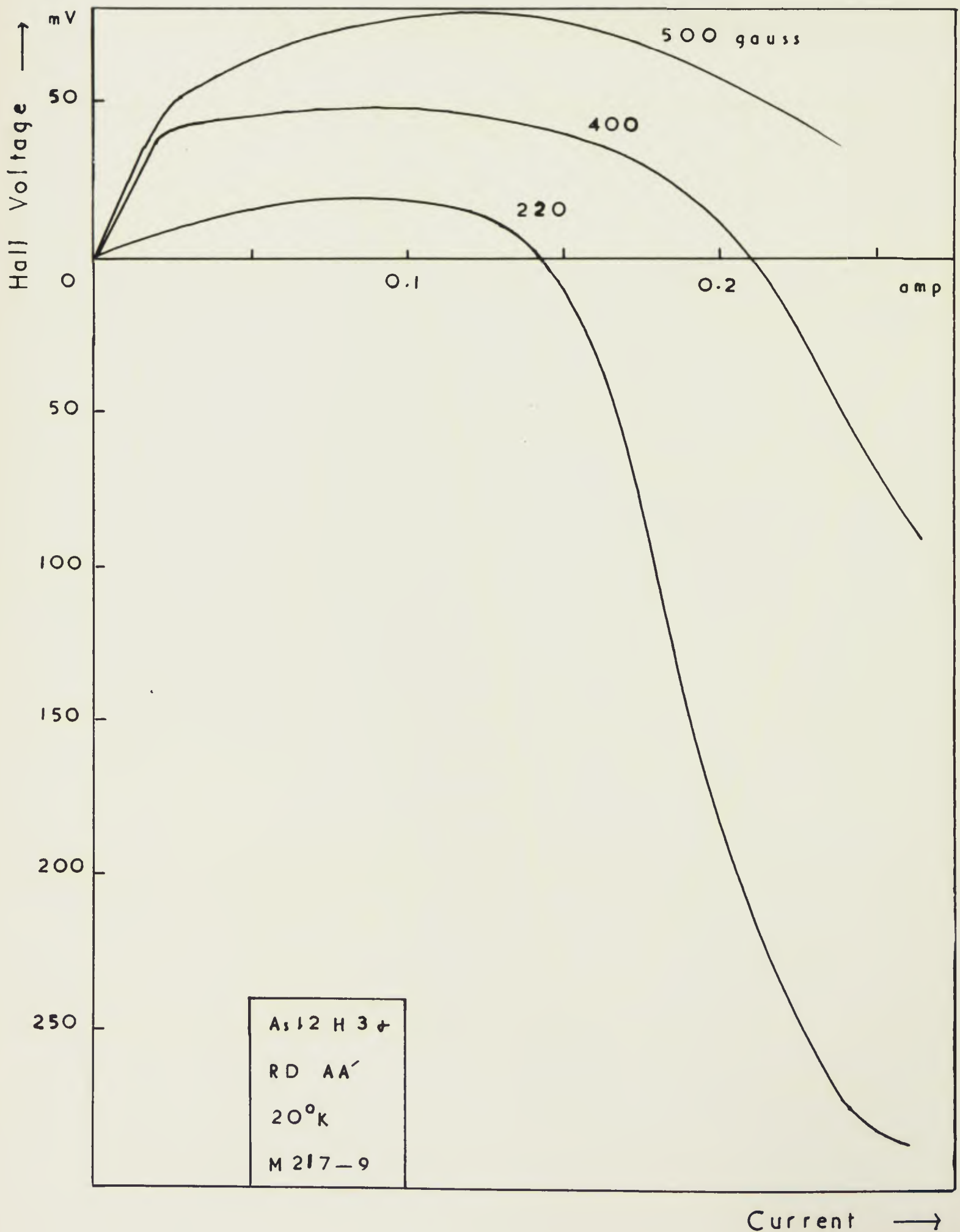
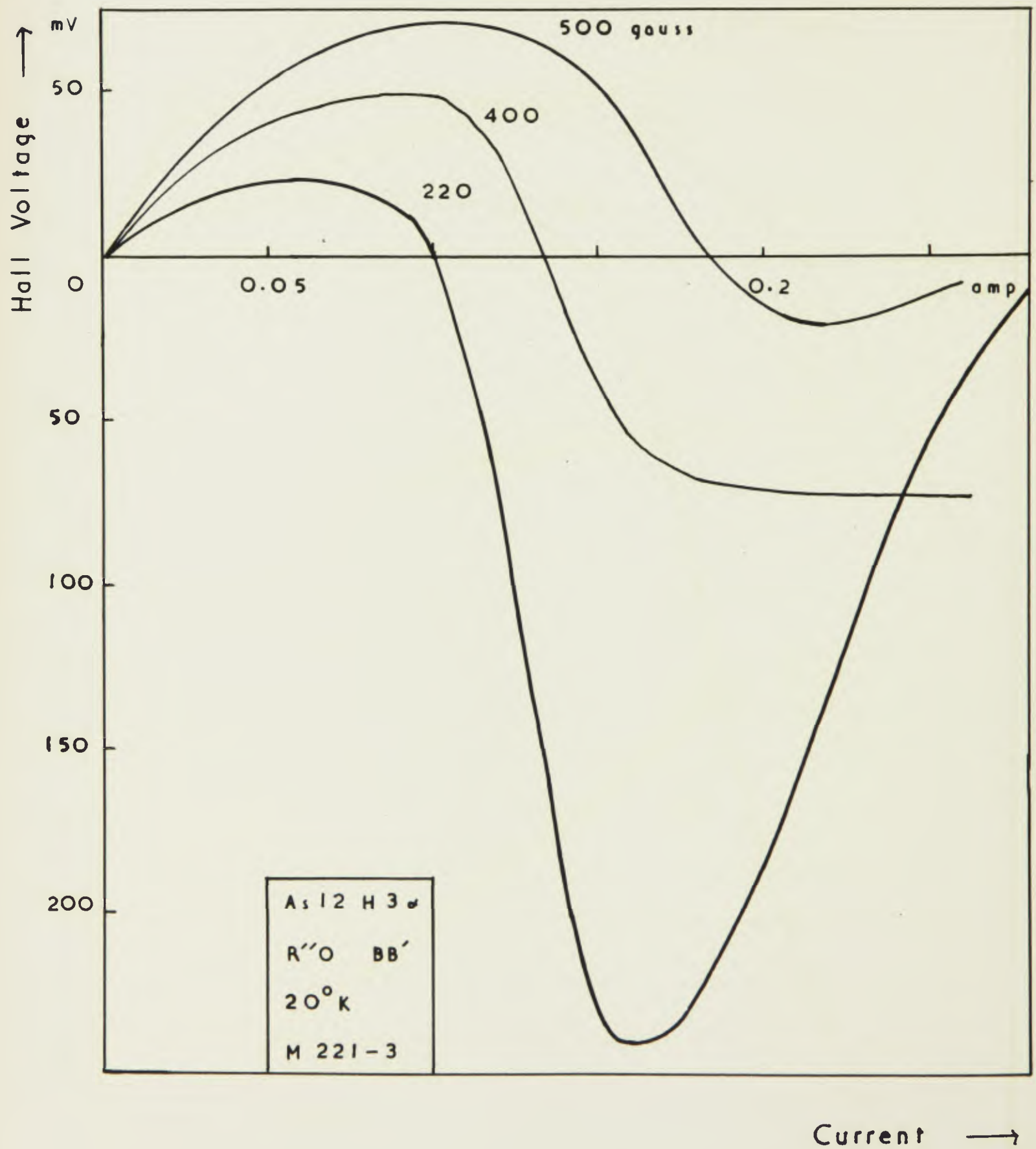




Figure 4.3.20

Hall Voltages at Three Magnetic Fields at 20°K



Figures 4.3.16, 4.3.17 show the equivalent Hall voltage curves for As_{12}H_3 at 77°K for four magnetic fields; figure 4.3.16 refers to the contact set BB' which is closer to the injecting contact R", figure 4.3.17 to the contact set AA'. These serve to confirm points (1) -- (3) noted from figure 4.3.15, but give added information about the second inversion.

(5) For low magnetic fields, the current required to produce the second inversion rises as the magnetic field is raised. This is consistent with the explanation given in terms of increased surface recombination impeding the change-over from electron to hole flow.

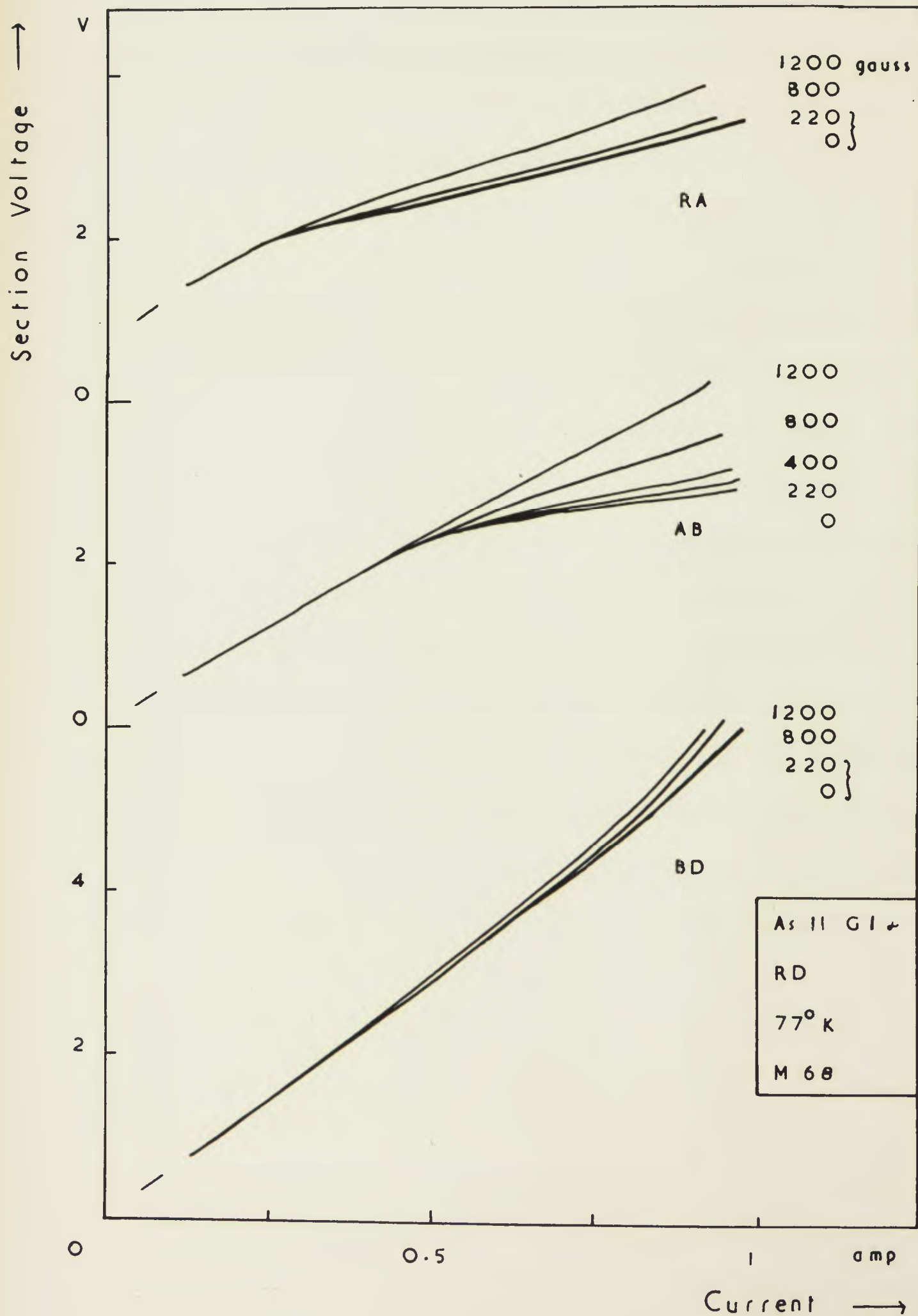
(6) At magnetic fields higher than 500 gauss this current trend is reversed and a rather discouraging feature occurs which is more clearly shown in figure 4.3.18. Here the second inversion current is plotted against the magnetic field for the two contact pairs. Then for the higher magnetic fields, the second inversion appears at the further contact first, which seems directly to contradict the hypothesis of holes penetrating steadily down the crystal. We prefer to consider this as evidence of severe disruption of the current flow, probably involving funnelled flow. This view is supported by the anticipated behaviour at low fields.

(7) The equivalent curves at 20°K , figures 4.3.19, 4.3.20 confirm points (1), (2) on the first inversion but give no evidence on the second inversion.



Figure 4.3.21

Sectional Voltages in Magnetic Fields



4.3.3.10 The effect of a magnetic field on the voltage distribution.

Figure 4.3.21 illustrates the effective magnetoresistance in double injection; it shows the distribution of the voltage along a crystal for a number of magnetic fields. We note:

(1) The effect is not a true magnetoresistance since a variation occurs not in the single carrier range but only where a non-linearity due to double injection is evident.

(2) The crystal conductivity is lower in a magnetic field in a double injection range and by 1200 gauss has virtually reverted to its single carrier value in the current region observed.

(3) The effect of applying 220 gauss is almost negligible, showing that measurement of the Hall voltage at this magnetic field gives a reasonable approximation to the field-free case, whereas 1 kgauss does not. This supports the conclusion of section 4.3.3.9 that at a field of 1 kgauss the double injection flow has been seriously disturbed.

4.3.4 Discussion of Results.

4.3.4.1 Aims.

It is intended at this stage to discuss how satisfactorily the aims of this experimental section have been achieved; we leave the attempted justification of these aims to the concluding section.

The initial aim was the explanation of the particular anomalies observed in the injection characteristic for a particular crystal, zinc-doped germanium, overcompensated with arsenic. A preliminary survey showed that double injection was being observed with appreciable minority carrier penetration. Such evidence was explained by two hypotheses:

- (1) The movement of a front of saturated hole-traps with rising current.
- (2) The change from an electron to a hole current with rising current, due to electron stagnation at the injecting contact.

From this juncture our aim has been to devise experiments which give information bearing directly on these hypotheses.

4.3.4.2 Measurements.

The measurements made, together with their specific relevance, have been summarised in the section dealing with the pattern of the results (section 4.3.2). We point out that four of the measurements do not seem to have been usefully employed in double injection study

up to now (apart from a preliminary paper by Melngailis and Rediker l.c. 454, 1962):

- (1) The positioning of several sets of potential contacts down the length of the crystal.
- (2) Injection from either end of the crystal using such contacts.
- (3) Examination of the potential difference between closely spaced contacts.
- (4) Investigation of the Hall voltage for injecting contacts.

A fifth new type of measurement was devised, the analysis of growth and decay transients, which may be used for all the contact sets and for Hall voltage investigation. Such measurements are not reported because those obtained raised more questions than they answered. The four new measurements, supplemented by more conventional current-voltage measurements, have been investigated with variation of specimen geometry, temperature, magnetic field and carrier concentration.

Three further types of experiment are available; detection of recombination, photoexcitation of carriers and analysis of thermoelectric properties. An investigation of the effect of injection on the thermoelectric properties of a crystal seems to be a new field where the results would be virtually undecypherable without full knowledge of the carrier flow parameters. Photoinjection by illumination with intrinsic radiation serves as an alternative method of producing holes; the effective injection ratio is controllable but would not be easy to determine numerically. As a

method of investigating double injection flow it has distinct possibilities. Illumination of the body of the crystal by extrinsic radiation might produce interesting results, analogous to the quenching of photoconductivity, due to the ejection of holes from trap sites. These experiments were not tried. Investigation of recombination radiation might be hoped to give information on the locality and mechanism of minority carrier recombination. Attempts at measuring this are described in chapter 7.

We deduce that the measurements conducted are all those that are most directly relevant to the current-voltage characteristic of a double injection diode. We proceed to compare the amount of information obtainable, the amount obtained and the amount presented here.

4.3.4.3 The Sample of Information Reproduced.

The variable parameters are:

- (1) Diode length.
- (2) Crystal area of cross-section.
- (3) Cross-section surface to perimeter ratio.
- (4) Concentration of deep trap sites in the material.
- (5) Compensation ratio.
- (6) Contact size.
- (7) Location of potential contacts (in pairs).
- (8) Temperature.

(9) Magnetic field strength.

(10) Current.

For all cases a voltage is to be measured between either adjacent or opposite contacts. Using 10 as the sampling number for each parameter, the total amount of obtainable information fills a 10^{th} order matrix, with $2 \cdot 10^{10}$ elements of information. In the experiments, over 600 injection curves, most plotted against current, were obtained, each giving a voltage reading at about 40 useful abscissa values, the information total being about 24,000 elements. This has involved changes in all the parameters listed above. Since the aim is to examine double injection in terms of the two hypotheses, front movement and flow inversion, we have two questions to ask:

(1) Are these hypotheses consistent with the data?

(2) Are all alternative hypotheses inconsistent with the data?

Supposing the answers to be in the affirmative, we have to determine the smallest volume of information which shows the hypotheses to be both sufficient and necessary; this then is the sample of results which should be presented. We now discuss the two questions.

4.3.4.4 The consistency between the data and the hypotheses.

4.3.4.4.1 Self-consistency of the data.

Data is no more useful than its self-consistency warrants. This refers particularly to repeatability. The injection i-V curve showed a change of $\sim 5\%$ from its first run after which repeatable readings

were obtained; since the resistance became higher and ohmic readings remained constant, this was put down to a forming effect at the contacts. For contact pairs which were very nearly opposite, the behaviour of the local potential was not consistent, varying once even for two successive switchings. Since these readings are extremely sensitive to the equipotential positions, this variation was not regarded as surprising. Hall voltage readings taken a month apart varied by less than 2%. The self-consistency of the readings was reckoned to be quite sufficient for the conclusions deduced.

4.3.4.4.2 Extraneous effects.

Two other effects were noted, which seemed to be independent of the hypotheses.

(1) A discontinuity appeared frequently but not invariably in the Hall voltage while at other times the curve was smooth over the same region. This was interpreted as due to changes, discontinuous or continuous, in the carrier flow streamlines; a corresponding discontinuity in the local potential was also noted.

(2) At 20°K and below measurement of the Hall voltage was made more complex by a change with a time constant of ~ 10 seconds. It appeared to give a slow decrease in the penetration of the hole saturation front. We deduce that a slow hole trap is responsible. Results plotted show the equilibrium value on increasing the current slowly.

4.3.4.4.3 Consistency of the hypotheses.

The aim has been to produce hypotheses consistent with all the experimental evidences available after making allowance for any mutual inconsistency in them. Experiments have been devised to test the hypotheses in the most effective ways; this has involved the use of multicontact specimens and in particular, injection from either end of such a specimen. It is reckoned that the hypotheses are fully consistent with all data obtained with the exception of one category, magnetic fields above 500 gauss. This objection has been dealt with in sections 4.3.3.9 and 4.3.3.10 where the conclusion is reached that the current flow is considerably disturbed in such a magnetic field, so that the results cannot be interpreted straightforwardly; the objection to the hypotheses is thus reduced.

Hence in answer to the first question of section 4.3.4.3., we conclude that the hypotheses are as consistent with the data as can reasonably be expected.

4.3.4.5 Inconsistency of Alternative Hypotheses.

So far the aim has been to show that the hypotheses are sufficient to explain the experimental data; in this considerable success has been claimed. It is however far more difficult to demonstrate their necessity. A large class of alternative explanations which rely on crystal inhomogeneity can be ruled out by the results from injection from either end; this is more effective than the demonstration

of consistent results from quite a number of specimens with a single injecting contact.

At the same time it is accepted that explanations basically similar to ours but differing in detail could be justified. In fact it has not been our aim to investigate the detailed nature or mechanism of either inversion for reasons to be given in section 4.4. On this topic, as on others, the work is speculative rather than exhaustive.

Finally it is apparent that the sample of data presented is not the perfect one, as seen with hindsight. What comparability there is in the data is a measure of the success of the anticipation commanded.

4.3.4.6 Experimental Results in the Literature.

As was described at the beginning of this chapter, work on double injection is at a relatively early stage of development. A number of instances of negative resistance have been published, but there has apparently been no intensive investigation at a fundamental level; results available have been reviewed by Laspert (l.c. 1964, 455). Here we only consider literature which has direct relevance to the ideas deduced from our experimental results.

In the first place there is the work of Stafcev et al. (l.c. 1959, 456; 1960, 457; 1961, 458; 1962, 459); this formed the starting point for the explanations and investigations reported here.

Experimental results are reported by Stafeev on the "modulation of the diffusion length", by which is meant the increase in the minority carrier diffusion length as the injection rate rises; the dependences of the current-voltage characteristic on temperature, magnetic field and intensity of illumination are determined, but no attempt is made to investigate the electric field or carrier concentrations and their variations down the length of the diodes. A point of confusion is that while Stafeev terms his specimens "long diodes", it is stated by Lampert that the ideas refer to short specimens; our analysis fully supports this. This may explain the lack of interest in Stafeev's work.

Secondly there is the work of Holonyak (l.c. 1962, 460). The circumstances in this case are slightly different from those employed in this work in that thin p-i-n structures were employed. Results showed negative resistances for high resistivity GaAs, Si and Ge, doping elements employed being Au, Zn, Cd, Co, Cu, Fe, Ni and Mn. The results were confined to V-i characteristics, under varied conditions of illumination, temperature and magnetic field. The success with thin specimens and with hole traps is in accordance with ideas put forward here.

Thirdly there is the paper of Melngailis and Rediker (l.c. 1962, 461) on indium antimonide. Here three sets of potential and Hall contacts were employed for use with an injecting contact, and sectional voltages were recorded showing the spread of minority carriers

down the crystal with increasing current. The explanation given is in terms of trap saturation following Stafeev (l.c. 1959, 462) and is thus similar to that given here. However Melngailis and Rediker give no Hall results, though these were available; we deduce that this because of difficulties in their interpretation, which we have tried to overcome.

Finally there is a considerable amount of literature from Japan on the effect of avalanche breakdown on the current flow in a crystal (M. Kikuchi, Y. Abe l.c. 1962, 463; 1962, 464; 1962, 465; M. Kikuchi, Y. Abe, P. Chong 1963, 466; K. Ando, M. C. Steele, M. A. Lampert l.c. 1962, 467; 1963, 468). We leave an analytical discussion of impact ionisation to chapter 5, and consider here just the effect of localized electron-hole production. At a physical constriction in a semiconductor crystal, which becomes a region of high electric field, avalanche multiplication may take place. Thus carriers are effectively injected into the body of the material. The results presented by the Japanese authors include those for the investigation of the voltage distributions down specimens incorporating a constriction, and also Hall voltage measurements. Phenomena observed include negative resistances and Hall voltage inversion as well as spontaneous oscillation. It is interesting to note that Kikuchi and Abe (l.c. 1962, 469) noted a second incipient Hall voltage inversion similar to that reported here, while the effect of a magnetic field is also similar. However in view of the uncertain conditions imposed

by the constriction, it is not felt that the results provide much confirmation for those given here. The explanation given to the Hall effect inversions in section 4.1.13 is rather more satisfactory than that given by Kikuchi and Abe (l.c. 1962, 470) and by Ando, Steele, Lampert (l.c. 1962, 471) which relies on an orthodox magnetic field variation of the carrier mobility ratio; our explanation in terms of the interrelation of hole and electron diffusive and field driven currents might be valid for "sagicons" (semiconductor oscillation generators by injection and constriction).

4.3.4.7 Conclusion to the Discussion of Experimental Results.

We conclude the discussion of experimental results by making the following points:

(1) As a method of studying recombination and trapping, double injection is a useful device which has been largely neglected.

(2) The new experimental techniques introduced here, namely the investigation of the voltage distribution and the Hall voltage, provide tools for investigating the fundamental carrier flow properties.

(3) N-type germanium containing zinc levels appears to give high hole penetration and hence is more suitable for demonstrating Hall voltage inversion than uncompensated material.

(4) The experimental results obtained are highly complex, giving high information gain compared with monotonic curves.

(5) Two hypotheses have been set out to explain the results; the advance of a hole-trap saturation front as the current is raised and the subsequent change from an electron to a hole dominated flow, each effect being noted first at the injecting contact and spreading down towards the other end.

(6) These hypotheses have proved consistent with stringent tests devised to examine them, with the exception of one special case.

(7) There is insufficient experimental evidence in the literature to confirm or to demand the modification of the hypotheses.

(8) The work presented here must be regarded as preliminary and merely serves to show the need for a comprehensive analysis.

4.4 Conclusions on Double Injection.

In this chapter, both computationally and experimentally, the discussion has been entirely in qualitative terms. No experimental error margins are quoted here or elsewhere because no measurements have any absolute significance. The primary cause of this is the lack of knowledge of the exact composition of the specimen material employed. As has already been indicated, the work on double injection represents a pilot scheme for a quantitative project on deep level trapping and recombination, where crystal growing facilities are available. An outline of the project might be:

- (1) Take high purity germanium, produce n-type, p-type low resistance specimens by doping with shallow impurities.
- (2) Measure mobilities in both materials for majority carriers by the Hall effect and for minority carriers by the drift mobility.
- (3) Measure minority carrier lifetimes.
- (4) Examine the Hall effect for injected carriers for multi-contact specimens and pulsed currents.
- (5) Attempt to fit results to computed curves for variable recombination and mobility parameters.
- (6) Add in deep level impurities and examine the difference in results due to trapping and extra recombination.
- (7) Repeat for ranges of deep level concentrations and impurities. This could be applied to materials other than Ge and Si. Besides a vast extension to recombination and trapping knowledge, the result might be a really effective negative resistance region.

Chapter 5

Impact Ionisation.

5.1 Introduction.

With the application of a high enough electric field to a semiconductor, a reversible breakdown takes place. In the case of diodes it was first explained in terms of a field induced excitement of electrons across the band-gap (Zener effect) but this explanation has been replaced in most cases by that of impact ionization and the formation of an avalanche (R. A. Smith, 1959, 501, p.273). Under some circumstances it is found that a negative resistance may accompany the breakdown. The aim in this chapter is to examine the conditions for this negative resistance with the aid of the analysis in chapters 1 and 2.

The line of approach to be taken is this. A distinction is drawn between two carrier phenomena with electron-hole production and single carrier production with impurity ionization. In each case filament and domain formation are discussed as secondary processes, but the cause of the instability is sought in the intrinsic properties of the crystals. Emphasis is placed on the single carrier case. Analysis of the ionization rate equations is to be made to show that previous treatments which have claimed to explain the negative resistance are physically inadequate to provide it. Two features of the system which are capable of providing the negative resistance will then be analysed and

the importance of compensation demonstrated in each case. Finally the physical and technological implications of current controlled negative resistance will be discussed.

5.2 Two Carrier Impact Ionization.

For a specimen of pure germanium or one of germanium doped with shallow donors at a temperature at which these are fully ionized, the only source of extra electrons for the conduction band is the valence band; ohmic electrical contacts are assumed so that double injection does not take place from a less pure region. As shown in section 4.1.5, for semiconductor applications Poisson's law prevents the introduction of an electron space-charge above the negligible level required to provide the electric field gradients. Impact ionization thus means the production of electron hole pairs by an energetic carrier. All extra carriers are assumed to be mobile, since trap-free germanium has been specified. The process is cumulative, for these extra carriers may also be accelerated to cause ionization, producing an avalanche.

Experimental results under these circumstances have been provided by J. B. Gunn (1956, 502); they show a negative resistance with increased conductivity. Referring back to section 1.2 we note that this negative resistance should be unstable against filament formation rather than domain formation. Impact ionization may be expected to provide extra carriers where the electric field is high, while current continuity demands that the carrier density be lower in regions of high

electric field. Gunn (l.c. 1956, 503) attempts an explanation in terms of hole accumulation at the negative contact in the planar flow case. The employment of a constriction throws light on the situation; at a constriction the electric field is automatically higher. Gunn noted that more striking results were obtained with contacts of higher spreading resistance; ionization within the spreading resistance would fit the case of apparently planar flow also. In practice avalanche injection whether effected at a contact spreading resistance or at a constriction is equivalent to double injection with a different boundary condition at the injection point. Thus the equations developed for double injection in section 4.1 are applicable in this case. Experimental results obtained in Japan on negative resistance and oscillation through the use of a constriction, (e.g. M. Kikuchi, Y. Abe 1962, 504), have been discussed under double injection in section 4.3.4.6. Double carrier impact ionisation thus requires no further treatment here.

5.3 Single Carrier Impact Ionization.

5.3.1 Special Features.

When a specimen of germanium contains unionized impurities sited in the forbidden gap, a new possible source of extra carriers exists. For n-type material the number of ionized donors depends on the depth of the impurity level, the density of states in the conduction band and the temperature (R. A. Smith l.c. 1959, 505, p.82). At a temperature low enough so that a significant reduction in the free carrier

density has been caused by the 'freezing out' of conduction electrons, impact ionisation of impurities may be observed. In the n-type case only one type of carrier, the electron, is involved. Experimental results on this were first obtained and interpreted by H. Sclar and E. Burstein (1957, 506); electric fields required for impurity impact ionization are low in germanium, down to 3 volts per cm. compared with those for pair production at 10^4 volts per cm.

A difference has been found in the behaviour of uncompensated and compensated specimens. Uncompensated material has been studied by Sclar, Burstein (l.c. 1957, 507) and by S. H. Koenig, C. R. Gunther Mohr (1957, 508); a steep rise in conductivity is observed but this is monotonic, giving no negative resistance. For compensated material, a current controlled negative resistance is observed (A. L. McWhorter, R. H. Rediker 1959, 509; 1960, 510; R. J. Phelan, W. F. Love 1964, 511). This negative resistance will be examined, with particular reference to the importance of compensation.

5.3.2 Domain and Filament Formation.

Domain formation is not applicable to single carrier ionization because the extra carriers are not free to migrate from the region of their production; this is because the space-charge cannot be neutralized by a migration of oppositely charged carriers (here the ions), as in the double carrier ionization case. Hence a constriction, real or imaginary, is not suitable for explaining a negative resistance.

The fundamental difference between one and two carrier ionization processes lies in the role played by space-charge neutralization; the significance of this seems to have first been noted by V. V. Paranjape (1961, 512).

If a crystal has a region of current controlled negative resistance, it was shown in section 2.17 that the i - V characteristic exhibits an area ABCD (figure 2.4b) in which filamentary flow is allowed, only the bounding lines AB, CD representing homogeneous flow. It was suggested that the steady state position should be determined by the stationary state of a suitably computed entropy production function, subject to conditions imposed by the load line, the path taken into the area, and the relationship between fluctuations and the minimum viable filament. From these the working i - V characteristic can be explained. Thus we see that filament formation is merely symptomatic of an intrinsic instability in the electrical resistance of the crystal.

5.3.3 The Ionization Rate Equation.

Intrinsic instability is to be considered here in terms of planar (i.e. non-filamentary) flow; the aim is to find which sorts of relationship between the local field strength, the ionisation cross-section and the mobility can provide a negative resistance and which cannot. Features found for planar instability will then be employed to discuss the possible subsequent formation of filaments.

We begin by deducing an ionization rate equation in terms of generation and recombination, following the notation of B. H. Koenig, R. D. Brown, W. Schillinger (1962, 513). A_T , A_I are generation rates due to thermal excitation and impact ionization, B_T , B_I are the equivalent recombination rates, N_D , N_A , n are the donor, acceptor and conduction electron concentrations. Then the rate of conduction electron production is given by:

$$\frac{dn}{dt} = nA_I(N_D - N_A - n) + A_T(N_D - N_A - n) - nB_T(N_A + n) - n^2B_I(N_A + n) \quad .$$

The density of states in the conduction band is incorporated into A_I , A_T . In the steady state the carrier concentration is constant, giving $dn/dt = 0$.

The ionization rate equation will be considered in two ways. Firstly we will examine the orthodox equilibrium criterion of P. J. Price (unpublished) as developed by Koenig, Brown, Schillinger (l.c. 1962, 514) and then introduce an approach suited to negative resistance breakdown.

The orthodox approach (Koenig, Brown, Schillinger l.c. 1962, 515) sets dn/dt to zero. Field induced recombination of electrons with donors, the Auger process, is neglected, so $B_I = 0$. The carrier density n is assumed low, so that terms in n^2 may be ignored, and $n \ll N_D - N_A$. Hence the rate equation gives:

$$n = \frac{A_T(N_D - N_A)}{B_T N_A - A_I(N_D - N_A)} \quad .$$

The breakdown criterion is then $B_T N_A = A_I (N_D - N_A)$. Examining the physics behind this criterion, we note that it simply defines the point at which the production and recombination rates no longer balance. The concentration becomes infinite because the limiting terms, those in n^2 , have been ignored. This approach is unable to define what happens at breakdown, or whether negative resistance develops.

To treat the behaviour in the breakdown region itself, a more exact analysis of the ionization rate equation is required. We return to the original equation, again setting dn/dt to zero and neglecting the field induced recombination process. Then:

$$\frac{dn}{dt} = nA_I(N_D - N_A - n) + A_T(N_D - N_A - n) - nB_T(N_A - n) = 0.$$

We choose to consider initially only highly compensated specimens; for these, putting $N_D - N_A = N$, we have $N_D, N_A \gg N$, so that the proportion of unionized donors is small even before impact ionization has taken place. Thus the thermal recombination rate $nB_T(N_A + n)$ may be taken as nB ; the density of ionized donors is considered as effectively independent of the state of ionization ($n < N \ll N_A$). The ionization rate equation becomes:

$$A_I n(N - n) + A_T(N - n) - nB = 0.$$

At zero applied voltage there is no field induced ionization; then by the principle of detailed balance, the thermal generation and recombination rates are equal:

$$A_T(N - n_0) - n_0 B = 0.$$

By substitution:

$$\frac{A_I}{B} \cdot n(n - N) + \frac{N}{N - n_0} \cdot (n - n_0) = 0 \quad .$$

The attainment of a negative resistance turns upon the variation of the coefficient A_I/B .

5.3.4 Classification of Physical Influences.

Referring back to section 1.2, the criterion for a current controlled negative resistance is:

either $1 - \frac{J}{\sigma} \cdot \frac{\partial \sigma}{\partial J} \leq 0 \quad ,$

or $\frac{\partial B}{\partial \sigma} \leq 0 \quad .$

Features of the electrical conductivity which may be relevant to the satisfaction of the criterion, and hence giving a negative resistance, are:

- (a) A rise in the lattice temperature may increase the number of carriers.
- (b) A rise in the electron temperature may increase the number of carriers.
- (c) A rise in the lattice temperature may increase the carrier mobility when this is governed by ionized impurity scattering as in heavily compensated material at temperatures suitable for impact ionization.
- (d) A rise in the electron temperature may increase the carrier mobility for similar conditions.

(e) An increased carrier density may reduce the mobility since it increases the degree of ionization.

(f) The carrier mobility may increase as a function of carrier density, electric field or current density for special reasons.

(g) The ionization rate may increase as a function of carrier density, electric field or current density for special reasons.

These features are favourable to breakdown with the exception of (e). As the compensation ratio is increased, the relative proportion of the donors which may be ionized in the breakdown is made smaller, so the unfavourable effect of (e) is reduced.

In the analysis made here, the mobility and the ionization rate coefficient A_I/B are taken to vary with the carrier density, the electric field and the current density. Three particular combinations are considered.

(1) Normal ionization.

A_I/B single-valued in E

A_I/B constant in n

μ single-valued in n

μ single-valued in E .

(2) A carrier density dependant ionization rate.

A_I/B single-valued in E

A_I/B single-valued in n

μ constant in E

μ constant in n .

(3) A current density dependent mobility.

A_I/B constant in E

A_I/B constant in n

μ single-valued in J

μ constant in n .

It is intended to show that no matter what single-valued functions are chosen, negative resistance is not possible in (1), but that suitable functions can provide a negative resistance for (2), (3). Since constancy is a special case of a single-valued function, it is clear that (1), (2) include all cases (except the uninteresting one A_I/B constant in E , single-valued in n), which may be examined with the criterion $\partial E/\partial \sigma \leq 0$. (3) is the simplest case incorporating a current density dependence and is to be examined with the criterion $1 - J/\sigma \cdot \partial \sigma/\partial J \leq 0$. We now investigate the three cases.

5.3.5.1 Case 1. Normal ionization.

Explanations given for the negative resistance by Rediker and McWhorter (l.c. 1960, 516) and by J. Yamashita (1961, 517) fall into this category. Rediker and McWhorter have suggested that neighbouring donor pairs act together to form the configuration of a singly ionized hydrogen molecule; this is to act as a powerful scattering agent until its ionization takes place; as total ionisation is reached the carrier mobility rises strongly. Thus the mobility is a strong function of the carrier density which itself is a strong function

1. The first part of the paper is devoted to a general discussion of the problem of the origin of life. It is shown that the problem is one of the most important and most difficult in the history of science.

2. The second part of the paper is devoted to a detailed discussion of the problem of the origin of life. It is shown that the problem is one of the most important and most difficult in the history of science.

3. The third part of the paper is devoted to a detailed discussion of the problem of the origin of life. It is shown that the problem is one of the most important and most difficult in the history of science.

4. The fourth part of the paper is devoted to a detailed discussion of the problem of the origin of life. It is shown that the problem is one of the most important and most difficult in the history of science.

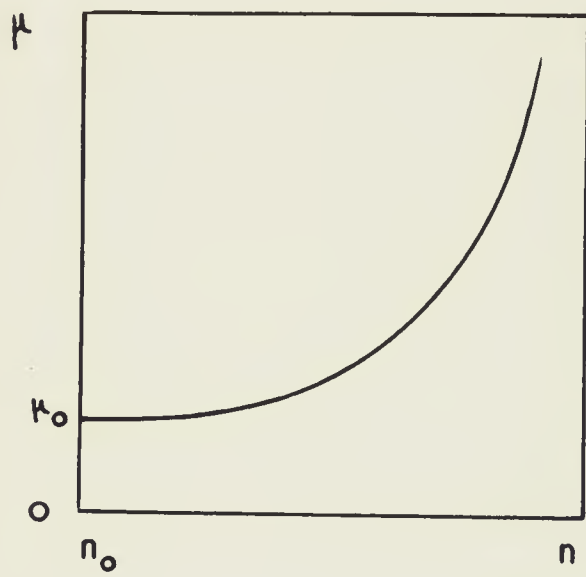
5. The fifth part of the paper is devoted to a detailed discussion of the problem of the origin of life. It is shown that the problem is one of the most important and most difficult in the history of science.

6. The sixth part of the paper is devoted to a detailed discussion of the problem of the origin of life. It is shown that the problem is one of the most important and most difficult in the history of science.

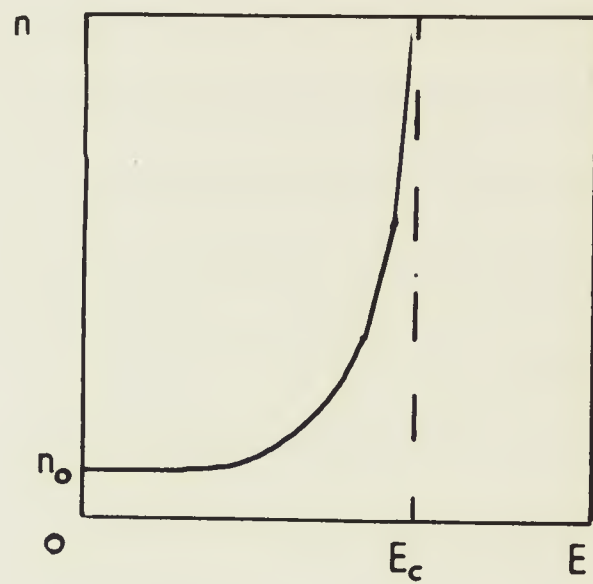
7. The seventh part of the paper is devoted to a detailed discussion of the problem of the origin of life. It is shown that the problem is one of the most important and most difficult in the history of science.

Figure 5.1

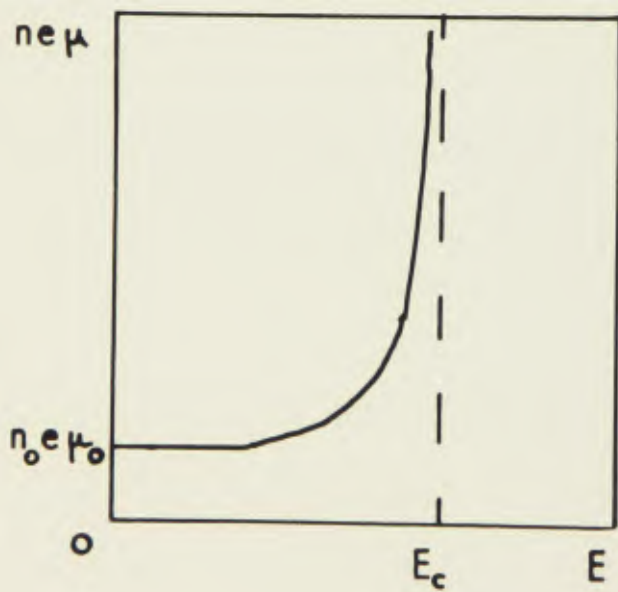
The Impossibility of Negative Resistance for Case 1



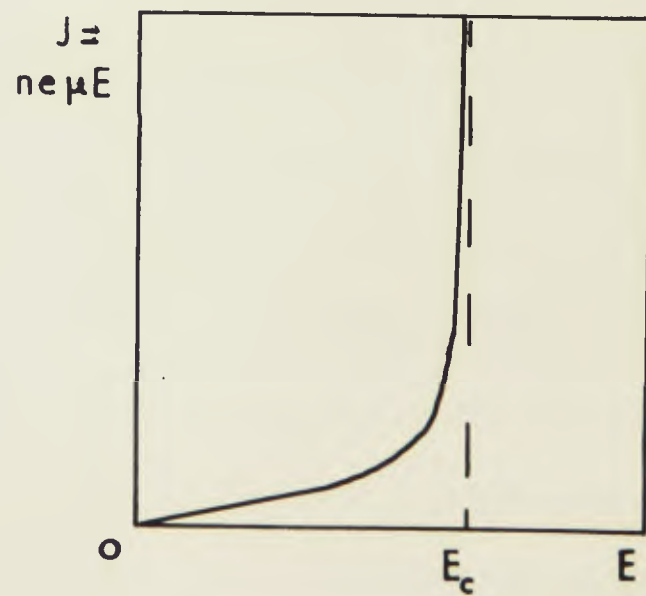
(a)



(b)



(c)



(d)

of the electric field. The model of Yamashita (l.c. 1961, 518) is similar in that a postulated interaction between conduction and impurity levels becomes smaller as the conduction electron density rises. Again the steep rise in mobility due to the increased carrier density is stated to lead to a negative resistance without any attempt to present computational evidence for this. T. Kurosawa has pointed out (1965, 519) that the explanation of Rediker and McWhorter (l.c. 1960, 520) is not in good agreement with experimental results on the variation of the breakdown intensity with the compensation ratio.

However we go further in claiming that for the physical processes built into explanations under case (1), negative resistance cannot in fact occur. A qualitative argument for this is given by figure 5.1. Single-valued curves for μ as a function of n , n as a function of E , the conductivity $ne\mu$ as a function of E and the current density $ne\mu E$ as a function of E are shown. It is clear that such single-valued curves cannot provide a characteristic which bends back on itself to give a current controlled negative resistance. The argument of Rediker and McWhorter (l.c. 1960, 521) is equivalent to trying to pull oneself up with one's own bootstrings.

This may be demonstrated more rigorously from the ionization rate equation. We put $A_I/B = KF(V)$ where V is the total potential difference applied to the crystal, assumed to be uniformly distributed and K is independent of V . We postulate that $F(V)$ is a single-valued function of V as required for case (1), so that plotting $F(V)$

is equivalent to plotting V on a distorted scale. This is consistent with the assumptions that the distribution function $f(v)$ is single-humped, that it becomes bodily displaced in the direction of the electric field as the field is raised and that the ionization cross-section is a regular function of the electron velocity. Distortion of the distribution function itself is to be considered in case 3. In this way it is determined that $dF(V)/dV$ is finite and positive for all V . The mobility is assumed to be a single-valued function of the carrier density for case (1). Thus:

$$\frac{d\sigma}{dn} = \frac{d}{dn} (ne\mu) = e\mu + ne \frac{d\mu}{dn} .$$

Hence $dn/d\sigma$ is finite and positive for all values of n .

The ionization rate equation for strongly compensated material was found in section 5.3.3 to be:

$$\frac{A_I}{B} \cdot n(n - N) + \frac{N}{N - n_0} \cdot (n - n_0) = 0 .$$

So:
$$F(V) = \frac{N}{K(N - n_0)} \cdot \frac{n - n_0}{n(N - n)} .$$

The negative resistance condition in the absence of current dependence is: $\frac{dV}{d\sigma} = 0$,

i.e.
$$\frac{dV}{d\sigma} = \left(\frac{dF(V)}{dV} \right)^{-1} \cdot \frac{dn}{d\sigma} \cdot \frac{N}{K(N - n_0)} \cdot \frac{d}{dn} \left\{ \frac{n - n_0}{n(N - n)} \right\} .$$

Applying the above conditions on $F(V)$, ,

$$n^2 - 2nn_0 + n_0N = 0 .$$

$$n = n_0 \pm \sqrt{n_0(n_0 - N)} .$$

But N is larger than n_0 in all circumstances, so no real turning points exist. Hence negative resistance is not possible under the circumstances of case (1). We proceed to examine case (2).

5.3.5.2 Case 2. A carrier density dependent ionization rate.

The conditions of case (2) are that the ionization rate coefficient should be single-valued in both V and n and that the mobility should be constant. It is not easy to show qualitatively that there is a difference in principle between a carrier density dependent mobility and a carrier density dependent ionization rate, so we pass directly to a mathematical formulation. The ionization rate coefficient is taken to vary as:

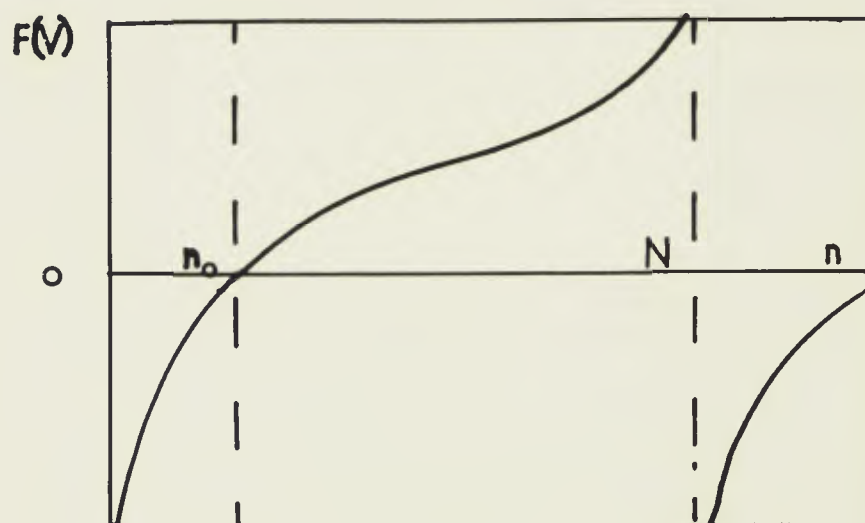
$$\frac{A}{B} = kn^{p-1}F(V) \quad .$$

p is chosen for convenience such that $p=1$ gives the carrier density independent situation. A case for such a variation may be made as follows; the donor site is known to have excited states which are of especial significance in recombination when they act as giant traps (M. Lax, 1959, 522); these excited states may be of significance to breakdown in uncompensated material due to Auger recombination processes, as described by Koenig, Brown, Schillinger (l.c. 1962, 523); such excited states are shown to have short lifetimes (10^{-6} seconds) but it remains true that the proportion of unionized donors excited at any time rises as the free carrier population rises. Ionization

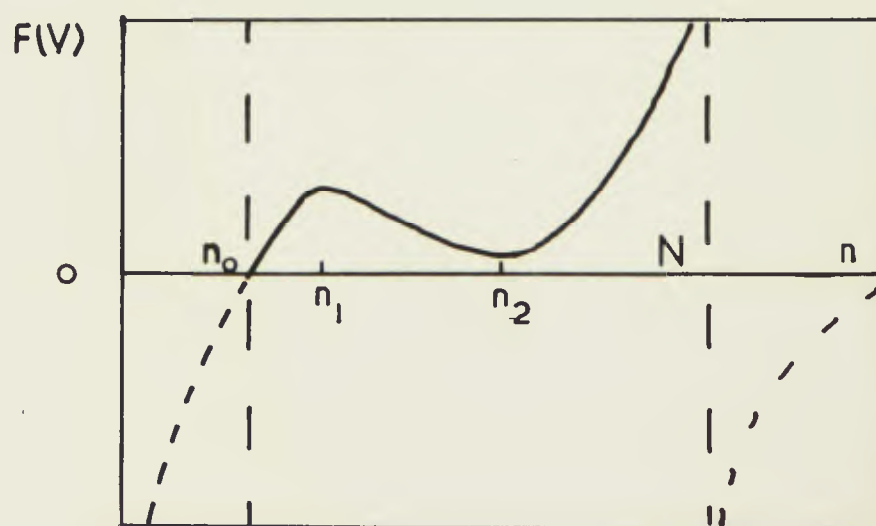


Figure 5.2

Curve forms for Case 2



(a)



(b)

of an excited donor is clearly possible for a larger proportion of the free carriers. If this process is significant it clearly gives rise to a carrier density dependent ionization cross-section. The particular form of expression chosen has calculational advantages over the more physically meaningful expression $k'(1+n)F(V)$.

We thus aim to investigate the behaviour of the equation:

$$kn^p(n-N)F(V) + \frac{N}{N-n_0}(n-n_0) = 0.$$

Rewriting:
$$F(V) = \frac{N}{k(N-n_0)} \frac{n-n_0}{n^p(N-n)},$$

we note singularities at $n=N$, $n=n_0$, $n=0$, the general form being given in figure 5.2. The only physically meaningful region is that between $n=n_0$ and $n=N$ in which impact ionisation takes place. For a negative resistance, $dV/d\tau$ is to pass through zero. Since we have assumed that $F(V)$ is a monotonic function of V , and that the mobility is constant, this is equivalent to the criterion $dF(V)/dn = 0$, as shown in figure 5.2b. The equation is to show two roots with $n_0 < n_1, n_2 < N$. Performing the differentiation:

$$\begin{aligned} \frac{d}{dn} F(V) &= \frac{N}{k(N-n_0)} \frac{1}{n^{2p}(n-N)^2} - pn^{p+1} + (Np+n_0p-N-n_0)n^p + \\ &\quad + n^{p-1}n_0Np = 0. \end{aligned}$$

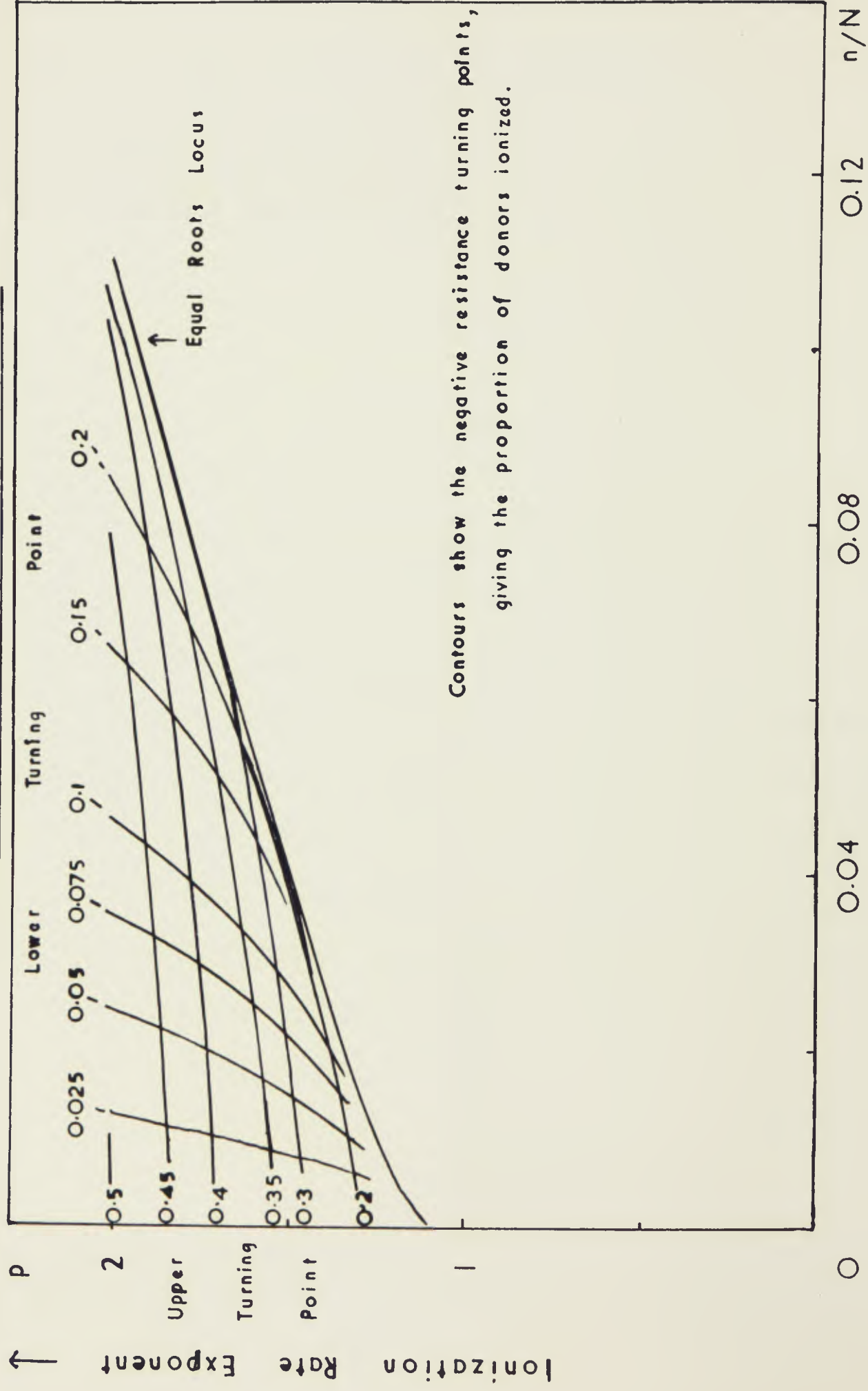
If $p \neq 0$, $n \neq N$, $n \neq n_0$,

$$n^2 - n\left\{N + n_0 - \frac{N-n_0}{p}\right\} + n_0N = 0.$$

$$n = \frac{N}{2} \left\{ 1 + \frac{n_0}{N} - \frac{1}{p} \left(1 - \frac{n_0}{N} \right) \pm \sqrt{\left[\left(1 - \frac{1}{p} \right) + \frac{n_0}{N} \left(1 + \frac{1}{p} \right) \right]^2 - 4 \frac{n_0}{N}} \right\}. \quad (5.1)$$

Figure 5.3 Contour Plot Showing Negative Resistance Region in Terms of the

Ionization Rate Coefficient Variation.



Two real roots exist if:

$$\left\{ \left(1 - \frac{1}{p}\right) + \frac{n_0}{N} \left(1 + \frac{1}{p}\right) \right\}^2 - 4 \frac{n_0}{N} > 0 .$$

This may be simplified to:

$$\frac{n_0}{N} < \left\{ \frac{p-1}{p+1} \right\}^2 .$$

For a negative resistance to be demonstrated, we require two real turning points inside the range $n_0 < n < N$. The negative resistance regions may be examined with reference to a contour plot as shown in figure 5.3. Here p and n_0/N constitute the axes and the marked area is the region where there exist two real turning points with $n_0 < n < N$. Figure 5.3 also shows contour plots for the two roots.

We are now in a position to interpret a negative resistance in terms of the ionization rate equation coefficient A_I/B of section 5.3.3. Orthodox treatment puts $A_I = F(V)$, $B = \text{constant}$, giving $p=1$. Reference to the diagram shows that negative resistance is not shown at $p=1$ for finite n_0 . However for a given initial ionization n_0/N , a negative resistance is shown if p is large enough, that is if the ionisation rate is strongly dependent enough on the carrier density. As p rises, so does the magnitude of the negative resistance region.

For a more rigorous analysis of partially compensated material, we may take:

$$\frac{A_I}{B} = k' \frac{1 + \alpha n}{1 + \beta n} F(V) ,$$

and carry out an analysis in terms of α, β instead of p . This does not

lend itself to qualitative analysis, as presented for p. However the diagram shows that a current controlled negative resistance is produced if the carrier density dependence of the ionization rate coefficient is strong enough. It is apparent that a favourable mobility variation at the same time can enhance the negative resistance, but ionization rate variation is still essential. Thus the conditions of case (2) are sufficient for the occurrence of negative resistance, while those of case (1) are not.

5.3.5.3 Case 3. A current density dependent mobility.

In the conditions of case (3), the ionization rate coefficients are specified to be independent of electric field and carrier density. This means that the carrier density is specified as constant. It would hence appear that case (3) has no relevance to impact ionization; what we mean in case (3) is that impact ionization is not really relevant to the negative resistance. Just as in case (2) a mobility variation was found to enhance a negative resistance, without being sufficient of itself to cause one without an ionization rate variation, so we examine the case where a carrier density increase contributes towards, but does not basically cause, a similar negative resistance. Since we consider the basic cause in case (3), the carrier density is taken to be constant. The extra effect of a density increase will be considered following this.

With a fixed number of carriers, a current controlled negative

resistance requires an increase in the carrier mobility. We may now apply the negative resistance criterion, $\partial V / \partial \mu \leq 0$ cannot be satisfied, since this would require that the mobility should go through the value infinity at some electric field. The criterion to be satisfied is thus:

$$1 - \frac{J}{\mu} \frac{\partial \mu}{\partial J} \leq 0 \quad .$$

A specific dependence of the conductivity on the current density seems bound to be rooted in the energy balance equation. In this way either the electrons become hot or the whole lattice does so. In this section we consider the hot electron situation.

If an electric field is applied to a crystal, the equilibrium state is determined by the energy and momentum balance equations. When the mobility is governed by the same scattering processes as the energy balance, as in the case of acoustic phonon scattering, the carrier velocity distribution function may be considered to be the equilibrium Boltzmann shape displaced proportionately to the electric field (J. M. Ziman 1960, 524, p.261). No carrier heating takes place and the mobility remains constant. When the mobility is governed by elastic processes such as carrier-carrier scattering, or ionized or neutral impurity scattering, the energy balance equation cannot be satisfied in this way. The velocity distribution function, besides being displaced as the electric field is raised, becomes distorted. If the carrier density is high enough for carrier-carrier scattering to be appreciable,

the distribution function may be represented by a Boltzmann shape at a temperature above the lattice temperature. This gives the temperature of the hot carrier system. We consider a system where a rise in input power causes a rise in carrier temperature which due to ionized impurity scattering or such results in a rise in the mobility. To determine whether this leads to a negative resistance, the energy balance and the mobility have to be examined as functions of the carrier heating.

The energy loss rate is given by an expression of the form:

$$P = EJ = \int_0^{\infty} \left\{ \left(\frac{dE}{dt} \right)_{ac} + \left(\frac{dE}{dt} \right)_{op} + \left(\frac{dE}{dt} \right)_{piez} + 0 \right\} f(E) dE .$$

where the terms represent acoustic mode, optical mode, piezoelectric and other energy loss rates. At the same time the current is determined by:

$$J = N \int_{-\infty}^{\infty} f(v) v dv ,$$

while the distribution function is determined by the generalized Boltzmann equation:

$$\left(\frac{\partial f}{\partial t} \right)_{scatter} = - \left(\frac{\partial f}{\partial t} \right)_{field} .$$

The application of approximate methods of tackling this problem has been reviewed by E. G. S. Paige (1964, 525). In view of the complexity of the problem and the difficulty found in dealing with the non-ohmic dependence, no attempt will be made here to look for specific criteria for negative resistance for authentic physical conditions.

Instead we recall that the aim in case (3) is to show that negative resistance is logically possible under the given conditions. Therefore we take the simple case of a hot carrier system with a well-defined temperature; the analysis is a more general version of that given by S. M. Kogan (1963, 526).

The temperature of the electron system is taken as T and that of the lattice as T_0 . The mobility is assumed to have a power law dependence on the carrier temperature; with the number of carriers specified as constant, this gives:

$$\sigma = aT^m.$$

Following Kogan (l.c. 1963, 527), we consider the energy transfer rate out of the electron system to be of the form:

$$P = bT^n(T - T_0).$$

But $\sigma = J/E$,

$$P = JE.$$

Hence
$$\frac{dE}{dJ} = \frac{1}{aT^m} \frac{(n - m + 1)T - (n - m)T_0}{(n + m + 1)T - (n + m)T_0}.$$

For a current controlled negative resistance the numerator is to go through zero so that:

$$T \geq \frac{m - n}{m - n - 1} T_0.$$

For ionized impurity scattering m may be taken as $3/2$; Kogan shows that for acoustic mode energy loss, n equals $1/2$, making negative resistance impossible, but for piezoelectric scattering n equals $-1/2$,

requiring $T \gg 2T_0$. Piezoelectric scattering is not relevant in germanium, but may be significant in indium antimonide. Nevertheless the argument shows that given suitable temperature dependences for the scattering processes, a negative resistance can be provided.

It is interesting to examine the effect of varying the compensation ratio for case (3). We assume that T_0 is chosen so that at this temperature ionized impurity scattering is predominant and m is $3/2$. For a negative resistance the electron temperature has to reach a certain critical value. The effect of compensation is to put up the temperature at which acoustic mode scattering becomes important, and hence to maintain m at $3/2$ to higher temperatures. Hence it may happen that for uncompensated material m drops below $3/2$ before the critical electron temperature for negative resistance is reached, but that for compensated material the critical temperature is attained. Thus in case (3) negative resistance is more probable in compensated material.

5.3.6 Variation in the Number of Carriers.

We now drop the limitation of a constant carrier density of case (3) and examine the effect of impact ionization on the negative resistance. In the case of hot carriers at a well-defined temperature it is reasonable to assume that the carrier density is a single-valued function of T and T_0 ; we take $n=h(T)$. Again from Kogan (l.c. 1963, 528), we have:

$$\sigma = aT^m h(T) \quad ,$$

$$P = bT^n h(T) \quad ,$$

$$\frac{dR}{dJ} = \frac{1}{\sigma} \cdot \frac{(n - m + 1)T - (n - m)T_0}{(m + n + 1)T - (m + n)T_0 + T(T - T_0) \frac{2}{h} \frac{dh}{dT}} \quad .$$

For a current controlled negative resistance, $dR/dJ < 0$. The possibility of dh/dT going through infinity has already been ruled out by the specification of a single-valued function, so the criterion

$$\text{is:} \quad T > \frac{m - n}{m - n - 1} T_0 \quad .$$

But this exactly repeats the condition for negative resistance obtained in the case of constant carrier density. It is thus apparent that the change in the number of carriers has no bearing on the attainment of a negative resistance for a hot carrier system.

It must be pointed out that in the above calculation, it has been assumed that the energy loss processes are independent of the carrier density. A contribution to a negative resistance can be made if an increase in carrier density reduces the average energy loss rate per carrier. Investigation of suitable mechanisms is beyond the scope of the present work.

5.3.7 Critique of an Analysis Provided by T. Kurosawa.

We are now in a position to examine the explanations of the negative resistance in impact ionization given in a recent paper by Kurosawa (l.c. 1965, 529). The following criticisms are put forward:

- (1) His method of deducing a negative resistance is fallacious.
- (2) His breakdown criterion is not useful.
- (3) His introduction of the optical phonon is unnecessary.

Yet in spite of these failings in the treatment, the paper is felt to be of considerable interest because several of the ideas for which an analysis has been attempted here, are implicitly included by Kurosawa.

Our most fundamental objection is to Kurosawa's method of deducing that a negative resistance occurs; it is not explicitly stated but appears to be this. He firstly determines a criterion for instability in the ionization process; secondly he calculates from this the critical field as a function of compensation; thirdly he calculates in a completely different way the sustaining field in a fully ionized specimen; because the sustaining field obtained is lower than the critical field, he reckons to have deduced that a negative resistance exists; so all that he then reckons to require is a mechanism which raises the critical field as the compensation is raised. We claim that in the absence of certain safeguards this technique is fallacious. Where the calculations for critical and sustaining fields are independent and accurate only to orders of magnitude, no reliance is to be placed on their relative values; an obvious deduction from Kurosawa's curves, which he does not draw, is that negative resistance is to be expected even in uncompensated specimens, which is experimentally incorrect. The necessary safeguard for this technique is a correspondence principle; the calculations should be positively related to each other in the low

compensation case through the application of the breakdown criterion. It is to avoid this sort of trouble that we have specified as the only satisfactory criterion for negative resistance $dS/dJ < 0$.

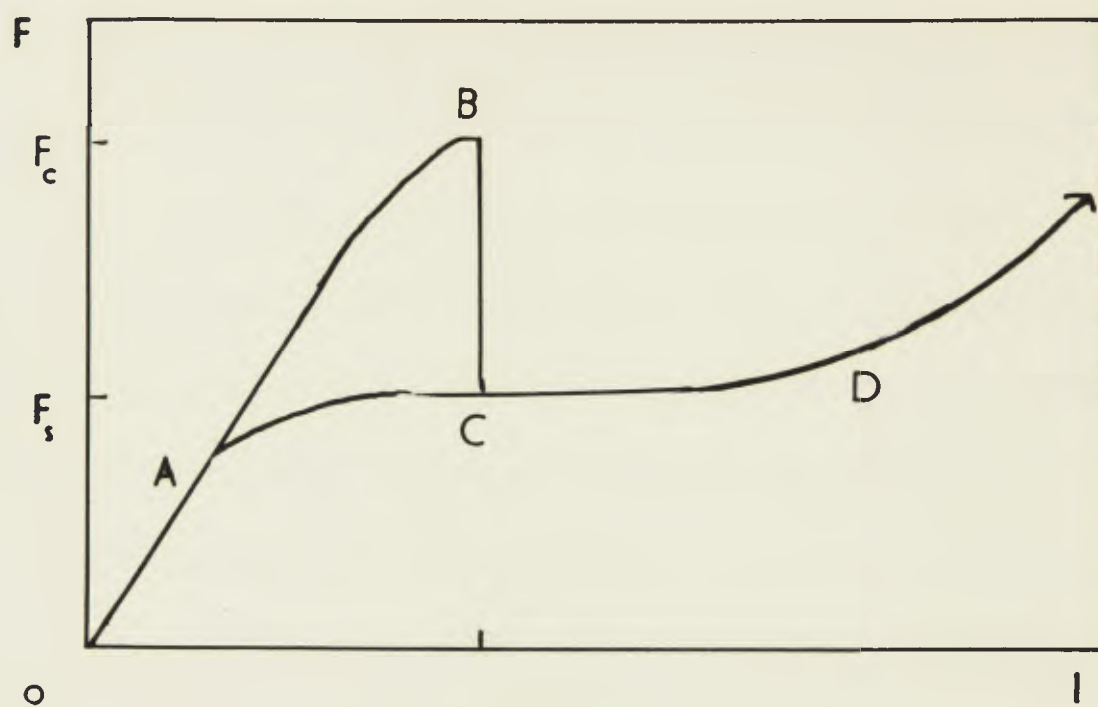
This leads to the second criticism of Kurosawa, namely that his breakdown criterion is not useful. He reckons both here and in an earlier paper (1962, 530), that the rate of ionization at low fields is lower than the rate of recombination, while beyond breakdown, the ionization rate is higher than the recombination rate; thus the breakdown criterion is when the two rates are equal (l.c. 1965, 531, p.1411, foot of column 1). This is ridiculous since the rates must adjust themselves to equality at every stable working point of the characteristic; this is the reasoning behind the ionization rate equation. In a negative resistance region it is possible to balance the rates at one field strength for more than one current; the negative resistance criterion is thus the existence of turning points, as has been employed in previous sections. Kurosawa should look for points where more than one solution to his equations is possible. The fact that with an extra degree of freedom, his results show consistent variations serves to demonstrate the dangers of computation since a neutral approach would give random results.

Our third objection is to Kurosawa's introduction of the optical phonon, firstly because it is unnecessary and secondly because its effect is unfavourable. We deduce that Kurosawa's interest in the optical phonon stems from a misconception as to how negative resistance

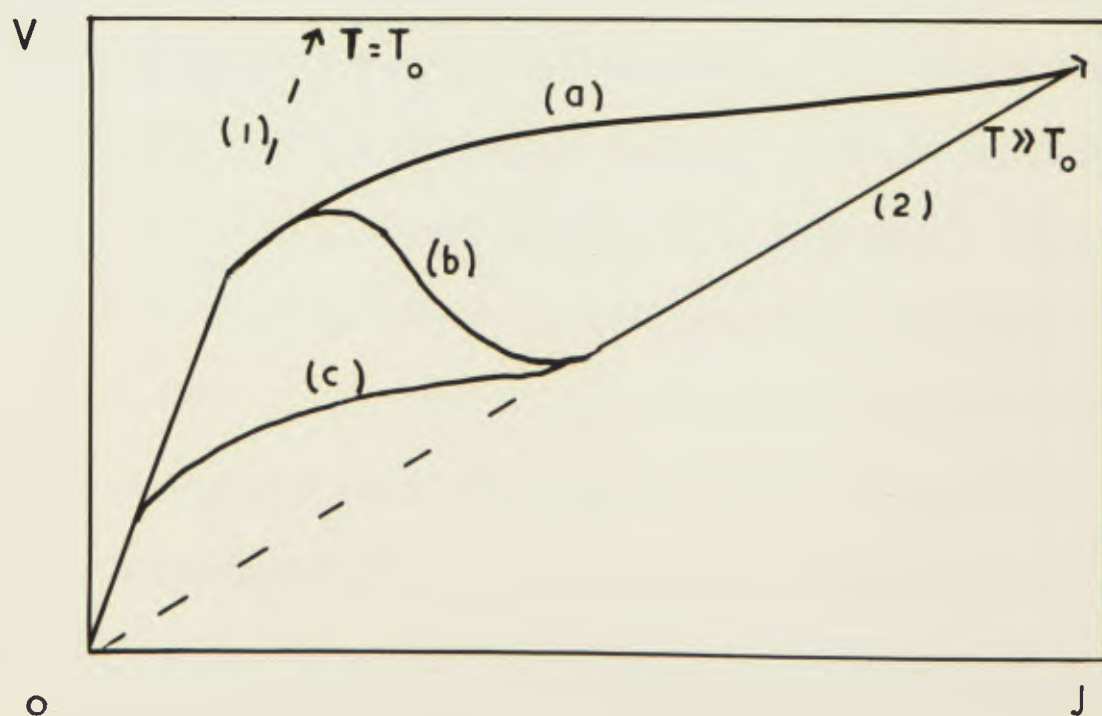


Figure 5.4

The Origin of Negative Resistance



(a)



(b)

arises. He considers it to arise in a spike (figure 5.4a); uncompensated material is to follow path OACD and equivalent compensated material OABCD. Hence it is necessary to find a mechanism which can "superheat" the system in compensated material only. Our interpretation of how negative resistance arises is shown in figure 5.4b. It represents the transfer from one characteristic to another, say from a low carrier density, ionized impurity scattering curve 1 to a high carrier density phonon scattering curve 2. The interesting feature is then whether the transfer is fast enough to provide a negative resistance as in (b) or not, (a), (c). The transfer itself takes place in all cases and not specifically in compensated material.

We may now examine the effect of introducing optical phonon scattering. The effect on the equilibrium temperature branch 1 which is dominated by ionized impurity scattering may be expected to be small, but the effect on the hot electron branch is bound to be considerable in the reduction of the mobility. This displaces curve 2 towards curve 1 and makes any negative resistance of smaller extent. Since it provides additional energy loss, the attainment of a higher electron temperature, necessary for transfer to take place, is made less likely. It is reasonable to deduce that the less optical phonon scattering there is, the more likely negative resistance is, and that the greater the degree of compensation (and hence the greater the separation of 1, 2), the more likely negative resistance is. It is not possible to deduce that there is a relationship between optical

phonon scattering and compensation in the way Kurosawa does. His calculation seems to be based on the inconsistency that optical phonon scattering is important for curve 1 but may be neglected for curve 2.

5.3.8 Hot Lattice Filaments.

So far all the features encouraging a current controlled negative resistance listed in section 5.3.4 have been examined except for those depending on a variation of the lattice temperature. The questions arising over heat transport at the specimen boundaries will be dealt with in chapter 7. Here we confine ourselves to the idea of a hot filament in an otherwise isothermal crystal.

Application of the principles of irreversible thermodynamics as developed in chapter 2 shows that a hot filament can exist provided that all forces and fluxes balance, under which circumstances the entropy production rate has a stationary value; this filament is stable if, at all neighbouring points, the entropy production rate is non-stationary. For the case of a hot filament, we ignore thermoelectric effects and consider the heat transport to be radial and the electric current flow to be axial.

Firstly we examine the radial temperature structure of a filament. An annular element at radius y is taken. Then the rate of production of Joule heat inside this radius is given by:

$$H(y) = \int_0^y 2\pi x \sigma(x, T) V^2 dx \quad .$$

With thermal conductivity k , the temperature drop across this element is:

$$\delta T = \frac{1}{2k(y,T)} \frac{H(y)}{y} \delta y \quad .$$

Thus the temperature at radius r is given in terms of the temperature at the centre of the filament by:

$$T(r) = T(0) - V^2 \int_0^r \left\{ \frac{k(y,T)}{y} \int_0^y \sigma(x,T) x dx \right\} dy \quad .$$

This is an implicit expression for the temperature as a function of the filament radius which is analytically intractable. We therefore choose to examine instead a simplified situation.

Unit length of a filament of radius y is considered at a uniform temperature ΔT above that of the main part of the lattice, surrounded by a narrow annulus of width x in which the temperature gradient is uniform. Then the Joule heating in the filament is $JV\pi y^2$ while that in the annulus may be neglected. The heat flow rate in the annulus is $2\pi yk\Delta T/x$. Since in equilibrium these are equal:

$$xy = \frac{2k\Delta T}{VJ} \quad .$$

It is fair to assume that under these circumstances virtually the whole current flows in the filament so that $yJ = I$. Numerical values are now put in to give the size of a viable filament to an order of magnitude. We take ΔT as 15°K to give complete ionization from 4°K , the thermal conductivity as $1 \text{ watt}/^\circ\text{K}/\text{cm.}$, which is a conservative estimate for germanium, indium antimonide and most semiconductors, (M. G. Holland, 1964, 532), the electric field as $3 \text{ volt}/\text{cm.}$ and a current of 1 amp. which is immense. This gives: $x = 10 \text{ cm.}$

To satisfy the original condition $y \gg x$, this requires a filament of radius 100 cm. in material of resistivity after breakdown $3 \cdot 10^4$ ohm cm. It is clear that the formation of a filament of hot lattice is out of the question because of the size of the thermal conductivity. Where heat transfer difficulties arise, it is more likely to be at the specimen boundaries, as chapter 7 will show.

This conclusion is in direct contradiction to that of S. H. Koenig and R. D. Brown (1959, 533); they have concluded that a hot filament gives a satisfactory explanation to their experimental results on germanium. The difference arises because they consider not the heat transport but the decay time constant; because this is found to be longer than the pulse rise time, they claim that a filament will form. An inconsistency is involved because while they find that the heating must be concentrated inside 10^{-2} of the cross-section, they calculate the time constant for the whole crystal. Because of the specific heat, filament formation is less likely under pulsed conditions, and the calculation is more difficult; thus examination of d.c. conditions is a more reliable approach. For these reasons we feel justified in not accepting the conclusions of Koenig and Brown.

5.3.9 Filament Flow in an Isothermal Lattice.

The three processes which can provide a current controlled single carrier negative resistance in a semiconductor crystal, obtained from the previous sections, may be summarized as:

- (1) A rise in the ionization cross-section as the carrier density increases.
- (2) A faster rate of rise in the mobility than in the average carrier energy loss rate.
- (3) A fall in the carrier average energy loss rate as the carrier density increases.

We consider the second process first.

In the second process the occurrence of non-linearity and possibly of negative resistance is due entirely to the properties of the carriers themselves; it is the change in scattering due to a change in the distribution function which effectively alters the mobility; the lattice itself is unaffected. Thus it is clear that there is no reason for special discrimination of any sort; the electrons carry their non-linearity with them as they diffuse through the crystal. The instability against filament flow is in electron velocity space and not in real lattice space; (see section 2.11). Furthermore in the absence of processes (1), (3) of this section, marginal lattice effects are non-cooperative; the main effect is that the mobility is lower where a higher degree of ionization exists due to higher scatter by an ionized than by a neutral impurity.

On the other hand for processes (1), (3) the cause of the non-linearity or negative resistance is localized in lattice space; a locally higher carrier density is associated with a locally higher degree of impurity ionization through Poisson's law. Hence lattice

filaments can exist. Whether or not they do may be determined from the principles of irreversible thermodynamics in exactly the way employed in the treatment of thermal lattice filaments. In fact in this case it was shown that diffusive processes are so strong that a thermal filament cannot be stable under normal crystal conditions. The stability of a carrier density dependent filament depends on the conditions in the filament walls, but basically the stronger the density dependence is, the more likely a negative resistance is and the more stable a filament will tend to be. In practice it seems fairly certain that a degree of current non-uniformity will set in under most conditions long before a negative resistance is manifested.

5.3.10 Application to Semiconductor Physics.

The aim of this chapter has been to find under what circumstances a current controlled negative resistance can or cannot take place. The analysis has been concerned with one physical process at a time with in most cases considerable oversimplification. The work done up to the present is thus a necessary preliminary survey for one of two purposes; the search for a new negative resistance, or the explanation of one which has been discovered. In view of the effort already put into both of these, it is surprising that no such survey seems to have appeared. This survey was undertaken for the purpose of explaining a newly discovered impact ionization negative resistance. It has subsequently been found that the phenomenon is not relevant to this chapter;

it will be discussed in chapter 7. Since this thesis is concerned with the whole subject of bulk negative resistance, the analysis of this chapter forms an integral part of the work; we proceed to examine briefly both the experimental results already available and the way ahead to the discovery of new examples of greater potential usefulness.

Experimental results already discovered will not be treated in any great detail. This is because such would require a detailed treatment of each scattering process, an estimation of their relative strengths and variations for a given material, and an elaborate computation. It is doubtful whether the results so far available justify this and it is well outside the scope of this work. We may however assign the phenomena observed into the categories of the above sections.

Results on germanium are considered first. A summary of experimental properties of the negative resistance with references for germanium is given by Kurosawa (l.c. 1965, 534). The three main features are:

- (1) It is observed only in p-type material.
- (2) It is observed only in compensated specimens.
- (3) The current flow after breakdown is filamentary.

The explanations of Rediker and McWhorter (l.c. 1960, 535) and of Yamashita (l.c. 1961, 536) were rejected in section 5.3.5.1 because they cannot lead to a negative resistance. It was accepted that the hot electron effect, which Kurosawa (l.c. 1965, 537) incorporates, can

give a negative resistance, but such a negative resistance is not localized in space. Since localized filaments do occur in germanium an explanation in terms of a local density dependent phenomenon is required. Localized thermal heating of the lattice was ruled out for normal sized specimens in section 5.3.8 because of the high thermal conductivity of the material. This leaves processes (1) and (3) of section 5.3.9, a carrier density dependent ionization cross-section or a carrier density dependent carrier energy loss rate. Neither of these processes seems to have been treated in the literature.

Secondly we consider experimental results on indium antimonide. A negative resistance has been reported by Phelan and Love (l.c. 1964, 538) connected with the occurrence of impact ionization in compensated material. Further details of this have been obtained by M. A. Kinch (1964, 539). It would appear from the Hall voltage readings of Kinch that this is a non-localized effect which would support an interpretation in terms of electron heating, along the lines considered by Kogan (l.c. 1963, 540). The contribution of compensation to this model, as explained in section 5.3.5.3 is in accordance with experiment.

5.3.11 Device Potential in Impact Ionization.

It would appear that a current controlled negative resistance has been largely written off as a device basis in favour of the voltage controlled variety. In practice the method of employment

of a stable negative resistance is relatively independent of its type; the source of difficulty is the provision of stability, against either filament or domain formation. It will be shown in the next chapter that domain formation is a rather more fundamental physical obstacle to stable operation than filament formation. Here we examine methods of stabilizing a current controlled negative resistance.

It has been noted that two types of current controlled negative resistance may be obtained, a localized one where the occurrence is due to the properties of the lattice and a non-localized one where the occurrence is due to the properties of the carriers themselves. Hence there is no filament formation with hot carriers. It would however be incorrect to assume that hysteresis and stability are associated directly with the presence and absence of filaments. The problem of hysteresis was considered in chapter 2; the important feature is the exact contour shape of the entropy production rate in the filament formation region. It was pointed out in 2.17 that it is the inhibition of the initial formation of the filament which causes the hysteresis. Where this is absent stable operation in the negative resistance region can be combined with filament growth and decay. There are no hysteresis problems of this kind with a non-localized negative resistance. A method of avoiding filament formation for a localized process is to produce a specimen of area of cross-section smaller than that required for filament initiation. Experimental results on very small cross-sectional area specimens

of epitaxially grown germanium films would be of interest. At the same time continued investigation of the scattering processes in indium antimonide and other semiconductors could provide a material with favourable hot carrier properties.

Three limiting features to the usefulness of impact ionization processes stand out. The employment of a single carrier process involves, for most semiconductor materials, the use of cooled specimens; the experimentally observed requirement of compensation shows the importance of ionized impurity scattering to the attainment of negative resistance, which involves the use of liquid helium temperatures; and the occurrence of recombination and trapping automatically limits the speed of response. There is thus much to be said for the development of a current controlled negative resistance in which a carrier density change does not take place, which would thus be limited only by the energy relaxation time. Carrier heating is one such possibility; the discovery of a process truly inverse to that employed in the Gunn effect is another (see section 6.4.1 for a discussion of the parallels). Such a system should be subject to neither filament nor domain instabilities.

Chapter 6

Voltage Controlled Negative Resistance.

6.1 Introduction.

Over the past few years interest in voltage controlled bulk negative resistance properties of semiconductors has grown considerably. It seems to have been generated mainly by a widely quoted paper of B. K. Ridley and T. B. Watkins (1961, 601) giving an analytical discussion of the possibility of employing band structure properties to give negative resistance, together with the almost simultaneous appearance of a basically similar paper by C. Hilsum (1962, 602). These papers, predicting negative resistance, domain formation and spontaneous oscillation, foreshadow in a remarkable way subsequent experimental developments. Bulk effects have been observed in germanium doped with gold (B. K. Ridley, R. G. Pratt 1965, 603), gallium arsenide --- the Gunn effect (J. B. Gunn 1964, 604; 1964, 605, p.199), gallium arsenide --- low frequency (A. Barraud 1963, 606), indium phosphide (Gunn l.c. 1964, 607), cadmium telluride (A. G. Foyt, A. L. McWhorter 1966, 608), cadmium sulphide (E. W. Boer, W. K. Wilhelm 1964, 609), and zinc sulphide (C. Fischler-Hasani, F. R. Williams 1965, 610).

There are two distinct but interrelated facets to an understanding of the experimental evidence. Firstly the negative resistance is to be explained in terms of the internal processes in the

crystal, such as band-structure, trapping and scattering. Secondly the mechanisms of domain formation, growth and movement require investigation. This double nature gives the phenomena their complexity. In this chapter we examine firstly some of the theoretical possibilities for the internal processes, and then the subject of domain formation. Experimentally an attempt has been made to devise methods of investigation such as the Hall effect which throw light on the internal processes, since basic measurements relate only to domain mechanics and prove equivocal at the fundamental level.

6.2 Crystal Features Leading to Voltage Controlled Negative Resistance.

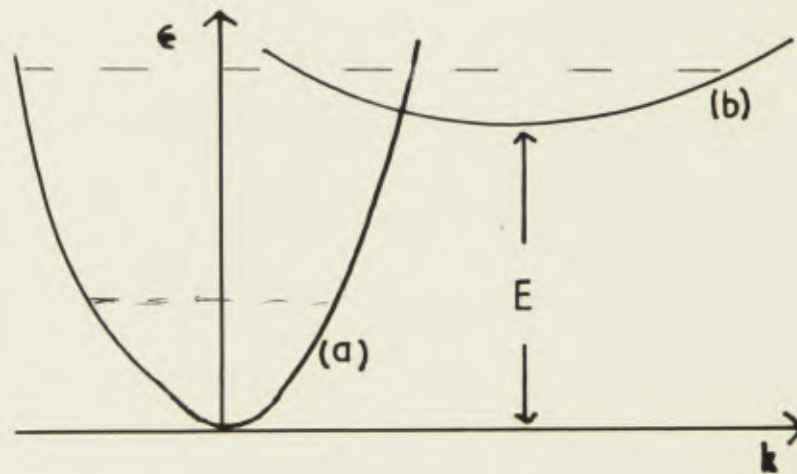
It seems convenient to classify processes for bulk differential voltage controlled negative resistance under three headings. These are basically the inverse processes to those giving the three cases for current controlled negative resistance in section 5.3.4 of the previous chapter, but due to the need for differences in emphasis, no attempt will be made to match the treatment. Parallels between the two types will be drawn in the concluding sections of this chapter. The three headings are:

- (1) Band-structure effects.
- (2) Effects due to an energy dependent capture cross-section.
- (3) Distribution function effects.

Figure 6.1

Sub-bands in a Semiconductor.

(From B.K.Ridley T.B.Watkins 611, 1961).



6.2.1 Band-Structure Effects.

This is a feature of prime importance; we summarize the discussion of Ridley and Watkins (l.c. 1961, 611).

The electrical conduction properties of a semiconductor are basically determined by the band structure; the conduction and valence bands consist of a number of sub-bands and for normal conduction only the sub-bands with the lowest carrier energy are effective. Energy gaps between sub-bands may exist, or a normal degeneracy may be split by the application of electric or magnetic fields or by strain. The effective mass of a carrier varies according to which sub-band it is in, due to varying band curvature. Hence the mean carrier mobility depends on the particular sub-band occupation figures. Application of an electric field produces carriers of higher energy; if these transfer to a lower mobility sub-band at a sufficiently high rate, the ensuing current may be lower for a higher applied field. Thus a differential negative resistance becomes possible. Diagrammatically the band structure is to be of the form of figure 6.1. Here (a) refers to some occupation probability $\ll \frac{1}{2}$ at an electron temperature T_{e1} and (b) to the probability at T_{e2} , $T_{e1} < T_{e2}$, defining these temperatures fairly vaguely.

After giving the above qualitative discussion, Ridley and Watkins (l.c. 1961, 612) have proceeded to make a quantitative analysis. Using the simplest possible distribution functions and scattering formulae, the conclusion drawn was moderately favourable. However

this calculation has been overshadowed by the discovery of microwave oscillations in gallium arsenide by Gunn (l.c. 1964, 613) and the explanation of the effect in terms of electron transfer between subbands. Experimental features have been accumulated rapidly and are summarized by Gunn (l.c. 1964, 614). Evidence for the explanation is summarized by P. N. Butcher and W. Fawcett (1965, 615), who give a set of calculations with specific reference to gallium arsenide and a detailed list of references. It is not our intention to investigate band structure possibilities here.

6.2.2 Effects due to an Energy Dependent Capture Cross-Section.

For the first and third headings of features which can cause a voltage controlled bulk differential negative resistance, we consider modulation of the mobility with the number of carriers held constant. In this section we consider a fixed mobility coupled to a reduction in the carrier density. This has been discussed by B. K. Ridley and T. B. Watkins (1961, 616; 1961, 617) and by B. K. Ridley and R. W. Pratt (1963, 618; 1965, 619; 1965, 620) in some detail, so that we shall do no more than refer to these papers. A coulomb barrier may exist at a partially compensated deep impurity site as described in section 1.3 so that only an energetic electron can penetrate it. Hence the relative proportions of free and captured electrons may change as the electric field is raised. If the capture rate is strongly dependent enough on the carrier velocity, the electron

density may fall sufficiently sharply as the electric field rises to give a negative resistance. An effect which is interpreted along these lines has been observed by B. K. Ridley and R. W. Pratt (1963, 621; l.c. 1965, 622) in gold-doped germanium. Again no analysis will be made here.

6.2.3 Distribution Function Effects.

Thirdly it may be possible to obtain a negative resistance through a change in the effective carrier mobility due to differential scattering. This situation is considered in the literature, especially with reference to plasma physics, but with a surprising lack of clarity as to what negative resistance criteria mean, as will become apparent. We intend to examine to what sort of extent scattering processes can or cannot lead to a negative resistance.

6.2.3.1 The Negative Resistance Criterion.

The negative resistance criterion of section 1.2 is given by:

$$\frac{dJ}{dE} = \frac{\sigma + E \frac{d\sigma}{dE}}{1 - \frac{J}{\sigma} \frac{d\sigma}{dJ}} .$$

In this chapter we are interested in a voltage controlled negative resistance, requiring (section 1.2) either:

$$(a) \quad \sigma + E \frac{d\sigma}{dE} \leq 0 ,$$

$$\text{or: } (b) \quad \frac{dJ}{d\sigma} \leq 0 .$$

In this section, our concern is for field dependent phenomena, so the criterion is (a). We require the current to fall over a range of increasing voltage; interest is thus confined to scattering processes which reduce the mobility as the electric field is raised.

6.2.3.2 The Effects of Electric Field and of Scattering.

The effect of an electric field on a carrier in a crystal of simple band structure is to accelerate it in the direction of the field. The energy and momentum thus gained are lost to the lattice through scattering. The distribution function represents a dynamic equilibrium and is given by Boltzmann's equation:

$$\left(\frac{\partial f}{\partial t} \right)_{\text{field}} + \left(\frac{\partial f}{\partial t} \right)_{\text{scatter}} = 0 \quad .$$

In the absence of an electric field it is often convenient to describe the carrier distribution function as approximating to a Maxwell-Boltzmann normal distribution function. Further, the effect of an electric field may be considered as a displacement of such a velocity function bodily in the direction of the electric field, (J. M. Ziman 1960, 623, p.261). When scattering anomalies take place, severe distortion of the distribution function may be expected.

A knowledge of the distribution function at a particular applied voltage gives the current flowing. We have:

$$\int f(\mathbf{v}) d\mathbf{v} = N \quad ,$$

$$\int f(\mathbf{v}) \mathbf{v} d\mathbf{v} = J/e \quad .$$

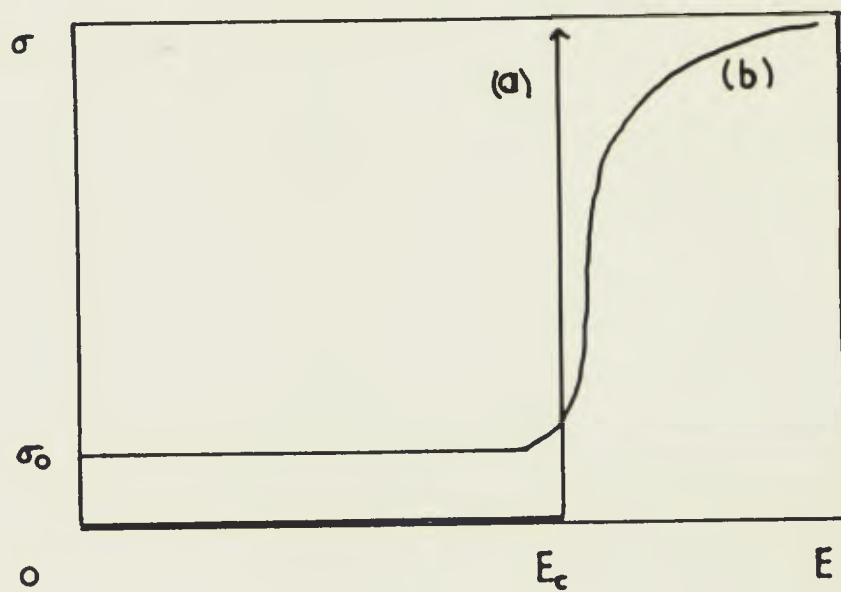
In the absence of an electric field, the distribution of velocities is symmetric about the zero point, so that the current is zero. The important feature is the development of the distribution function as the electric field is raised.

It is important to distinguish between the effects of elastic and inelastic collisions on the distribution function. Elastic scattering presents a paradox; if all collisions are elastic, there is no energy loss and hence there should be no resistance; since as many carriers are scattered into forward moving states as out of them, the scattering seems ineffective. The solution may be seen as follows: consider a crystal with an electric field applied, in which both elastic and inelastic scattering processes occur; neglect all elastic scattering first, and work out the distribution function for the carrier velocities for the inelastic scattering; this will give a displacement of the peak in the field direction which on integration yields a forward current; now switch in the elastic scattering; since elastic scattering affects all carriers and more carriers are moving forwards, elastic scattering reduces this forward bias so that the Boltzmann equation is satisfied. This is the physical basis to the use of energy and momentum balance equations as described in section 5.3.5.3. It seems that a change in the energy balance can have a considerable effect on the distribution even when the momentum balance is mainly determined by elastic processes.

We now postulate a strongly varying scattering probability and

Figure 6.2

Scattering Functions



consider its effect on the distribution function for different conditions.

6.2.3.3 Dominant Inelastic Scattering with Threshold.

The type of scattering function to be considered is shown in figure 6.2. When we take a situation to its logical extreme, as we do in this section, it is to clarify issues for subsequent sections; thus we take the scattering threshold to be quite sharp (curve (a) rather than (b)). The vital questions are then seen to be: what happens to the electrons which are scattered at the threshold and hence 'fall off the end' of the distribution; and still more important for negative resistance, what happens to the distribution function as it moves up with the electric field?

Here we consider the case where all scattering processes can be ignored except one inelastic process with a sharp threshold at carrier energy E_c which removes all energy and momentum. It is immediately clear that the carriers which are removed at energy E_c are fed back into the system at energy zero. Thus the energy distribution function matches at energies zero and E_c and has a rectangular shape. Since thermal velocities and extraneous scattering are being ignored, the velocity distribution is similar. Hence the current has a fixed value corresponding to the moment of the area; raising the applied voltage has no effect except to cause the merry-go-round to revolve faster. The differential electrical resistance

THE JOURNAL OF THE
ROYAL ANTHROPOLOGICAL INSTITUTE
OF GREAT BRITAIN AND IRELAND
VOLUME 14, PART 1, 1904
PUBLISHED BY THE INSTITUTE
11, BEDFORD SQUARE, LONDON, W.C.1



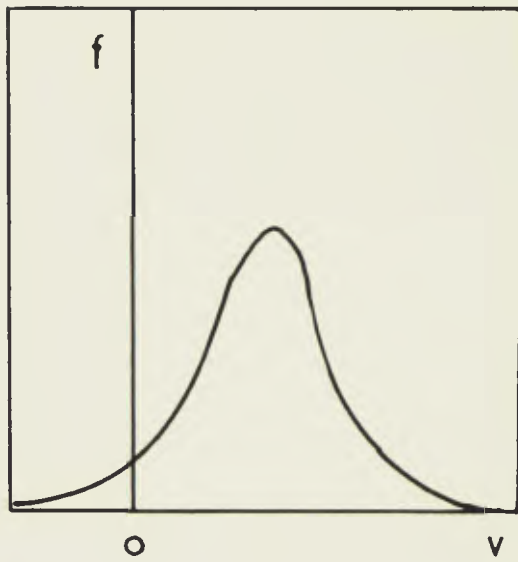
THE JOURNAL OF THE
ROYAL ANTHROPOLOGICAL INSTITUTE
OF GREAT BRITAIN AND IRELAND
VOLUME 14, PART 2, 1904
PUBLISHED BY THE INSTITUTE
11, BEDFORD SQUARE, LONDON, W.C.1



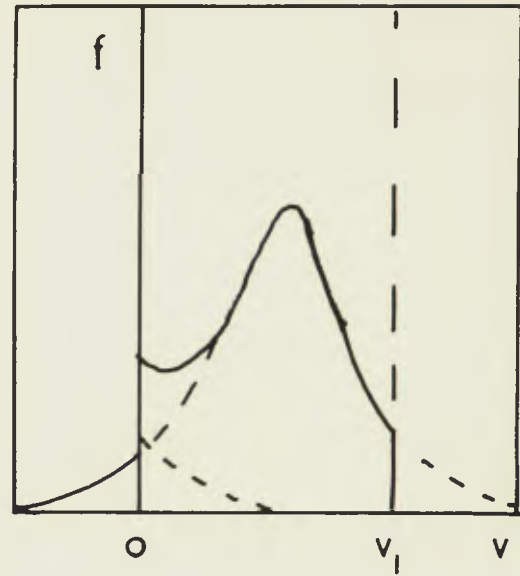
THE JOURNAL OF THE
ROYAL ANTHROPOLOGICAL INSTITUTE
OF GREAT BRITAIN AND IRELAND
VOLUME 14, PART 3, 1904
PUBLISHED BY THE INSTITUTE
11, BEDFORD SQUARE, LONDON, W.C.1

Figure 6.3

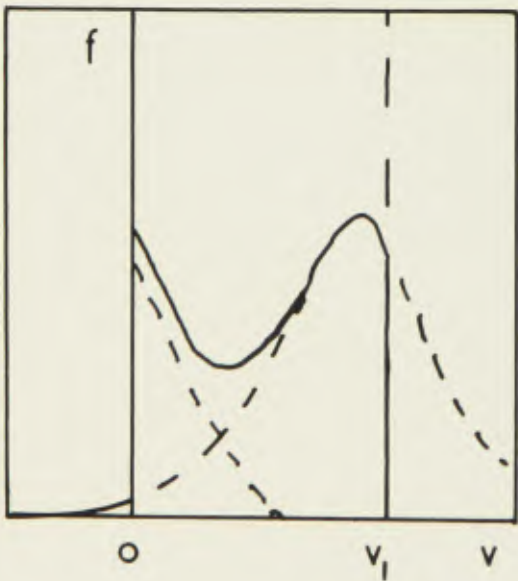
Distribution Function Distortion



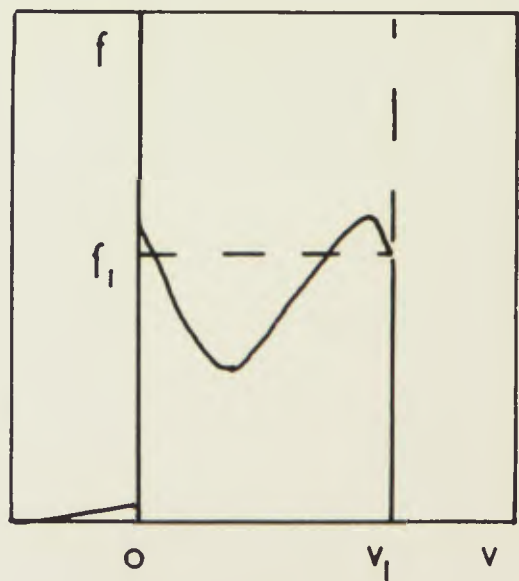
(a)



(b)



(c)



(d)

in this case is neither positive nor negative but infinite.

6.2.3.4 Non-Dominant Inelastic Scattering with Threshold.

We now consider a crystal where an inelastic process of the same kind as in the last section occurs, but where ordinary scattering determines the shape of the distribution function. There is no longer any need for exact correspondence at energies 0 and E_c because of the variation of other scattering processes with carrier energy. The situation is made still more complex by the need to transfer to a velocity picture for analysis of the current flow. Illustrative diagrams are shown in figure 6.3; (a) shows an orthodox distribution function; (b) shows a distribution function after the consideration of the scattering threshold; (c) shows the effect of increasing the electric field marginally. It is apparent that a negative resistance is possible if the advancement of the function minus the contribution from the leading element causes the moment of the total area to fall.

A simple calculation may be made as follows, the difference in current being taken for a marginal advance δv :

$$\begin{aligned}
 1/e \cdot \delta J &= \int_{-v_1}^{v_1} f(v - \delta v) v dv - \int_{-v_1}^{v_1} f(v) v dv \quad ; \\
 &= - \int_{-v_1}^{v_1} \frac{\delta f}{\delta v} \delta v v dv \quad . \\
 \frac{1}{e} \frac{dJ}{dv} &= - \int_{-v_1}^{v_1} \frac{df}{dv} v dv \quad ; \\
 &= - \left[f v \right]_{-v_1}^{v_1} + \int_{-v_1}^{v_1} f dv \quad ; \\
 &= N - f_1 v_1 \quad .
 \end{aligned}$$

Hence
$$\frac{\delta J}{\delta E} = \frac{\delta v}{\delta E} (N - f_1 v_1) e ,$$

where N is the total number of carriers in unit volume. A negative resistance is given if the dotted rectangular area of figure (d) is larger than the total area of the diagram covered by the carriers. In physical terms the negative resistance criterion is this: does the moment about the zero axis in the velocity picture of the extra carrier area scattered for a marginal increase in voltage exceed the moment due to the marginal increase in velocity of the rest of the distribution? At its simplest this criterion is given by the equation calculated: $1/e \cdot \delta J = N \delta v - f_1 v_1 \delta v \leq 0$.

In the diagrams it has been assumed that the difference in scatter rates out of energy E_c and into energy zero is exactly countered by the scattering of carriers by the threshold at a forward velocity lower than v_1 . This is to a lesser extent implicit in the calculation. It is accepted that a scattering threshold in practice is not sharp; a shallow one must make negative resistance less likely since the fast carriers remaining unscattered would provide an important contribution to the positive differential current. A full detailed treatment of the Boltzmann equation, including the correlation of energy and velocity effects, is required for any useful quantitative criterion for negative resistance. What we have tried to show is firstly how a double humped distribution arises, and secondly why a double hump can be favourable to the observation of a negative resistance.

6.2.3.5 Perfect Elastic Scattering with Threshold.

So far the threshold has been considered as one for inelastic scattering at which the whole of the carrier energy is removed. Here we examine the case where, at the threshold, elastic scattering takes place and assume that the redistribution is perfect, in that there is no persistence of forward velocity. As explained in section 6.2.3.2, there must also be strong inelastic scattering especially at this energy, if the situation is not to be self-contradictory.

The carriers scattered at the threshold are not returned to the origin as in sections 6.2.3.3, 6.2.3.4 but equally to all velocities in the range $-v_1$ to v_1 . In this way a double humped distribution is not produced, but the moment of the scattered carriers about the zero line is zero as in previous cases. Hence the criterion for negative resistance is a relative moments rule, as in 6.2.2.4 and at its simplest is given by:

$$\frac{dJ}{dE} = \frac{dv}{dE} (N - f_1 v_1) e \quad .$$

6.2.3.6 Elastic Scattering with Threshold and Persistence of Velocity.

Finally we consider the case of carriers scattered elastically at a threshold for which the forward momentum is not totally destroyed. We consider the case in which the carriers are scattered into the same distribution as that already existing. This clearly makes negative resistance far less likely than for isotropic scattering.

We chose to consider the special case of an exponential distribution function, for reasons which will soon become clear.

$$f = Be^{ax} ,$$

$$N = \int_{-\infty}^{\infty} Be^{ax} dx = \frac{B}{a} e^{av} ,$$

$$J = \int_{-\infty}^{\infty} Be^{ax} x dx = a(v - \frac{1}{a})N ,$$

$$J' = \int_{-\infty}^{\infty} Be^{a(x-r)} x dx = a(v - \frac{1}{a})Ne^{-ar} ,$$

$$\Delta J = J(1 - e^{-ar}) .$$

As r tends to zero, so does ΔJ . Hence for these conditions an exponential shape is neutral for negative resistance; to display one, a still steeper region at the threshold is required. Thus under the given conditions an exponential shape is critical for negative resistance.

6.2.3.7 Methods of Producing a Mobility Variation.

The discussion so far has been in idealized terms without any mention of physical processes. Processes which may cause a disturbance of the distribution function are:

- (1) Anomalies in normal scattering mechanisms.
- (2) Inter-valley exchange.
- (3) Acoustic phonon production.
- (4) Optical phonon production.
- (5) Plasmon interaction.

6.2.3.7.1 Anomalies in normal scattering mechanisms.

The normal scattering mechanisms are considered as those without any well-defined thresholds in terms of carrier energy or velocity; they include ionized impurity scattering, neutral impurity scattering, dislocation scattering, deformation potential scattering and carrier-carrier scattering. Cross-sections for these have some general dependence on the carrier energy, so it is possible for a change from one regime to another to take place. Reference back to section 5.3.5.3 shows that for:

$$\sigma = aT^m,$$

$$P = bT^n,$$

the negative resistance criterion is:

$$\frac{dR}{dJ} = \frac{1}{aT^m} \frac{n-m}{n+m} < 0.$$

For the voltage controlled type, we require:

$$n > m,$$

$$n + m < 0.$$

It is thus essential that m should be negative, ruling out ionized impurity scattering. In practice it does not seem possible to satisfy these conditions for normal scattering mechanisms for voltage controlled negative resistance; this is not really surprising since it has been shown (section 5.3.5.3) that normal mechanisms are reasonably favourable to current controlled negative resistance.

6.2.3.7.2 Inter-valley exchange.

A scattering mechanism incorporating a threshold is described by G. Weinreich, T. R. Sanders and D. L. White (1959, 627). It is based on a capture-ejection mechanism at an ionized impurity site of the type described in section 4.3; if the lower levels of a multi-level acceptor impurity are compensated, a further electron may be captured in an excited state if it has sufficient energy to overcome the repulsive coulomb barrier. It is an elastic mechanism in capture-ejection, but momentum is destroyed if in germanium or silicon a change of valley takes place. The development in time of an electron wavefunction results in a redistribution of the valley probability so that the length of time of the electron-impurity interaction determines the scatter rate. The process comes under the heading of 6.2.3.5 as a perfect elastic process with threshold. Since inter-valley scattering is small compared with intra-valley scattering in n-type uncompensated germanium (Weinreich, Sanders, White l.c. 1959, 625), the scattering cross-section introduced at the threshold is not necessarily significant. The possibility of observing non-ohmic effects or even a negative resistance seem higher in material with carefully chosen impurity composition. Such material has been discussed in section 6.2.2.

6.2.3.7.3 Acoustic phonon production.

The scattering of carriers by the absorption and emission of acoustic phonons, as may be treated by a deformation potential, has been considered under normal processes. Anomalous behaviour is possible however if the carrier drift velocity comes up to the velocity of sound in the material. This has been investigated in terms of the attenuation of an acoustic wave in a semiconductor; in cadmium sulphide it is found that acoustic wave amplification occurs when the electric field passes through a critical value (A. R. Hutson, J. H. McFee, D. L. White 1961, 626). It is necessary to use a semiconductor, $n < 10^{18}$ to limit the current density and hence the heating at the required field strengths (H. N. Spector 1962, 627). The results of Hutson, McFee, White (l.c. 1961, 628) do not show the effect of passing through the amplification threshold on the voltage-current characteristic; since power is being transferred from the electron system to the acoustic wave, the mobility is progressively reduced as the electric field is raised. The question arises whether this can lead to a negative resistance. The process may come under the heading of section 6.2.3.4 as a non-dominant inelastic process with threshold; the analysis there shows that if the scattering cross-section at the sound velocity is strong enough, the distribution function may be distorted sufficiently to produce a negative resistance. So far acoustic amplification has been demonstrated but not electrical negative resistance (Hutson, McFee, White l.c. 1961, 629).

6.2.3.7.4 Optical phonon production.

It is generally reckoned that optical phonon momentum scattering at low temperatures is unimportant, though it may have an appreciable effect on the energy scattering (R. G. S. Paige, 1964, 630, p.183); this is because of the high energy quantum involved. However it seems that extremely high electric fields require optical phonon production as the limiting factor to the current (R. A. Smith 1959, 631, p.162). The situation then resembles that of section 6.2.3.3. The possibility of obtaining a negative resistance in an intermediate condition is to be answered by the analysis of section 6.2.3.4.

6.2.3.7.5 Plasmon production.

Free electrons in a crystal with fixed ionized donors form a single component plasma system. Fluctuations of space-charge density have a natural frequency given for an idealized band-structure by the expression (L. Tonks, I. Langmuir):(1929, 632)

$$\omega_p^2 = \frac{ne^2}{\epsilon_0 m^*} \quad .$$

This motion is quantized in units of $\hbar\omega_p$ denoted as plasmons. It is apparent that if a carrier has a kinetic energy such that:

$$\frac{1}{2}m^*v^2 > \hbar\omega_p \quad ,$$

it can transfer energy to the plasmon system if a coupling mechanism exists. Such a plasmon system is damped by acoustic and optical phonon scattering (D. Bohm, E. P. Gross 1949, 633). D. E. McGumber

and A. G. Chynoweth claim (1964, 634) that the emission of plasmons is stimulated in the neighbourhood of ionized impurities; this is said to give rise to an 'instability'. However there seems no reason why as the electric field is raised, equilibrium should not be maintained, with the plasmon density rising steadily in order to balance the rates of production and decay at each field strength. The use of the term 'instability' for this is as misleading as was Kurosawa's use of it in a similar situation in impact ionisation (see section 5.3.7). It is thus possible that as the electric field is raised a strong extra inelastic scattering mechanism is switched in as the drift velocity approaches the plasma velocity. Whether this gives rise to a negative resistance is determined by the effect on the velocity distribution function as considered in section 6.2.3.5. Such a negative resistance, due to a distortion of the distribution function by a scattering resonance, has been observed in a gas plasma (S. Ohara, 1963, 635).

6.2.3.8 Critique of Papers by T. Musha and F. Yoshida.

The treatment given in the sections above has been in idealized physical terms and may seem inconclusive. The importance of an understanding of the basis of a voltage controlled negative resistance is well illustrated by the misconceptions in papers by T. Musha (1962, 636) and T. Musha, F. Yoshida (1964, 637), in which negative resistances are predicted for systems with orthodox distributions.

In the first paper Musha deduces from his mathematical equations that any low frequency a.c. modulation upon a d.c. bias voltage produces a reverse phase a.c. current; thus he shows in his diagram that the distribution function moves backwards for an extra increment of electric field. Now he fails to conclude that this gives a negative resistance in all cases, which is clearly ridiculous, is not certain. The cause of the fallacy seems to be an error in sign in his writing of the Boltzmann equation; if the third term on the LHS of his equation (1) is corrected to be positive, the negative resistance disappears and the analysis becomes useless. The same error is carried into the second paper. In this a complicated mathematical treatment for a high frequency modulation field, completed by computation, results in a negative resistance. Correction of the error requires the sign of $f_{1,3}^{(1)}$ to be reversed. It seems unlikely that computation for this case would result in a negative resistance, simply because insufficient physical ingredients have been put into the treatment to yield one. This paper has been quoted by McCumber and Chynoweth (l.c. 1964, 638) to support their reasoning on drifted plasma. We feel that the papers of Musha (l.c. 1962, 639) and Musha and Yoshida (l.c. 1964, 640) cannot lend authority to any physical consideration. This underlines the need for a physical understanding of the features of a calculation before intensive computation is undertaken; it represents a justification for the analysis of the preceding sections.

6.2.3.9 Summary.

The discussion of the distortion of the distribution function as a method of obtaining a negative resistance will not be extended further. We have examined firstly what sort of progressive distortion with applied voltage is necessary for different categories of scattering and secondly what physical processes are available to provide it. The conclusion reached is that if such scattering processes are invoked as causes of negative resistance, it is essential to work in terms of the distortion of the velocity distribution function; in the literature this has gone widely unrecognized, as shown by the employment of shifted Maxwellians by Musha and Yoshida (l.c. 1964, 641) and McCumber and Chynoweth (l.c. 1964, 642). Further we have shown that the extra scattering process must have a cross-section whose velocity variation is both sharp and strong, if a negative resistance is to be shown. Whilst this may be observed in a gas plasma (Ohara, l.c. 1963, 643), the possibility of obtaining a large useful region of negative resistance for a solid state crystal thus seems fairly remote. It can only be advantageous however to know what is required in the search for negative resistance, or what explanations are plausible if one is observed.

This completes our analysis of methods of obtaining voltage controlled negative resistance in the solid state. We now move on to a consideration of what happens when it is procured.

6.3 Domain Formation.

6.3.1 Introduction.

In chapter 1 we showed that a bulk negative resistance is unstable against the formation of domains or filaments. In chapter 2 the steady state domain size was considered and it was shown that consideration of irreversible thermodynamics is relevant to the situation. In both places mobility curves of the form of figure 6.4a were postulated. We now examine the nature of the domain walls and their contribution to the steady state current flow.

An interesting feature of the domain structure is the possibility of domain movement; this is an experimentally attested phenomenon (Barraud l.c. 1963, 644; Ridley and Pratt l.c. 1963, 645; Gunn l.c. 1964, 646; Fischler-Hazoni and Williams l.c. 1964, 647). The velocity of movement is of significance; a method of estimating it has been provided by B. K. Ridley (1965, 648), based on his interpretation of the principle of minimum entropy production with which we have already disagreed. Our analysis here places more emphasis on investigating possible stable domain configurations than on matching theory to current experimental results.

Before writing down the carrier flow equations, a qualitative picture of the domain structure will be given. The mobility-field relationship of figure 6.4 (a) may be changed into a 'local' voltage-current curve, figure 6.4 (b); by 'local' is meant that this curve is relevant to the material itself, but does not necessarily apply

1. The first part of the paper is devoted to a general discussion of the problem of the existence of a solution of the system of equations

which are satisfied by the functions u and v in the domain G .

2. In the second part the author considers the case when the functions u and v are harmonic in the domain G .

3. In the third part the author considers the case when the functions u and v are biharmonic in the domain G .

4. In the fourth part the author considers the case when the functions u and v are triharmonic in the domain G .

5. In the fifth part the author considers the case when the functions u and v are tetraharmonic in the domain G .

6. In the sixth part the author considers the case when the functions u and v are pentaharmonic in the domain G .

7. In the seventh part the author considers the case when the functions u and v are hexaharmonic in the domain G .

8. In the eighth part the author considers the case when the functions u and v are heptaharmonic in the domain G .

9. In the ninth part the author considers the case when the functions u and v are octaharmonic in the domain G .

10. In the tenth part the author considers the case when the functions u and v are nonaharmonic in the domain G .

11. In the eleventh part the author considers the case when the functions u and v are decaharmonic in the domain G .

12. In the twelfth part the author considers the case when the functions u and v are undecaharmonic in the domain G .

13. In the thirteenth part the author considers the case when the functions u and v are duodecapharmonic in the domain G .

14. In the fourteenth part the author considers the case when the functions u and v are tridecapharmonic in the domain G .

15. In the fifteenth part the author considers the case when the functions u and v are tetradecapharmonic in the domain G .

16. In the sixteenth part the author considers the case when the functions u and v are pentadecapharmonic in the domain G .

17. In the seventeenth part the author considers the case when the functions u and v are hexadecapharmonic in the domain G .

18. In the eighteenth part the author considers the case when the functions u and v are heptadecapharmonic in the domain G .

19. In the nineteenth part the author considers the case when the functions u and v are octadecapharmonic in the domain G .

20. In the twentieth part the author considers the case when the functions u and v are enneadecapharmonic in the domain G .

21. In the twenty-first part the author considers the case when the functions u and v are vigintapharmonic in the domain G .

22. In the twenty-second part the author considers the case when the functions u and v are unvigintapharmonic in the domain G .

23. In the twenty-third part the author considers the case when the functions u and v are trigintapharmonic in the domain G .

24. In the twenty-fourth part the author considers the case when the functions u and v are untrigintapharmonic in the domain G .

25. In the twenty-fifth part the author considers the case when the functions u and v are triunvigintapharmonic in the domain G .

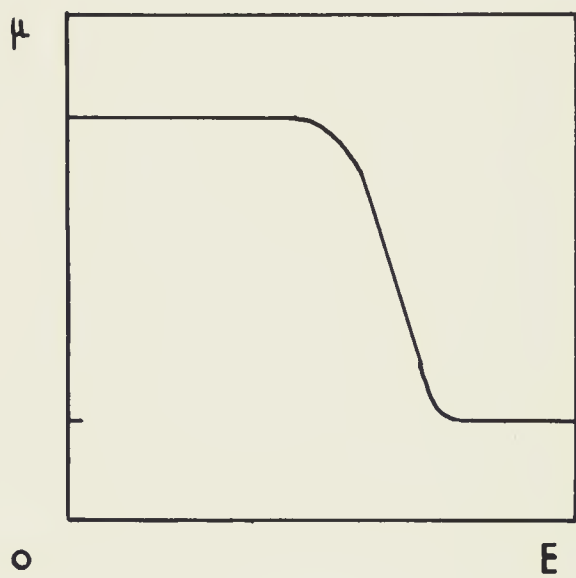
26. In the twenty-sixth part the author considers the case when the functions u and v are quadrigintapharmonic in the domain G .

27. In the twenty-seventh part the author considers the case when the functions u and v are unquadrigintapharmonic in the domain G .

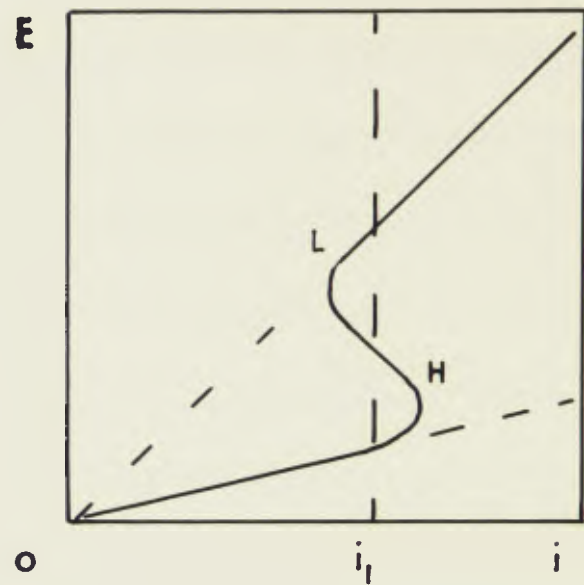
28. In the twenty-eighth part the author considers the case when the functions u and v are quinquagintapharmonic in the domain G .

Figure 6.4

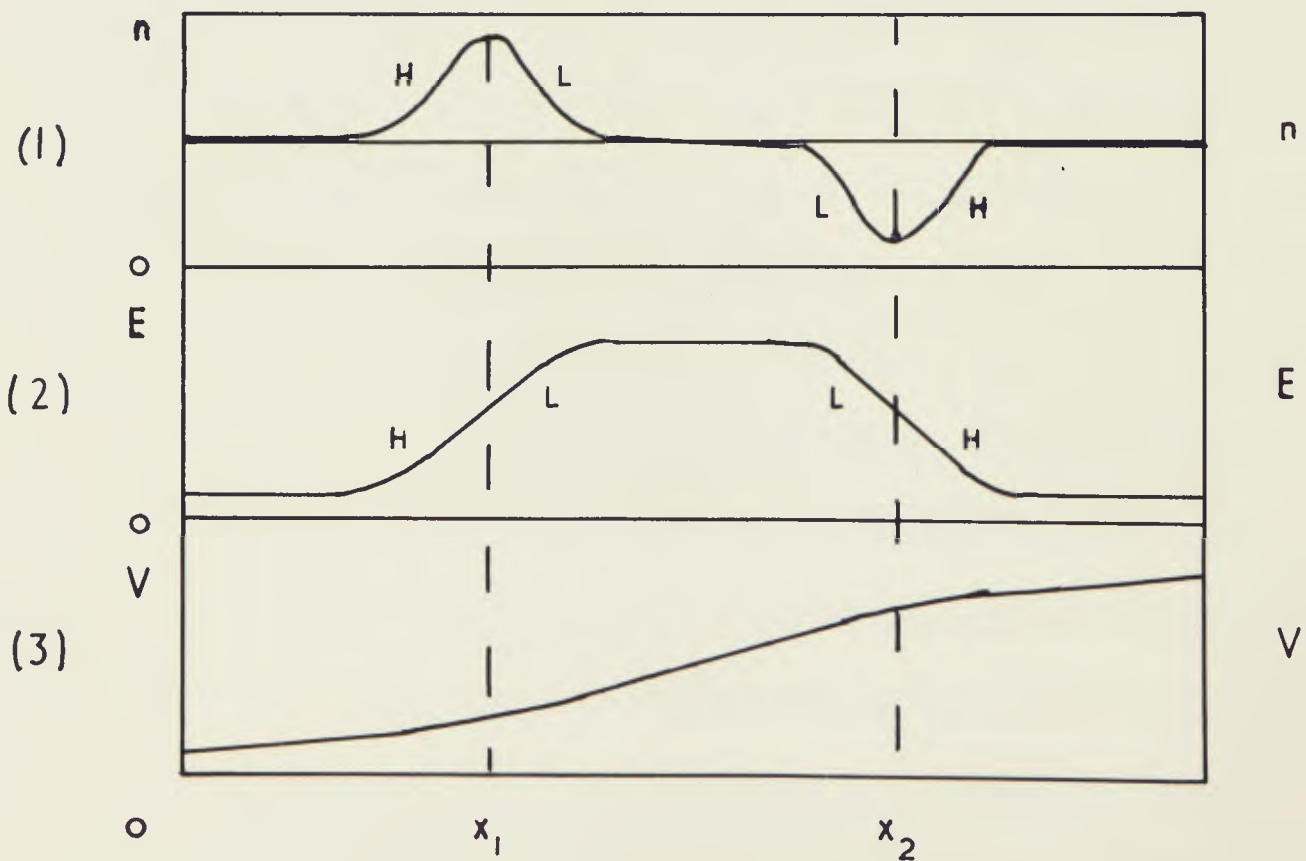
Domain Representation



(a)



(b)



(c)

to the measured $V-i$ curve due to the distortion provided by domains. Figure 6.4 (c) shows three spacial representations of a domain: (1) space-charge, (2) the electric field, (3) the potential. These are interrelated by differentiation. We consider operation at a fixed current i_1 . Reference to figure 6.4 (b) shows that over the two regions marked H in figure 6.4 (c) (2) the field driven current is higher than the true value i_1 , while over the regions marked L it is lower. In fact the effect of the negative resistance is to try to heap up the space-charge into a narrower region. This is opposed by diffusion. But whereas diffusion acts proportionately to the space-charge spacial distribution, the electric field acts according to the negative resistance shape of the $V-i$ characteristic of figure 6.4 (b). This means that the equilibrium shape for the domain space-charge is determined by the 'local' $V-i$ characteristic. But it also means that a force is available to drive the domain down the crystal. In fact it can be intuitively seen that a given characteristic still permits an equilibrium domain shape to be found to correspond to any required domain velocity relative to the drift velocity (within limits). Yet what determines the equilibrium velocity?

6.3.2 Equations of Carrier Flow.

We now examine the carrier flow equations for the domain. We consider n-type material with an equilibrium free carrier concentration n_0 , the space-charge density being n . Planar flow is assumed.

The current density is given by:

$$j = e(n + n_0)F \mu(F) + kT \mu(F) \frac{\partial n}{\partial x} \quad (1)$$

Carrier continuity requires:

$$\frac{\partial j}{\partial x} = e \frac{\partial n}{\partial t} \quad (2)$$

Poisson's equation for the space-charge is:

$$\frac{\partial F}{\partial x} = - \frac{ne}{\epsilon \epsilon_0} \quad (3)$$

Substitute:

$$F = G(u) = G(x - vt) \quad ,$$

so that G represents the waveform and v the domain velocity. After elimination of n , writing $\dot{G} = dG/du$,

$$j_0 + \epsilon \epsilon_0 v \dot{G} = (en_0 - \epsilon \epsilon_0 \dot{G})G \mu - \frac{kT}{e} \epsilon \epsilon_0 \ddot{G} \quad (4)$$

This is a second order differential equation to be solved for G in terms of the constants j_0 and v . The two constants of integration can be employed to satisfy boundary conditions. As in the formal theory of collective oscillations in crystals, we first assume cyclical boundary conditions; we have in fact a kind of single carrier plasma vibration. The two constants may be employed to match the zeroth and second differentials; matching the first differentials effectively "quantizes" the motion for an oscillatory phenomenon since this means fitting in an exact number of domains into the ring. Higher order derivatives match since the equation is of second order. We thus postulate that for a fixed relation between μ and E and for a chosen current j_0 , there exists a set of allowed

values of v , each giving an exact number of domains and each having its own field distribution G . Integration under G over the crystal gives the total applied voltage V . Hence a unique macroscopic current-voltage characteristic is not defined --- as considered in the section on irreversible thermodynamics, a whole set of domain shapes still seems to be available; the physics of the domain wall space-charge seems merely to control the velocity of domain movement.

We must point out that the term 'domain structure' is being used in a wider sense than is generally found; it is not restricted to flat-topped dipole layers, but may refer to any kind of oscillatory shape, with the sinusoidal form of Ridley (l.c. 1965, 649) being just one example. In fact for arbitrary choice of j_0 and v , the domain is oscillatory, and the flat-topped structure is a special limiting case. Since the domain space-charge is itself acted upon by the electric field, the domain velocity v is measured with respect to the electron drift velocity.

6.3.3 Extended Use of Butcher's Equal Areas Rule.

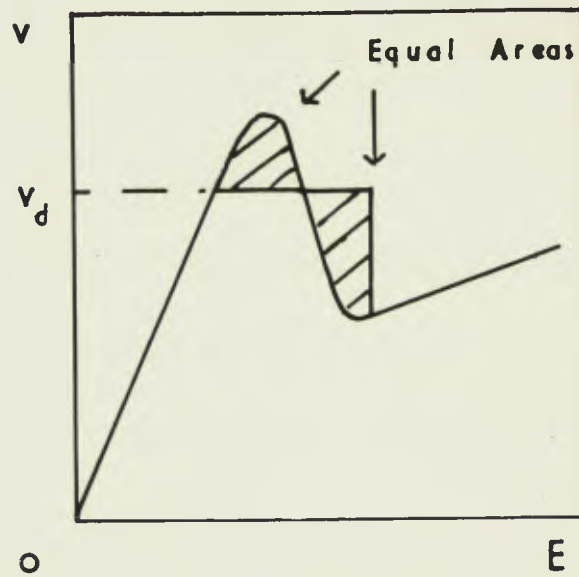
Solution of the domain structure equation (4) of this section under fixed boundary conditions is not a trivial problem. However P. N. Butcher (1965, 650) has developed an interesting technique for the situation where the diffusion coefficient D is assumed to be constant; the result is the 'equal areas' rule as applied to the case of the flat-topped domain. We wish to suggest that this technique

...the ... of ...
...the ... of ...
...the ... of ...

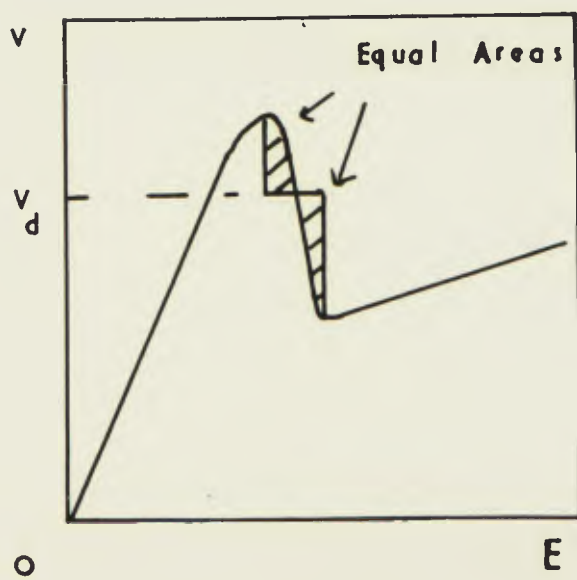


Figure 6.5

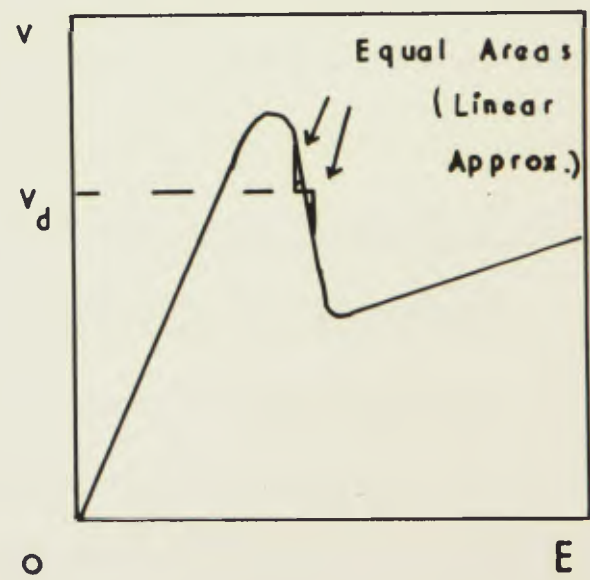
The Equal Areas Rule



(a) Butcher's Discrete Domains



(b) Continuous Domains



(c) Small Signal Domains

is also applicable to a generalized domain structure. The boundary condition to which Butcher works is that the electric field outside the domain is to give the chosen steady current j_0 . For an oscillatory solution, this boundary condition may be relaxed. A fixed base line corresponding to an arbitrarily chosen current j_0 for equation (4) exists, from which the equal areas rule is applied; just as in the triangular domain case, the area computation is stopped as soon as the equal areas point is reached (figure 6.5 a), so in this case truncation takes place on both sides (figure 6.5 b). In this way, instead of there being just one domain configuration and hence just one operating voltage at one current j_0 , there are now an infinite range of possibilities. At the same time, the imposition of cyclic boundary conditions restricts the working to just a few of these. Cyclic boundary conditions are equivalent to the employment of completely passive electrodes connected to a constant voltage source. This is a reasonable parallel to the experimental use of a dry cell connected between the ends of a crystal.

6.3.4 Small Signal Solutions.

6.3.4.1 Equal Areas.

It is instructive to examine the small signal oscillation qualities of the Gunn effect. We employ the equal areas rule in the region shown in figure 6.5 c. The calculation starts from the equation derived by Butcher (l.c. 1965, 651), rewritten in our notation as:

$$\frac{\delta n}{n_0} - \log(1 + \frac{\delta n}{n_0}) = \frac{e\epsilon_0}{en_0D} \int (v(E) - v_0) dE .$$

This equation is obtained by the elimination of x between (1), (5) and subsequent integration. Here we have effectively specified the base line to be given by $v_0 = v_D$. We assume a linear relationship between $v(E)$ and E with a negative slope, and apply a small voltage swing E_1 about v_0, E_0 ; this gives the local characteristic as:

$$v(E) - v_0 = -\alpha(E - E_0) ;$$

the boundary condition is that $\delta n = 0$ at $E = E_0 \pm E_1$,

$$\text{So: } \frac{(\delta n)^2}{2n_0^2} = \frac{\alpha e \epsilon_0}{2n_0 e D} \left\{ (E_0 - E)^2 + E_1^2 \right\} ;$$

$$\frac{e \delta n}{e \epsilon_0} = -\frac{dE}{dx} = \left(\frac{\alpha e n_0}{e \epsilon_0 D} \right)^{\frac{1}{2}} \left\{ -(E_0 - E)^2 + E_1^2 \right\}^{\frac{1}{2}} ;$$

$$E = E_0 - E_1 \sin \left\{ \left(\frac{\alpha e n_0}{e \epsilon_0 D} \right)^{\frac{1}{2}} x \right\} .$$

This gives the wavelength of the oscillation as:

$$\lambda = 2\pi \sqrt{\frac{\mu}{\alpha} \frac{kT e \epsilon_0}{n_0 e^2}} = \sqrt{\frac{\mu}{\alpha}} \cdot 2.2 \cdot 10^{-7} \text{ M.K.S. units at } 77^\circ \text{K.}$$

Estimating the field and mobility parameters from P. N. Robson, S. M. Mahrous (1965, 652), we take the domain velocity as 10^5 metres per second and $\sqrt{\mu/\alpha}$ as 5. Then:

$$\lambda = 1 \text{ micron,}$$

$$v = 100 \text{ gc./sec.}$$

This gives a frequency for small signal conditions, and represents an upper frequency limit when working on the steepest portion of the natural characteristic. It is clear that for operation about a point

where the slope α is shallower, or for larger signals where it is effectively so, the wavelength is greater and the frequency of operation lower. This frequency is similar to the estimated sub-band relaxation time of 10^{-12} secs. (J. E. Carroll 1965, 653).

From the above analysis it appears that the current-voltage characteristic of a specimen is not uniquely determined. At a given applied voltage the current may take one of a range of values, each with its own domain structure. However the assumption of cyclic boundary conditions causes the allowed currents to take discrete values; then the crystal runs in the mode which is preferentially triggered. Experimentally, crystals have been found to run at frequencies between 0.06 gc./sec. (J. S. Heeks 1966, 654) and 10 gc./sec. (B. W. Hakki, S. Knight 1966, 655); these frequencies are not always correlated with total length and in many cases represent smoothly oscillating domain structures rather than individual flat-topped domains (Foyt, McWhorter l.c. 1966, 656; Hakki, Knight l.c. 1966, 657). Thus there is some experimental evidence for the oscillatory domain structures suggested in these sections. Carrier doping inhomogeneities in the bulk of the crystal have been shown to exert a strong influence on the triggering of domains (D. E. McCumber, A. G. Chynoweth 1966, 658). We point out that if the slower oscillatory tendencies can be suppressed, frequencies up to 100 gc./sec. seem possible, according to the small signal calculation, without any need for particularly small specimens.

6.3.4.2 Ridley's Sinusoidal Solution.

The space-charge solution obtained above under conditions suitable for the application of the equal areas rule is:

$$\delta n = \sqrt{\frac{\alpha \epsilon \epsilon_0 n_0}{eD}} E_1 \cos \left\{ \left(\frac{\alpha n_0}{\epsilon \epsilon_0 D} \right)^{\frac{1}{2}} x \right\} ;$$

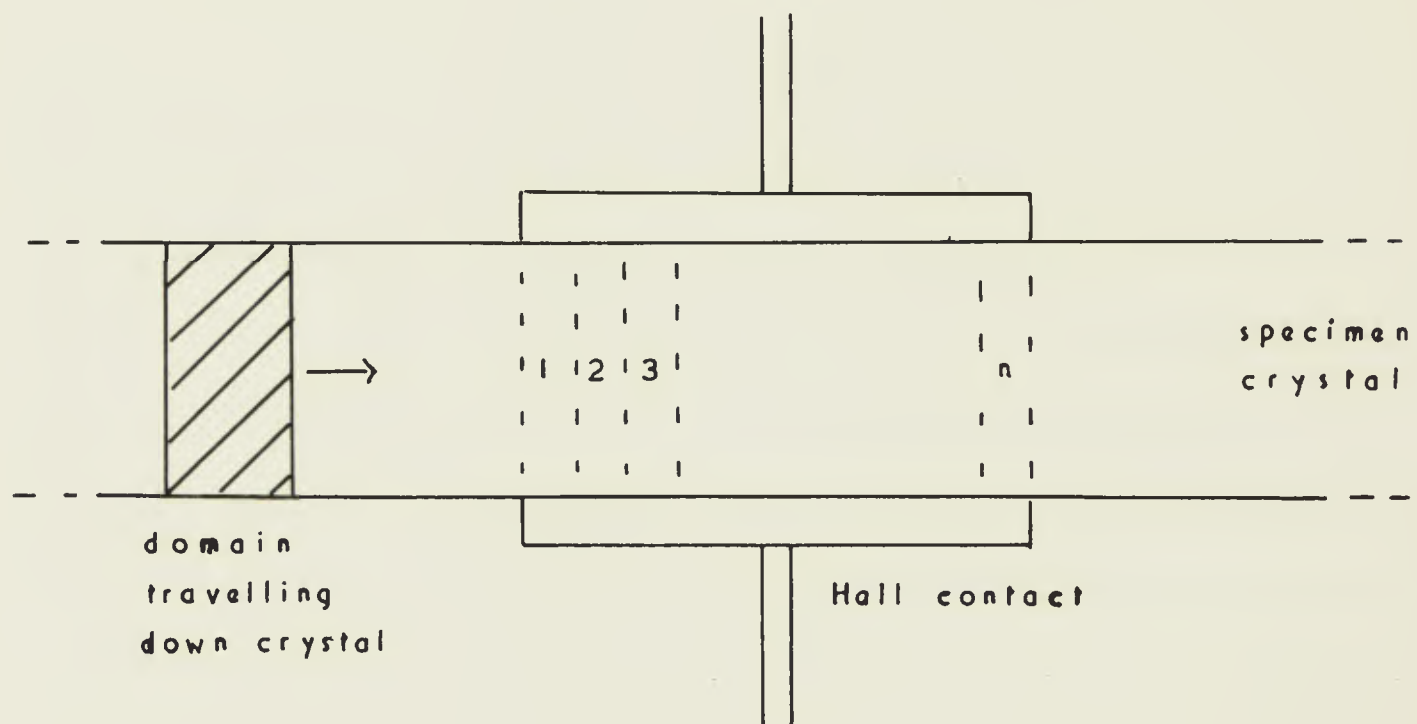
or $\delta n = \delta n_0 \cos(kx) ,$

$$k = \left(\frac{\alpha n_0}{\epsilon \epsilon_0 D} \right)^{\frac{1}{2}} .$$

Then the wavelength of the small signal oscillation is fixed by the properties of the crystal itself, in particular by the negative slope resistance α . If a non-linear negative slope is considered, then α is a function of E and the integration cannot proceed. Ridley has also considered the small signal case (l.c. 1965, 659); he has postulated a sinusoidal space-charge variation and attempted to attach a meaning to k , the exponent. However he allows k to vary down the crystal and suggests that the oscillation is controlled by the region where k has its largest value. We feel that for our set of conditions, which are within the scope of Ridley's work, we have shown that a sinusoidal solution is only valid if k is fixed by the crystal parameters and is constant down the crystal; then Ridley's deductions are invalid because of a self-inconsistency; he is using a crystal whose voltage-current characteristic is completely different at different places and different times. This is borne out by the interpretation given to figure 6.4 in section 6.3.1.

Figure 6:6

Division of Region between
Hall Contacts into Elements



6.3.5 Hall Effect Techniques.

Experimentally, results which were apparently due to domain formation were found to be caused by a thermal effect, as will be shown in the next chapter. However it is felt worth while to discuss Hall effect techniques as means of examining domain structure.

In theory the Hall effect differentiates between a change in carrier density and a change in mobility. This feature is of relevance to the slow moving domains of B. K. Ridley and R. G. Pratt (l.c. 1963, 660; l.c. 1965, 661) associated with carrier trapping. The question which arises concerns the value of the Hall voltage shown by a set of domains when the domain wavelength is small compared with the physical size of the Hall probes.

We consider the Hall contacts to be regions where the conductivity is far higher than in the specimen body; also potential differences across the contact interface are neglected. While this is probably not strictly valid, the experimental phenomenon of large area contacts acting as short circuits shows that it is not unwarranted (E. H. Putley 1960, 662, p.45). Thus the boundary condition at the Hall contacts for a potentiometer reading of the Hall voltage is that no net current should flow; the Hall voltage is such that the currents which flow across the specimen in various places due to the carrier density variation in the domains exactly compensate each other. We divide the portion of specimen between the contacts into regions as in figure 6.6. Then the common Hall voltage is given by:

$$V_H = \frac{A}{n_1} + \frac{i_1}{n_1 e \mu_1} = \frac{A}{n_2} + \frac{i_2}{n_2 e \mu_2} = \dots = \frac{A}{n_r} + \frac{i_r}{n_r e \mu_r} ,$$

$$i_1 dS_1 + i_2 dS_2 + i_3 dS_3 + \dots + i_r dS_r = 0 .$$

Eliminating the currents:

$$V_H = A \frac{\sum_j \mu_j dS_j}{\sum_j n_j \mu_j dS_j} .$$

For a sharply defined single domain d in a normal crystal r , of proportionate size $S_d : S_r$, we have:

$$V_H = A \frac{S_d \mu_d + S_r \mu_r}{S_d \mu_d n_d + S_r \mu_r n_r} .$$

For a domain where the mobility varies and the carrier density is unaltered:

$$V_H = \frac{A}{n} ,$$

but for one where the carrier density varies and the mobility is unaltered:

$$V_H = A \frac{S_d + S_r}{S_d n_d + S_r n_r} .$$

In the latter case the observed conductivity is:

$$\sigma = e \mu \frac{S_d + S_r}{S_d / n_d + S_r / n_r} .$$

Thus the effect of the domain structure on the Hall voltage is the inverse of that on the conductivity in terms of the number of carriers. The change in the Hall voltage produced by a narrow high

resistance low carrier density domain is evidently rather small, while the change produced by a low mobility domain is zero. So the Hall effect is not a suitable method for differentiating between domain types unless either the domain size is comparable with the size of the Hall contacts or the high field part of the domain is an appreciable fraction of the inter-domain distance.

In cases where a Hall voltage reading is useful, any unrepeatability in the domain waveform makes a d.c. reading unreliable. This is because of the unknown contribution of the potential drop between the Hall contacts which is normally eliminated by reversing the magnetic field. An a.c. technique overcomes this, in which the magnetic field is required to swing several times during the passage of a domain through the contact region. Then the Hall voltage swing is examined as the domain structure passes between the Hall contacts. Application of an oscillating magnetic field will be described in section 7.5.4.

6.4 Parallels between Current and Voltage Controlled Negative Resistance.

6.4.1 Intrinsic Instability.

In section 6.2 at the beginning of this chapter, it was suggested that the three cases treated for impact ionisation correspond to the three sections of 6.2 on voltage controlled negative resistance. This point merits elaboration.

The three cases for impact ionization are given as:

5.3.5.1 Normal ionization.

5.3.5.2 A carrier density dependent ionization rate.

5.3.5.3 A current density dependent mobility.

The three sections for voltage controlled negative resistance are given as:

6.2.1 Band-structure effects.

6.2.2 Effects due to an energy dependent capture cross-section.

6.2.3 Distribution function effects.

Then the correspondence is to be one-to-one in the order shown; it is more a correspondence for convenience than a strict mathematical inversion of roles between carrier mobilities and carrier densities.

In the first instance we note that while a band-structure effect in the voltage controlled case, as in gallium arsenide, does produce a negative resistance, the whole effort of the argument of section 5.3.5.1 is to show that the inverse band-structure, whether in terms of carrier mobility or density, does not produce one. This is simply because a complete inversion has not taken place; in the current controlled case, a current density dependence of the inverse band-structure is required for negative resistance, but only a voltage dependence is provided. In the second and third instances a true inversion of conditions is provided so that in these cases negative resistance is consistently available.

6.4.2 Filament and Domain Structure.

Current controlled negative resistance was treated before the voltage controlled type because of its greater simplicity. This is due to space-charge movement in the voltage controlled type. The space-charge is a direct result of Poisson's Law. In the case of filaments it seems possible to neglect this and its flow in the local electric field. In the case of domains, the flow of the space-charge associated with the domain wall is an integral part of the phenomenon. Any attempt to employ a voltage controlled bulk negative resistance for amplification purposes must harness this domain flow rather than attempt to prevent it. In this context the development of a small signal sinusoidal solution in section 6.3.4 may have considerable implications in terms of an amplification characteristic, especially for the frequency response and band-width.

This comparison of the causes and effects of the two types of negative resistance concludes the treatment of fundamental solid state properties which may lead to either. We now turn to consider surface and size effects.

Chapter 7

Effects Induced by Specimen Heating.

7.1 Introduction.

The analyses in the two previous chapters on impact ionization and on voltage controlled negative resistance were undertaken with the specific aim of explaining experimental results which seemed to display both of these phenomena. In spite of the use of an approach of maximum generality, in which all processes leading to a negative resistance seemed to be being treated, at least in principle, it did not appear to be possible to fit the results into the pattern in a way consistent with other known properties. It was only at a relatively late stage that a chink in the logic was discovered which can provide an adequate explanation. In previous chapters it has been shown that it would require unreasonably large specimens for thermal filaments or domains of appreciable effect to be stable. It was the realization that there can be irregularities in the heat transport at the specimen boundaries which completed the picture. As a direct result, the aim of this work has been widened from an investigation of the particular negative resistance properties of overcompensated zinc-doped germanium to a more comprehensive study of bulk negative resistance in semiconductors.

7.2 Theory.

The impossibility of setting up an appreciable temperature gradient in a semiconductor at a low temperature has already been demonstrated (section 5.3.3). In a non-infinite specimen, the heat transport has to be maintained across the boundaries; to obtain good thermal flow, it is convenient to immerse the specimen directly in the liquid coolant. The rate of heat flow which can be sustained by direct contact without any apparent temperature gradient is high, but if this limit is exceeded, a new phenomenon sets in; this is film boiling. These points are made by P. C. Eastman and W. R. Datars (1963, 701); our analysis follows theirs.

The effect produced by a heat production rate which requires the boundary transfer limit to be exceeded, is the formation of a gaseous layer around the specimen surface; this layer or film has a relatively low thermal conductivity, so that a considerable temperature gradient across it is required to transfer the heat being produced in the sample. The specimen rises in temperature at this juncture, causing a change in its resistance. Thus at the onset of film boiling, a sudden transition along the external circuit load line takes place, following the specimen resistance until a stable operating point is found. Because of the large specific heats involved (S. H. Koenig, R. D. Brown 1959, 702), this change is relatively slow. It may thus be expected to be possible to follow the resistance changes due to variations in both carrier density and mobility as the

temperature of the specimen changes. Where no stable operating point on the load line exists, the result must be disintegration of the specimen.

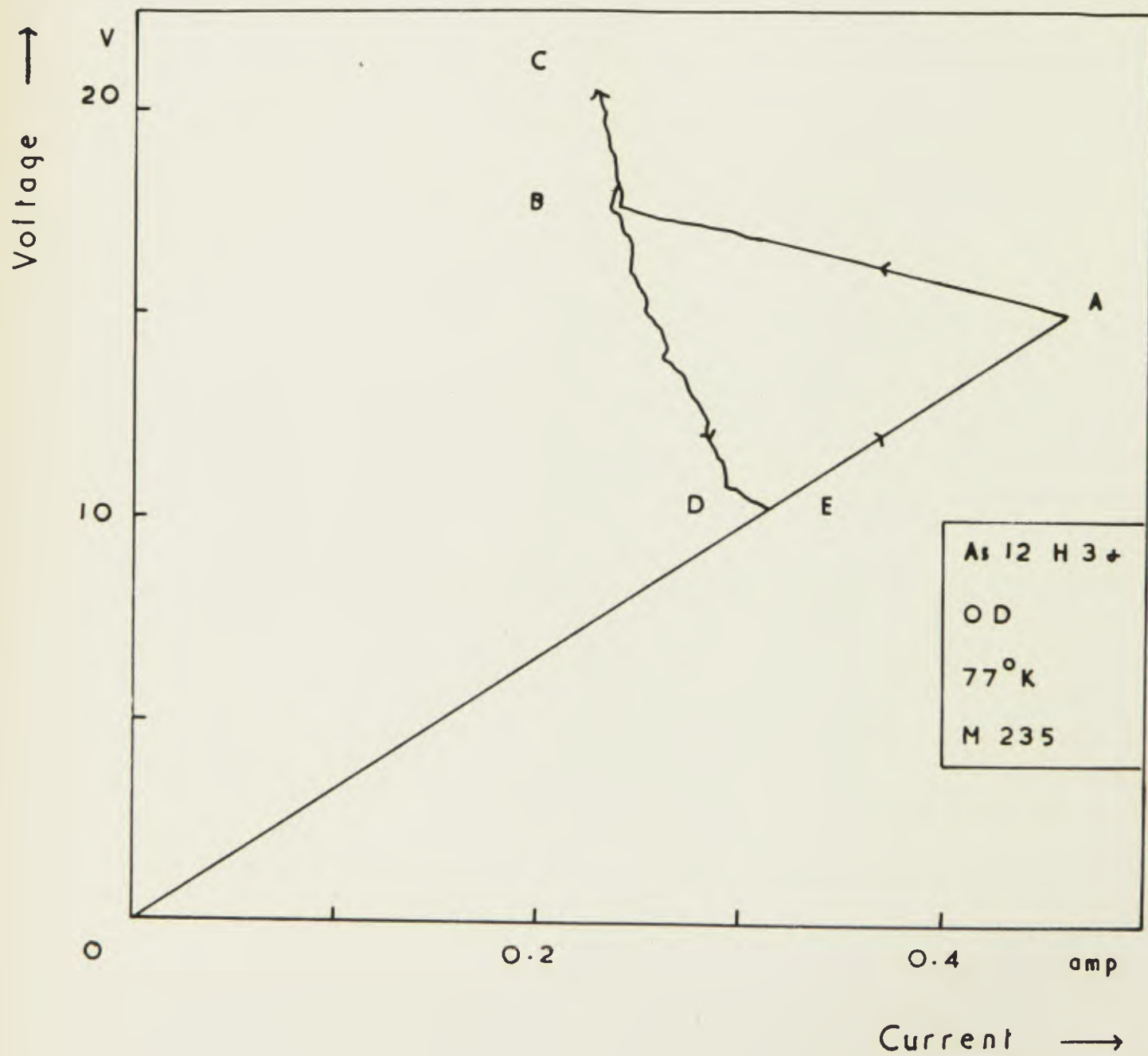
A critical parameter is the power dissipated in the specimen per unit surface area. We consider a rectangular specimen of cross-sectional dimensions a , b and ignore the ends. Then the parameter is (Eastman, Datars, l.c. 1963, 703):

$$W = \frac{KE^2}{\rho} \frac{ab}{2(a+b)} \quad .$$

The constant K depends on the operating conditions. Considerable differences between different coolants may be expected due to influences such as different latent heats of vaporization and different surface tensions. Variation of the temperature by the reduction of the pressure over the coolant may be expected to have some effect on W due to the temperature dependence of the rate coefficients; in the case of liquid helium at the superfluid transition, a considerable effect may be anticipated (C. Nanney 1962, 704). The surface treatment is a feature which is likely to affect K ; however for carefully prepared semiconductor crystal surfaces, any changes in the microscopic surface structure produced by dislocation, scratching or etching, though they are of considerable interest in terms of surface currents, should have a marginal influence on the value of K .

Figure 7.1

Ohmic Characteristic Showing Hysteresis



7.3 Experimental Results.

Experimentally a negative resistance has been observed over the temperature range 14°K — 77°K . It was first noted under conditions of double injection; after its observation without injection, all subsequent examination was undertaken under the simpler conditions of ohmic flow. Results under d.c. application will be considered before those for pulsed or oscillatory currents.

An example of a current-voltage characteristic is given in figure 7.1. The arrows show the hysteresis of the cycle. AB and DE are spontaneous non-reversible transitions along the load line and constitute the negative resistances. These are to be explained as the result of film formation and film disappearance. The branch BC is near vertical for all specimens, showing a negative slope in most cases.

To show up the internal processes, the Hall effect has been measured at the transition. The Hall voltage is shown in figure 7.2 for 220 gauss with corresponding notation. We consider a straightforward interpretation in terms of the number of carriers n and mobility μ , with conductivity $\sigma = ne\mu$ and single carrier Hall coefficient $R_H = 1/ne$. At the transition $A \rightarrow B$ the number of carriers has risen by 15% while the mobility has been reduced by a factor of 2.3. Over a range of temperatures the values shown in the table of figure 7.3 are obtained. The hydrogen reading is complicated by the ionization of frozen donor sites and will be dealt with in more



Figure 7.2

Recording of Hall Voltage Showing Hysteresis

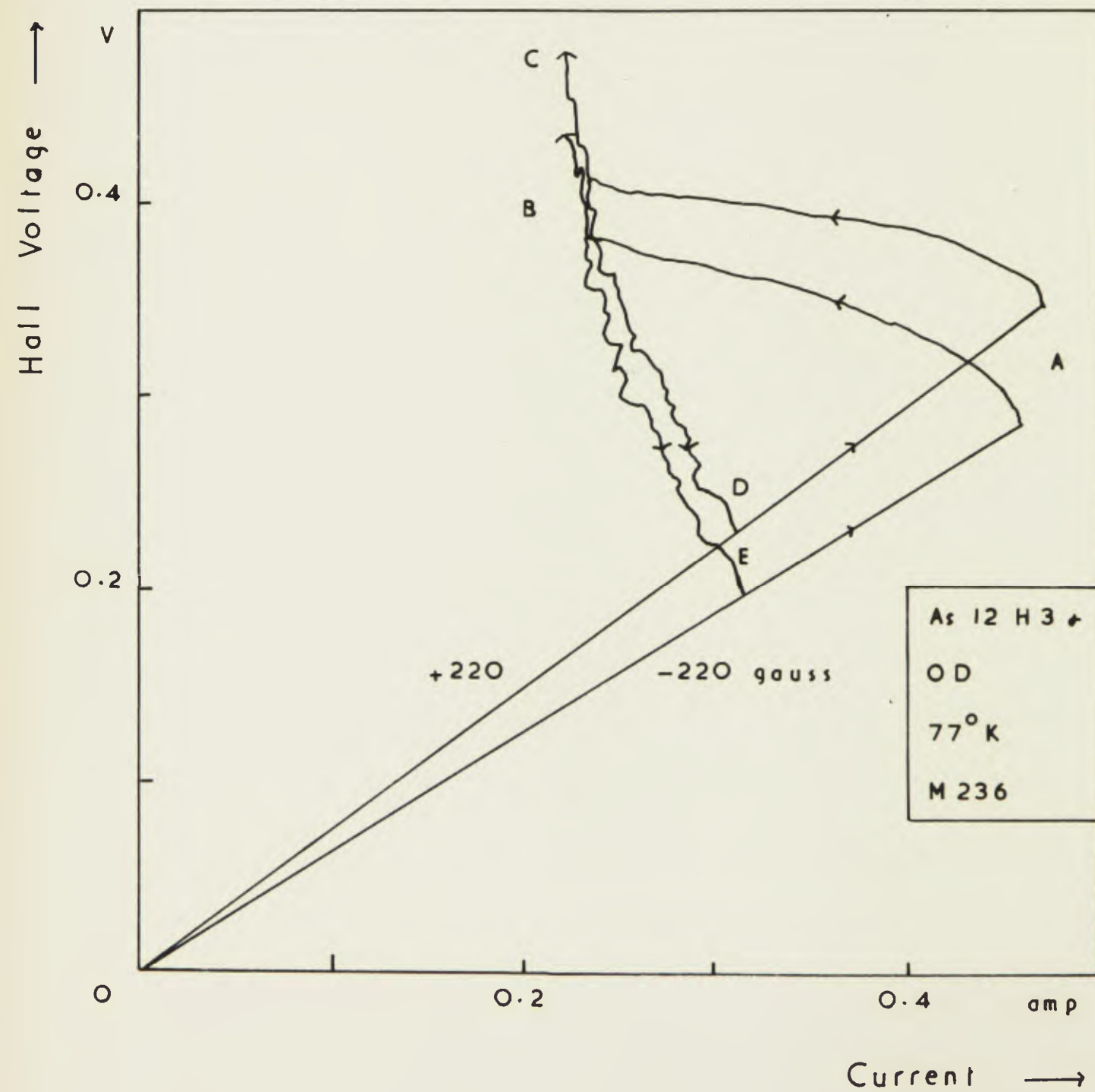


Figure 7.3

Temperature Dependence of Breakdown

		<u>As₁₂H₃</u>	<u>OD</u>	<u>220 gauss</u>			
Temperature	°K	77	70	64	20	20*	
Current	amp.	1	0.45	0.42	0.40	0.19	0.50
Current	amp.	2	0.26	0.24	0.23		0.31
Hall voltage	mV.	1	28.2	27.2	31.7	35.6	41.5
Hall voltage	mV.	2	14.2	13.6	15.0		13.8
Potential drop	V.	1	4.3	3.8	4.0	6.4	5.6
Potential drop	V.	2	4.8	4.3	4.4		6.0
n	10 ¹⁵	1	2.81	2.72	2.22	0.94	2.12
n	10 ¹⁵	2	3.23	3.11	2.70		3.97
μ	10 ³	1	7.81	8.54	9.46	6.64	8.83
μ	10 ³	2	3.52	3.78	4.07		2.75
Factor in n			1.15	1.14	1.21	4.21	
Factor in μ			0.45	0.44	0.43	0.41	

1 refers to the values immediately prior to breakdown.

2 refers to the values immediately after breakdown.

* refers to the highest current reached (point D' of figure 4.3.1).

detail later. The column headed 20 " relates to the situation immediately after ionisation has taken place, and corresponds roughly to the mobility maximum.

7.4 Elimination of Alternative Explanations.

Several alternative explanations of the negative resistance have been rejected from experimental evidence. There is the possibility that it might be a contact effect, associated with the injection of holes. In this case the straightforward interpretation of the Hall voltage just given would no longer be valid. Set against this are five points:

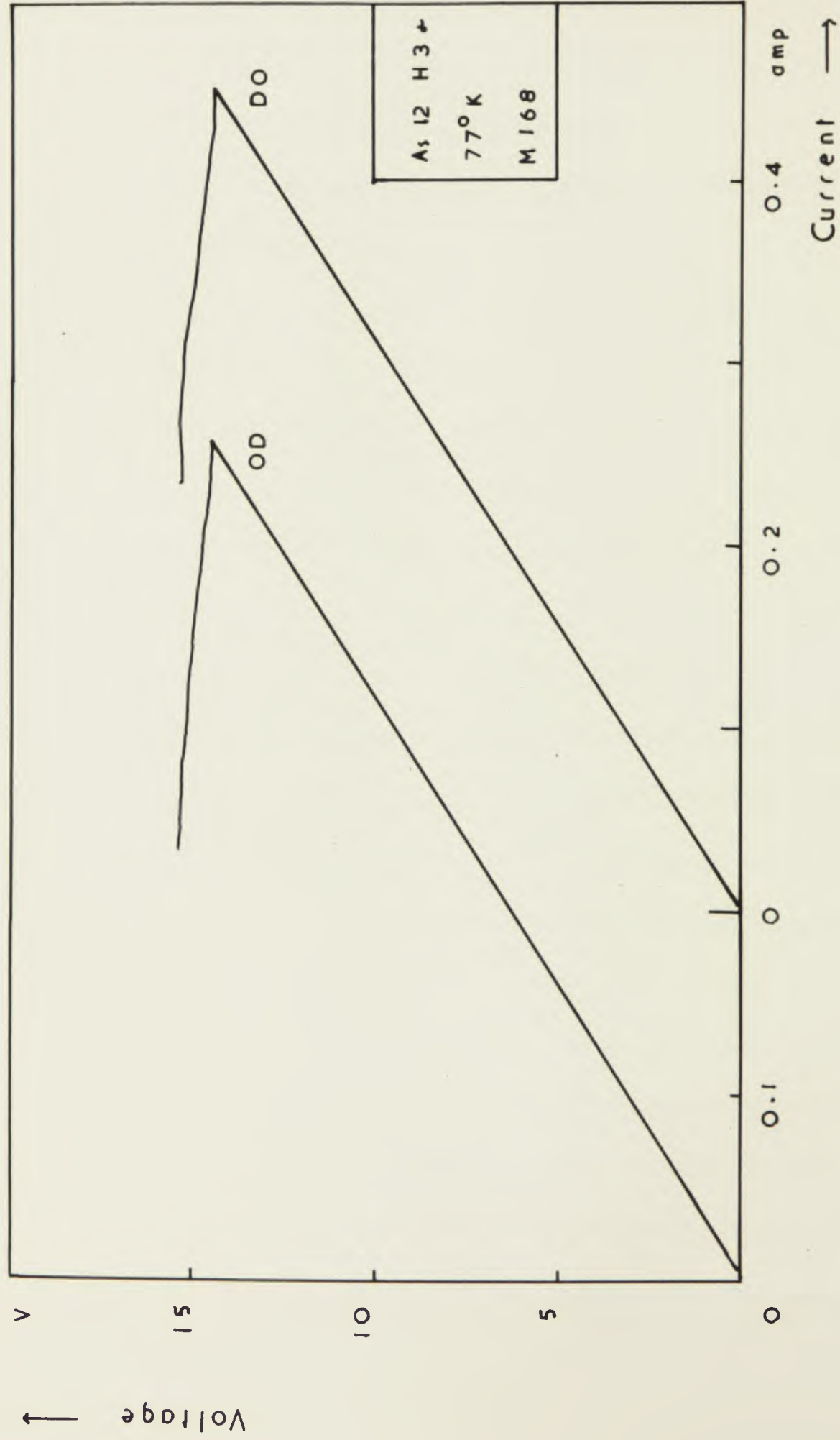
(1) Results for injecting contacts on this material have already been examined in chapter 4. Injection effects are explained there but the sudden interruption by the negative resistance being considered here did not seem to be a feature of injection.

(2) Potential and Hall voltage measurements at several probe positions on an ohmic specimen show that the whole crystal is involved to the same extent at the transition, whereas the main deduction from similar measurements in injection has been that minority carriers spread down the crystal from the injecting contact.

(3) The i - V characteristics of well-constructed ohmic specimens are linear up to the transition point as in figure 7.1; a sudden spontaneous repeatable change in the contact properties is unlikely.

TABLE I		TABLE II	
1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20
21	22	23	24
25	26	27	28
29	30	31	32
33	34	35	36
37	38	39	40
41	42	43	44
45	46	47	48
49	50	51	52
53	54	55	56
57	58	59	60
61	62	63	64
65	66	67	68
69	70	71	72
73	74	75	76
77	78	79	80
81	82	83	84
85	86	87	88
89	90	91	92
93	94	95	96
97	98	99	100

Figure 7.4 Reversibility of Ohmic Characteristic



(4) The characteristic is identical for currents flowing through a specimen in each direction; the reproducibility for this as shown by the example in figure 7.4 is remarkable. The probability of a single pair of injecting contacts having similar properties is small, that of a number of pairs negligible.

(5) The physical nature of the contact seems to have little influence. Pairing contacts of different sizes made no difference between different current directions; also the use of sulphur-doped indium instead of tin-lead as a contact material had no observable effect on the breakdown.

It is reckoned that this evidence is strong enough to eliminate injection.

As a second possible cause of the negative resistance we consider electrical surface currents. Relevant evidence is as follows:

(1) B. K. Ridley and R. G. Pratt have estimated (1963, 705) that the surface current for a carefully prepared surface of germanium is of the order 10^{-3} amp. per volt per cm. This is negligible compared with the current employed which for figure 7.1 would give 1 amp. per volt per cm.

(2) Specimens cut from raw material show properties similar to those with carefully polished and etched surfaces. Preparation of the surface would be expected to have a very large effect on the electrical surface states and hence on the surface current. The corresponding effect on the heat transfer should be marginal.

(3) A specimen showed similar properties before and after immersion in an etching solution; the change in the transition point was wholly consistent with the change in the physical dimensions during etching. If surface electrical flow was the cause of the transition, a considerable effect would have been anticipated.

Again it seems possible to eliminate this feature as a cause of the negative resistance.

It has already been shown that thermal filaments cannot form inside the bulk of the specimens because of the high thermal conductivity of semiconductor materials in this temperature range; a similar calculation may be made for domains. We consider a crystal containing a domain which has absorbed virtually the whole applied voltage of 10 volts, running at 1 amp./sq. mm. current density. The thermal conductivity is taken as before as 1 watt/°K/cm. from M. G. Holland (1964, 706). Then the heat production rate is 10 joules/sec./sq. mm.; even at 200°K temperature excess above which the material would be intrinsic and of low resistivity, the length of the zone of temperature gradient and hence the minimum total specimen length would have to be 5 cm; this is unlikely to be realized, and does not correspond with observed effects in 1 mm. length specimens.

We are thus led to a consideration of the intrinsic processes of a bulk semiconductor which can lead to a negative resistance of either voltage or current controlled types. This analysis has been given in general terms in chapters 5, 6. The experimental results

fail to fit into the pattern on two counts; firstly the special properties required of a semiconductor material to display a negative resistance are not those which the specimens studied may be expected to have over the ranges of temperature and composition employed; and secondly the phenomena observed vary in ways inconsistent with explanations in terms of carrier velocity or density thresholds. Rather than show why every possible mechanism of this type is to be eliminated, which would be difficult if not impossible, we turn to consider the evidence which strongly indicates that film boiling is taking place.

7.5 Evidence for Film Boiling.

Evidence obtained which supports the explanation of film boiling may be classified under the following headings:

- (1) The applicability of the power dissipation per unit surface area formulae.
- (2) The observation of breakdown in various materials at a similar power dissipation.
- (3) Instances of specimen disintegration.
- (4) The observation of fluctuations.
- (5) The effect of a pulsed field on breakdown.
- (6) The detection of emitted radiation.

1. Generalized Linear Model (GLM)

Definition: A statistical model that extends the linear model to include non-normal distributions and link functions.

$$Y = \mu + \epsilon$$

$$\mu = X\beta$$

Where Y is the response variable, X is the design matrix, β is the vector of parameters, and ϵ is the error term.

The GLM framework is defined by three components:

1. Random Component: The response variable Y is assumed to follow a distribution from the exponential family, such as the normal, binomial, or Poisson distribution.
2. Systematic Component: The mean of the response variable is modeled as a linear combination of the predictors, $\mu = X\beta$.
3. Link Function: A function that connects the linear predictor to the mean of the response variable, $\eta = X\beta$ to μ . Common link functions include the identity link for normal distributions, the logit link for binomial distributions, and the log link for Poisson distributions.

The GLM framework provides a unified approach to modeling a wide range of data types. It allows for the estimation of parameters using maximum likelihood methods and provides a framework for testing hypotheses about the parameters. The GLM framework is a powerful tool for analyzing data that do not conform to the assumptions of the standard linear model.

Figure 7.5

Breakdown Statistics for Overcompensated Germanium

Specimen	units	type	As11G1+	As12H1+	As13H2+	As12H3+	As8E1+
Surface area	sq.cm.		0.895	0.556	0.548	0.355	0.26
Length	cm.		1.335	1.050	1.016	1.027	0.10
Cross-section	sq.cm.		0.027	0.018	0.018	0.0074	0.075
Breakdown	volts	R	12.5	9.2	9.2	10.2	
voltage		O		15	15	14.5	5.6
Breakdown	amp.	R	2.0	2.2	2.35	1.28	
current		O		1.55	1.55	0.9	0.98
Power/unit	watts/	R	27.9	36.4	39.4	36.7	
surface area	sq.cm.	O		41.7	42.4	36.7	21.1
Voltage/unit	volts/	R	9.36	8.77	9.05	9.93	
length	cm.	O		14.3	14.8	14.1	56.0
Current/unit	amp./	R	72.1	125.3	130.3	172.1	
cross-section	sq.cm.	O		88.3	86.1	121.3	6.5

7.5.1 The Power Dissipation Formula.

The criterion for film boiling was written in section 7.2 as:

$$W = \frac{K^2}{\rho} \frac{ab}{2(a+b)} > W_c .$$

This may be written in terms of the total applied voltage and current

as:
$$W = \frac{K'V_1}{2l(a+b)} > W_c .$$

The results available are not suitable for a detailed investigation of this kind of relationship or its temperature or coolant variation. For that purpose instead of a semiconducting crystal specimen, a long resistance wire would be far more satisfactory. In our case serious uncertainty is caused by the existence of up to nine contacts on the specimen faces each with a copper lead; copper wires act as cooling fins and raise the required power (Eastman, *Datara* l.c. 1965, 707). Also a certain amount of the voltage is developed across the current contacts; the question arises whether it is better to employ the total power delivered to the specimen, or the power delivered to a specific region; the latter may be in error in view of the high thermal conductivity. For injection specimens it seems incorrect to treat anything except the whole.

In view of these problems, no very great self-consistency is expected for the results. Our purpose is comparison of the criterion of power dissipation per unit surface area with two other possible criteria for breakdown, current per unit cross-section and voltage per unit length. In the table of figure 7.5 the values of these

Figure 7.6

Breakdown Statistics for Different Materials at 77° K

Specimen	units	2UC6H1 _d	1UC1H2 _d	3UC1H1 _d	InSb1H3 _d
Length	cm.	0.8	0.73	0.75	0.91
Pot. length	cm.	0.266	0.49	0.39	
Width	cm.	0.1	0.09	0.1	0.20
Depth	cm.	0.1	0.09	0.1	0.06
i	amp.	0.2	0.5	1.0	
V _H	mV.	175	55	6.9	
V _L	V.	4.0	4.1	1.4	
n (10 ¹⁵)	/c.c.	0.31	2.7	40	0.06
μ (10 ³)	cm./V.	26	16	4.4	
i _{breakdown}	amp.	0.49	0.67	1.5	1.07
V _{breakdown}	V.	15.2	9.4	4.4	9.5
Power/unit	watts/				
surface area	sq.cm.	23.2	23.9	22.0	21.0
Voltage/unit	V./				
length	cm.	19.0	12.9	5.7	10.42
Current/unit	amp./				
cross-section	sq.cm.	49	83	150	89.2

expressions are compared for a number of different specimens at 77° K. It is apparent that the power dissipation criterion is the most satisfactory.

7.5.2 Dependence on Material.

The table of figure 7.6 shows results at 77° K for different materials. Those used were three specimens of uncompensated germanium and also one of indium antimonide. It is virtually impossible to imagine a mechanism other than film boiling which could work consistently over such a range of materials and temperatures.

7.5.3 Specimen Disintegration.

On a number of occasions early during the investigation it was found that specimen contacts became detached during the breakdown. An empirical method of preventing this was found to be to provide a shallow enough load line, so that the change in current was matched by a small voltage change --- to have in effect a near constant voltage source. In this way the power dissipation in the specimen fell sharply on breakdown, sometimes to one third of its previous value. It was this reduction in power dissipation which caused bulk thermal effects to be rejected for so long. Specimens were found to be permanently damaged if run at voltages far above breakdown.

7.5.4 Fluctuations.

In chapter 6 it has been shown that the formation of steady or travelling domains is to be expected for a voltage controlled negative resistance. For this reason a search was made for domain structure, with apparently favourable results. A voltage fluctuation was found with a mean frequency of about 40 cycles per second and a shape resembling that of noise. It was most easily observed between Hall contacts in the absence of a magnetic field where the peak to peak amplitude was 3 mV.; these fluctuations were completely absent before breakdown took place. It was felt to be important to determine whether the 'domains' were caused by a reduced carrier density or a reduced carrier mobility. Hence a 10 gauss a.c. magnetic field was applied at a frequency large compared with 40 cycles. A Hall voltage swing of 0.5 mV. was obtained at 1000 c/s with the circuitry described in section 3.9. The deduction made was that the amplitude of the Hall voltage was independent of the fluctuations and hence that the fluctuations did not represent changes in the local carrier density. The argument of section 6.3.5 shows that this conclusion needs qualification to deal with narrow domains. Similar fluctuations have been reported by Eastman, Datars (l.c. 1963, 708) and by Nanney (l.c. 1962, 709) and are easily explained as mobility variations due to changes in temperature caused in turn by turbulence in the gas film. The structure and band-width of the noise may be expected to be partially determined by bubble separation

Figure 7.7

Input Power for Pulsed Breakdown

77° K Au2B2 +

Frequency	Duty Cycle	Power at Breakdown
c/s	%	watts
1000	84	6.4
1000	80	7.2
1000	70	6.8
1000	60	7.2
1000	50	7.2
1000	40	6.4
1000	30	7.9
1000	22	7.6
200	85	6.0
200	80	6.4
200	70	5.9
200	60	6.1
200	50	6.6
200	40	6.7
200	30	7.1
200	20	6.8
5000	50	8.1
2000	50	7.0
1000	50	6.6
500	50	6.5
200	50	6.7
100	50	6.8
50	50	6.6
20	50	6.2
10	50	5.5

from the film. The process was observed in an unsilvered glass dewar; the rate of bubble separation from the crystal was high but not excessive; no visual change was noted at breakdown in contrast to Manney's reported observation in liquid helium (l.c. 1962, 710).

7.5.5 Pulsed Breakdown.

Effects for which high electric fields or high current densities are required may be observed without specimen heating through the use of pulsed fields. Then the effect of heating may be demonstrated by increasing the pulse on-off ratio i.e. the duty cycle. This is illustrated by H. J. Goldsmit, J. M. Corsan in the case of the Esaki effect in bismuth (1964, 711).

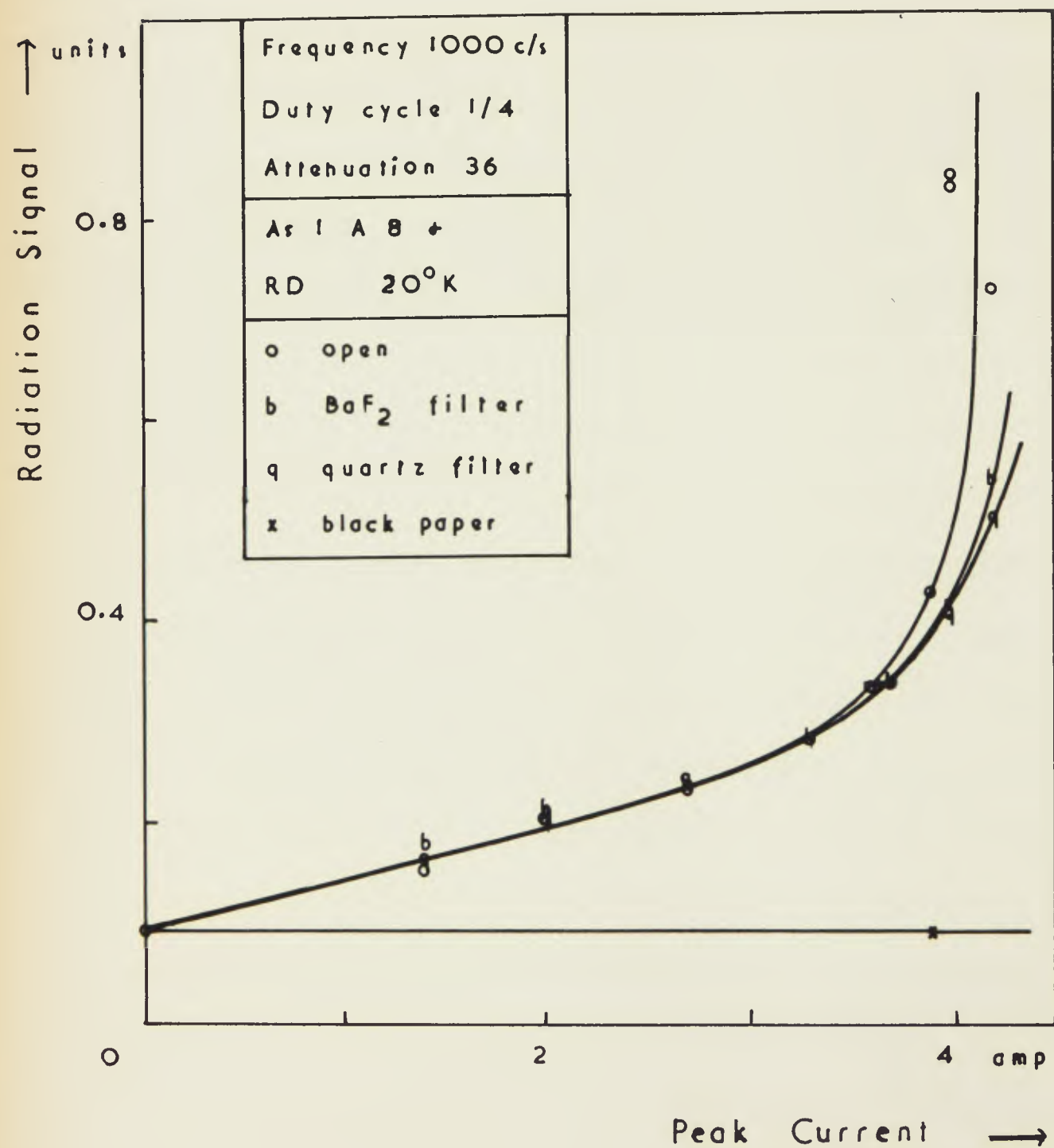
Results obtained for the germanium specimen As₂B₂ are shown in the table of figure 7.7. These values were obtained with pulse circuitry with a rise time of $\sim 1/10$ msec., employing Newmarket 404 transistors in the arrangement of figure 3.6, AUY 10's being limited to 0.75 amp. Uncertainty in estimation of the duty cycle accounts for most of the discrepancies in the power figures.

While this evidence is consistent with film boiling as an explanation, it is equally consistent with any effect with a time constant of 0.1 --- 1 seconds. Trapping mechanisms of this sort are known (R. G. Pratt, B. K. Ridley 1965, 712; C. Fischler-Hazoni, F. E. Williams 1965, 713).



Figure 7.8

Radiation Signal Passed by Filters



7.5.6 Radiation.

It is fitting that this thesis should be concluded with a description of the project which was initially to have been the basic topic of the work, a study of recombination radiation in overcompensated zinc-doped germanium. It shows how easily an apparently important effect can be relegated to the realm of trivia.

The strategy for the investigation of extrinsic recombination at the zinc site for the 0.09 eV. transition has been set out in section 1.5. The apparatus employed is as described in sections 3.6, 3.7, 3.8. A set of measurements for diode As1A8_a taken with the generation-detection system is shown in figure 7.8. The line marked 'black paper' shows that there was no observable change in the noise level arising from a rise in the current passing through the diode; thus problems over electrical pick-up had been overcome. For the linear rise up to 3 amp. peak current, any proportion of signal removed by a quartz filter is not detectable. This signal may thus be assumed to be due to intrinsic recombination radiation at a wavelength of the order of 1.7μ . Points above 3.3 amp. peak represent the post-breakdown region, the lower ones being taken on the return part of the hysteresis. It is clear that a considerable proportion of the radiation in the breakdown case is being absorbed by the quartz filter and is hence of a longer wavelength than 2μ .

Diode As1A15_a gave the following figures just beyond breakdown; frequency 1000 c/s; duty cycle 5/12; peak current 3.7 amp.; detector attenuator 50; total radiation figure 0.73; radiation passed by BaF₂ filter 0.46; radiation passed by quartz filter 0.11; zero level 0.09 . Here the total signal at breakdown was almost 30 times higher than for diode As1A8_a so that intrinsic recombination radiation may be neglected in comparison to breakdown radiation; this is demonstrated by the negligible proportion passed by the quartz filter. The barium fluoride filter, with a cut off at 13 μ , passed just on half of the radiation. From the shape of the BaF₂ absorption edge reported (W. Kaiser, W. G. Spitzer, R. H. Kaiser, L. E. Howarth 1962, 714) it was estimated that if the radiation represented a single sharp line, its wavelength should be $14 \pm 0.5 \mu$, equivalent to 0.09 eV., which is exactly the upper zinc level transition energy which it was hoped to see. It was however rather disconcerting that diodes showed different proportions of post-breakdown radiation transmitted by barium fluoride.

It was then decided to make a study of the spectral composition of this radiation. This involved the construction of the optical dewar described in section 3.8. Earthing and electrical shielding problems were found to be still more severe in this case, but nevertheless a valid signal was obtained for diode As8C1_a immersed in liquid nitrogen in the post-breakdown range. This signal was not large enough to pass through a Leiss spectrometer, which had been

calibrated with the water vapour spectrum for the purpose. Two alternatives were available; either to press ahead to improve the sensitivity and optical alignment so as to get a spectral distribution, or to put optical measurements into abeyance until a fuller understanding of the electrical properties of the breakdown was available; in this way it was hoped to design specimens of far higher radiative output which would not require immersion in a coolant.

The second alternative was chosen and there can be no doubt that the choice was correct. It is now apparent that the rise in long wavelength radiation output produced after breakdown is easily explained by a rise in specimen temperature. It was estimated at the time that the figures above can be produced by a specimen at 200° K, which no longer seems unreasonable. The variation in spectral distribution as observed with the BaF₂ filter is explained by differences in the post-breakdown stabilised temperature for different specimens.

The initial project was to find whether or not radiative transitions between the upper zinc level and the valence band can be observed. Due to a diversion of effort into other channels, an upper limit to the probability has not been provided. However the discussion of section 1.3 suggests that zinc is a relatively unsuitable choice for the investigation of deep level long wavelength extrinsic radiative recombination, and our experimental evidence supports this.

7.6 Conclusions on Heating.

From the above evidence, there seems little doubt that film boiling is the cause of the negative resistances noted at the breakdown. For a specimen in liquid hydrogen, the first trend is the ionization of the remaining donors, giving an apparent "impact ionisation" negative resistance. The resistivity continues to fall until around 50°K , when the mobility maximum is reached. After this, as for specimens starting in liquid nitrogen, the resistivity rises due to a falling, phonon dominated mobility. Finally the intrinsic region may be reached. In any instance the stable operating point is determined by the surface heat transport rate and the load line.

So much then for the anomalous experimental negative resistance.

Chapter 8

Conclusion.

It has been shown in section 7.5.6 how the initial aim, the investigation of extrinsic recombination radiation was generalized because of anomalous results, so as to be a survey of negative resistance. Thus the aim has been a comprehensive survey of all processes which might cause the effects observed and not a comprehensive survey of all processes which might in general lead to a negative resistance. The possibility of using quantum anomalies in high magnetic fields has not been considered; similarly there is no treatment of effects which rely on crystal anisotropy such as negative effective mass (H. Kromer 1961, 801) or applications of the Esaki effect (L. Esaki 1962, 802; H. J. Goldsmid, J. M. Coraen 1964, 803).

In the discussion of the available literature, disagreement has been registered on a number of occasions. However an attempt has been made to confine criticism to those instances where a treatment which surmounts the problems is presented for substitution. Similarly this work must be considered not in terms of absolute validity, but as to whether something more accurate is reasonably within reach.

Several objections to the work may be raised. The first is whether it is within the field of low temperature semiconductor

physics. We have just pointed out that the aim has been to examine those processes which might be relevant to a current-voltage characteristic observed between 20°K and 77°K . If the subsequent discussion is relevant to other materials or at other temperatures then so much the better. Similarly it may be claimed that we have chased hares instead of keeping to the main path. But progress in research has to be chased where it can be scented, rather than down a prearranged trail; then the catch is its own justification.

Thirdly the work may be said to be speculative when it should be quantitative. The aim has been to deal with the philosophy of physics, in particular for irreversible thermodynamics and for negative resistance. In the literature for these fields, this has been neglected; as has been indicated, authors have plunged in intuitively, even into elaborate computation, without examining the basic picture, and the rest of the world seems to be following their lead. At least an attempt has been made here to chart the philosophy, even if the result is not altogether unobjectionable.

Finally objection may be raised to the overall length. The difficulty has been the diffuse nature of the subject of negative resistance and the fast rate of growth of interest in it at a superficial level. As noted in the preface, an attempt has been made to keep the length down. We try to conform to this spirit by closing the discussion at this point.

APPENDIXProgramme PCS F

FARBA01
PCSTF ZETA→

```

begin library A0,A6;
real pa,aa,at,b,s,knc,kps,F,c,bote,emu,gt,h,e,j,jhsc,djll,
    djul,dl,hh,df,dfh,df1,dfs,fpu,d,pjf,ll,ul,nl,u,w,x,z,
    bb,v;
integer i,p,pp,ppl,ppu,l,q,m,mm,mt,ll,ul,t,k,kk,hf,hg,hf1,
    hf2,n,nn,ls,sl,trip,figa,figc,figd,fige;
array f,r,int[0:400];

procedure span(yy,L7B,L9B,L10B,L12B);
value yy;
real yy;
label L7B,L9B,L10B,L12B;
begin r[p]:= (axat+bxw)x(aa+w);
    int[p]:= int[p-1]+(r[p-1]+r[p])/2xh;
    df:= hx(jx(yy-(aa+w)/(aa+2xw))+cxint[p])/F;
    x:= z:= f[p+1]:= f[p-1]+2xdf;
    dfs:= df/h;
    if z<0 then goto L7B;
    l:= 0;
    L11: w:= z;
    r[p+1]:= (sxat+bxw)x(aa+w);
    int[p+1]:= int[p]+(r[p]+r[p+1])/2xh;
    dfh:= hx(jx(yy-(aa+w)/(aa+2xw))+cxint[p])/F;
    z:= f[p]+(df+dfh)/2;
    dfs:= (df+dfh)/2/h;
    if z<0 then goto L7B;
    l:= l+1;
    if abs(z-w)<exf[p] then goto L1;
    if l<5 then goto L11;
    if mx5+5>1+lX(m-p-1) then goto L11;
    hf:= hfX2;
    if hf>33 then goto L9B;
    goto L12B;
    L1: w:= f[p+1]:= (4xz+x)/5;
    dfs:= (w-f[p])/h;
    if w<0 then goto L7B;
    if trip=1 then
    begin if dfs>0 then goto L10B;
    end;
    l:= l+1;
end;

```



```

procedure bridge(y,L9,span);
value y;
real y;
label L9;
procedure span;
begin real jea;
jea:= (aa+y)/(aa+kpa);
L12: m:= 10*hf;
h:= hh/hf/32;
f[0]:= f[m+1]:= w:= y;
gt:= int[0]:= 0;
l:= p:= sl:= pp:= 0;
df:= hx*jx(jea-(aa+y)/(aa+2xy))/F;
r[0]:= (sxat+bxy)x(aa+y);
x:= z:= f[1]:= w+df;
for l:= 1 step 1 until 8 do
begin w:= z;
dfs:= df/h;
if z<0 then goto L7;
r[1]:= (sxat+bzw)x(aa+w);
int[1]:= (r[0]+r[1])/2*xh;
dfh:= hx(jx(jea-(aa+w)/(aa+2xw))+cxint[1])/F;
z:= y+(df+dfh)/2;
if abs(z-w)<exy then goto L3;
end;
hf:= hf*2;
if hf>33 then goto L9;
goto L12;
L3: dfs:= (z-y)/h;
w:= f[1]:= z;
if w<0 then goto L7;
if trip=1 then
begin if dfs>0 then goto L10;
end;
l:= l+1;
p:= 1;
for mm:= 1 step 1 until 5 do
begin span(jea,L7,L9,L10,L12);
f[1]:= f[2];
h:= 2*xh;
end;
for p:= 1 step 1 until m-1 do
begin span(jea,L7,L9,L10,L12);
gt:= gt+abs(z-x)/(f[p]+n5);
end;
if trip=1 then

```

```

begin pp:= p:= m;
jhsc:= 0;
sl:= 1;
span(jea,L7,L9,L10,L12);
pp:= pp+1;
goto L13;
end;
r[m]:= (sxat+bxw)x(aa+w);
int[m]:= int[m-1]+(r[m-1]+r[m])/2xh;
jhsc:= jx(w/(knc-aa)-kpa/(aa+kpa))+cxint[m];
if 1>3xm then hf:= hf*2 else if gt<0.01xexm then hf:= hf+4
    else if gt<0.3xexm then hf:= hf+2;
goto L8;
L10: sl:= 1;
pp:= p;
goto L13;
L7: sl:= -1;
f[p+1]:= 0;
pp:= p+1;
L13: for mm:= p+1 step 1 until m do f[mm]:= 0;
jhsc:= 0;
if 1>3xp+9 then hf:= hf*2 else if gt<0.3xexp then hf:= hf+2;
L8: if hf<hg then hf:= hg;
if hf>33 then goto L9;
end;

open(20);
open(30);
figa:= format([3s-d.ddddddn-nd]);
figc:= format([sssndd]);
figd:= format([ssssnddd]);
fige:= format([s-d.ddddn-nd]);
n:= read(20);
for nn:= 1 step 1 until n do
begin aa:= read(20);
at:= read(20);
b:= read(20);
s:= read(20);
knc:= read(20);
kpa:= read(20);
F:= read(20);
c:= read(20);
bote:= read(20);
emu:= read(20);
e:= read(20);
t:= read(20);

```

```

ll:= read(20);
d:= read(20);
k:= read(20);
trip:= 1;
hh:= read(20);
hf:= hg:= read(20);
for kk:= 1 step 1 until k do
begin switch sw:= G1,G2,G3;
goto sw[read(20)];
G1: ll:= read(20);
G2: d:= read(20);
G3: pjf:= read(20);
j:= read(20);

L2: m:= 10xhf;
hf1:= hf;
pa:= l1xpjf;
ll:= df1:= 0;
li:= ppl:= 0;
djll:= 0;
for q:= 1 step 1 until 7 do
begin bridge(pa,L9A,span);
if sl=-1 or jhsc<0 then
begin ll:= pa;
hf1:= hf;
li:= 1;
ppl:= pp;
df1:= dfs;
djll:= jhsc;
pa:=dxpa;
end
else goto L4;
end;
writetext(30,[failure*for*pa]);
write(30,figa,pa);
write(30,figd,p);
write(30,figd,i);
write(30,figc,hf);
writetext(30,[c]);
goto L9A;
L4: ul:= pa;
hfu:= hf;
ppu:= pp;
fpu:= f[pp];
ui:= 1;
djul:= jhsc;

```



```

if trip=1 then writetext(30,[[10s]nl[[14s]ll[[14s]ul[[13s]dfl
  [[13s]fpu[8s]ppl[[3s]ppu[[3s]hfl[[3s]hfu[5s]ll[[6s]ul[[cc]]])
  else writetext(30,[[10s]nl[[14s]ll[[14s]ul[[12s]djll[[12s]
  djul[8s]ppl[[3s]ls[[4s]hfl[[3s]hfu[5s]ll[[6s]ul[[cc]]]);
write(30,figa,pa);
write(30,figa,ll);
write(30,figa,ul);
if trip=1 then
begin write(30,figa,dfl);
write(30,figa,fpu);
write(30,figc,ppl);
write(30,figc,ppu);
end
else
begin write(30,figa,djll);
write(30,figa,djul);
write(30,figc,ppl);
write(30,figc,q);
end;
write(30,figc,hfl);
write(30,figc,hfu);
write(30,figd,ll);
write(30,figd,ul);
writetext(30,[[c]]);

for q:= 1 step 1 until t do
begin dl:= (ul-ll)/((if q=1 and ll#0 and d>4 then d-1
  else 10));
ls:= 1;
u:= nl:= if djll=0 then (ul+ll)x0.5 else (ulxdjll-llxdjul)
  /(djll-djul);
hf:= if hfl>hfu then hfl else hfu;
L5: bridge(u,L9A,span);
if f[m+1]=0 then
begin trip:= 2;
writetext(30,[[ccc]]);
pjf:= 1;
goto L2;
end;
if sl=-1 or jhsc<0 then
begin ll:= u;
djll:= jhsc;
hfl:= hf;
ppl:= pp;
dfl:= dfs;
li:= 1;
u:= u+dl;

```

```

ls:= ls+1;
goto L5;
end;
if ls#1 then
begin ul:= u;
hfu:= hf;
ppu:= pp;
fpu:= f[pp];
ui:= 1;
djul:= jhsc;
end
else
begin bb:= 1;
L6: ul:= u;
hfu:= hf;
ppu:= pp;
fpu:= f[pp];
ui:= 1;
djul:= jhsc;
u:= u-dl;
bridge(u,L9A,span);
ls:= ls+1;
if f[m+1]=0 then
begin trip:= 2;
writetext(30,[[ccc]]);
pjf:= 1;
goto L2;
end;
if sl=1 or jhsc>0 then goto L6;
ll:= u;
hfl:= hf;
ppl:= pp;
dfl:= dfs;
li:= 1;
djll:= jhsc;
end;

write(30,figa,nl);
write(30,figa,ll);
write(30,figa,ul);
if trip=1 then
begin write(30,figa,dfl);
write(30,figa,fpu);
write(30,figc,ppl);
write(30,figc,ppu);
end
else

```

```

begin write(30,f1ga,dj11);
write(30,f1ga,djul);
write(30,f1gc,pp1);
write(30,f1gc,ls);
end;

write(30,f1gc,hf1);
write(30,f1gc,hfu);
write(30,f1gd,li);
write(30,f1gd,ui);
writetext(30,[[c]]);
end;

if sl=-1 or jhsc<0 then writetext(30,[[L[c]]]) else writetext
(30,[[U[c]]]);
mm:= m+10;
mt:= m;
for q:= mm step mm until mt do
begin write(30,f1ge,f[q]);
if f[q]=0 then mt:= q;
end;
writetext(30,[[c]]);
for q:= mm step mm until mt do write(30,f1ge,j/emu/
(aa+2xf[q]));
v:= 0;
writetext(30,[[c]]);
for q:= mm step mm until mt-1 do write(30,f1ge,botex(jxaa/
(aa+2xf[q])-Px(f[q-1]-f[q+1])/h)/(aa+2xf[q]));
writetext(30,[[cc]]);
for q:= 0 step 1 until mt do v:= v+j/emu/(aa+2xf[q]);
v:= v+j/emu/(aa-at)x(m-mt);
write(30,f1ga,j);
write(30,f1ga,(nl+aa)/(aa+kpa));
write(30,f1ga,w/(knc-aa));
write(30,f1ga,j/emu/(aa+2xf[0]));
write(30,f1ga,vxhhx10/(m+1));
writetext(30,[[ccc]]);
end;
L9A: end;
close(30);
close(20);
end→

```


Variable Nomenclature.

* indicates a parameter to be assigned an initial value from a data tape.

aa	*	n-type donor impurity concentration;
at	*	hole-trap concentration;
b	*	bimolecular recombination coefficient;
bb		dummy;
bridge		procedure to traverse the crystal;
bete	*	magnetic field coefficient, B/te for Hall voltage value;
c	*	recombination factor;
d	*	factor for use with initial value in search for lower limit;
df		free hole concentration gradient at position p;
dfr		free hole concentration gradient at projected position p+1;
dfl		free hole concentration gradient where zero concentration is reached at lower limit;
dfrs		working value of dfl;
djll		error in current at cathode, produced by lower limit anode free hole concentration;
djul		error in current at cathode, produced by upper limit anode free hole concentration;
dl		interpolation increment in finding new bounds for anode free hole concentration;
e	*	error margin for predictor-corrector process;
emu	*	output factor for voltage, $e\mu$;
f		array holding free hole concentration at fixed crystal points;
F	*	diffusion factor, $1/\mu kT$;
figa		output format;
figb		output format;
figd		output format;
fige		output format;

fpu free hole concentration where zero concentration gradient
 is reached at upper limit;
gt cumulative rounding error in prediction-correction;
h step length;
hf * factor determining number of steps;
hfl value of hf required by lower limit;
hfu value of hf required by upper limit;
hg * initial value of hf, retained;
hh * crystal length divided by ten;
i count for number of loops in bridge;
int array holding values of recombination integral at fixed
 crystal points;
j * value of current;
jea electron leakage current at anode;
jhsc extra current which should flow across cathode due to hole
 concentration there above what cathode will accept;
k * number of values of current for computation;
kk working value of k;
knc * cathode electron concentration;
kpa * anode hole concentration;
l working value of step number;
ll total number of corrections in deriving lower limit;
ll * lower limit value of anode free hole concentration;
ls number of interpolations in deriving new limits;
m number of steps;
mm $m \div 10$ (integer division);
mt working value of mm;
n * number of parameter sets for computation;
nl interpolated value of free hole concentration at anode;
nn loop variable for n;
p position variable during bridge procedure;
pa starting free hole concentration at anode;

pfj * multiplying factor for pa from result of previous calculation (for working set of iV points);

pp position where bridge procedure is abandoned;

ppl lower limit value of pp;

ppu upper limit value of pp;

q main loop variable;

r array holding values of recombination function at fixed crystal points;

s * trap recombination coefficient;

sl store holding reason for bridge abandonment;

span procedure for single step;

sw * switch to determine what set of initial values is to be read in;

t * number of decimal places to be obtained;

trip criterion for boundary condition for trap saturation front;

u free hole concentration at anode at bridge call;

ui total number of corrections in deriving upper limit;

ul upper limit value of anode free hole concentration;

v total voltage;

w current value for free hole concentration during bridge;

x initial value for free hole concentration during bridge;

y value of anode free hole concentration local to bridge;

yy value of anode free hole concentration local to span;

z next value for free hole concentration during bridge;

References.

- Abe Y., Kikuchi M. (463)
1962 J. Phys. Soc. Jap. 17, 241.
- Abe Y., Kikuchi M. (464)
1962 J. Phys. Soc. Jap. 17, 881.
- Abe Y., Kikuchi M. (415, 465, 469, 470, 504)
1962 J. Phys. Soc. Jap. 17, 1268.
- Abe Y., Kikuchi M., Chong P. (466)
1963 Jap. J. Appl. Phys. 2, 440.
- Ando K., Steele M.C., Lampert M.A. (412, 467, 471)
1962 J. Phys. Soc. Jap. 17, 1729.
- Ando K., Steele M.C., Lampert M.A. (413, 468)
1963 J. Phys. Soc. Jap. 18, 591.
- Archey W.B. (303)
1964 Proc. I.E.E.E. 52, 1657.
- Arthur J.B., Gibson A.F., Gunn J.B. (432)
1956 Proc. Phys. Soc. 69B, 697.
- Ashley K.L., Milnes A.G. (410, 438)
1964 J. Appl. Phys. 35, 369.
- Banbury P.C., Henisch H.K., Many A. (450)
1953 Proc. Phys. Soc. 66A, 753.
- Barraud A. (606, 644)
1963 C. R. Acad. Sci. Paris 256, 3632.
- Baur M., Jordan J.R., Jordan P.C., Mayer J.E. (238)
1965 Ann. Phys. 35, 96.
- Beaumont C.F.A. (223)
1965 Am. J. Phys. 33, 547.
- Bevacqua S.F., Holonyak N., Ing S.W., Thomas R.C. (409)
1962 Phys. Rev. Lett. 8, 426.

- Blount E.I. (245)
1963 Phys. Rev. 131, 2354.
- Boer K.W., Wilhelm W.E. (609)
1964 Phys. Stat. Sol. 4, 237.
- Bohm D., Gross E.P. (633)
1949 Phys. Rev. 75, 1864.
- Brown R.D., Koenig S.H., (533, 702)
1959 J. Phys. Chem. Sol. 10, 201.
- Brown R.D., Koenig S.H. (108, 114)
1960 Phys. Rev. Lett. 4, 170.
- Brown R.D., Koenig S.H., Schillinger W. (P5)(513, 514, 515, 523)
1962 Phys. Rev. 128, 1668.
- Bube R.H. (437)
1960 Photoconductivity of Solids, Wiley, New York.
- Burgess R.E. (440)
1960 Intern. Conf. Semicond. Phys., Prague p.818.
- Burstein E., Solar N. (506, 507)
1957 J. Phys. Chem. Sol. 2, 1.
- Butcher P.N. (650, 651)
1965 Phys. Lett. 19, 546.
- Butcher P.N., Fawcett W. (615)
1965 Phys. Lett. 17, 216.
- Carroll J.E. (653)
1965 Electr. Lett. 1, 189.
- Chong P., Abe Y., Kikuchi M. (466)
1963 Jap. J. Appl. Phys. 2, 440.
- Chynoweth A.G., McCumber D.E. (634, 638, 642)
1964 Phys. Rev. Lett. 13, 651.
- Chynoweth A.G., McCumber D.E. (658)
1966 I.E.E.E. Trans. El. Dev. ED13, 4.
- Corsan J.M., Goldsmit H.J. (711, 803)
1964 Phys. Lett. 8, 221.

- Datars W.R., Eastman P.C. (P7)(701, 703, 707, 708)
1963 Canad. J. Phys. 41, 161.
- de Groot S.R. (P2)(201, 205, 209, 234, 235, 236)
1959 Thermodynamics of Irreversible Processes, North
Holland Publishing Company, Amsterdam.
- Domenicali C.A. (240)
1953 Phys. Rev. 92, 877.
- Domenicali C.A. (241)
1954 J. Appl. Phys. 25, 1310.
- Dorn W.S., McCracken D.D. (452)
1964 Numerical Methods and Fortran Programming,
Wiley, New York.
- Dunlap W.C. (443)
1957 An Introduction to Semiconductors, Wiley,
New York.
- Eastman P.C., Datars W.R. (P7)(701, 703, 707, 708)
1963 Canad. J. Phys. 41, 161.
- Ehrenfest P., Ehrenfest T. (204)
1911 Encyc. Math. Weiss. 4, 32.
- Esaki L. (802)
1962 Phys. Rev. Lett. 8, 4.
- Fawcett W., Butcher P.N. (615)
1965 Phys. Lett. 17, 216.
- Fischler-Hazoni C., Williams F.E. (610, 647, 713)
1965 Phys. Rev. 139, A583.
- Flood W.F., Haynes J.R., Lax M. (109)
1959 J. Phys. Chem. Sol. 8, 392.
- Fowler R.H. (449)
1936 Statistical Mechanics, 2nd Edition, Cambridge U.P.
- Fox L. (451)
1962 Numerical Solution of Ordinary and Partial
Differential Equations, Pergamon Press, Oxford.

- Foyt A.G., McWhorter A.L. (608, 656)
1966 I.E.E.E. Trans. on El. Dev. ED13, 79.
- Gibson A.F., Arthur J.B., Gunn J.B. (432)
1956 Proc. Phys. Soc. 69B, 697.
- Gippius A.A., Vavilov V.S. (113)
1965 Sov. Phys. Sol. St. 6, 1873.
- Goldsmid H.J., Corsan J.M. (711, 803)
1964 Phys. Lett. 8, 221.
- Gross E.P., Bohm D. (633)
1949 Phys. Rev. 75, 1864.
- Gunn J.B. (502, 503)
1956 Proc. Phys. Soc. 69B, 781.
- Gunn J.B. (423, 426)
1958 J. Electr. and Contr. 4, 17.
- Gunn J.B. (604, 607, 613)
1964 I.B.M. Journal Res. Dev. 8, 141.
- Gunn J.B. (605, 614, 646)
1964 Plasma Physics in Solids, Dunod, Paris.
- Gunn J.B., Arthur J.B., Gibson A.F. (432)
1956 Proc. Phys. Soc. 69B, 697.
- Gunther-Mohr G.R., Koenig S.H. (508)
1957 J. Phys. Chem. Sol. 2, 268.
- Gurney R.W., Mott N.F. (402)
1940 Electron Processes in Ionic Crystals, Oxford U.P.
- Hakki B.W., Knight S. (655, 657)
1966 I.E.E.E. Trans. El. Dev. ED13, 94.
- Haynes J.R., Lax M., Flood W.F. (109)
1959 J. Phys. Chem. Sol. 8, 392.
- Heeks J.S. (654)
1966 I.E.E.E. Trans. El. Dev. ED13, 68.
- Honisch H.K. (405)
1957 Rectifying Semiconductor Contacts, Oxford U.P.

- Henisch H.K., Banbury P.C., Many A. (450)
1953 Proc. Phys. Soc. 66A, 753.
- Hilsun C. (602)
1962 Proc. I.R.E. 50, 185.
- Holland M.G. (532, 706)
1964 Phys. Rev. 134, A471.
- Holonyak N. (408, 460)
1962 Proc. I.R.E. 50, 2421.
- Holonyak N., Ing S.W., Thomas R.C., Bevacqua S.F. (409)
1962 Phys. Rev. Lett. 8, 426.
- Howard N.R., Johnson G.W. (427)
1965 Sol. St. El. 8, 275.
- Howarth L.E., Kaiser W., Spitzer W.G., Kaiser R.H. (714)
1962 Phys. Rev. 127, 1950.
- Hurle D.T.J., Mullin J.B., Pike E.R. (210, 242)
1965 J. Chem. Phys. 42, 1651.
- Hurle D.T.J., Mullin J.B., Pike E.R. (243)
1965 J. Chem. Phys. 43, 3420.
- Hutson A.R., McFee J.H., White D.L. (626, 628, 629)
1961 Phys. Rev. Lett. 7, 237.
- Ing S.W., Holonyak N., Thomas R.C., Bevacqua S.F. (409)
1962 Phys. Rev. Lett. 8, 426.
- Ivanov Y.L., Yuhnevich A.V. (112)
1965 Sov. Phys. Sol. St. 6, 2965.
- Jeans J. (225, 227, 228)
1925 Electricity and Magnetism, Cambridge U.P.
- Jeffreys H. (230, 232)
1961 Theory of Probability, Oxford U.P.
- Johnson G.W., Howard N.R. (427)
1965 Sol. St. El. 8, 275.
- Jordan J.R., Baur M., Jordan P.C., Mayer J.E. (238)
1965 Ann. Phys. 35, 96.

- Jordan P.C., Baur M., Jordan J.R., Mayer J.E. (238)
1965 Ann. Phys. 35, 96.
- Kaiser W., Spitzer W.G., Kaiser R.H., Howarth L.E. (714)
1962 Phys. Rev. 127, 1950.
- Kaiser R.H., Spitzer W.G., Kaiser W., Howarth L.E. (714)
1962 Phys. Rev. 127, 1950.
- Kalashnikov S.G., Morozov A.I. (117)
1961 Sov. Phys. Sol. St. 3, 2520.
- Karakushan E.I., Stafeev V.I., (420, 458)
1961 Sov. Phys. Sol. St. 3, 493.
- Karakushan E.I., Stafeev V.I. (421, 459)
1962 Sov. Phys. Sol. St. 3, 1476.
- Karpova I.V., Pokrovskii Y.E. (111)
1965 Sov. Phys. Sol. St. 6, 2905.
- Kikuchi M., Abe Y. (463)
1962 J. Phys. Soc. Jap. 17, 241.
- Kikuchi M., Abe Y. (464)
1962 J. Phys. Soc. Jap. 17, 381.
- Kikuchi M., Abe Y. (415, 465, 469, 470, 504)
1962 J. Phys. Soc. Jap. 17, 1268.
- Kikuchi M., Abe Y., Cheng P. (466)
1963 Jap. J. Appl. Phys. 2, 440.
- Kinch M.A. (539)
1964 Thesis, Oxford.
- Kittel C. (215)
1958 Elements of Statistical Physics, Wiley, New York.
- Klein M.J. (212)
1955 Phys. Rev. 98, 1736.
- Klein M.J. (213, 214)
1959 Proc. of the Intern. School of Phys. "Enrico Fermi" on the Thermodynamics of Irreversible Processes, p.198, Bologna.

- Klein M.J., Meijer P.H.E. (211)
1954 Phys. Rev. 96, 250.
- Knight S., Hakki B.W. (655, 657)
1966 I.E.E.E. Trans. El. Dev. ED13, 94.
- Koenig S.H., Brown R.D. (533, 702)
1959 J. Phys. Chem. Sol. 10, 201.
- Koenig S.H., Brown R.D. (108, 114)
1960 Phys. Rev. Lett. 4, 170.
- Koenig S.H., Brown R.D., Schillinger W. (P5)(513, 514, 515, 523)
1962 Phys. Rev. 128, 1668.
- Koenig S.H., Gunther-Mohr G.R. (508)
1957 J. Phys. Chem. Sol. 2, 268.
- Kogan E.M. (526, 527, 528, 540)
1963 Sov. Phys. Sol. St. 4, 1813.
- Kohn W. (103)
1957 Sol. St. Phys. Vol. 5 Academic Press, New York.
- Komatsubara K., Kurono H. (414)
1962 J. Phys. Soc. Jap. 17, 883.
- Kramareva S.A., Stafeev V.I. (417, 419, 457)
1960 Sov. Phys. Sol. St. 2, 350.
- Kromer H. (801)
1961 Prog. in Semicond. 4, 1 Heywood, London.
- Kurono H., Komatsubara K. (414)
1962 J. Phys. Soc. Jap. 17, 883.
- Kurosawa T. (530)
1962 Jap. J. Appl. Phys. 1, 365.
- Kurosawa T. (519, 529, 531, 534, 537)
1965 J. Phys. Soc. Jap. 20, 1405.
- Lampert M.A. (P4)(401, 403, 404, 406, 407, 430, 455)
1964 Rep. on Prog. in Phys. 329 27 Phys. Soc.,
London.

- Lampert M.A., Steele M.C., Ando K.
1962 J. Phys. Soc. Jap. 17, 1729. (412, 467, 471)
- Lampert M.A., Steele M.C., Ando K.
1963 J. Phys. Soc. Jap. 18, 591. (413, 468)
- Landauer R., Swanson J.
1953 Phys. Rev. 91, 555. (444, 448)
- Langmuir I., Tonks L.
1929 Phys. Rev. 33, 195. (632)
- Lax M.
1959 J. Phys. Chem. Sol. 8, 66. (107, 522)
- Lax M., Haynes J.R., Flood W.F.
1959 J. Phys. Chem. Sol. 8, 392. (109)
- Lindmayer J., Slobodsky A.
1963 Sol. St. El. 6, 495. (434)
- Love W.F., Phelan R.J.
1964 Phys. Rev. 133, A1134. (511, 538)
- McCracken D.D., Dorn W.S.
1964 Numerical Methods and Fortran Programming,
Wiley, New York. (452)
- McCumber D.E., Chynoweth A.G.
1964 Phys. Rev. Lett. 13, 651. (634, 638, 642)
- McCumber D.E., Chynoweth A.G.
1966 I.E.E.E. Trans. El. Dev. ED13, 4. (658)
- McFee J.H., Hutson A.R., White D.L.
1961 Phys. Rev. Lett. 7, 237. (626, 628, 629)
- McWhorter A.L., Foyt A.G.
1966 I.E.E.E. Trans. El. Dev. ED13, 79. (608, 656)
- McWhorter A.L., Rediker R.H.
1959 Proc. I.R.E. 47, 1207. (509)
- McWhorter A.L., Rediker R.H.
1960 Intern. Conf. Semicond. Phys., Prague p.134. (510, 516, 520, 521, 535)

- Mahrous S.M., Robson P.N. (652)
 1965 Proc. Joint Symp. Microwave Appl. Semicond.,
 I.R.E., I.E.E. Suppl. Vol.2, 32/1.
- Many A., Banbury P.C., Henisch H.K. (450)
 1953 Proc. Phys. Soc. 66A, 753.
- Mayer J.E., Baur M., Jordan J.R., Jordan P.C. (238)
 1965 Ann. Phys. 35, 96.
- Mazur P. (229)
 1959 Proc. of the Intern. School of Phys. "Enrico
 Fermi" on the Thermodynamics of Irreversible
 Processes, p.167, Bologna.
- Meijer P.H.E., Klein M.J. (211)
 1954 Phys. Rev. 96, 250.
- Melngailis I., Rediker R.H. (411, 454, 461)
 1962 J. Appl. Phys. 33, 1892.
- Mills J.R., Rollin B.V., Russell J.P. (305)
 1960 Cryogenics 1, 75.
- Milnes A.G., Ashley K.L. (410, 438)
 1964 J. Appl. Phys. 35, 369.
- Morozov A.I., Kalashnikov S.G. (117)
 1961 Sov. Phys. Sol. St. 3, 2520.
- Mott N.F., Gurney R.W. (402)
 1940 Electronic Processes in Ionic Crystals, Oxford U.P.
- Mullin J.B., Hurle D.T.J., Pike E.R. (210, 242)
 1965 J. Chem. Phys. 42, 1651.
- Mullin J.B., Hurle D.T.J., Pike E.R. (243)
 1965 J. Chem. Phys. 43, 3420.
- Musha T. (636, 639)
 1962 Phys. Fluids 5, 1310.
- Musha T., Yoshida F. (637, 640, 641)
 1964 Phys. Rev. 133, A1303.

- Nannay C. (704, 709, 710)
1962 Appl. Phys. Lett. 1, 71.
- Ohara S. (635, 643)
1963 J. Phys. Soc. Jap. 18, 852.
- Paige E.G.S. (453, 525, 630)
1964 Prog. in Semicond. Vol.8 Temple Press, London.
- Pankova J.I. (105)
1965 Prog. in Semicond. Vol.9, 47, Temple Press, London.
- Paranjape V.V. (512)
1961 Proc. Phys. Soc. 78, 516.
- Petford A.D. (302)
1958 Thesis, Oxford.
- Phelan R.J., Love W.F. (511, 538)
1964 Phys. Rev. 133, A1134.
- Pike E.R., Hurle D.T.J., Mullin J.B. (210, 242)
1965 J. Chem. Phys. 42, 1651.
- Pike E.R., Hurle D.T.J., Mullin J.B. (243)
1965 J. Chem. Phys. 43, 3420.
- Pokrovskii Y.E., Karpova I.V. (111)
1965 Sov. Phys. Sol. St. 6, 2905.
- Pokrovskii Y.E., Svislanova K.I. (110)
1964 Sov. Phys. Sol. St. 6, 13.
- Pratt R.G., Ridley B.K. (618, 705)
1963 Proc. Phys. Soc. 81, 996.
- Pratt R.G., Ridley B.K. (621, 645, 660)
1963 Phys. Lett. 4, 300.
- Pratt R.G., Ridley B.K. (620, 712)
1965 J. Phys. Chem. Sol. 26, 11.
- Pratt R.G., Ridley B.K. (603, 622, 661)
1965 J. Phys. Chem. Sol. 26, 21.
- Pratt R.G., Ridley B.K. (619)
1965 Proc. Phys. Soc. 85, 293.

- Prigogine I. (244)
1955 Introduction to Thermodynamics of Irreversible
Processes, C.C. Thomas, Springfield,
Illinois, USA.
- Prigogine I., Severne G. (237)
1963 Phys. Lett. 6, 177.
- Putley E.H. (301, 662)
1960 The Hall Effect and Related Phenomena,
Butterworth, London.
- Read W.T., Shockley W. (436)
1952 Phys. Rev. 87, 835.
- Rediker R.H., McWhorter A.L. (509)
1959 Proc. I.R.E. 47, 1207.
- Rediker R.H., McWhorter A.L. (510, 516, 520, 521, 535)
1960 Intern. Conf. Semicond. Phys., Prague p.134.
- Rediker R.H., Melngailis I. (411, 454, 461)
1962 J. Appl. Phys. 33, 1892.
- Ridley B.K. (P1)(101, 202, 203, 207, 208, 216)
1963 Proc. Phys. Soc. 82, 954.
- Ridley B.K. (648, 649, 659)
1965 Proc. Phys. Soc. 86, 637.
- Ridley B.K., Pratt R.G. (618, 705)
1963 Proc. Phys. Soc. 81, 996.
- Ridley B.K., Pratt R.G. (621, 645, 660)
1963 Phys. Lett. 4, 300.
- Ridley B.K., Pratt R.G. (620, 712)
1965 J. Phys. Chem. Sol. 26, 11.
- Ridley B.K., Pratt R.G. (603, 622, 661)
1965 J. Phys. Chem. Sol. 26, 21.
- Ridley B.K., Pratt R.G. (619)
1965 Proc. Phys. Soc. 85, 293.

- Ridley B.K., Watkins T.B. (P6)(102, 601, 611, 612)
1961 Proc. Phys. Soc. 78, 293.
- Ridley B.K., Watkins T.B. (106, 616)
1961 Proc. Phys. Soc. 78, 710.
- Ridley B.K., Watkins T.B. (617)
1961 J. Phys. Chem. Sol. 22, 155.
- Robson P.N., Mahrous S.M. (652)
1965 Proc. Joint Symp. on Microwave Appl. Semicond.,
I.R.E., I.E.E. Suppl. Vol.2, 32/1.
- Rollin B.V., Mills J.R., Russell J.P. (305)
1960 Cryogenics 1, 75.
- Rollin B.V., Russell J.P. (307, 309)
1963 Proc. Phys. Soc. 81, 571.
- Rowell J.M. (308)
1961 Thesis, Oxford.
- Russell J.P. (304, 306)
1962 Thesis, Oxford.
- Russell J.P., Rollin B.V. (307, 309)
1963 Proc. Phys. Soc. 81, 571.
- Russell J.P., Rollin B.V., Mills J.R. (305)
1960 Cryogenics 1, 75.
- Sanders T.R., Weinreich G., White H.G. (624, 625)
1959 Phys. Rev. 114, 33.
- Schillinger W., Koenig S.H., Brown R.D. (P5)(513, 514, 515, 523)
1962 Phys. Rev. 128, 1668.
- Sclar N., Burstein E. (506, 507)
1957 J. Phys. Chem. Sol. 2, 1.
- Severne G., Prigogine I. (237)
1963 Phys. Lett. 6, 177.
- Shockley W. (447)
1950 Electrons and Holes in Semiconductors,
Van Nostrand, New York.

- Shockley W., Read W.T. (436)
1952 Phys. Rev. 87, 835.
- Shockley W., Suhl H. (439, 445)
1949 Phys. Rev. 75, 1617.
- Slobodskoy A., Lindmayer J. (434)
1963 Sol. St. El. 6, 495.
- Smith R.A. (422, 424, 425, 431, 446, 501, 505, 631)
1959 Semiconductors, Cambridge U.P.
- Sondaevskii V.P., Stafeev V.I. (429)
1964 Sov. Phys. Sol. St. 6, 63.
- Spector H.N. (627)
1962 Phys. Rev. 127, 1084.
- Spitzer W.G., Kaiser W., Kaiser R.H., Howarth L.E. (714)
1962 Phys. Rev. 127, 1950.
- Stafeev V.I. (416, 418, 428, 456, 462)
1959 Sov. Phys. Sol. St. 1, 763.
- Stafeev V.I., Karakushan E.I. (420, 458)
1961 Sov. Phys. Sol. St. 3, 493.
- Stafeev V.I., Karakushan E.I. (421, 459)
1962 Sov. Phys. Sol. St. 3, 1476.
- Stafeev V.I., Kramareva S.A. (417, 419, 457)
1960 Sov. Phys. Sol. St. 2, 350.
- Stafeev V.I., Sondaevskii V.P. (429)
1964 Sov. Phys. Sol. St. 6, 63.
- Steele M.C., Ando K., Lampert M.A. (412, 467, 471)
1962 J. Phys. Soc. Jap. 17, 1729.
- Steele M.C., Ando K., Lampert M.A. (413, 468)
1963 J. Phys. Soc. Jap. 18, 591.
- Suhl H., Shockley W. (439, 445)
1949 Phys. Rev. 75, 1617.
- Swistunova K.I., Pokrovskii Y.E. (110)
1964 Sov. Phys. Sol. St. 6, 13.

- Swanson J., Landauer R. (444, 448)
 1953 Phys. Rev. 91, 555.
- ter Haar D. (219)
 1954 Elements of Statistical Mechanics, Rinehart,
 New York.
- Thomas R.C., Holonyak N., Ing S.W., Bevacqua S.F. (409)
 1962 Phys. Rev. Lett. 8, 426.
- Tonks L., Langmuir I. (632)
 1929 Phys. Rev. 33, 195.
- Tyler W.W. (442)
 1954 Phys. Rev. 96, 226.
- Tyler W.W. (104)
 1959 J. Phys. Chem. Sol. 8, 59.
- Tyler W.W., Woodbury H.H. (115, 116, 441)
 1956 Phys. Rev. 102, 647.
- van Vliet K.M. (435)
 1966 Sol. St. El. 9, 185.
- Vavilov V.S., Gippius A.A. (113)
 1965 Sov. Phys. Sol. St. 6, 1873.
- von Mises R. (231, 233)
 1964 Mathematical Theory of Probability and
 Statistics, Academic Press, New York.
- Watkins T.B., Ridley B.K. (P6)(102, 601, 611, 612)
 1961 Proc. Phys. Soc. 78, 293.
- Watkins T.B., Ridley B.K. (106, 616)
 1961 Proc. Phys. Soc. 78, 710.
- Watkins T.B., Ridley B.K. (617)
 1961 J. Phys. Chem. Sol. 22, 155.
- Webster W.M. (433)
 1954 Proc. I.R.E. 42, 914.
- Weinreich G., Sanders T.R., White H.G. (624, 625)
 1959 Phys. Rev. 114, 33.

- White D.L., Hutson A.R., McFee J.H. (626, 628, 629)
1961 Phys. Rev. Lett. 7, 237.
- White H.G., Weinreich G., Sanders T.R. (624, 625)
1959 Phys. Rev. 114, 33.
- Wilhelm W.E., Boer K.W. (609)
1964 Phys. Stat. Sol. 4, 237.
- Williams F.E., Fischler-Hazoni C. (610, 647, 713)
1965 Phys. Rev. 139, A583.
- Woodbury H.H., Tyler W.W. (115, 116, 441)
1956 Phys. Rev. 102, 647.
- Yamashita J. (517, 518, 536)
1961 J. Phys. Soc. Jap. 16, 720.
- Yoshida F., Musha T. (637, 640, 641)
1964 Phys. Rev. 133, A1303.
- Yukhnovich A.V., Ivanov Y.L. (112)
1965 Sov. Phys. Sol. St. 6, 2965.
- Ziman J.M. (P3)(217, 221, 224, 226, 239)
1956 Canad. J. Phys. 34, 1256.
- Ziman J.M. (206, 218, 220, 222, 524, 623)
1960 Electrons and Phonons, Oxford U.P.

Acknowledgements.

I wish to thank:

Professor B. Bleaney, F. R. S., for permitting me to carry out this
work in the Clarendon Laboratory;

Professor L. Fox for allowing me time on the University KDF9 computer;

Dr. M. A. Finch, Mr. G. D. Peckett and Mr. J. H. Morrissy for assist-
ance in the design and construction of apparatus
and in the investigation and interpretation of
results;

Dr. P. N. Butcher of R. R. E., Malvern, Dr. B. K. Midley of Essex
University and Mr. G. H. B. Thompson for discus-
sions which gave rise to particular sections;

The Science Research Council for a research studentship;

and finally my supervisor Dr. B. V. Rollin for assistance in this work.

C. O. Newton