DIFUSÃO DE ENXOFRE EM ARSENETO DE GÁLIO POR PROCESSAMENTO TÉRMICO RÁPIDO

AUTOR : ALEXANDRE SANSIGOLO LUJAN

ORIENTADOR : PROF. DR. FRANCISCO CARLOS DE PRINCE

Dissertação apresentada no Instituto de Física "Gleb Wataghin" da UNICAMP para a obtenção do Título de "Mestre em Ciências".

UNIVERSIDADE ESTADUAL DE CAMPINAS - UNICAMP

INSTITUTO DE FÍSICA "GLEB WATAGHIN"
DEPARTAMENTO DE FÍSICA APLICADA
LABORATÓRIO DE PESQUISAS EM DISPOSITIVOS - LPD

Janeiro de 1991

Este Exemplar corresponde à redação final da Tese defendida pelo aluno Alexandre Lujan no IF-UNICAMP, e aprovada pela comissão julgadora.

23-07-91

UNICAMP
BIBLIOTECA CENTRAL

CM00013738-1

UNIDADE I F
Nº CHAMADA T/UNICAMP/TM
L968d
V. ..... EX. .....
TOMBO BC/ 14294
PROC. 308/91
C [ ] D [X]
PRECO CR$ 500,00
DATA 7-5-91
Nº CPD .....

M

I. FÍSICA - UNICAMP
nº ...
nº ...
nº tombo. TM1826
23/01/92

À minha esposa, Fátima, aos meus pais, Nelson e Maria Lúcia, à minha avó, Paula, e ao meu irmão, Guilherme.

AGRADECIMENTOS

À minha esposa, Fátima por todo o carinho e apoio para a realização deste trabalho.

Aos meus pais, Nelson e Maria Lúcia, à minha avo, Paula, ao meu irmão Guilherme, pelo carinho e apoio.

Ao meu orientador, Chico, pela sua orientação, dedicação e paciência.

Ao Celso e Roberto, pela ajuda na montagem do forno, e principalmente pela amizade.

Ao Emílio pela montagem dos controladores de potência e termômetro.

À Cristiane pelas medidas de efeito Hall e companheirismo nos congressos.

Ao Régis da Telebrás pelas medidas de C-V.

Ao Pedrinho e Laerte da oficina mecânica pela confecção das peças do forno.

Ao Rogério, Totó e Elaine pelas ajudas nas fotogravações e evaporações de metais.

Ao André, Eduardo, Zé Luiz e Márcia da administração.

À Dina e Eduardo da biblioteca.

À FAPESP e Telebras pelas bolsas concedidas.

À todos os outros que ajudaram direta ou indiretamente neste trabalho.

RESUMO

Neste trabalho nos desenvolvemos uma nova técnica de difusão de S em GaAs. A técnica usa um processador térmico rápido (RTP), para difusões de tempo curto (15-90s). Camadas com alta concentração ($3\times10^{18}cm^{-3}$) e alta mobilidade ($2000cm^2/Vs$) foram obtidas usando esta técnica. FET's de arseneto de gálio foram fabricados e caracterizados. Transcondutâncias de 160mS/mm, para um comprimento de "gate" de 3μm, e resistências série de 1Ωmm foram obtidas. Os resultados mostram que a técnica desenvolvida é capaz de produzir dispositivos para uso prático.

ABSTRACT

In this work we develope a new diffusion technique of S in GaAs. The technique uses a rapid thermal processor (RTP), for a very short time diffusions (15-90s). High concentration ($3\times10^{18}cm^{-3}$) and high mobility ($2000cm^2/Vs$) layers were obtained using this technique. Gallium arsenide FET's were fabricated and characterized. Transcondutances of 160mS/mm, for gate length of 3μm, and series resistances per gate width of 1Ωmm were obtained. The results show that the technique developed is capable of yield devices for practical use.

SUMÁRIO

CAPÍTULO I - INTRODUÇÃO

1.1 - Motivação 1
1.2 - Materiais e Técnivas Usados nos Dispositivos Rápidos ... 1
1.3 - Trabalho Desenvolvido 2
1.4 - Técnica de Difusão Desenvolvida 4
1.5 - Transistores MESFET 5
1.6 - Referências 5

CAPÍTULO II - FORNO RÁPIDO

2.1 - Introdução 6
2.2 - Mecanismos de Aquecimento em um Forno 6
2.3 - Tipos de Fornos Rápidos 7
2.4 - Aplicação do Processamento Térmico Rápido (RTP) 9
2.5 - Projeto do Forno 9
2.5.1 - Parte Mecânica do Forno 10
2.5.1.1 - Câmara de Aquecimento 12
2.5.1.2 - Refrigeração da Câmara 13
2.5.1.3 - Controle da Atmosfera no Interior do Tubo de Quatzo 14
2.5.2 - Circuitos Eletrônicos 15
2.5.2.1 - Controlador de Potência das Lâmpadas 17
2.5.2.2 - Conversor A/D e D/A e Microcomputador 17
2.5.2.3 - Sensor de Temperatura 19
2.5.3 - Programa de Controle 21
2.6 - Testes e Caracterização do Forno 24
2.7 - Referências 28

CAPÍTULO III - DIFUSÃO DE IMPUREZAS NO GaAs

3.1 - Introdução 29
3.2 - Semicondutores e Impurezas 30

3.2.1 - Importância da Difusão em Semicondutores 33
3.3 - Defeitos Pontuais em Semicondutores Compostos 36
3.3.1 - Tipos de Defeitos Pontuais em Semicondutores Compostos 37
3.4 - Lei de Ação de Massas 40
3.4.1 - Reações de Formação de Defeitos Pontuais 42
3.5 - Diagramas de Fase 46
3.5.1 - Diagramas Binários 48
3.5.2 - Diagramas Ternários 50
3.6 - Difusão em Semicondutores 53
3.6.1 - Mecanismos de Difusão em Semicondutores 57
3.6.2 - Cálculo do Coeficiente de Difusão 62
3.7 - Difusão de Enxofre em Arseneto de Gálio 67
3.7.1 - Diagramas de Fase, Pressões de Vapor de Equilíbrio e Concentrações de Vacâncias no GaAs 67
3.7.2 - Trabalhos sobre a Difusão de Enxofre em Arseneto de Gálio 70
3.7.3 - Mecanismo Físico de Difusão de Enxofre em GaAs 74
3.8 - Conclusões 81
3.9 - Referências 81

CAPÍTULO IV - DIFUSÃO DE ENXOFRE EM GaAs : EXPERIMENTO

4.1 - Introdução 85
4.2 - Projeto do Porta-Amostra 85
4.2.1 - Projeto do Suporte de Grafite 88
4.2.2 - Suporte de Quartzo 91
4.3 - Ciclo Térmico 93
4.4 - Escolha e Preparação das Amostras 97
4.5 - Escolha e Preparação da Fonte de Difusão 98
4.6 - Determinação das Faixas de Temperaturas Utilizáveis 103
4.7 - Caracterização das Amostras 106
4.8 - Resultados Obtidos 113
4.9 - Análise dos Resultados 125
4.10 - Conclusões 132
4.11 - Referências 133

CAPÍTULO V - TRANSISTORES DE EFEITO DE CAMPO DE GaAs

5.1 - Introdução 134
5.2 - Aspectos Gerais 136
5.3 - Característica Velocidade x Campo dos Elétrons no GaAs . 139
5.4 - Saturação de Corrente no MESFET de GaAs 141
5.5 - Modelo do Canal Parcialmente Saturado com Perfil Não Uniforme de Dopagem 149
5.5.1 - Região Linear 151
5.5.2 - Região de Saturação 152
5.6 - Conclusões 158
5.7 - Referências 158

CAPÍTULO VI - FABRICAÇÃO E TESTES DOS TRANSISTORES

6.1 - Introdução 160
6.2 - Características Físicas dos Transistores Fabricados 160
6.2.1 - Escolha dos Cristais Difundidos 162
6.3 - Tecnologia Empregada na Fabricação 163
6.3.1 - Etapas de Processo 163
6.3.2 - Fotogravação Normal 164
6.3.3 - Fotogravação para "Lift-Off" 166
6.4 - Formação das "Mesas" 168
6.4.1 - Marcas de Alinhamento 169
6.4.2 - "Mesas" 170
6.5 - Metalização de Contato Ôhmico 172
6.5.1 - Metalização Utilizada 173
6.5.2 - "Alloy" de Contato 174
6.6 - Metalização de "Gate" 174
6.6.1 - Ajuste da Corrente de Saturação 175
6.6.2 - Metalização Utilizada para o "Gate" 176
6.7 - Metalização dos "Pads" 178
6.8 - Testes e Caracterização dos Transistores 179
6.9 - Conclusões 185
6.10 - Referências 186

CAPÍTULO VII - CONCLUSÕES E TRABALHOS PROPOSTOS

7.1 - Introdução 188
7.2 - Conclusões 188
7.3 - Trabalhos Propostos 190
7.3.1 - Difusão de Zinco em GaAs 190
7.3.2 - Difusão de Enxofre e Zinco em InP 191
7.3.3 - Difusão de Zinco em InGaAs 192
7.3.4 - Difusão Localizada de Enxofre em GaAs 192
7.3.5 - Desenvolvimento de uma Fonte Sólida de Enxofre para Difusão 193
7.3.6 - "Alloy" de AuGeNi 194
7.4 - Referências 194

APÊNDICE

Programas Desenvolvidos para Análise das Difusões e dos Transistores 196
Programa COEF_DIF 197
Programa ANALISBM 199
Programa FINDVPO 201
Programa SIGMA 203

CAPITULO I

INTRODUÇÃO

1.1 - MOTIVAÇÃO

O estudo de dispositivos rápidos e de suas técnicas de fabricação tem, a cada dia que passa, uma maior importância. Entende-se por dispositivos rápidos, transistores, diodos e circuitos integrados que operam em frequências maiores que a dos dispositivos convencionais. Devido a essa característica os dispositivos rápidos encontram aplicações típicas em circuitos de alta frequência de operação para equipamentos de telecomunicações e circuitos lógicos de alta velocidade para computadores digitais, entre outras. São dessas aplicações que vem a importância de se estudar os dispositivos rápidos e de aprimorar as suas técnicas de fabricação. Neste trabalho vamos apresentar uma nova técnica de difusão de enxofre em arseneto de gálio por processamento térmico rapido, e a sua aplicação à fabricação de um transistor MESFET para uso em circuitos de alta velocidade.

1.2 - MATERIAIS E TÉCNICAS USADOS NOS DISPOSITIVOS RÁPIDOS

Os dispositivos rápidos são geralmente fabricados usando-se materiais semicondutores formados por elementos químicos dos grupos IIIB e VB da tabela periódica. A esses materiais damos o nome de semicondutores compostos do grupo III-V. Destes materiais podemos citar como exemplo o arseneto de gálio (GaAs), o fosfeto de índio (InP) ou compostos desses dois (por exemplo, InGaAs, InGaAsP, etc.). A escolha desses materiais na fabricação ao invés de silício (Si) ou germânio (Ge), que são usados nos dispositivos convencionais, se deve ao fato de que os portadores de corrente possuem maior mobilidade em baixo campo nos compostos III-V do que

no silício ou germânio (essa regra não é geral, temos por exemplo o GaP, que possue uma mobilidade para os elétrons de 110 cm^2/Vs, que é menor do que a do silício, que é de 1500 cm^2/Vs, para amostras não dopadas [1]). A maior mobilidade nos compostos III-V é responsavel pela maior frequência de operação dos dispositivos, além disso ela ainda acarreta em um maior produto "potência x velocidade". O significado desse produto é que um dispositivo de silício (que possui menor produto "potência x velocidade") irá consumir maior potência do que um dispositivo feito com um composto III-V para operar na mesma frequência que este último opera. Resumidamente podemos dizer que dispositivos fabricados com compostos III-V além da maior velocidade de operação ainda possuem um menor consumo de energia.

As técnicas de fabricação utilizadas nos dispositivos rápidos são as técnicas tradicionais, sendo que entre elas podemos citar a difusão, implantação iônica, recozimento, crescimento e deposições diversas [2].

A difusão e a implantação iônica seguida do recozimento são processos básicos de fabricação de dispositivos, no qual uma certa região do substrato recebe impurezas que modificam algumas de suas propriedades físicas. Esses dois processos podem ser usados, por exemplo, na fabricação de uma junção PN em um diodo e/ou na formação da região de canal em um transistor do tipo de efeito de campo (FET). Já o crescimento é usado para a formação de camadas de composição química ou dopagem (concentração de impurezas) diferentes entre si. As deposições são processos usados na formação de contatos e interligações entre dispositivos. É claro que a fabricação de um determinado dispositivo não usa apenas uma das técnicas enumeradas acima, mas sim uma combinação delas.

1.3 - Trabalho Desenvolvido

O principal trabalho desenvolvido foi o estudo da difusão de enxofre em arseneto de gálio por processamento térmico rápido.

Neste processo usamos tempos curtos de difusão, da ordem de um décimo do tempo usado nos processos convencionais.

O estudo das difusões compreendeu o seu desenvolvimento e posterior caracterização. Estudamos principalmente as características das camadas difundidas pelo processamento rápido, tais como profundidade da camada difundida, mobilidade dos portadores de corrente na camada, concentração superficial, e perfil de difusão, entre outras.

O objetivo inicial do estudo era verificar a possibilidade de se usar o processamento térmico rápido na difusão e testar a aplicação desse processo, caso ele fosse realizável, na fabricação de transistores MESFET. Como veremos no capitulo 4, este processo é possível de ser realizado e leva a resultados muito interessantes.

Os materiais escolhidos para a difusão rápida foram o arseneto de galio como cristal que recebe a difusão e o enxofre como material a ser difundido. O motivo da escolha do arseneto de galio é devido a vários fatores. Podemos apontar em primeiro lugar a grande disponibilidade do arseneto de gálio na forma semi-isolante, que é particularmente adequada para a fabricação dos transistores MESFET. Outros fatores que influenciaram essa escolha são o uso tradicional do arseneto de gálio na fabricação de dispositivos rápidos, e portanto a grande quantidade de trabalhos e referências do uso desse material. O arseneto de gálio ainda possibilita a integração monolítica de MESFET's com dispositivos opto-eletrônicos, o que é de grande interesse para a fabricação de dispositivos para telecomunicações.

A escolha do enxofre como material a ser difundido é devido ao fato desse material ser um dopante do tipo N, ou seja, ele altera a concentração de cargas no semicondutor gerando elétrons. Apenas os elétrons possuem alta mobilidade no GaAs (as camadas tipo P possuem portadores, lacunas, de baixa mobilidade), o que nos limita ao uso de dopantes tipo N. Entre todos os dopantes tipo N (selênio, enxofre, telúrio, silício, etc.) o enxofre é o que possui a maior difusidade, sendo portanto mais adequado em processamentos rápidos.

Para a realização das difusões rápidas foi necessário a

construção de um forno para o aquecimento rapido das amostras a serem difundidas. A necessidade do aquecimento rápido é explicada no próximo capitulo, junto com a descrição do projeto do forno.

1.4 - Técnica de Difusão Desenvolvida

O processo de difusão é um fenômeno físico onde introduzimos uma certa quantidade de impurezas em um material. Para que isso ocorra em geral é necessário que a amostra que irá receber as impurezas, e as próprias impurezas sejam aquecidas.

Normalmente a difusão é feita no interior de um forno, no qual a fonte de impurezas para a difusão e a amostra que recebe a difusão estão próximos um do outro no interior de uma ampola selada e evacuada (essa técnica é conhecida por difusão em ampola selada) [3]. As temperaturas normalmente usadas são de 600°C a 1300°C, com tempos de difusão variando de 10 minutos a alguns dias, conforme o tipo de impureza e substrato usados, e a profundidade que desejamos que as impurezas atinjam.

A tecnica por nós desenvolvida é diferente da tradicional em varios aspectos. A principal caracteristica que a difere das tradicionais é o uso de tempos curtos de difusão. Os máximos tempos de difusão usados foram de 90 segundos. Para que ocorra a difusão com pequenos tempos foi necessário o uso de temperaturas ligeiramente maiores, sendo essa a segunda caracteristica diferente do nosso processo. A última característica diferente é o uso do metodo de difusão em tubo aberto. Neste método de difusão a amostra e a fonte de difusão não estão lacrados no interior de uma ampola evacuada. Eles estão colocados em um suporte (não lacrado) feito especialmente para recebê-los [4]. Esse suporte permite o uso de amostras grandes no processo de difusão, o que é conveniente para a fabricação de circuitos integrados.

1.5 - Transistores MESFET

Uma vez desenvolvido o processo de difusão este foi aplicado na fabricação de transistores MESFET, que são basicamente transistores de efeito de campo de junção (JFET) que usam uma barreira Schottky ao invés de uma junção PN. Este tipo de transistor é particularmente adequado para a fabricação usando o arseneto de gálio. Isto é devido primeiro à facilidade de formação da barreira Schottky no GaAs, o que facilita a fabricação, e segundo à alta frequência de operação alcançada pelos MESFET's de GaAs.

Como veremos no capitulo 6, os transistores fabricados a partir de cristais difundidos apresentaram boas caracteristicas elétricas, tais quais baixa resistência série de dreno e fonte e alta transcondutância, que são devidas às características das camadas ativas obtidas pela difusão por processamento térmico rapido.

1.6 - Referências

[1] - S. M. Sze, *Physics of Semiconductor Devices*, 2ª edição, JONH WILEY & SONS, New York, 1981, apêndice G.

[2] - J. H. Magerlein, et al., "Characterization of GaAs Self Aligned Refractory Gate Metal-semiconductor Field-effect Transistor (MESFET) Integrated Circuits", *J. Appl. Phys.*, vol. 61, n. 8, p. 3080 (1987).

[3] - H. Matino, "Reproducible Sulfur Diffusion into GaAs", *Solid-State Electronics*, vol. 17, p. 35 (1974).

[4] - F. C. Prince, M. Oren, e M. Lam, "High Mobility GaAs Layers Obtained by Open Tube Sulfur Diffusion", *Appl. Phys. Lett.*, vol. 48, n. 8, p. 546 (1986).

CAPITULO II

FORNO RÁPIDO

2.1 - Introdução

Para a realização das difusões rápidas precisamos de alguma forma conseguir aquecer rapidamente a amostra que irá receber a difusão e a fonte de difusão. Como já foi dito no capitulo 1, temos tempos de difusão da ordem de dezenas de segundos (de 30 a 90 segundos), precisamos, portanto, aquecer a amostra partindo de uma baixa temperatura (no máximo 400°C) até a temperatura de difusão, que como veremos é de 830°C a 900°C, em tempos dessa ordem. Admitindo um aquecimento de 600°C (de 300°C a 900°C) em 30 segundos, teriamos uma taxa de aquecimento de 20°C por segundo. É necessario também que o resfriamento da amostra seja feito a uma taxa comparável à de aquecimento.

A necessidade do uso de aquecimento rápido no processo de difusão rápida pode ser entendida quando observamos que para uma dada temperatura de difusão também irá ocorrer a difusão para temperaturas ligeiramente inferiores. Sem o uso de aquecimento rápido podemos gastar tanto tempo em temperaturas inferirores quanto o tempo gasto na difusão. Portanto durante o estágio final de aquecimento da amostra já temos a difusão. Isso faz com que as caracteristicas das camadas difundidas não dependam apenas da temperatura ou do tempo de difusão. Elas irão depender também do aquecimento da amostra, o que torna difícil a analise dos resultados obtidos.

2.2 - Mecanismos de Aquecimento em um Forno

Em um forno podemos ter três mecanismos diferentes de aquecimento, que são condução térmica, convecção e radiação.

Desses três mecanismos os dois primeiros são bastante lentos, já que envolvem o transporte de calor pela matéria, sendo rápido apenas o mecanismo de radiação.

Em um forno convencional a amostra é aquecida por um gas quente existente no interior do forno, que é por sua vez aquecido, normalmente, por um resistência elétrica. Temos dessa forma o aquecimento da amostra sendo feito principalmente por condução e convecção do gás no interior do forno (a resistência elétrica não é uma boa fonte de radiação). Esse fato aliado à grande massa térmica do gas no interior do forno irá levar à baixas taxas de aquecimento (tipicamente taxas máximas de 1 °C por segundo).

Em um forno rápido deve predominar o mecanismo de aquecimento por radiação, pois ele é o único mecanismo rápido de aquecimento. Dessa forma, no forno rápido a amostra é aquecida absorvendo a radiação proveniente da fonte de aquecimento, ou fonte de radiação. Necessitamos, portanto, no forno rápido de uma eficiente fonte de radiação para que o aquecimento seja o mais rápido possivel. Tipicamente podemos atingir taxas de aquecimento da ordem de 400 °C por segundo, e taxas de resfriamento tão altas quanto 100 °C por segundo, já que o interior do forno está a uma temperatura bem menor que a amostra (não temos nenhum gás aquecido e apenas a amostra é aquecida pela radiação).

2.3 - Tipos de Fornos Rápidos

As fontes de radiação normalmente usadas para o aquecimento rapido são lasers, fontes de micro-ondas, resistência de grafite e lâmpadas de arco-voltaico ou tungstênio-halogênio. Dentre todas as fontes listadas as três mais usadas são a resistência de grafite e as lâmpadas de arco-voltaico ou tungstênio-halogênio, sendo esta última de uso predominante em sistemas comerciais de aquecimento rápido.

No forno com resistência de grafite temos uma câmara onde é feito vácuo, e onde é colocada a resistência de grafite. Fazemos circular uma corrente elétrica pelo grafite, de tal forma a

aquecê-lo à temperatura desejada. A amostra que será aquecida é colocada tambem no interior da câmara e é separada do grafite por um anteparo opaco à radiação emitida pelo grafite. Quando queremos aquecer a amostra retiramos o anteparo e aproximamos a amostra do grafite, fazendo que ela seja aquecida pela radiação emitida pelo grafite.

Podemos apontar duas desvantagens desse tipo de forno. A primeira é que o grafite é mantido permanentemente aquecido. A sua taxa de aquecimento é pequena, daí o fato de se usar um anteparo à radiação para controlar o aquecimento da amostra. Isto aliado ao fato de que a temperatura da amostra é no máximo igual à do grafite implica na impossibilidade de se realizar ciclos de temperatura com vários "steps", onde a amostra é mantida a valores intermediários de temperatura antes do aquecimento final. A segunda desvantagem é a necessidade de um sistema de vácuo no forno, para que o grafite não oxide. O uso de uma atmosfera inerte poderia evitar essa oxidação, porem ela poderia prejudicar o ciclo térmico da amostra, absolvendo calor do grafite e mantendo por sua vez a amostra permanentemente aquecida.

Outro tipo de forno é o de aquecimento por radiação de lâmpadas de tungsténio-halogênio. Neste tipo de forno a amostra é aquecida pela radiação emitida pelas lâmpadas. Estas possuem uma alta eficiência, transformando cerca de 65% de potência elétrica recebida em radiação. Neste forno a câmara onde e colocada a amostra pode ser feita de placas de quartzo, transparentes à radiação. Dessa forma podemos colocar as lâmpadas do lado de fora da câmara, isolando assim a amostra a ser aquecida da fonte de radiação. Conseguimos assim utilizar qualquer tipo de atmosfera no interior da câmara e manter a contaminação da câmara a níveis bem inferiores. O forno de aquecimento por radiação de lâmpadas não sofre das desvantagens do forno com resistência de grafite. Não é necessario o sistema de vácuo para o funcionamento do forno (a não ser que se queira o aquecimento em vácuo) e podemos ainda através de um controle sobre a potência das lâmpadas realizar "steps" no processo de aquecimento, com um controle preciso do tempo e temperatura de cada "step". Em uma forma mais sofisticada desse controle poderíamos até controlar a taxa de subida da temperatura.

2.4 - Aplicação do Processamento Térmico Rápido (RTP)

As primeiras experiências de que se tem notícia que usaram aquecimento rápido foram feitas por cientistas russos na década de 70. Eles usaram um laser como fonte de radiação para aquecer localmente um "wafer". Aos tratamentos térmicos de curto tempo de duração é dado o nome de processamento térmico rápido, ou RTP.

Hoje em dia, com o advento dos fornos rápidos, o RTP tem tido cada vez mais uma maior aplicação em cristais semicondutores [1]. Inicialmente o uso de RTP era restrito apenas a trabalhos de ativação de íons implantados [2], porém com o passar do tempo o seu uso expandiu-se, atingindo outras áreas do processamento térmico, tais quais difusões rápidas [3,4], formação de contatos ôhmicos [5] ("alloy"), formação de silicetos [6], e mais recentemente o recozimento de filmes supercondutores de alto Tc (temperatura crítica) [7] tem sido feito por RTP.

Alguns processos tem sido desenvolvidos unicamente devido ao uso de RTP. Podemos citar como exemplo a formação de siliceto de titânio auto-alinhado. Entre as vantagens oferecidas pelo RTP temos em processos de recozimento a ausência de difusão dos íons implantados, a obtenção de junções rasas em processos de difusão e a obtenção de contatos ôhmicos de menor resistência em processos de "alloy".

2.5 - Projeto do Forno

No nosso projeto optamos pelo forno de aquecimento rápido com lâmpadas de tungstênio-halogênio devido à sua maior simplicidade (ausência de sistema de vácuo), maior versatilidade (melhor controle de temperatura e possibilidade de criação de "steps" de aquecimento) e maior disponibilidade de material (lâmpadas, tubo de quartzo, etc.) para a montagem do forno.

No projeto utilizamos 18 lâmpadas de 190 milímetros de comprimento, cada lâmpada com uma potência de 1000 W (110 volts), perfazendo um total de 18 kW de potência.

O controle de temperatura adotado é do tipo PID, sendo realizado através de um microcomputador acoplado ao forno. Foi implementado também um controle manual de temperatura.

O projeto do forno pode ser dividido em três partes, que são a parte mecânica do forno, os circuitos eletrônicos e o programa de controle do forno.

A parte mecânica compreende a câmara de aquecimento onde é colocada a amostra e onde estão as lâmpadas, os circuitos de refrigeração da câmara e os suportes e o rack utilizados na montagem dos sistemas.

Os circuitos eletrônicos são formados pelos controladores de potencia das lâmpadas, pelo microcomputador para controle de temperatura, pelos conversores A/D e D/A para a medida e controle da temperatura e pelos circuitos de controle, proteção e diagnóstico do forno.

A terceira parte é o programa de controle, que é responsavel pelo funcionamento automático do forno. O programa é encarregado de controlar a temperatura do forno e fazer diagnósticos de um possivel mau funcionamento do sistema.

Segue uma descrição detalhada de cada parte do sistema.

2.5.1 - Parte Mecânica do Forno

Todo o forno, com exceção do microcomputador, seu teclado e monitor de video, está montado no interior de um rack, de 515 x 395 x 565 mm de tamanho. No painel do rack estão localizados os principais controles do forno e a porta para o interior da câmara de aquecimento. No painel traseiro do forno estão localizadas as entradas e saídas de gases para a câmara, a água para a refrigeração, os cabos de alimentação elétrica e o barramento para a ligação do forno com o microcomputador.

No interior do rack estão colocadas a câmara de aquecimento, as válvulas de controle do circuito de refrigeração e todos os circuitos eletrônicos. A fixação da câmara e dos circuitos ao rack é feita através de barras metálicas presas às paredes deste.

As figuras 2.1 e 2.2 mostram respectivamente uma vista do

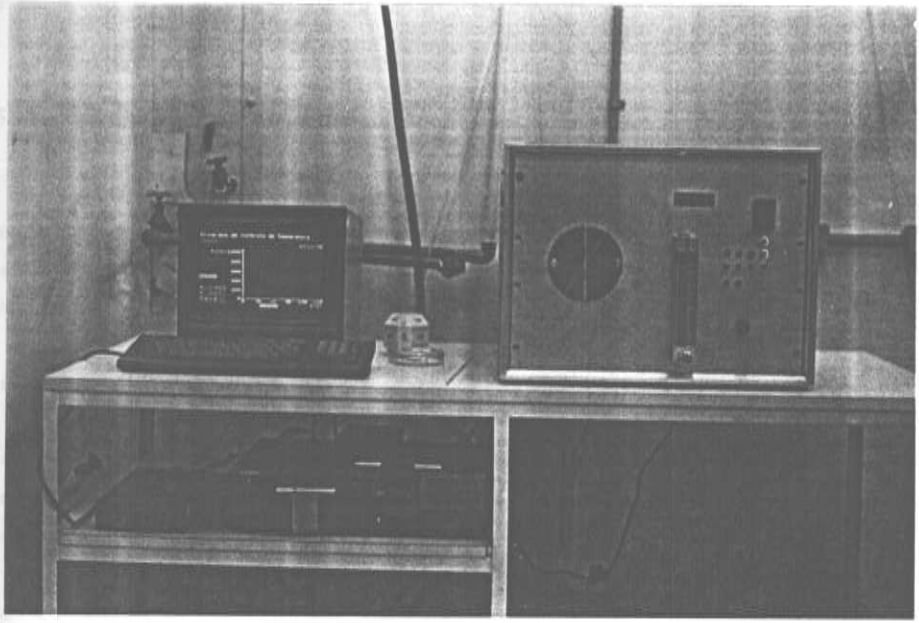

Figura 2.1 - Fotografia mostrando o forno e o microcomputador de controle com o seu terminal de vídeo e teclado.

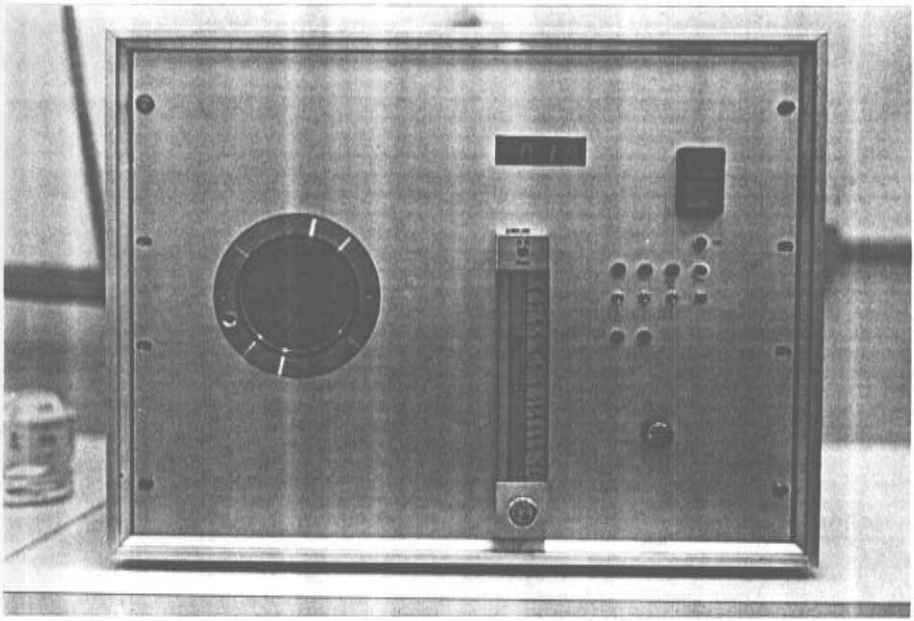

Figura 2.2 - Fotografia do painel do forno mostrando os principais controles e a entrada para a câmara de aquecimento.

forno e do microcomputador, e o painel do forno com os controles.

2.5.1.1 - Câmara de Aquecimento

A câmara de aquecimento é onde colocamos as amostras a serem aquecidas, e onde estão as lâmpadas de tungsténio-halogénio. Ela é formada por quatro flanges circulares (de diâmetro 190mm), por um tubo de quartzo de comprimento aproximado de 400mm por 80mm de diâmetro interno (com espessura de parede de 2,5mm) e por dois refletores.

O tubo de quartzo é totalmente vedado e a amostra é colocada no seu interior. O diâmetro interno do tubo de quartzo (80mm) permite que se use amostras de até três polegadas de diâmetro. As lâmpadas que fazem o aquecimento da amostra estão colocadas fora do tubo de quartzo, radialmente ao tubo. Essa disposição garante uma iluminação uniforme na amostra, produzindo um aquecimento homogêneo desta.

Existem apenas quatro entradas para o interior do tubo de quartzo. Uma das entradas serve para injetar nitrogênio no interior do tubo (com a finalidade de criar uma atmosfera inerte). A segunda entrada é por onde é feita a exaustão do nitrogênio injetado. Finalmente temos duas portas que dão acesso ao interior do tubo. A primeira, porta frontal, está localizada no painel do rack, e serve para introduzirmos a amostra no interior da câmara. A segunda, porta traseira, está localizada no interior do rack, e é por onde é feita a entrada do sensor de temperatura do forno.

As flanges estão divididas em dois tipos. Temos duas flanges de suporte do tubo de quartzo e duas flanges de suporte dos soquetes das lâmpadas. A disposição das flanges é a que segue. A primeira flange e de suporte do tubo. A segunda flange está a 40mm da primeira e e de suporte dos soquetes das lâmpadas. A terceira flange está a 240mm da segunda e é também de suporte dos soquetes (o espaço de 240mm é necessário para a colocação de uma lâmpada, 190mm, e dois soquetes, de aproximadamente 25mm cada um). A última flange é de suporte do tubo e está a 40mm de distância da terceira. As flanges estão presas entre si através de barras

cilíndricas (de 10mm de diâmetro).

A primeira e a última flange (suporte do tubo) servem para a fixação e vedação do tubo de quartzo. Essas flanges ainda servem de suporte para a câmara de aquecimento, e é nelas que vão presas quatro barras (duas em cada flange) que fixam a câmara no interior do rack. Nessas flanges ainda estão colocadas as portas e a entrada e saída de nitrogênio para o interior do tubo.

A segunda e terceira flange servem de suporte para os soquetes das lâmpadas. Elas são furadas (furos de 90mm de diâmetro) para permitir a passagem do tubo de quartzo. Essas flanges ainda possuem vários furos para a passagem dos cabos dos soquetes e para a entrada de ar comprimido para a refrigeração das lâmpadas.

A câmara de aquecimento fica completa com os refletores. Eles são formados por duas chapas de latão cromado em forma de semi-círculo. Os refletores cobrem toda a área entre a segunda e a terceira flange, onde estão as lâmpadas. Eles são presos à câmara nessas duas flanges.

A função dos refletores é coletar a luz emitida pelas lâmpadas na direção oposta à amostra, e refletir essa luz de volta para a amostra, tornando mais eficiente o aquecimento.

2.5.1.2 - Refrigeração da Câmara

A câmara de aquecimento recebe duas refrigerações diferentes. Elas são à água e ar comprimido. O seu uso evita que tenhamos transferência de calor da câmara para o interior do rack e protege as peças que compõe a câmara de um eventual aquecimento excessivo.

A refrigeração à água é usada nos dois refletores e nas duas flanges que contém os soquetes das lâmpadas. Seu uso evita que os refletores super-aqueçam (eles recebem diretamente a radiação das lâmpadas) e mantém os soquetes das lâmpadas refrigerados. As flanges de suporte do tubo não recebem radiação das lâmpadas, não sendo necessário, portanto, a sua refrigeração.

O ar comprimido é usado para a refrigeração das lâmpadas. Ele é injetado no interior da câmara (onde estão as lâmpadas,

entre o tubo de quartzo e os refletores) através dos furos existentes nas flanges de suporte dos soquetes. Seu uso é para manter a temperatura dos contatos das lâmpadas em um valor seguro para a sua operação, pois durante o aquecimento da amostra, sem o ar comprimido, a temperatura dos contatos pode atingir 450°C, o que é prejudicial para as lâmpadas, podendo acasionar a sua queima.

A temperatura máxima dos contatos recomendada pelo fabricante é de 250°C. Com o uso do ar comprimido conseguimos manter a temperatura ligeiramente abaixo de 250°C, o que evita a queima das lâmpadas.

Tanto o ar comprimido quanto a água possuem válvulas para ligar ou desligar o seu fluxo. Dessa forma podemos desligar as refrigerações quando a câmara está fria. O controle das válvulas é feito através de duas chaves (independentes) existentes no painel do rack, ou automaticamente através de dois sensores de temperatura, um colocado no refletor, que controla a válvula de água, e o outro colocado próximo às lâmpadas, que controla a válvula de ar comprimido.

2.5.1.3 - Controle da Atmosfera no Interior do Tubo de Quartzo

Como dito anteriormente, o tubo de quartzo é totalmente vedado em suas extremidades e é feita uma entrada de nitrogênio no seu interior para a criação de uma atmosfera inerte. Na realidade pode-se usar qualquer gás inerte para a criação dessa atmosfera. Poderíamos usar, por exemplo, argônio ao invés de nitrogênio.

O controle do fluxo de gás para o interior do tubo é feito através de uma válvula agulha e um fluxômetro, ambos no painel do rack. A válvula agulha serve para o ajuste do fluxo gasoso, e o fluxômetro para a medida deste. Podemos controlar o fluxo desde 0,1 l/min. até 8 l/min.

2.5.2 - Circuitos Eletrônicos

Os circuitos eletrônicos são responsaveis por todo o funcionamento do forno. Sua principal função é o controle da potência das lâmpadas e a medição da temperatura na amostra. Os circuitos ainda realizam outras funções, como por exemplo teste do forno, controle da refrigeração, etc. Na figura 2.3 podemos ver uma fotografia mostrando os principais circuitos.

Na figura 2.4 vemos um diagrama de blocos de todo o sistema eletrônico do forno. O sistema está dividido em seis blocos, que são o controlador de potência das lâmpadas, o microcomputador para controle de temperatura, os conversores A/D e D/A do microcomputador, o controlador das válvulas, os sensores de proteção e os circuitos de teste das lâmpadas.

De todos os blocos listados acima podemos apontar como o principal o controlador de potência das lâmpadas. Ele é responsável pela potência entregue às lâmpadas, e portanto pela temperatura do forno.

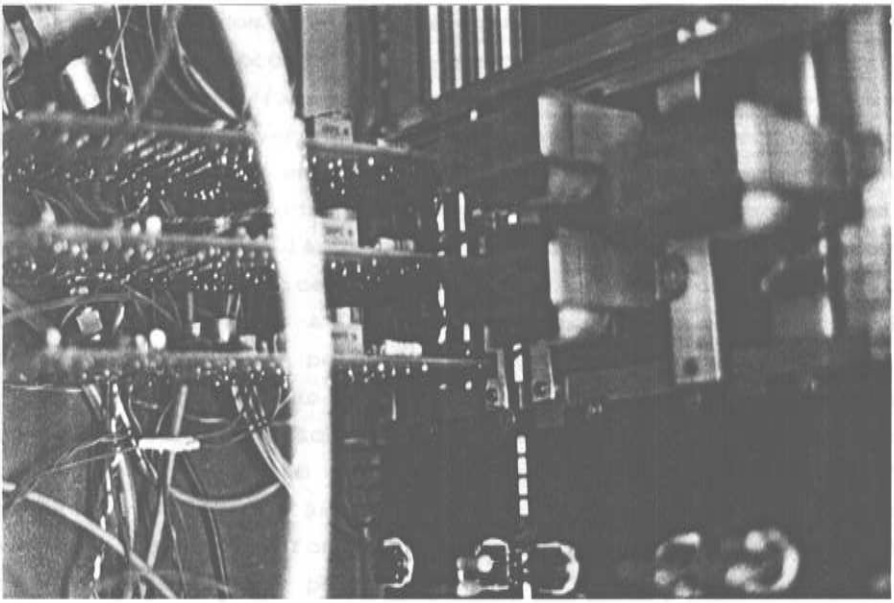

Figura 2.3 - Fotografia mostrando os principais circuitos eletrônicos do forno.

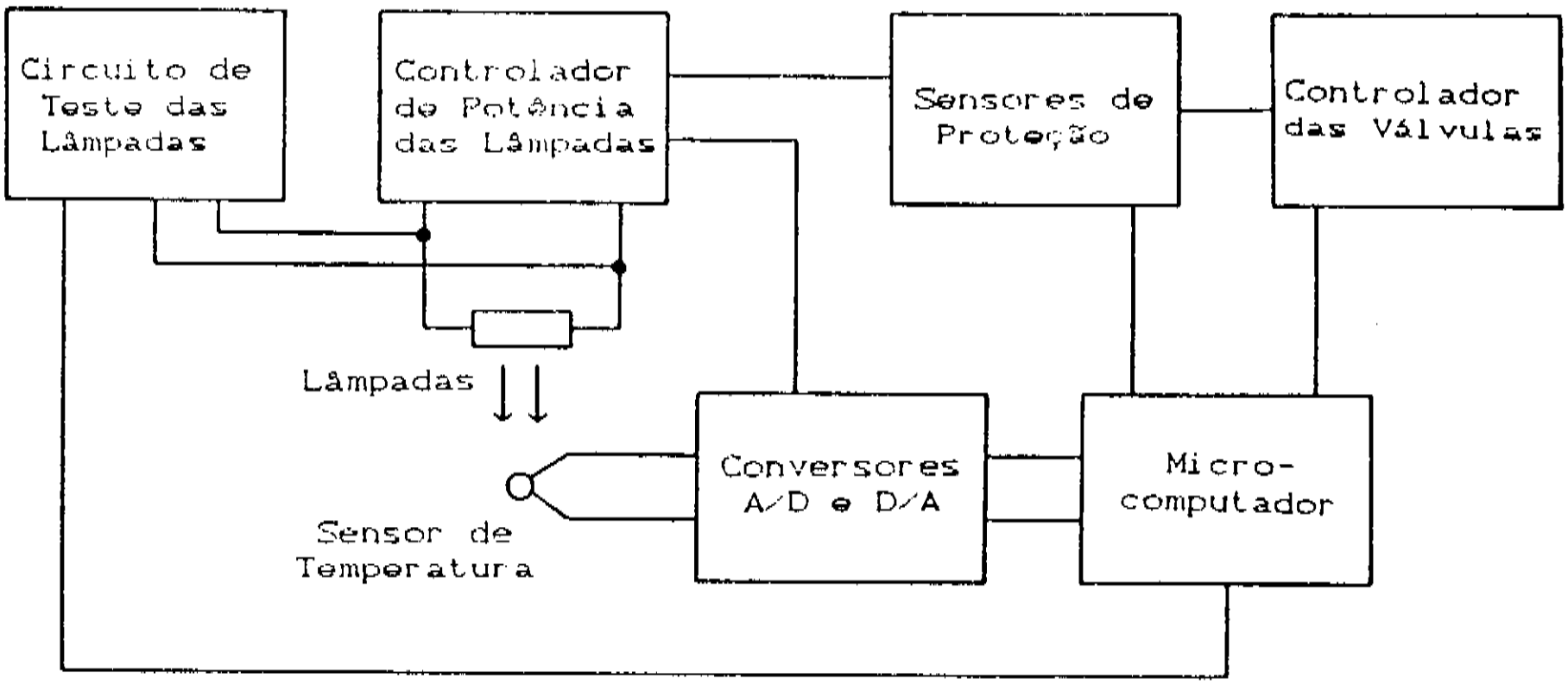

Figura 2.4 - Diagrama em blocos dos circuitos eletrônicos.

Uma parte dos circuitos é do tipo digital e é formada basicamente pelo microcomputador e pelos conversores A/D e D/A (os outros blocos também fazem uso de circuitos digitais, porém em uma menor escala). O conversor A/D funciona na medida de temperatura pelo microcomputador, enquanto que o D/A é ligado ao controlador de potência das lâmpadas, informando a este qual deve ser a potência entregue às lâmpadas. O microcomputador é responsável pelo controle automático da temperatura. Ele faz a medida da temperatura (conversor A/D) e usando um algoritmo do tipo PID calcula qual deve ser a poténcia entregue às lâmpadas.

Fora a informação da temperatura do forno o microcomputador recebe informações sobre o funcionamento das lâmpadas e sobre as condições de operação do forno, como por exemplo porta aberta, sobre-aquecimento nas lâmpadas ou nos refletores, entre outras. Essas informações são fornecidas respectivamente pelos circuitos de teste das lâmpadas e pelos sensores de proteção.

Os sensores de proteção além de gerarem informações para o microcomputador ainda controlam automaticamente as válvulas de refrigeração (água e ar comprimido), ligando-as caso seja

detectado um aquecimento nas lâmpadas ou nos refletores. Os sensores de proteção podem ainda desativar os controladores de potência, desligando as lâmpadas, caso alguma condição anormal seja detectada, como um super-aquecimento nos refletores, por exemplo. Fora o controle automático das válvulas, temos ainda o controle manual e pelo microcomputador, que pode através de sua programação controlar as válvulas.

2.5.2.1 - Controlador de Potência das Lâmpadas

A alimentação das lâmpadas (energia elétrica) é trifásica. Temos seis lâmpadas em cada fase da rede elétrica, como mostra a figura 2.5. Isso faz necessario o uso de três circuitos controladores, um em cada fase da rede. Cada circuito tira a sua propria alimentação da fase onde está ligado.

O controle da potência sobre as lâmpadas é feito com o uso de TRIACS. O tipo de controle feito é o do ângulo de condução dos TRIACS (veja a figura 2.6), tanto no semi-ciclo positivo quanto no negativo da rede elétrica. O disparo dos TRIACS é feito através de um "trem de pulsos" gerado pelo circuito.

Cada circuito possui duas entradas. Uma delas e a entrada da tensão de controle, que varia de 0 a 1 volt, correspondendo a um ângulo de condução do TRIAC de 0° a 180° (um semi-ciclo), que por sua vez corresponde a 0% a 100% de potência sobre as lâmpadas. A segunda entrada é do tipo lógico e inibe o funcionamento do circuito ao ser levada para o nível lógico "1".

A tensão de controle pode ser gerada pelo conversor D/A ou por um potenciômetro existente no painel do forno, que é o controle manual de potência sobre as lâmpadas. A seleção da fonte de tensão de controle é feita por uma chave, também localizada no painel do forno.

2.5.2.2 - Conversor A/D e D/A e Microcomputador

Os conversores A/D e D/A e o microcomputador compõem a

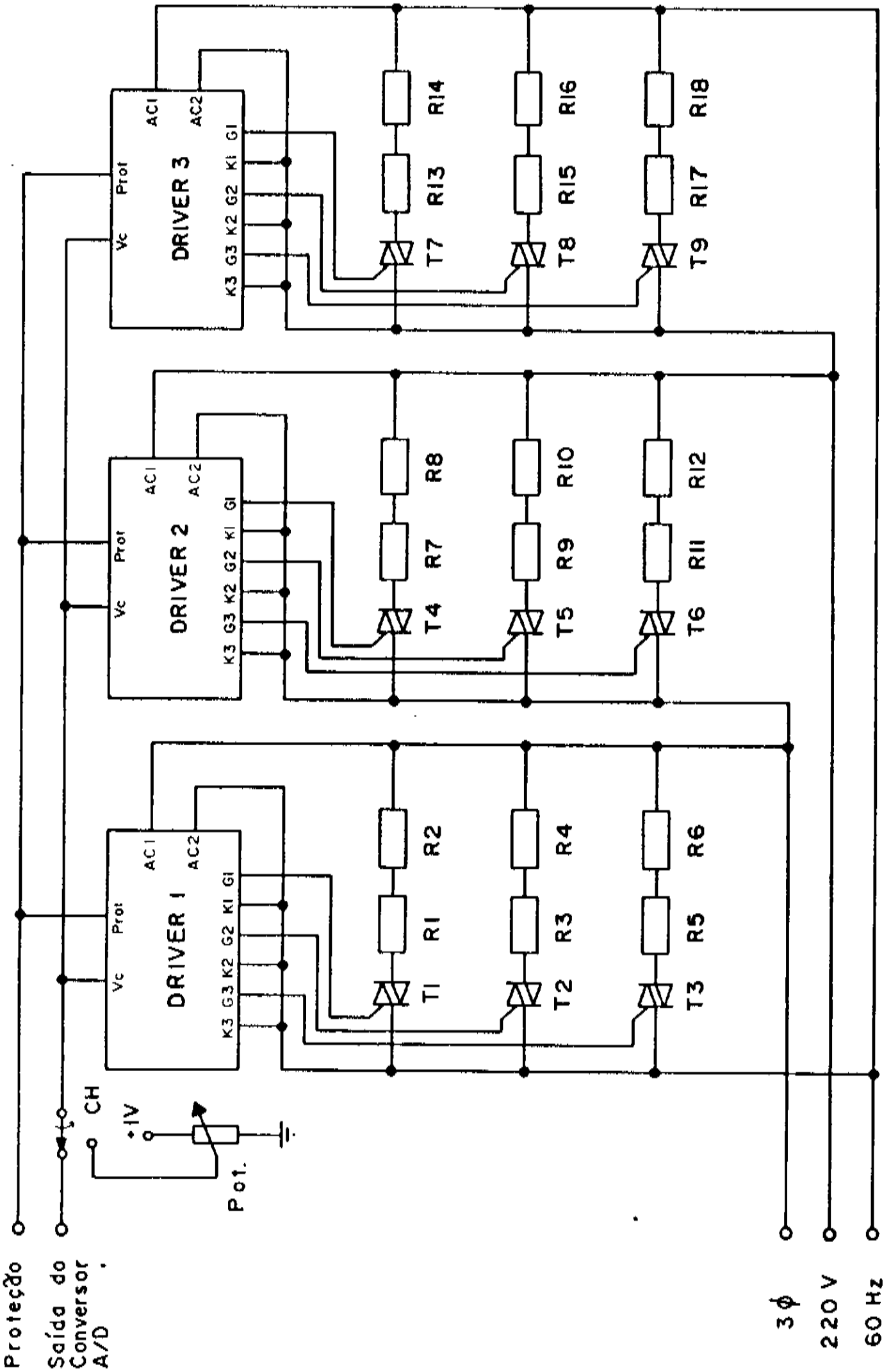

Figura 2.5 - Ligação dos controladores e lâmpadas à rede elétrica.

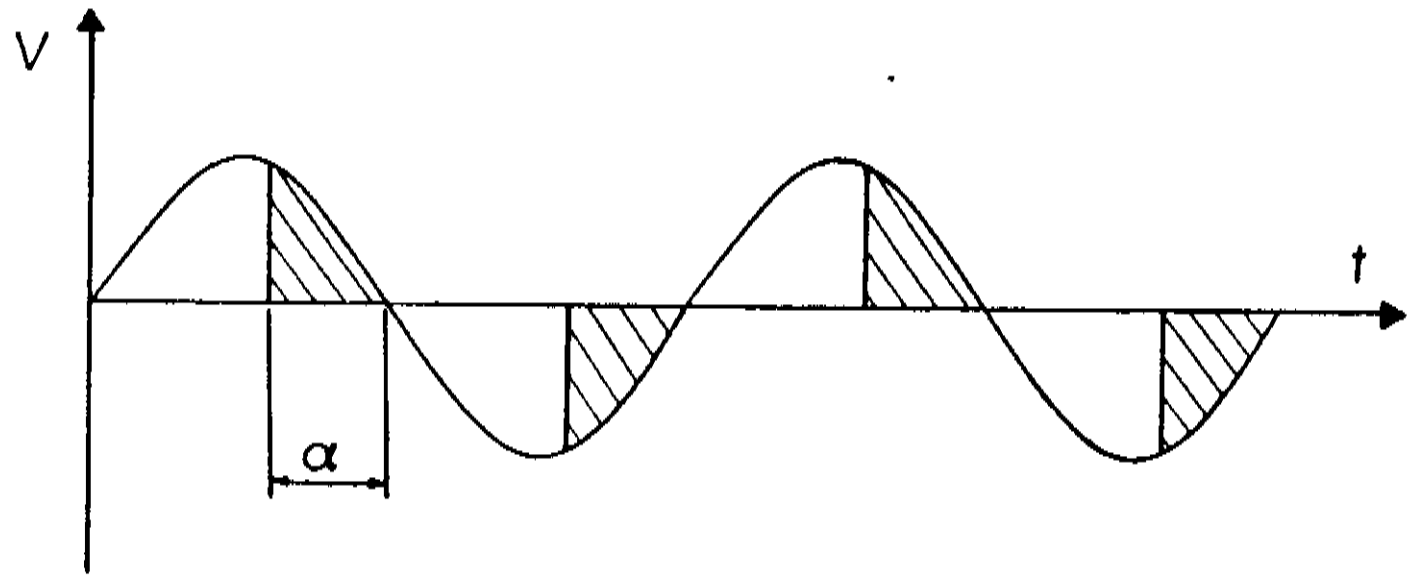

Figura 2.6 - Forma de onda da tensão sobre as lâmpadas. O ângulo α pode ser variado de 0° a 180° pelos controladores.

parte digital do circuito eletrônico. Eles são responsáveis pelo controle automático de temperatura.

O microcomputador usado é do tipo comercial de 8 bits. O conversor D/A usado é também de 8 bits, o que produz uma resolução de 4 mV na tensão de controle, correspondendo a uma resolução de menos que 1° no controle do ângulo de condução dos TRIACS. Já o conversor A/D usado é de 12 bits, o que dá uma resolução de 0,25°C na medida da temperatura.

Alem dos conversores A/D e D/A ainda estão ligados ao microcomputador uma unidade de disco, onde está armazenado o programa de controle, um teclado e um monitor de vídeo (figura 2.7), por onde podemos programar as características do processo (temperatura, tempo, ciclo de pré-aquecimento, etc.) e visualizar o ciclo térmico através de um gráfico temperatura x tempo mostrado no monitor.

2.5.2.3 - Sensor de Temperatura

Existe um grande problema associado à vantagem oferecida pelo forno rápido de alta taxa de variação de temperatura. Este problema é a medida de temperatura na amostra e surge do fato de

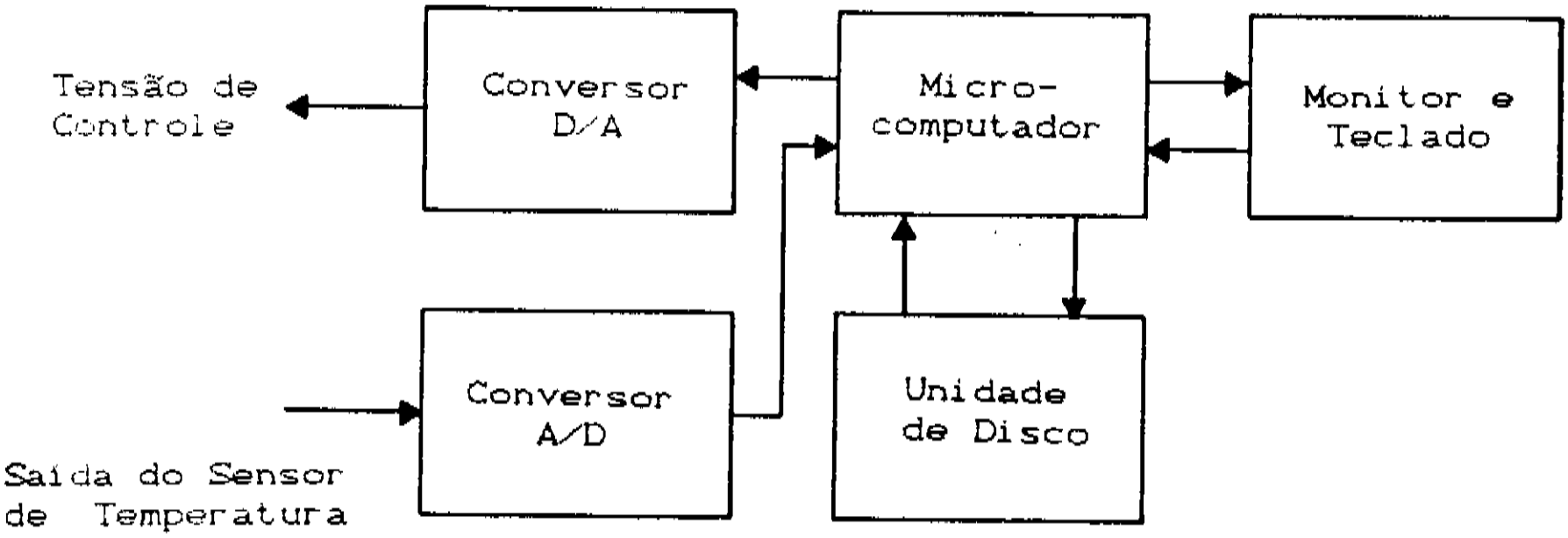

Figura 2.7 - Diagrama em blocos das ligações entre os conversores A/D e D/A e o microcomputador e periféricos.

apenas a amostra estar aquecida pelas lâmpadas. Podemos ter diferenças de temperatura tão altas quanto 700°C entre a amostra e as paredes do interior do forno. Isso implica em que a medida da temperatura deve ser feita diretamente na amostra, ou em algum material ou sensor, colocado no interior do forno, que simule o "comportamento térmico" da amostra. Temos como exemplo o caso em que um "wafer" de silício é aquecido diretamente pelas lâmpadas. Nesse caso o sensor de temperatura usado é um pequeno pedaço de silício (também na forma de "wafer"), fixado no interior do forno, cuja temperatura será aproximadamente igual à do "wafer" a ser aquecido.

No nosso caso usamos como sensor de temperatura um disco de grafite de 34 mm de diâmetro por 1,5 mm de espessura. Esse disco simula o "comportamento térmico" do porta-amostra onde é feita a difusão (veja o capítulo 4) e é nele que é feita a medida da temperatura.

O sensor de temperatura propriamente dito é um termopar do tipo K (cromel-alumel) formado por fios de 0,25mm de diâmetro. A junção dos dois metais é uma pequena bola de 0,5mm de diâmetro que vai alojada em um furo lateral no grafite, furo este de 0,7mm de diâmetro por 5mm de profundidade. Dessa forma o termopar é fixado

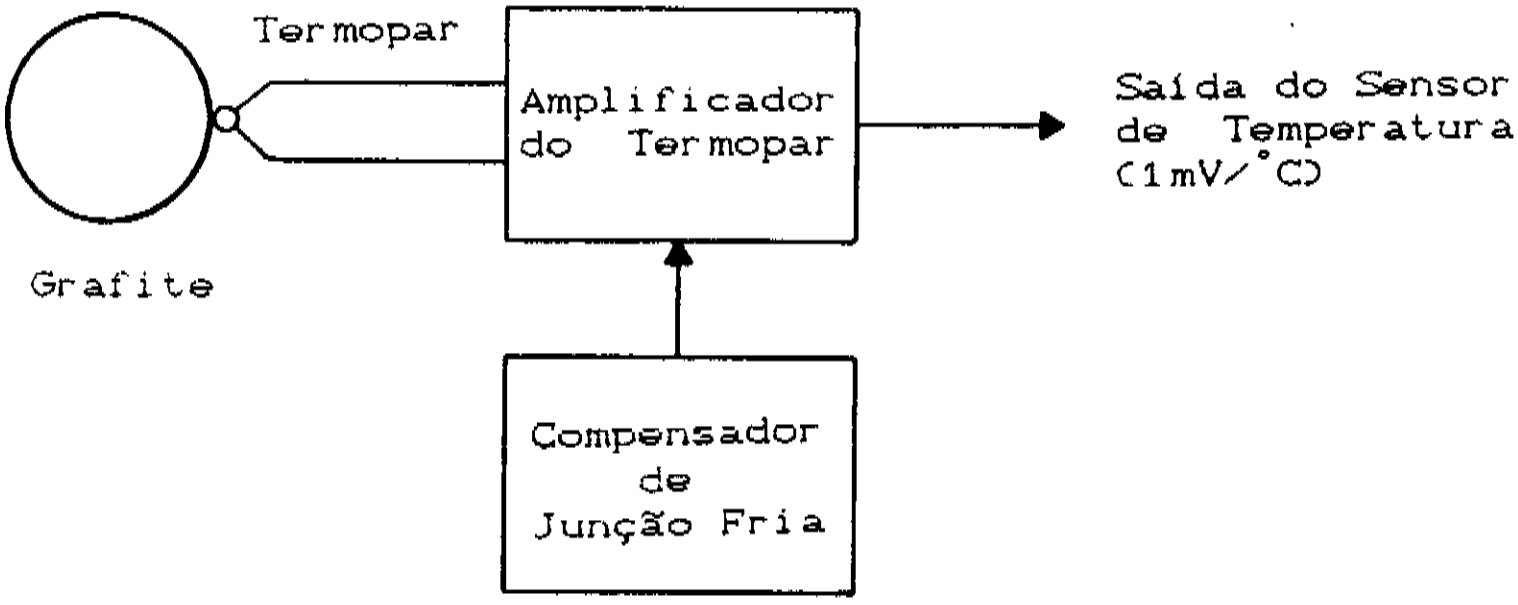

Figura 2.8 - Sistema de medição de temperatura.

no grafite, fazendo um bom contato térmico e medindo adequadamente a temperatura do grafite. O tempo de resposta do termopar é minimizado com o uso da pequena bola na junção dos dois metais.

O termopar como sensor de temperatura irá gerar uma tensão elétrica, que a menos de uma pequena correção, é diretamente proporcional à temperatura. Para termopares do tipo K a tensão gerada é de 40μV/°C. Essa tensão deve ser amplificada antes de ser entregue às outras partes do circuito, devido ao seu baixo nível (para 1000°C teríamos uma tensão de 40mV). O circuito que faz essa amplificação é o amplificador do termopar e compensador de junção fria, figura 2.8. O compensador de junção fria é um circuito que gera uma referência de tensão para a correta medida de temperatura pelo termopar. Na saída do amplificador temos um sinal de tensão de 1mV/°C, que é usado pelo conversor A/D na leitura da temperatura pelo microcomputador e por um voltímetro digital no painel do forno para a indicação da temperatura.

2.5.3 - Programa de Controle

As funções do programa de controle são basicamente o controle automático da temperatura do forno e o diagnóstico de um possível mau funcionamento do forno.

O controle automático de temperatura é implementado com o uso de um algoritmo do tipo PID digital recursivo [8]. Esse algoritimo calcula a diferença entre a potência que deve ser aplicada sobre as lâmpadas e a potência aplicada anteriormente (interação anterior do programa), corrigindo-a e ajustando assim a temperatura do forno. A diferença de potência é calculada em função do erro de temperatura (diferença entre a temperatura desejada e a temperatura do forno) e das constantes P, I e D usadas no algorítmo.

Na atual versão do programa não é feito nenhum controle da taxa de subida da temperatura e foi implementado, além do ciclo normal de aquecimento, um ciclo de pré-aquecimento. O programa permite que se escolha as duas temperaturas (aquecimento e pré-aquecimento) e tempos nessas temperaturas de acordo com a vontade do operador, questionando-o os seus valores. São feitas apenas duas restrições. A primeira é que a temperatura de pre-aquecimento deve ser menor que 600°C, e a segunda é que o tempo em altas temperaturas (maiores que 700°C) deve ser menor que dois minutos.

Ao ser executado o programa de controle realiza duas tarefas antes de iniciar o controle do forno. A primeira é a inicialização e carga das constantes P, I e D e a segunda tarefa é requisitar ao operador as caracteristicas do processo de aquecimento. A seguir o programa irá inicializar o video gráfico para traçar o ciclo térmico da amostra (gráfico temperatura x tempo) e irá esperar um comando do operador para iniciar o aquecimento. Além do gráfico temperatura x tempo é mostrado no video a temperatura do forno, o tempo decorrido desde o inicio do processo de aquecimento e a taxa de variação da temperatura. A figura 2.9 mostra uma fotografia da tela gráfica do microcomputador.

O programa de controle foi escrito em duas linguagens diferentes. Ele foi escrito parte em BASIC (residente no microcomputador) e parte em ASSEMBLER (Z-80, que é a CPU usada pelo microcomputador). O trecho do programa em BASIC realiza as inicializações, pergunta ao operador as caracteristicas do processo e inicializa o video gráfico para uso pelo programa em

ASSEMBLER. O programa em ASSEMBLER faz o controle da temperatura do forno, recebendo as características do processo e as constantes P, I e D do programa em BASIC. Além do controle de temperatura o programa em ASSEMBLER irá controlar o vídeo gráfico, mantendo-o atualizado com as informações sobre o processo (temperatura, tempo e taxa) e traçando o gráfico de temperatura x tempo. O uso de ASSEMBLER no controle foi necessário devido à alta taxa de variação de temperatura do forno. O "software" de controle deve ser rápido o suficiente para responder às variações de temperatura do forno, sendo necessário, portanto, uma linguagem de rápida execução, como o ASSEMBLER, que foi utilizado.

O diagnóstico e teste do forno é uma das opções do programa. As rotinas de teste ativam todas as lâmpadas do forno e circuitos de controle, e concluído o teste o programa irá apontar os circuitos defeituosos, e as lâmpadas e TRIACS queimados, caso existam.

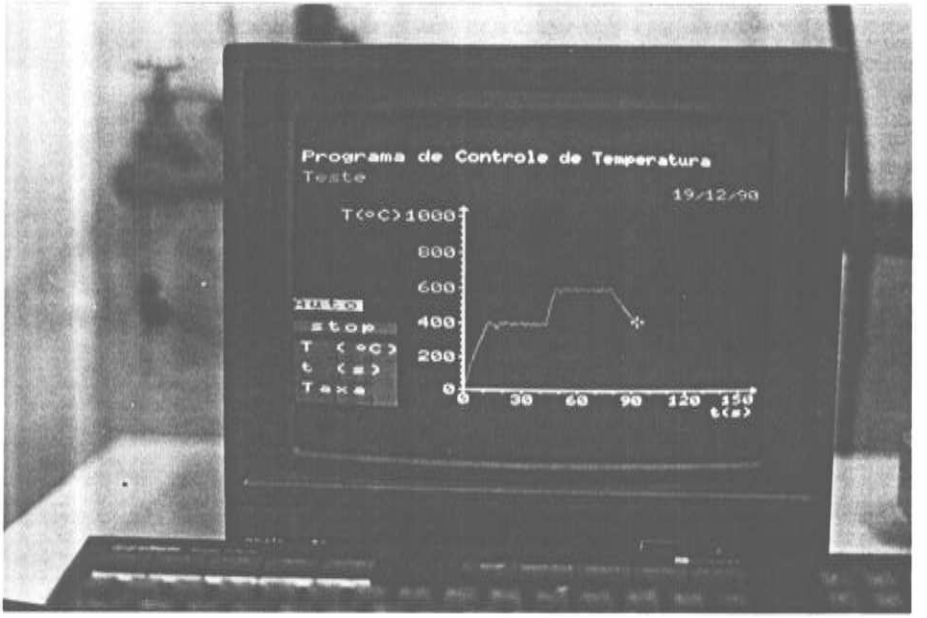

Figura 2.9 - Fotografia da tela gráfica do microcomputador.

2.6 - Testes e Caracterização do Forno

Uma vez terminada a montagem do forno foram feitos testes com a finalidade de caracterizar a sua performace. Os testes feitos foram de três tipos diferentes. Os primeiros testes realizados foram de máxima temperatura atingida, sendo seguidos pelos testes de aquecimento e resfriamento do forno. Os últimos testes realizados foram os de controle de temperatura pelo microcomputador. Os testes de máxima temperatura, aquecimento e resfriamento foram feitos com dois sensores diferentes. Foram usados um "wafer" de silício de duas polegadas de diâmetro por 0,5mm de espessura e o disco de grafite de 34mm de diâmetro. Os testes de controle de temperatura foram feitos apenas no grafite. Em todos os testes foi usado um termopar tipo K como sensor (item 2.5.2.3).

A medida de máxima temperatura foi feita aplicando a máxima potência sobre as lâmpadas e esperando que a temperatura atingisse o seu máximo valor, quando então foi feita a medida. Mesmo após a temperatura ter atingido o seu máximo valor ainda observamos uma pequena taxa de subida de 1 °C por segundo, devida ao aquecimento do gás interior e das paredes do forno (tubo de quartzo). Tanto para o grafite quanto para o silício foram registradas máximas temperaturas quase iguais. Para o grafite a máxima temperatura atingida foi de 1000°C, enquanto que para o silício esta foi de 960°C. A pequena diferença é devida em parte ao proprio material em que foi feita a medida (o grafite deve absorver melhor a radiação das lâmpadas) e em parte ao melhor contato térmico do termopar no grafite, já que este está alojado em um furo na sua lateral (no silício o termopar foi colocado na superfície da lâmina).

A medida da máxima taxa de subida da temperatura foi feita também com a máxima potência das lâmpadas, sempre iniciando o aquecimento com o forno a uma temperatura inferior a 100°C. Como era esperado observamos que a taxa de subida da temperatura cai em função do próprio aumento da temperatura. Dessa forma temos uma máxima taxa de 125°C/s (silício) para uma temperatura de 250°C, ao passo que essa taxa cai para apenas 11°C/s para 910°C. Outro

efeito esperado era que o grafite deveria apresentar uma menor taxa de aquecimento, como realmente ocorreu, devido à sua maior massa. Na figura 2.10 vemos o gráfico das taxas de aquecimento do silício e do grafite em função da temperatura. Podemos verificar no gráfico que as máximas taxas conseguidas foram de 125°C/s para o silício e 67°C/s para o grafite. Embora essas taxas possam parecer baixas para um forno rápido (os modelos comerciais podem atingir taxas de até 300°C/s) elas estão coerentes com a geometria da câmara de aquecimento. Devido à disposição radial das lâmpadas, estas ficaram distantes da amostra (em média 60mm), o que leva à taxas não muito altas. Isto não impede e nem tão pouco prejudica o processo de difusão, pois como visto necessitamos de taxas de aquecimento maiores que 20°C/s.

As medidas da taxa de resfriamento foram feitas depois de um aquecimento do forno a uma temperatura maior que 920°C. Na figura 2.11 vemos os resultados dessas medidas. Como era esperado observamos que à medida que a temperatura cai a taxa de resfriamento também diminui, já que o resfriamento é feito por radiação (corpo negro) da amostra. Podemos observar que novamente

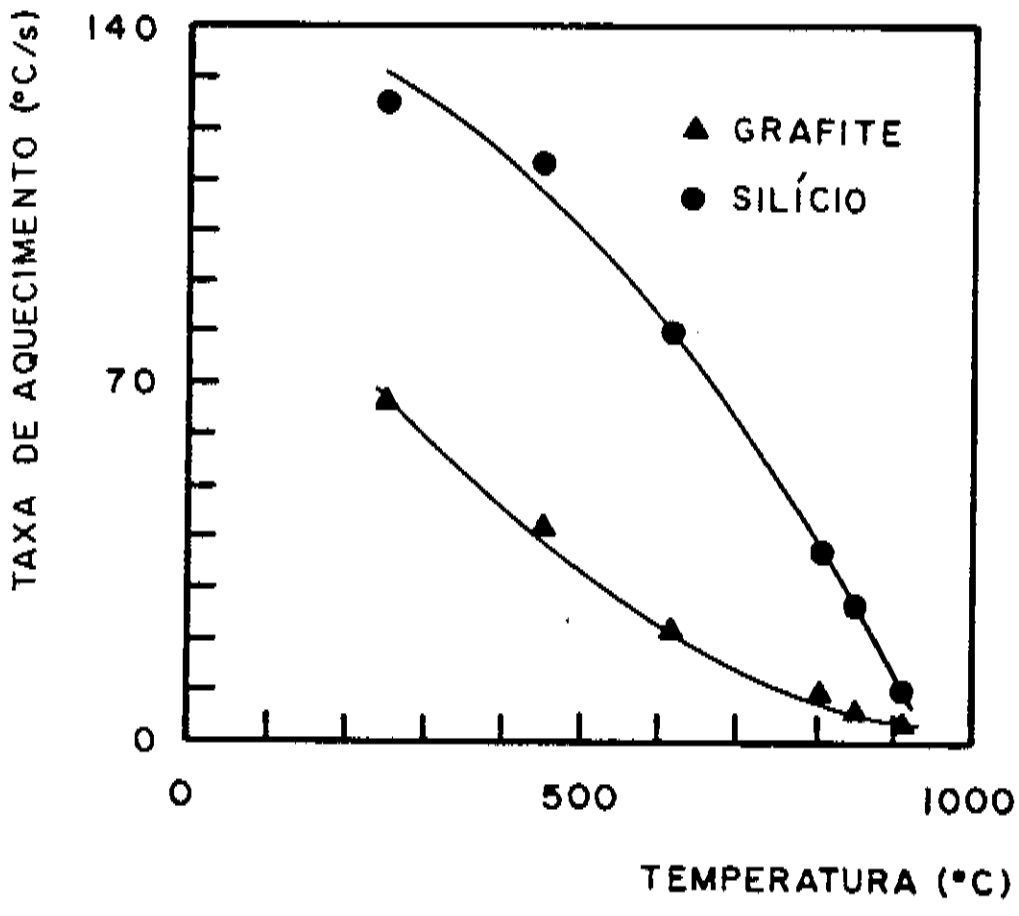

Figura 2.10 - Taxa de aquecimento em função da temperatura para o disco de grafite e para as lâminas de silício.

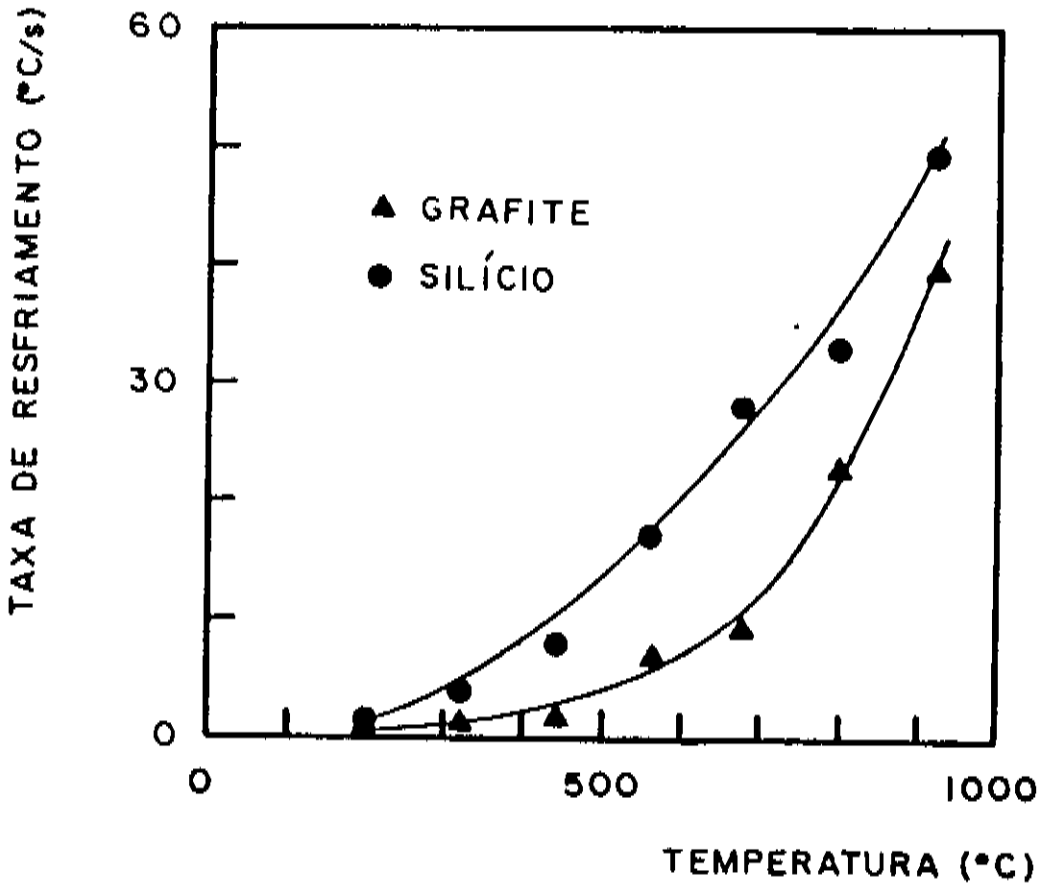

Figura 2.11 - Taxa de resfriamento em função da temperatura para o disco de grafite e para as lâminas de silício.

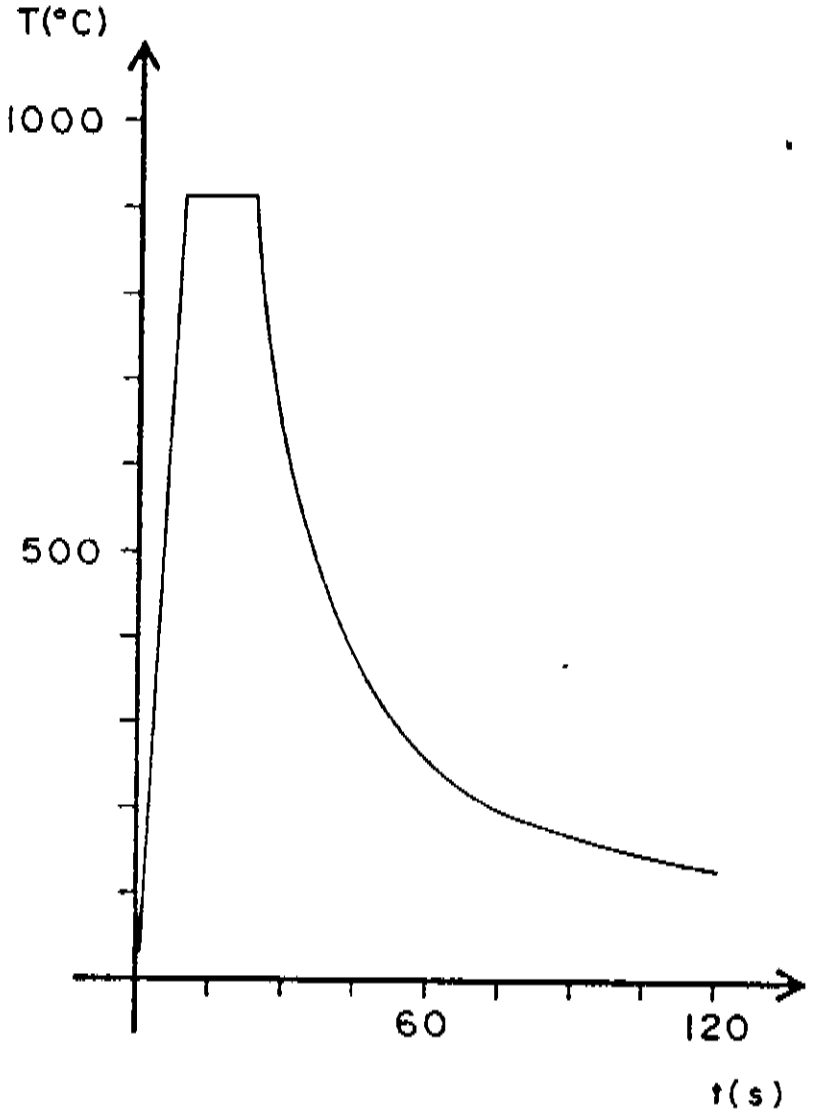

Figura 2.12 - Ciclo de aquecimento no silício.

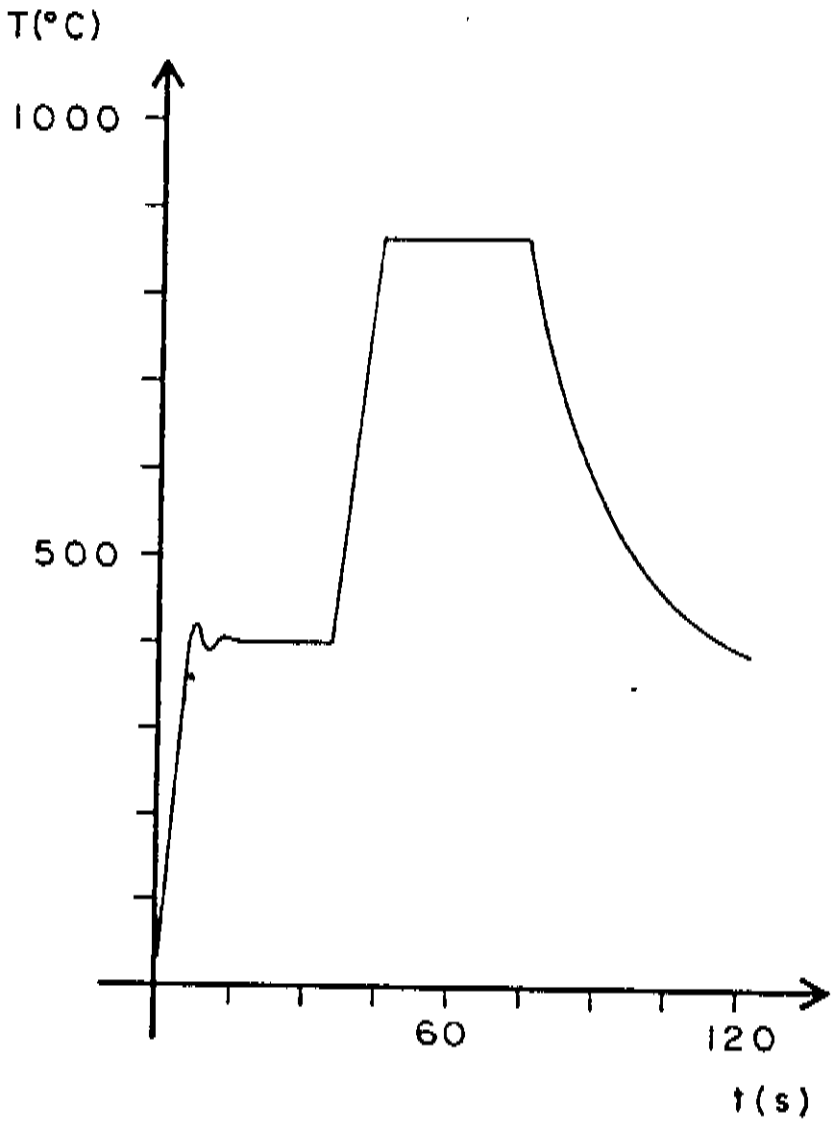

Figura 2.13 - Ciclo de aquecimento no grafite.

o grafite produziu menores taxas de variação da temperatura (devido à sua maior massa térmica). Observamos ainda que as taxas de resfriamento são comparáveis às de aquecimento, 50°C/s para o silício (920°C) e 40°C/s para o grafite (também a 920°C).

É importante dizer que as taxas de resfriamento dependem do tempo em que o forno é mantido aquecido. Isto é devido ao fato de que as paredes do forno tendem a aquecer-se com o passar do tempo, aquecendo por sua vez a amostra quando as lâmpadas forem desligadas. Portanto quanto mais tempo o forno é mantido aquecido mais as paredes aquecerão e será maior o aquecimento da amostra pelas paredes e menores serão as taxas de resfriamento.

Nas figuras 2.12 e 2.13 podemos ver dois ciclos de aquecimento, o primeiro feito no silício e com temperatura final de 896°C, e o segundo feito no grafite com um ciclo de pré-aquecimento a 400°C e temperatura final de 920°C.

Os testes do controle de temperatura foram feitos com o microcomputador ligado ao forno. Os testes foram realizados com o objetivo de verificar a eficiência do microcomputador no controle

de temperatura e ajustar as constantes P, I e D para o valor que melhor otimizasse as variações de temperatura. Após os ajustes conseguimos uma máxima variação de temperatura de ± 4°C em torno da temperatura programada, o que é um bom valor para um forno do tipo rápido.

Em resumo podemos dizer que os objetivos estabelecidos no projeto do forno foram alcançados, o que possibilitou, como veremos no capítulo 4, a realização das difusões.

2.7 - REFERÊNCIAS

[1] - Sandra Leavitt, "RTP : On the Edge of Acceptance", *Semiconductor International*, Março de 1987, p. 64 (1987).

[2] - B. Tell, K. F. Brown-Goebeler e C. L. Cheng, "Rapid Thermal Annealing of Elevated-Temperature Silicon Implants in InP", *Appl. Phys. Lett.*, vol. 52, n. 4 , p. 299 (1988).

[3] - S. K. Ghandhi, R. T. Huang e J. M. Borrego, "Fabrication of GaAs Tunnel Junctions by a Rapid Thermal Diffusion Process", *Appl. Phys. Lett.*, vol. 48, n. 6 , p. 415 (1986).

[4] - S. D. Lester et al., "Pulse Diffusion of Ge into GaAs", *Appl. Phys. Lett.*, vol. 48, n. 16 , p. 1063 (1986).

[5] - C. L. Chen, L. J. Mahoney, J. D. Woodhouse, M. C. Finn e P. M. Nitishin, "Ohmic Contacts to N-Type GaAs Using High-Temperature Rapid Thermal Annealing for Self-Aligned Processing", *Appl. Phys. Lett.*, vol. 50, n. 17 , p. 1179 (1987).

[6] - E. Ma, B. S. Lim, M-A Nicolet, N. S. Alui e A. H. Handi, "Growth of WSi_2 in Phosphorous-Implanted W/<Si> Couples", *Journal of Electronic Materials*, vol. 17, n. 3, p. 207 (1988).

[7] - M. Aslam et al., "Rapid Thermal Annealing of YBaCuO Films on Si and SiO_2 Substrates", *Appl. Phys. Lett.*, vol. 53, n. 2, p. 153 (1988).

[8] - Rolf Isermann, *Digital Control Systems*, SPRINGER-VERLAG Berlin, New York, 1981.

CAPÍTULO III

DIFUSÃO DE IMPUREZAS NO GaAs

3.1 - INTRODUÇÃO

A difusão é um fenômeno muito comum na natureza. Como um exemplo temos a difusão de uma gota de tinta que cai em um copo com água. A tinta da gota irá misturar-se com a água até que seja atingida uma uniformidade na mistura. Isso ocorrerá mesmo que não se tenha nenhuma agitação na água. A mistura acontece devido ao movimento aleatório de moléculas de água e tinta, que tende a dissipar qualquer gradiente de concentração na mistura. Outro exemplo é a difusão de vapor de álcool no ar, que ocorre quando abrimos uma garrafa contendo álcool. Conseguimos sentir o cheiro do álcool, pois parte do líquido evapora e suas moléculas na forma de vapor difundem através do ar, atingindo nosso olfato. Novamente isso ocorrerá mesmo que o ar esteja parado, sem nenhum movimento.

Podemos ter as mais variadas formas de difusão. Podemos ter difusão em um líquido, como é o caso da gota de tinta na água, podemos ter difusão em um gás, que é o caso do vapor de álcool no ar, ou podemos ter difusão em um sólido, como por exemplo cobre difundindo no germânio sólido. O material que difunde também pode apresentar-se em várias fases (externas). No caso da gota de tinta, a sua fase externa é líquida. No caso do álcool temos o material que difunde na fase externa gasosa, e no caso do cobre, a fase é sólida. Podemos ter por fim a difusão de materiais de fase diferente da do meio onde eles difundem, por exemplo vapor de fósforo (gás) difundindo no silício.

No caso do material que recebe a difusão ser um sólido, este pode ainda ser amorfo ou cristalino. Um caso específico e de grande interesse do ponto de vista tecnológico e científico é a difusão em um sólido cristalino semicondutor. Esse é o caso que iremos discutir nos próximos ítens. Antes porém vamos discutir rapidamente a estrutura, propriedades elétricas e influência de

impurezas em um cristal semicondutor, quais são os tipos de defeitos que ocorrem no semicondutor, quais são as relações entre as concentrações dos vários defeitos, e quais são as faixas de concentrações onde podemos ter o semicondutor sólido com os seus defeitos característicos e com alguma impureza dissolvida. Todo esse estudo se faz necessário para que possamos compreender os mecanismos de difusão em um semicondutor. Sempre que possível iremos usar o arseneto de gálio (GaAs) como exemplo, já que ele é o semicondutor que nos interessa.

3.2 - Semicondutores e Impurezas

Como já dissemos anteriormente os semicondutores são sólidos cristalinos, e como tais são formados por um arranjo periódico de átomos. Podemos ter semicondutores formados com apenas um tipo de átomo, por exemplo átomos do grupo IVB da tabela periódica, silício ou germânio, que são os semicondutores simples; ou podemos ter os formados por mais de um tipo de átomo. Nesse caso temos os semicondutores compostos, que são chamados de binários, ternários ou quaternários caso tenhamos dois, três ou quatro elementos diferentes na sua formação. Entre os semicondutores binários podemos citar os formados por elementos da coluna IIIB e VB da tabela periódica. Temos como exemplos GaAs, InP, GaP e InSb, entre outros.

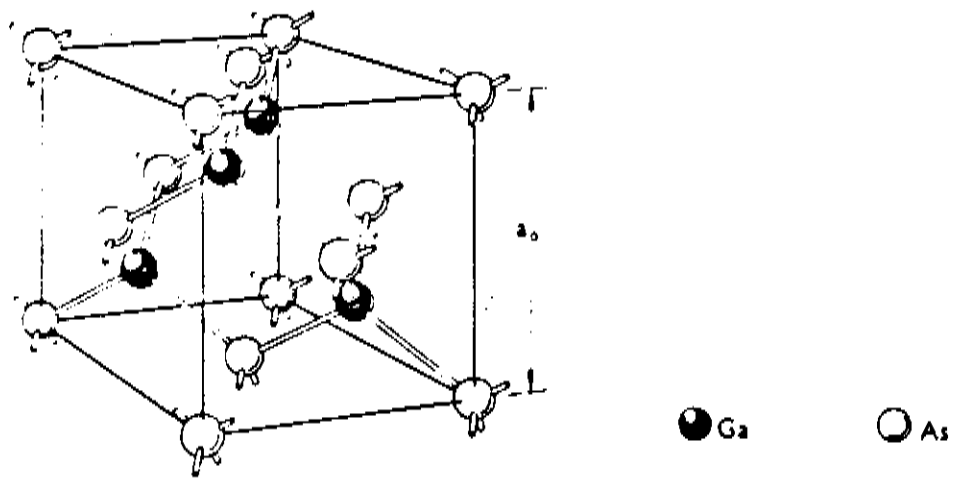

Figura 3.1 - Estrutura atômica do GaAs. As esferas pretas representam os átomos de gálio, e as brancas os átomos de arsênio.

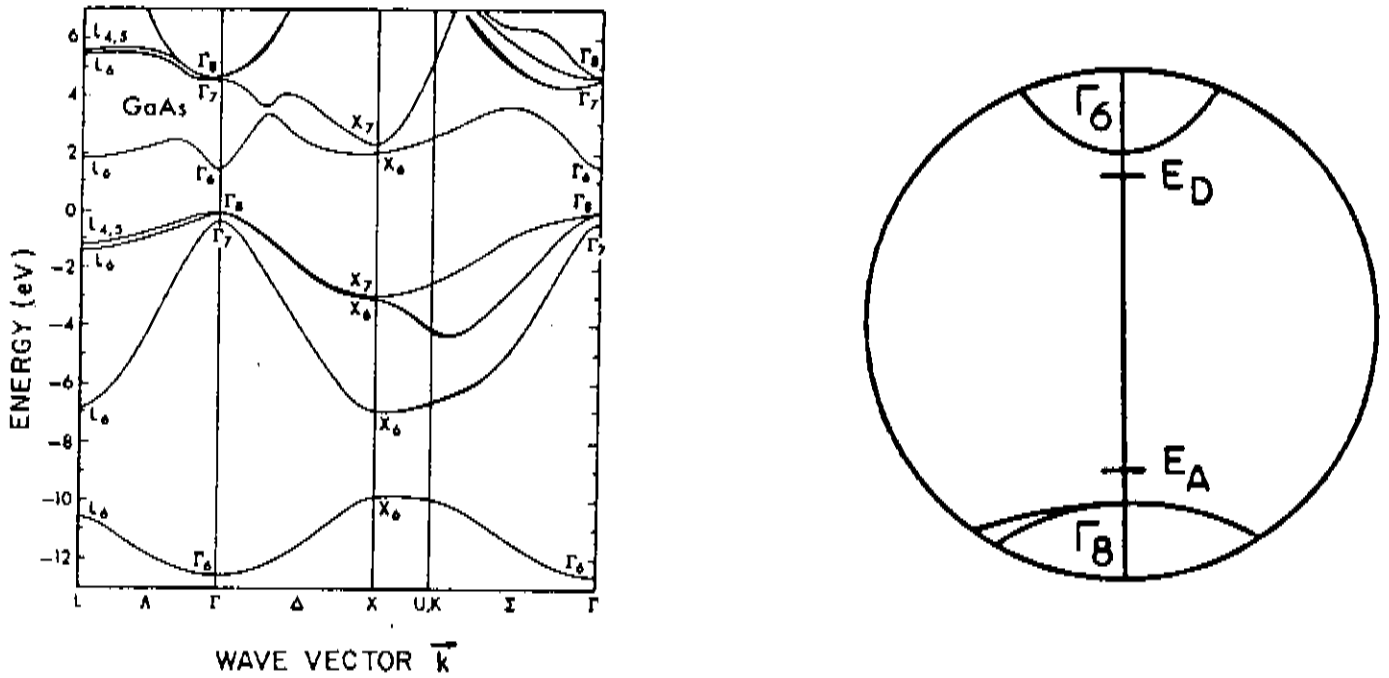

Figura 3.2 - Diagrama de bandas do GaAs com a representação dos níveis de impurezas [2].

Na figura 3.1 podemos ver a estrutura do arseneto de gálio [1]. Observe que temos dois tipos de átomos, os representados pelas esferas brancas, átomos de arsênio, e os representados pelas esferas pretas, gálio. Observe ainda que os átomos de um tipo ligam-se a quatro átomos de outro tipo. Essa estrutura recebe o nome de "zincblend", e a não ser pelos diferentes tipos de átomos, ela é idêntica à estrutura do diamante. Uma característica marcante deste tipo de estrutura é que ela é bastante "aberta", ou seja, é grande o espaço entre os átomos, espaço esse que é chamado de interstício. A consequência de termos grandes interstícios é que o cristal terá um densidade atômica pequena, e poderemos ter atomos nos intersticios.

Na figura 3.2. podemos ver o diagrama de bandas do GaAs [2]. Esse diagrama nos mostra os niveis de energia permitidos para os elétrons no interior do cristal em função do seu vetor de onda $\vec{k}$. Podemos observar que não existe nenhum nível de energia possivel entre os níveis dos pontos Γ_8 e Γ_6. Essa e uma importante caracteristica de um semicondutor, a existência de uma faixa de energias proibidas, ou como normalmente é chamado, a existência de um "gap" de energia. Esse "gap" geralmente tem uma energia da ordem de alguns elétron-volt (eV). Aos níveis de energia acima de Γ_6 é dado o nome de banda de condução, enquanto que aos niveis

abaixo de Γ_8 é dado o nome de banda de valência.

Na temperatura OK um semicondutor puro se comportara como um isolante perfeito, pois teremos toda a banda de valência ocupada e a de condução vazia. Com o aumento da temperatura alguns elétrons da banda de valência ganharão energia suficiente para passar para a banda de condução, tornando o cristal condutor. O fato de termos a banda de valência quase que totalmente preenchida nos leva a usar o conceito de lacunas, que são cargas ficticias positivas e móveis na banda de valência (isso simplifica o tratamento matematico do transporte de corrente pelo semicondutor). Dessa forma podemos dizer que quando um elétron deixa a banda de valência é criada uma lacuna em seu lugar, e que essa lacuna é responsável pela condução da banda de valência.

À temperaturas diferentes de OK temos portanto igual numero de elétrons na banda de condução e lacunas na banda de valência. Devido a isso chamamos esse semicondutor puro de intrinseco. Existe uma forma de alterar significativamente esse equilibrio entre elétrons e lacunas, tornando um tipo de portador de corrente predominante no cristal. Essa forma e a introdução de impurezas no cristal. Começamos com um exemplo onde trocamos um atomo do cristal por outro que possui um elétron a mais na sua camada de valência. Esse elétron a mais estará fracamente ligado ao nucleo. Com um pouco de energia termica recebida ele ira tornar-se um elétron livre passando para a banda de condução. Dessa forma conseguimos colocar elétrons na banda de condução sem gerar lacunas na banda de valência. Já se trocarmos o atomo do cristal por um com um elétron a menos em sua camada de valência, teremos uma ligação covalente faltando. Com uma pequena excitação termica um elétron de um atomo vizinho poderá preencher essa ligação, que possui um nivel de energia ligeiramente maior que o do topo da banda de valência. Ao fazer isso esse elétron tera gerado uma lacuna (ele deixa a banda de valência), sem ter gerado uma carga livre na banda de condução. Os niveis de energia introduzidos pelas impurezas no diagrama de bandas podem ser vistos na figura 3.2, E_D, que e o nível de energia do elétron extra de valência, e E_A, que e o da ligação covalente aberta.

Se a quantidade de impurezas adicionadas ao cristal for

Elemento	Energia (meV)
S	6,10
Se	5,89
Te	5,80
Si	5,81
Ge	6,08
Sn	6,00

TABELA 3.1 - Energias de ionização para doadores em GaAs.

muito grande, maior que a quantidade de elétrons ou lacunas geradas termicamente no cristal intriseco, teremos predominantemente apenas um tipo de portador de corrente no cristal, elétrons ou lacunas. Nesse caso o semicondutor passa a ser chamado de extrinseco, e as impurezas que fornecem elétrons de doadoras e as impurezas que recebem elétrons gerando lacunas de aceitadoras. Ao semicondutor com predominância de elétrons é dado o nome de tipo N, enquanto que o com predominância de lacunas de tipo P.

As impurezas que receberam ou doaram eletrons são chamadas de ionizadas, já que elas receberam ou perderam uma carga e não são mais neutras. Em geral a energia de ionização de uma impureza é pequena, de alguns meV a algumas dezenas de meV. Essas são as chamadas de impurezas rasas, em comparação com impurezas de mais alta energia de ionização, as chamadas de impurezas profundas. Na tabela 3.1 podemos ver os valores da energia de ionização para doadores em GaAs. O baixo valor de energia é responsavel por uma ionização quase que total a temperaturas muito baixas, o que é uma interessante propriedade de doadores no GaAs.

3.2.1 - Importância da Difusão em Semicondutores

Como vimos anteriormente, a introdução de impurezas em um

semicondutor pode alterar suas propriedades elétricas, e conforme o tipo de impureza usada podemos obter semicondutores apenas com elétrons de condução, tipo N, ou apenas com lacunas de condução, tipo P. Vários dispositivos eletrônicos são formados usando-se esses dois tipos de semicondutor. Temos por exemplo o diodo retificador, que é um dispositivo que conduz corrente elétrica em apenas um sentido. Ele é formado pela junção de um semicondutor tipo N com um tipo P. Outro exemplo é o transistor bipolar, que é formado pela sobreposição de camadas N-P-N ou P-N-P. A propriedade de interesse em um transistor é o efeito de amplificação de corrente. Isso torna grande o seu uso em circuitos amplificadores, comutadores e osciladores. Temos também os transistores de efeito de campo, ou FET's, que são constituidos por duas camadas, uma condutora de corrente, tipo N, e outra para o controle do fluxo de corrente, tipo P. As propriedades desse transistor são bem semelhantes as do bipolar, com algumas vantagens que são a operação em frequências mais altas e a maior estabilidade térmica. Temos por fim os modernos circuitos integrados, que formam a base para a construção de modernos computadores digitais. Esses circuitos são um arranjo complexo de milhares ou até milhões de dispositivos, cada um formado por pelo menos uma camada tipo P ou N.

Já deve ter sido percebido que introdução de impurezas em um semicondutor é um processo de importância vital para a fabricação de dispositivos eletrônicos. Existem três formas diferentes de introduzirmos as impurezas no semicondutor, durante o crescimento de uma camada semicondutora, por implantação iônica, ou por difusão. Os varios processos de introdução de impurezas recebem o nome generico de dopagem.

O crescimento de uma camada e um processo muito usado na fabricação de lasers semicondutores e fotodetetores. Esse processo implica na deposição epitaxial de uma camada cristalina sob um substrato semicondutor, com a mesma orientação deste. A deposição pode ser feita a partir de uma fase líquida ou gasosa, sendo que podemos introduzir impurezas nessa fase, depositando a camada com a dopagem desejada. Esse tipo de processo permite um fácil controle da dopagem (concentração e tipo), além de produzir camadas com dopagem uniforme. As desvantagens são os equipamentos

para o crescimento de camadas, que são sofisticados e caros.

A segunda forma de se dopar um cristal é através de implantação iônica. Neste processo as impurezas são introduzidas no interior de uma câmara de vácuo na forma de vapor. Esse vapor será ionizado e acelerado por um alto potencial elétrico (da ordem de dezenas de milhares de volts). Dessa forma teremos as impurezas na forma de íons a alta velocidade. Estes são dirigidos para o cristal, chocando-se contra a sua superfície. Ao fazerem isso os íons serão desacelerados, parando a alguma profundidade abaixo da superfície do cristal. A distribuição dos íons em profundidade não será uniforme, o que é uma desvantagem do processo. A maior concentração será a alguma profundidade que pode ser determinada teoricamente e experimentalmente. A vantagem da implantação iônica é que conseguimos controlar precisamente a concentração de dopantes, além de ser possível a sua introdução além da máxima solubilidade sólida no cristal. Existem outras desvantagens além da já exposta, que são a necessidade de um implantador de íons, que é um equipamento de custo muito elevado (da ordem de um milhão de dólares), e a criação de danos ao cristal, pelo bombardeio de íons altamente energéticos. Isso faz necessário um ciclo de aquecimento do cristal ("annealing") após o implante, que tem a finalidade de eliminar os danos e ativar os íons (estes passarão a se ligar aos átomos do cristal, dopando-o efetivamente).

Por fim o último processo é a difusão, que consiste na introdução dos dopantes no cristal pelo processo físico de difusão, partindo de alguma fase externa, sólida, líquida ou gasosa, em contato com o semicondutor, e que contenha os dopantes. A difusão é um processo ativado termicamente, e como tal é de se esperar relativamente altas velocidades de difusão (penetração) dos dopantes em altas temperaturas. Qualquer dopante será imóvel em um semicondutor à temperatura ambiente. Concluímos que é necessário que o cristal e a fonte fonte de dopantes estejam aquecidos para que ocorra a difusão. A vantagem desse processo é a sua simplicidade, pois necessitamos apenas de um forno e uma fonte de dopantes para a sua realização. A desvantagem é o uso de altas temperaturas no processo, o que pode ser prejudicial ao cristal, pois facilita a introdução de impurezas indesejadas que podem

também difundir, e com o uso de alta temperatura podemos criar uma grande concentração de defeitos pontuais no cristal, o que também é indesejável.

Um cristal tipo N receberá diferentes denominações de acordo com o seu grau de dopagem. Por exemplo, o GaAs tipo N de alta resistividade (baixa concentração de dopantes), maior que 1Ωcm, será chamado de N^-. Isso corresponde a uma concentração de dopantes menor que $10^{15}cm^{-3}$. Já um cristal de média resistividade, entre 1Ωcm e 10^{-3}Ωcm, será simplesmente N, e a faixa de dopagem equivalente será de $10^{15}cm^{-3}$ a $10^{18}cm^{-3}$. Um cristal de baixa resistividade, menor que 10^{-3}Ωcm, será chamado de tipo N^+, e a sua dopagem será maior que $10^{18}cm^{-3}$. Observe que essas concentrações de dopantes são pequenas, pois o cristal tem uma densidade da ordem de 10^{22}átomos/cm^3. Novamente todos os três processos de dopagem permitem obter cristais com os vários graus de dopagem, porém o crescimento epitaxial e o implante iônico permitem maior liberdade de escolha da concentração e espécie de dopantes.

3.3 - Defeitos Pontuais em Semicondutores Compostos

Como vimos no ítem anterior, um átomo de uma impureza dopante deve ocupar uma posição de um átomo da rede cristalina de um semicondutor, para que ela possa efetivamente dopar o cristal, fornecendo-lhe um elétron, ou capturando um elétron e deixando uma lacuna livre. Para que o cristal seja dopado ele deve ter ou devem ser geradas posições vagas na rede cristalina, que possam ser ocupadas pelas impurezas. Essas posições vagas não são nada mais do que um dos tipos de defeitos pontuais que podem ocorrer em um semicondutor.

Existem outros tipos de defeitos pontuais, além de posições vagas da rede, também chamadas de vacâncias. Podemos ter por exemplo átomos em posições intersticiais, ou átomos de impurezas em posições substitucionais ou intersticiais. Podemos ter também a associação de defeitos. Em um semicondutor composto temos outros defeitos possíveis, por exemplo um átomo de um elemento ocupando a

posição de um átomo do outro elemento que forma o semicondutor. Os vários tipos de defeitos pontuais são vistos a seguir.

3.3.1 - Tipos de Defeitos Pontuais em Semicondutores Compostos

Vamos estudar agora com mais detalhes os defeitos pontuais, porém introduziremos antes uma notação muito útil. Chamamos o semicondutor composto de AB, A significando o elemento do grupo IIIB da tabela periódica e B o elemento do grupo VB. Uma vacância será indicada por V_A, onde a letra A indica que essa vacância é do elemento A. Temos da mesma forma V_B, que é uma vacância do elemento B. Um átomo intersticial será indicado por A_i, se for um átomo do elemento A, ou B_i se for do elemento B. Um átomo A na posição de um átomo B será indicado por A_B.

Os defeitos pontuais em um semicondutor puro são classificados por Kendall [3] como simétricos ou assimétricos. O defeitos simétricos seriam aqueles que envolveriam apenas um tipo de defeito, todos os átomos intersticiais ou apenas vacâncias, enquanto que os assimétricos envolveriam tipos diferentes de defeitos.

Temos os seguintes tipos de defeitos simétricos em um semicondutor composto:

1 - Vacâncias ou defeitos Schottky, V_A ou V_B.
2 - Intersticiais, A_i ou B_i.
3 - Anti-estruturas, A_B ou B_A.

O primeiro tipo de defeito, as vacâncias, são simplesmente posições vagas da rede cristalina, deixadas pelos átomos A ou B. Essas vacâncias podem se associar aos pares, devido a uma atração mútua. Esses pares são chamados de divacâncias, e podemos ter três tipos diferentes, V_AV_B, que é uma associação de uma vacância de cada elemento, e V_AV_A ou V_BV_B, que seriam uma associação de vacâncias de um mesmo elemento. O segundo tipo de defeito simétrico, intersticiais, são átomos A ou B ocupando uma das posições vazias da rede, ou interstícios. Como no caso de

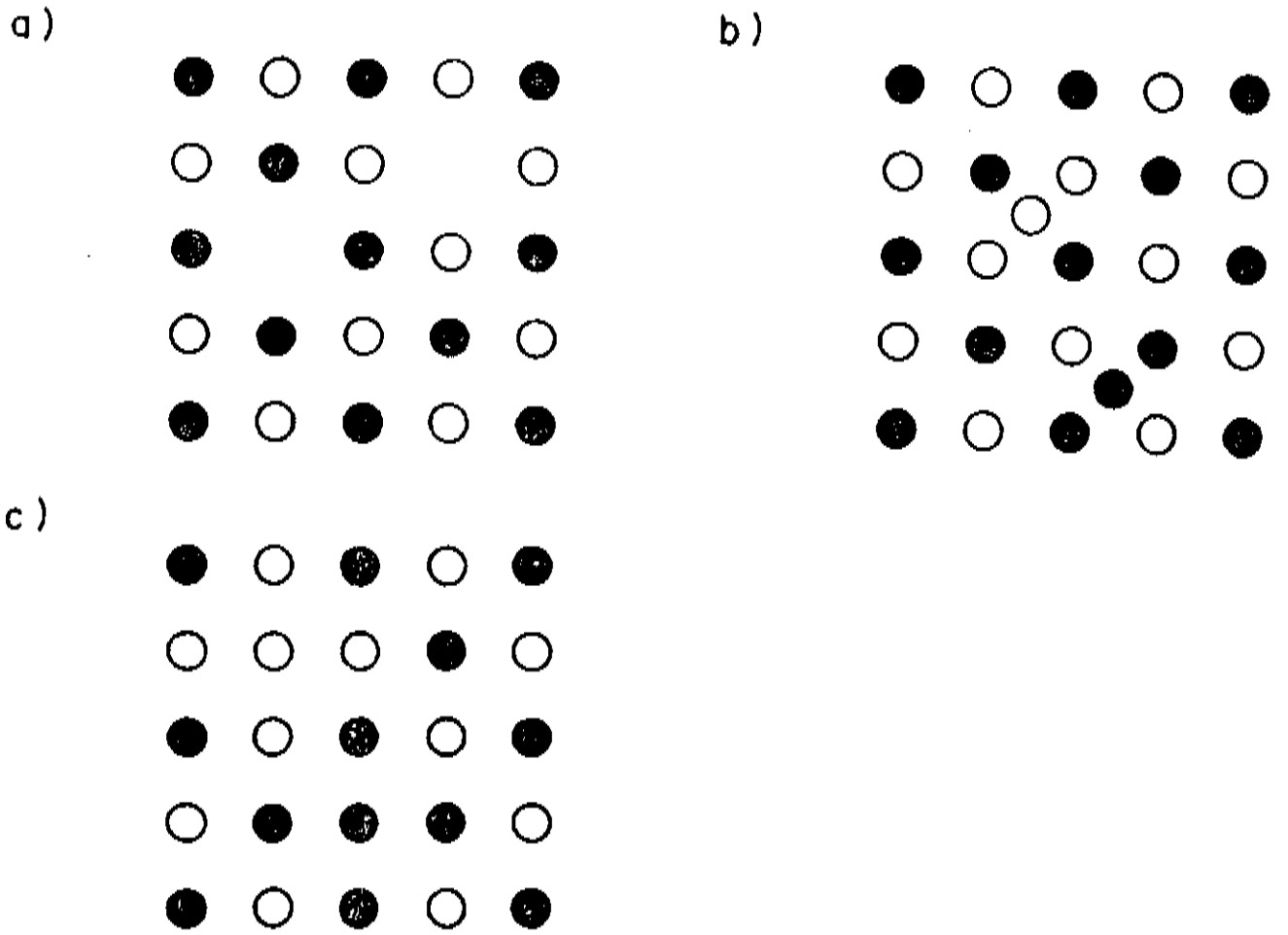

Figura 3.3 - Defeitos simétricos em um semicondutor composto. a) Vacâncias, b) Átomos intersticiais, e c) Anti-estruturas.

vacâncias, os átomos intersticiais podem atrair-se mutuamente, formando associações. Dessa forma teríamos um complexo $A_i B_i$, como mais uma nova forma de defeito pontual. O terceiro tipo são as anti-estruturas, e são formadas por um átomo A ocupando uma vacância B, ou um átomo B ocupando uma vacância A. Estes também podem associar-se aos pares, formando $A_B B_A$. Todos os tipos de defeitos simétricos podem ser vistos na figura 3.3.

Os defeitos assimétricos são classificados como:

4 - Frenkel ou vacância-intersticial, V_A e A_i; ou V_B e B_i.
5 - Vacância - anti-estrutura, V_A e A_B; ou V_B e B_A.
6 - Intersticial - anti-estrutura, A_i e B_A; ou B_i e A_B.

O defeito do tipo Frenkel é a formação de uma vacância com a emissão de um átomo intersticial, que ocupava antes uma posição substitucional. Temos assim um átomo A em sua posição na rede, que passou para um interstício, virando A_i, e deixando uma vacância, V_A; ou da mesma forma temos uma vacância deixada por um átomo B

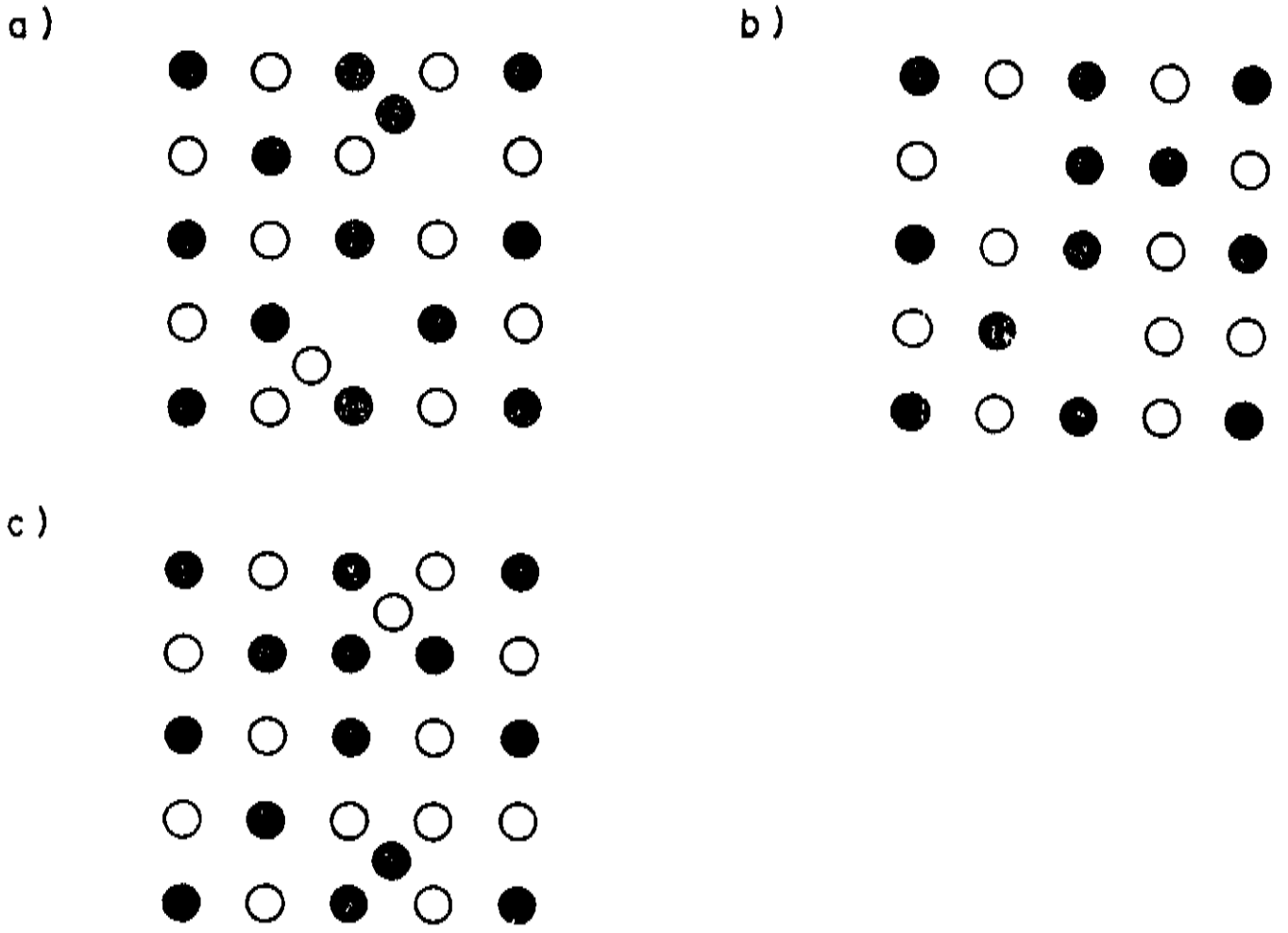

Figura 3.4 - Defeitos assimétricos em um semicondutor composto. a) Vacância-intersticial, b) Vacância - anti-estrutura, e c) Intersticial - anti-estrutura.

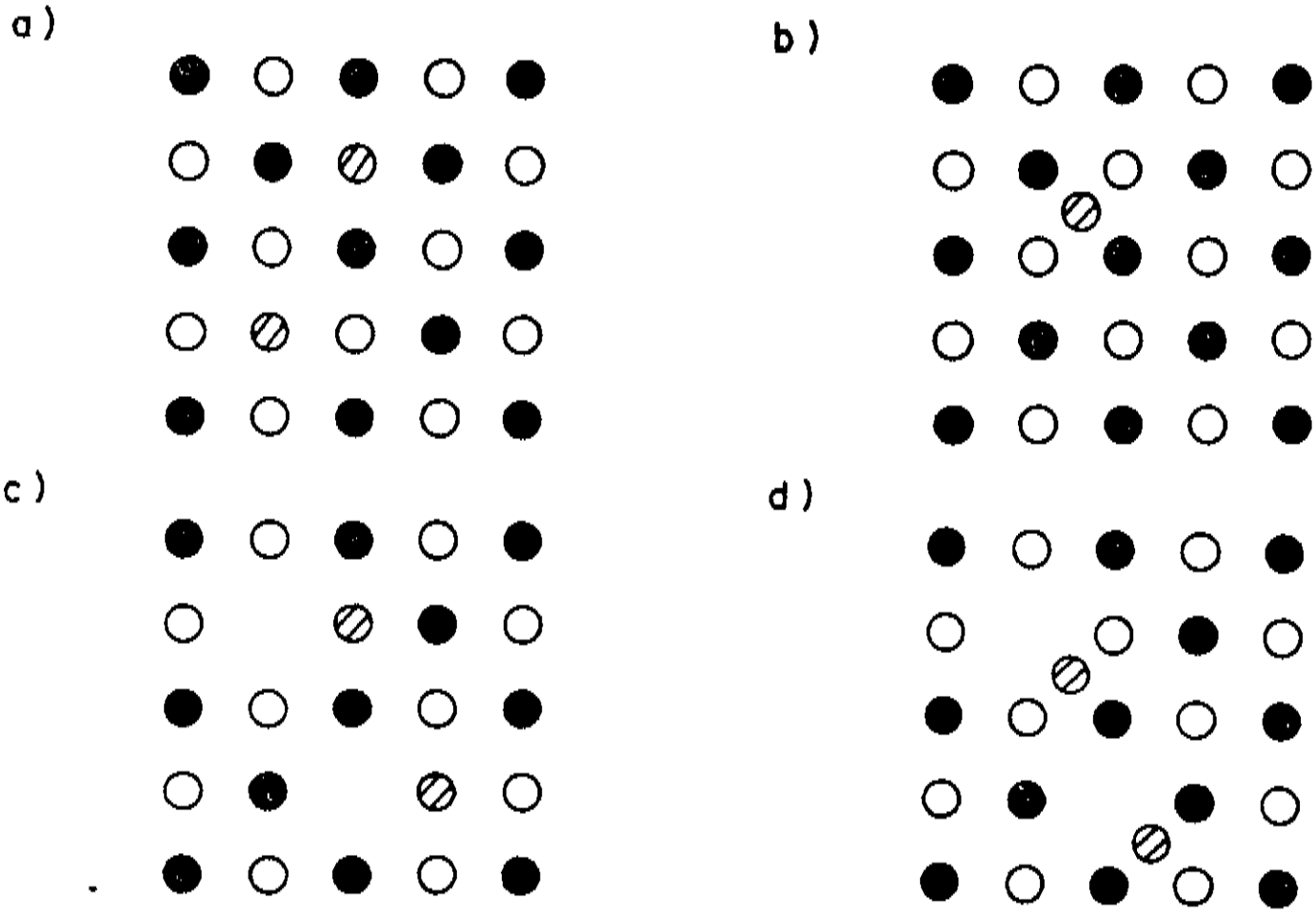

Figura 3.5 - Defeitos devidos a impurezas em um semicondutor composto. a) Impureza substitucional, b) Impureza intersticial, c) Complexo impureza substitucional - vacância, e d) Complexo impureza intersticial - vacância.

que tornou-se intersticial. A associação pode ocorrer na forma $A_i V_A$ ou $A_i V_B$, com complexos similares com B_i. O defeito vacância - anti-estrutura é formado pela migração de um átomo A, que estava em sua posição da rede, para a posição de uma vacância B vizinha. Ficamos dessa forma com um átomo A na posição de um B, com uma vacância A. O complexo possível neste caso é $A_B V_A$ ou $B_A V_B$. O último defeito, intersticial - anti-estrutura, é formado por um átomo B que empurrou um átomo A de sua posição, ocupando B a posição deixada por A. Ficamos assim com uma anti-estrutura, B_A, mais um intersticial, A_i. Esses defeitos são vistos na figura 3.4.

Um último grupo de defeitos pontuais são os gerados por impurezas, que iremos representar por I. Eles são:

7 - Impureza substitucional, I_A ou I_B.
8 - Impureza intersticial, I_i.
9 - Complexo impureza substitucional - vacância, $I_A V_B$; ou $I_B V_A$.
10 - Complexo impureza intersticial - vacância, $I_i V_A$; ou $I_i V_B$.

O primeiro tipo é um átomo de impureza ocupando uma posição da rede, que pode ser de um átomo A ou B. É esse o tipo de defeito que devemos esperar que seja gerado pela difusão. O segundo tipo é simplesmente um átomo de impureza numa posição intersticial. Os dois últimos tipos são associações de defeitos do cristal e de impurezas. Pelo menos para o complexo impureza substitucional - vacância existem evidências de sua ocorrência com doadores do grupo VIB em semicondutores III-V [4]. Um exemplo disso é a associação Te_{Sb} com vacâncias V_{In} no InSb [5]. Na figura 3.5 podemos ver os vários defeitos devidos às impurezas.

3.4 - Lei de Ação de Massas

Vimos no ítem anterior quais são os defeitos pontuais que ocorrem em semicondutores compostos. Vamos agora estudar como se relacionam as concentrações de vários defeitos presentes simultaneamente no semicondutor [3,6,7]. Precisamos primeiro

introduzir dois conceitos, o de pseudo-reações químicas e o de ação de massas.

Em geral podemos escrever reações químicas como,

$$aA + bB \rightarrow cC + dD \qquad (R1)$$

onde A e B são os reagentes e C e D os produtos da reação. As proporções em que cada reagente A e B são consumidos são dadas pelos seus coeficientes estequiométricos, respectivamente a e b, enquanto que as proporções formadas de C e D são dadas por c e d.

Da mesma forma que tratamos uma reação química acima, poderemos tratar a reação de formação de uma vacância. Por analogia chamamos essa reação de "pseudo-reação química", já que não há na realidade reações químicas ocorrendo. A reação de formação de uma vacância de arsênio no GaAs é,

$$As_{As} \rightarrow \frac{1}{2} As_2(g) + V_{As} \qquad (R2)$$

Temos no lado esquerdo da reação um átomo de arsênio na sua posição correta (As_{As}) na superfície do cristal. Esse átomo desprende-se da superfície passando à forma de vapor, que aqui seria As_2, deixando em seu lugar uma vacância, V_{As} (esta é uma das formas de gerar vacâncias em um cristal, a partir da sua superfície). Usando reações semelhantes a R2 podemos representar a formação de todos os tipos de defeitos pontuais.

O último conceito que veremos antes de estudar as pseudo-reações é o de ação de massas. Voltando à reação R1 observamos que ela pode ocorrer nas duas direções, dos reagentes formando produtos, ou dos produtos formando reagentes. No estado de equilíbrio as velocidades de reação em ambos os sentidos serão iguais e existirá um equilíbrio entre as concentrações de reagentes e produtos. Uma forma de relacionar essas concentrações é através da constante de equilíbrio, K(T), dada por,

$$K(T) = \frac{[C]^c\,[D]^d}{[A]^a\,[B]^b} \qquad (3.1)$$

onde os valores entre colchetes referem-se às concetrações de

reagentes e produtos em equilíbrio, e os expoentes são os respectivos coeficientes estequiométricos. Da equação (3.1) podemos observar que para K maior que um a concentração de produtos será maior que a de reagentes, enquanto que para K menor que um teremos mais reagentes do que produtos. O valor de K indica portanto a direção preferencial em que a reação irá ocorrer. Seu valor é uma função da temperatura e energia livre de formação de produtos, ΔG, e é dado por,

$$\Delta G = - RT \ln(K) \qquad (3.2)$$

onde T é a temperatura e R a constante dos gases.

3.4.1 - Reações de Formação de Defeitos Pontuais

Tendo visto o que são pseudo-reações e a lei de ação de massas passamos agora a estudar as pseudo-reações de formação de defeitos pontuais.

A primeira pseudo-reação, que passaremos a chamar simplesmente de reação, já foi vista e é a de formação de uma vacância de arsênio no GaAs,

$$As_{As} \rightarrow \frac{1}{2} As_2(g) + V_{As} \qquad (R3)$$

Para essa reação temos,

$$K = \frac{[V_{As}] \; p_{As_2}^{1/2}}{[As_{As}]} \qquad (3.3)$$

Como a concentração de átomos de arsênio em suas devidas posições é grande e praticamente constante, ela pode ser absorvida na constante de equilíbrio, e temos assim,

$$K_1 = [V_{As}]\ P_{As_2}^{1/2} \qquad (3.4)$$

Observe que a concentração de átomos gasosos de As_2 é dada simplesmente pela sua pressão parcial de vapor. Caso a espécie gasosa a ser formada na reação for As_4 teremos,

$$As_{As} \rightarrow \frac{1}{4}\ As_4(g) + V_{As} \qquad (R4)$$

com a respectiva constante de equilíbrio,

$$K_2 = [V_{As}]\ P_{As_4}^{1/4} \qquad (3.5)$$

Das equações (3.4) e (3.5) podemos observar que a concentração de vacâncias de arsênio será inversamente proporcional a uma potência da pressão parcial externa de arsênio, a uma temperatura constante.

Para a formação de vacâncias de gálio temos analogamente,

$$Ga_{Ga} \rightarrow Ga(l) + V_{Ga} \qquad (R5)$$

com a única diferença que a fase externa do gálio é líquida e não sólida como no arsênio. A constante de equilíbrio será,

$$K_3 = [V_{Ga}]\ [Ga(l)] \qquad (3.6)$$

O produto das concentrações de vacâncias pode ser obtido através da reação de decomposição do GaAs,

$$GaAs(s) \rightarrow Ga(l) + \frac{1}{2}\ As_2(g) \qquad (R6)$$

com a respectiva constante de equilíbrio,

$$K_4 = [Ga(l)]\ P_{As_2}^{1/2} \qquad (3.7)$$

Estamos supondo que a espécie gasosa do arsênio é As_2, e que da mesma forma que na formação de vacâncias de gálio, a sua espécie na fase externa é gálio líquido. Substituindo o valor de [Ga(l)]

de (3.6) e P_{As_2} de (3.4) em (3.7), obtemos,

$$[V_{As}]\ [V_{Ga}] = \frac{K_2\ K_3}{K_4} = K_5 \qquad (3.8)$$

e observamos que independente de qual seja a espécie gasosa de arsênio, o produto das concentrações de vacâncias é constante e apenas uma função da temperatura.

Outra reação importante é a formação de divacâncias de gálio, V_{2Ga}, que são a associação de vacâncias aos pares,

$$2\ V_{Ga} \rightarrow V_{2Ga} \qquad (R7)$$

com a constante de equilíbrio,

$$K_6 = \frac{[V_{2Ga}]}{[V_{Ga}]^2} \qquad (3.9)$$

Dessa equação concluímos que a concentração de divacâncias varia com o quadrado da concentração de vacâncias de gálio.

Da mesma forma que descrevemos as reações de formação de vacâncias podemos descrever a formação de átomos intersticiais e o seu produto de concentrações. Podemos mostrar que a concentração de um determinado átomo intersticial é proporcional a alguma potência da sua pressão de vapor ou concentração na fase externa, e que o produto da concentração de intersticiais é constante.

A incorporação de impurezas também pode ser descrita na forma de uma reação. Supondo que a impureza se apresenta na fase externa na forma gasosa como I_2, e que ela irá ocupar a posição de um átomo de arsênio, teremos,

$$I_2(g) + 2\ V_{As} \rightarrow 2\ I_{As} \qquad (R8)$$

com a respectiva constante de equilíbrio,

$$K_7 = \frac{[I_{As}]^2}{P_{I_2}\ [V_{As}]^2} \qquad (3.10)$$

podemos substituir em (3.10) o valor de $[V_{As}]$ de (3.4), e achamos,

$$[I_{As}]^2 \frac{P_{As_2}}{P_{I_2}} = K_7 K_1^2 = K_8 \qquad (3.11)$$

A equação (3.11) nos mostra que para uma temperatura constante um aumento da pressão parcial da impureza (fase vapor) aumenta a concentração de impurezas em posições substitucionais no cristal. Já um aumento da pressão de As_2 irá diminuir a concentração de impurezas, uma vez que esse aumento da pressão ocasiona uma diminuição na concentração de vacâncias de arsênio, e portanto a impureza terá um menor número de posições para ocupar.

Por fim temos as chamadas reações de ionização. Uma impureza em uma posição substitucional pode não dopar efetivamente a rede cristalina. Isso pode acontecer, por exemplo, em uma baixa temperatura do cristal, o que implica em uma baixa energia térmica que pode ser insuficiente para ionizar a impureza. A reação de ionização de uma impureza doadora é,

$$I_{As} \rightarrow I_{As}^{+} + e \qquad (R9)$$

onde I_{As}^{+} denota uma impureza doadora ionizada, e e um elétron. A constante de equilíbrio para essa reação é dada por,

$$K_9 = \frac{[I_{As}^{+}]\, n}{[I_{As}]} \qquad (3.12)$$

onde n é a concentração de elétrons. A equação acima nos mostra um efeito interessante. Suponhamos que por algum meio conseguimos aumentar a concentração de elétrons no cristal, introduzindo uma segunda impureza doadora, por exemplo. O valor da concentração da primeira impureza, $[I_{As}]$ pode ser considerado constante, supondo que a segunda impureza não afeta essa concentração. Da equação (3.12) vemos que o produto da concentração de elétrons e das impurezas ionizadas deve ser constante. Dessa forma como n aumenta, $[I_{As}^{+}]$ deve diminuir. Concluimos que a ionização de uma impureza é afetada pela presença ou não de outras impurezas, quer sejam elas doadoras ou aceitadoras.

Vacâncias e átomos intersticiais também podem ser ioniza-

dos. Em compostos III-V esperamos que as vacâncias do elemento III sejam negativamente carregadas. Isso deve-se ao fato de que quando retiramos um átomo do elemento III, este deixa quatro ligações covalentes abertas (quatro elétrons), enquanto que o átomo retirado leva consigo três elétrons. Ficamos dessa forma com um elétron livre junto à vacância. Com um raciocínio equivalente podemos concluir que as vacâncias do elemento V serão positivamente carregadas. Átomos intersticiais irão tornar-se carregados numa tentativa de atingir a configuração de um gás nobre, que seria a mais estável. Assim átomos intersticiais do elemento V ou de impurezas doadoras devem ser negativamente carregados, enquanto que os do elemento III ou de impurezas aceitadoras seriam positivamente carregados. Embora esperemos essas várias ionizações dos defeitos, elas não são observadas no GaAs, razão pela qual não iremos entrar nos detalhes das suas reações de ionização. Na realidade entre os elementos III-V poucos materiais apresentam algum defeito simples ionizado. Apenas os defeitos complexos irão mostrar algum grau de ionização. Um exemplo será discutido no ítem 3.7, a ionização de uma associação enxofre (doador raso) com uma vacância de gálio.

3.5 - Diagramas de Fase

Vimos no ítem anterior que em equilíbrio com o cristal deve haver um vapor externo dos elementos IIIB e VB, supondo um semicondutor binário. Vimos ainda que essa pressão de vapor externa irá determinar as concentrações de vacâncias no equilíbrio. Dessa forma alterando essa pressão podemos modificar as concentrações de vacâncias, o que equivale a uma pequena mudança da estequiometria do cristal. Obviamente a mudança na estequiometria não pode ser muito grande, pois isso implicaria na fusão de uma parte do cristal, formando novas fases sólidas ou líquidas, cujas composições e estruturas não lembrariam em nada a do semicondutor original. Um exemplo é o GaAs, onde retiramos continuamente o arsênio do cristal a alta temperatura, deixando

apenas o gálio. Podemos chegar a um ponto extremo onde temos apenas gálio e arsênio líquido e nenhum GaAs sólido. Isso nos leva a fazer uma pergunta, quais são os valores ou faixas de concentrações possíveis de vacâncias que podem ser conseguidas mudando-se a pressão de vapor sobre o cristal sem que haja a formação de uma nova fase sólida ou líquida? Essa é uma questão muito importante, pois durante a difusão não desejamos a formação de novas fases, quer sejam sólidas ou líquidas. Desejamos apenas obter o semicondutor dopado com a impureza, o que vem a ser a mesma fase que a inicial, o semicondutor antes da difusão.

A resposta à questão anterior é dada pelos chamados diagramas de fase [6,7,8]. Esses diagramas nos dão o número de fases de um sistema, quais são essas fases e suas composições a partir da concentração de cada componente do sistema e da temperatura. O diagrama de fase também irá nos mostrar qual é o número de graus de liberdade do sistema, ou seja, o número de variáveis do sistema que devemos escolher para fixar o seu estado. Esse conceito será explicado a seguir.

O número de graus de liberdade, F, é achado da regra de fase de Gibbs, que o relaciona com o número de fases do sistema, P, e o número de componentes do sistema, C. A regra de fase nos dá,

$$F = C - P + 2 \qquad (3.13)$$

Entendemos por uma fase do sistema como uma porção deste, homogênea e com um contorno definido. O número de fases é simplesmente achado contando-se as porções homogêneas do sistema, tomando-se o cuidado de não contar várias vezes uma fase descontinua. Já o número de componentes é achado do menor número de compostos ou elementos químicos usados para descrever a composição química de cada fase do sistema. Para um sistema contendo água e o seu vapor (H_2O líquida e H_2O vapor) o número de componentes é um, apenas H_2O. Para um sistema contendo arsênio e gálio, que formam o semicondutor GaAs, o número de componentes é dois, e cada componente é um dos dois elementos químicos, ao invés de ser um composto, H_2O, como no exemplo anterior.

Voltando ao exemplo da água em equilíbrio com o seu vapor, o valor de C é igual a um e P igual a dois, o que dá F igual a um. Temos portanto um grau de liberdade, e podemos fixar a nosso gosto um parâmetro do sistema, a pressão ou a temperatura. Supondo que escolhemos um valor para a temperatura, a pressão do sistema estará imediatamente fixada, e será nesse caso a pressão de vapor da água na temperatura que escolhemos. Do mesmo modo se fixarmos a pressão a temperatura estará imediatamente determinada e será a correspondente à pressão de vapor escolhida.

O número de graus de liberdade do sistema é muito importante quando estamos escolhendo a composição de uma fonte de difusão (fase externa ao semicondutor que contém a impureza a difundir). Podemos escolher de tal forma a composição da fonte que caimos numa região do diagrama que resulta em F igual a zero. Em uma região desse tipo as pressões de vapor dos componentes da fonte serão fixas, e supondo que possamos variar com um pouco de liberdade a composição sem sair fora da região, as pressões serão independentes da composição. A importância disso ficará mais clara no item 3.6, onde mostraremos que o coeficiente de difusão (que indica a facilidade com que uma impureza difunde em um material) depende da concentração de vacâncias, que depende por sua vez das pressões externas de vapor. Dessa forma em uma região de F igual a zero o coeficiente de difusão será fixo, e as difusões poderão ser realizadas com grande reprodutibilidade.

3.5.1 - Diagramas Binários

Os diagramas binários são os formados pela mistura de dois componentes. Temos portanto C igual a dois, e supondo que fixamos a pressão, ou seja usamos um dos graus de liberdade, o número de graus restante será,

$$F = 3 - P \qquad (3.14)$$

Na figura 3.6a podemos ver um diagrama binário dos componentes A e B. Nesse diagrama a porcentagem do componente A é

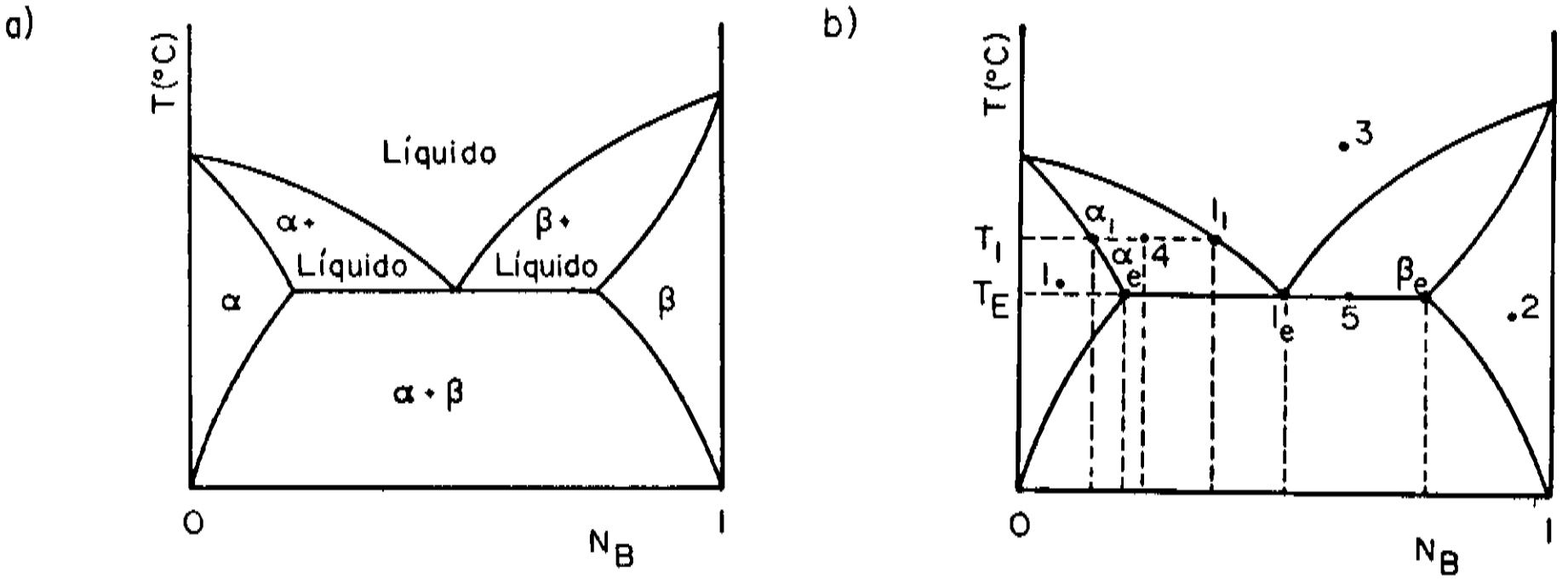

Figura 3.6 - Diagrama binário. a) Fases que ocorrem em cada região, b) Composição das diversas fases.

colocada no eixo x, e no eixo y temos a temperatura do sistema. A porcentagem de B será simplesmente dada por 100% - A%. No diagrama a região onde uma ou mais fases coexistem juntamente é separada das demais regiões, onde grupos diferentes de fases vão coexistir, por linhas cheias. Observe que no diagrama da figura 3.6a os elementos A e B possuem solubilidade parcial, ou seja, não podemos dissolver indefinidamente um no outro, sem que seja formada uma nova fase. Isso ocorre entre as regiões α e β, onde temos a região $\alpha+\beta$, que são as duas fases (α e β) misturadas. Observe ainda a existência de duas regiões, $\alpha+l$ e $\beta+l$, onde temos respectivamente a presença conjunta da fase α mais um líquido, e da fase β mais um líquido.

Agora suponha que o sistema está em uma região onde temos apenas uma fase, α, β, ou líquido. A regra de fase nos dá dois graus de liberdade, e poderemos escolher a temperatura e composição da fase. Na figura 3.6b temos os pontos 1, 2, e 3 como exemplos. Já se estivermos em uma região de duas fases, como por exemplo o ponto 4 da figura 3.6b, o número de graus de liberdade será um. Nesse caso fixando a temperatura no valor T_1 como mostra a figura, a composição das fases α e líquida estará imediatamente

fixada e corresponderá respectivamente aos pontos α_1 e l_1 no diagrama. Observe que uma reta isotérmica liga os pontos l_1, 4 e α_1. Essa reta que determina a composição das varias fases é chamada de "tie line". Da mesma forma se fixarmos a composição da fase α em α_1, a da fase líquida e a temperatura estarão fixadas. Temos por fim o caso de três fases que coexistem, as fases α, β e o líquido. O número de graus de liberdade é zero, e as composições das fases serão as dos pontos α_e, β_e e l_e da figura, e a temperatura correspondente será T_e, a chamada temperatura eutética. O ponto 5 na figura é um exemplo onde podemos ter 3 fases.

3.5.2 - Diagramas Ternários

Vimos acima qual é a utilidade dos diagramas de fase para representar a concentração de vacâncias em um semicondutor e a presença de várias fases em um sistema. Passamos agora a estudar os diagramas ternarios [8], formados por três elementos, e que possuem grande importância no estudo da difusão em semicondutores compostos.

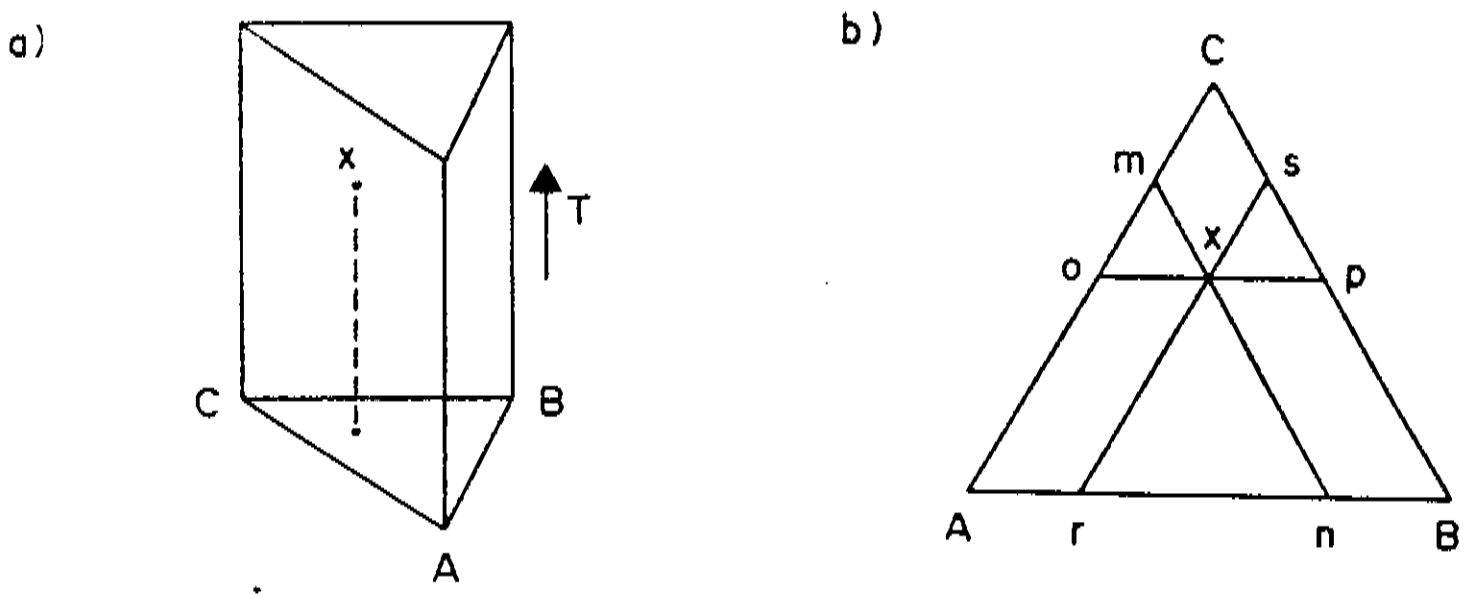

Figura 3.7 - Diagrama ternário. a) Representação do diagrama com o uso de um prisma tridimensional. No sentido da altura temos a temperatura. b) Projeção do ponto x na base do prisma, de onde podemos ver a composição do ponto.

O diagrama de fase ternário tem a forma de um prisma triangular no espaço tridimensional, como é mostrado na figura 3.7a. A base desse prisma é a composição das fases ou do sistema, sendo que em cada vértice desse triângulo temos apenas um elemento puro. A composição de um ponto x qualquer é obtida traçando-se retas paralelas aos lados do triângulo e que passam pelo ponto. Dessa forma a porcentagem de um elemento A é obtida pelo comprimento da reta traçada, do ponto x ao lado do triângulo oposto ao vértice A, como mostra a figura 3.7b, onde o ponto foi projetado na base do prisma. Temos dessa forma,

$$\%A = xs = xp = nB = mC \qquad (3.15a)$$

$$\%B = xo = xm = rA = sC \qquad (3.15b)$$

$$\%C = xr = xn = oA = pB \qquad (3.15c)$$

A temperatura é colocada no eixo vertical, no sentido da altura do prisma.

Já que o diagrama de fase ternário deve ser representado no espaço bidimensional (a superfície de uma folha de papel), optamos pela representação de seções isotérmicas, como mostra a figura 3.8a onde vemos o diagrama do Ga-As-Zn (este foi tomado

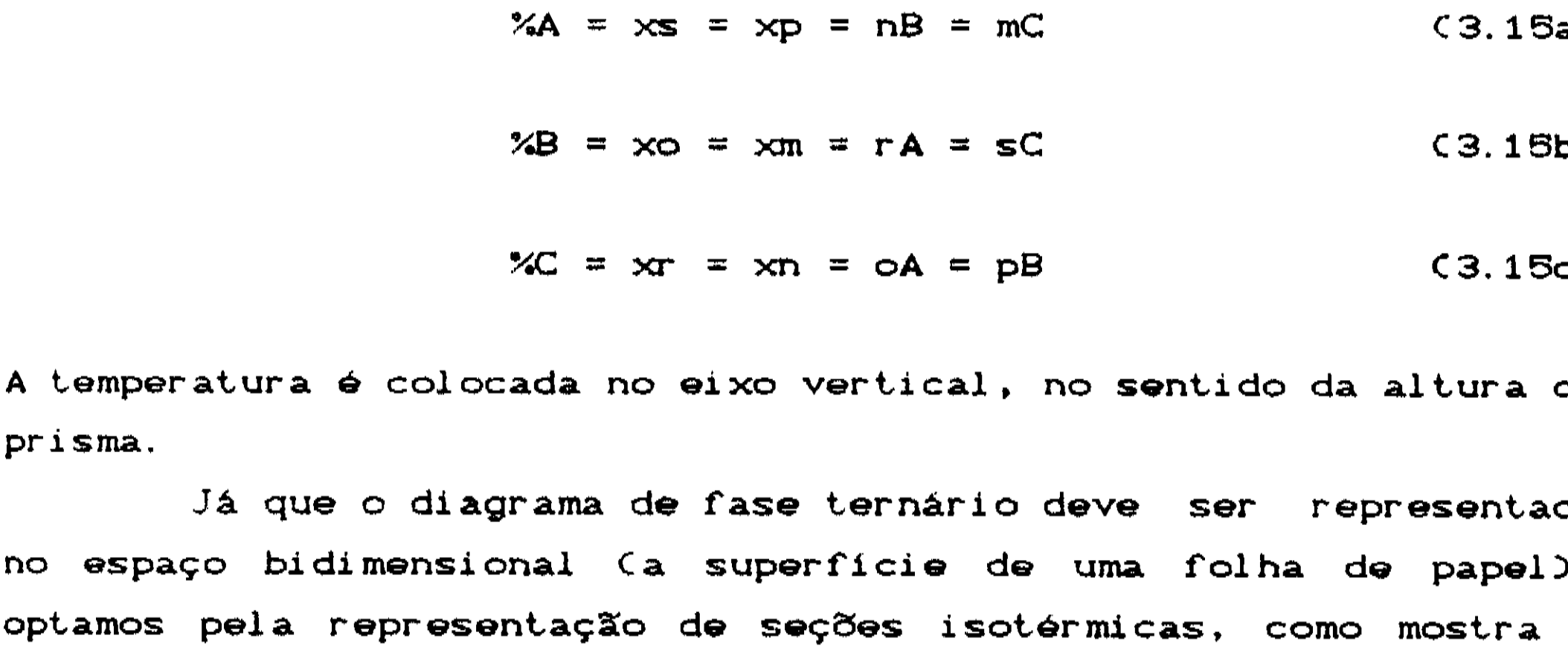

Figura 3.8 - Diagrama ternário do Ga-As-Zn [9]. a) Projeção de isotérmicas. b) Seção isotérmica mostrando a composição do GaAs em equilíbrio com o zinco.

como exemplo devido à grande quantidade de informações existentes). Podemos também representar a projeção da intersecção de um plano isotérmico com as superfícies que separam as fases, como mostrado na figura 3.8b.

A regra de fase agora nos dá para um sistema com três componentes um número de graus de liberdade de,

$$F = 5 - P \qquad (3.16)$$

Para uma região de duas fases, no diagrama ternário, as "tielines" também são isotérmicas, porém há uma complicação adicional. No diagrama binário uma isotérmica é simplesmente uma linha horizontal no diagrama, enquanto que no ternário uma isotérmica é um plano. Teremos portanto que determinar experimentalmente qual é a direção da linha, o que não é necessário no diagrama binário. Para o caso de regiões de três fases teremos "tie"-triângulos e os vértices desses triângulos irão determinar a composição de cada fase.

Voltando à figura 3.8b, vemos uma "tieline" que passa pelo ponto de composição inicial da fonte de difusão, ponto x na região B. Essa "tieline" irá cruzar a linha limite entre a região A, que contém apenas uma fase líquida, e a região B no ponto x'. Esse ponto será a composição da fase líquida que existe na região B. A "tieline" também cruza a linha limite entre a região B e a região de GaAs sólido com zinco em solução sólida (uma única fase), ponto x''. Esse ponto será a composição da fase sólida da região B, que é GaAs com zinco em solução sólida. Se a amostra de GaAs que recebe a difusão é muito pequena, de tal forma que a sua massa não influi na composição da fonte, se a levarmos em consideração como parte da fonte de difusão, poderemos dizer que após um tempo muito grande de difusão a amostra irá atingir a composição do ponto x'' (composição da fase sólida). Como veremos mais à frente, isso só acontecerá para tempos infinitamente grandes, porém a superfície da amostra prontamente atinge uma concentração de equilíbrio, que será a do ponto x''. Podemos concluir que o diagrama de fase ternário é útil para determinarmos qual será a concentração superficial de impurezas em uma amostra após a difusão, desde que

conheçamos a composição da fonte, a temperatura de difusão e o diagrama de fases, que nem sempre é conhecido.

3.6 - Difusão em Semicondutores

Vimos na introdução que a difusão deve dissipar qualquer gradiente de concentração dentro de um meio único. Vimos também que o movimento aleatório de átomos ou moléculas do meio e do material que difunde é responsável pela dissipação dos gradientes de concentração. Resta uma pergunta, qual é a "força" que dissipa os gradientes de concentração, impedindo as moléculas ou átomos do material que difunde de se concentrarem em uma pequena região?, ou ainda qual é a propriedade do meio que nos indicará se ocorre ou não a difusão? As respostas são dadas pela termodinâmica, que nos diz que um sistema isolado só estará em equilíbrio quando a sua entropia for a máxima possível. Entendemos a entropia como uma medida do grau de organização de um sistema. Assim um sistema ordenado terá uma entropia menor que o mesmo sistema num estado de menor ordenação ou maior desordem. Para o exemplo na introdução, uma gota de tinta em um copo com água, o estado inicial do sistema, a gota de tinta concentrada após ter pingado na água, é um estado de menor entropia do que o estado final, a tinta uniformemente distribuida na água, já que o estado inicial tem uma maior organização, pois a tinta e a água estão restritas a regiões bem definidas do copo. A termodinâmica nos diz ainda que o processo de mistura da tinta e da água é um processo irreversível, pois de nenhuma forma conseguimos, uma vez misturadas, separar a tinta da água. Por ser um processo irreversível, o estado final terá maior entropia que o inicial, e portanto será o estado de equilíbrio.

Voltamos agora ao caso de um semicondutor em contato com alguma fase externa que contém uma impureza. Vamos supor que a concentração da impureza nessa fase seja uniforme e que no semicondutor seja zero. Não seria esse o estado de equilíbrio do sistema, já que a distribuição de impurezas é uniforme nas duas

fases? (o semicondutor forma uma das fases do sistema). Qual será o motivo que leva parte das impurezas a penetrarem no cristal? As respostas são de novo dadas pela termodinâmica, porém vamos primeiro definir o que é potencial químico, μ. Vamos supor uma certa quantidade de moles de impureza, n_i, que será dissolvida em algum solvente, n_s. Supomos ainda a temperatura e pressão constantes. A energia livre de Gibbs por mol de impureza no seu estado puro será g_i. Definimos assim o potencial químico da impureza, μ_i, como a sua energia livre em uma solução, e temos,

$$\mu_i = g_i + RT \ln \gamma_i X_i \qquad (3.17)$$

onde R é a constante dos gases, T a temperatura, γ_i é o coeficiente de atividade, que vale um para uma solução ideal, onde nesse caso dizemos que a solução segue a lei de Raoult, e para o caso de γ_i constante e diferente de um temos a lei de Henry, e por fim X_i é a fração molar da impureza que é dada por,

$$X_i = \frac{n_i}{n_i + n_s} \qquad (3.18)$$

Como X_i é menor que um, o potencial químico da impureza em solução sempre será menor que a sua energia livre. Isso vale também para a impureza em solubilidade sólida no semicondutor.

A termodinâmica nos mostra que um sistema está em equilíbrio quando o potencial químico de cada elemento é o mesmo em todas as fases, e mais ainda, quando é uniforme em qualquer ponto de qualquer fase. Isso explica porque a impureza deve penetrar no semicondutor e difundir através dele. Isso deve acontecer para que o sistema semicondutor mais fase externa com impureza atinja o equilíbrio, que é conseguido quando uma parte da impureza penetra no semicondutor e difunde através dele resultando em uma distribuição uniforme de impurezas, garantindo um mesmo potencial químico na fase externa e no semicondutor.

Antes de ser atingido o equilíbrio termodinâmico no semicondutor deve haver um fluxo de impurezas neste, J_x. Definindo x como a profundidade, medida a partir da superfície, e sendo C(x) a concentração de impurezas em função da profundidade, temos,

$$J_x = - D \frac{\partial C(x)}{\partial x} \qquad (3.19)$$

onde D é uma constante de proporcionalidade chamada de coeficiente de difusão. Sendo J dado em $cm^{-2}s^{-1}$ e $\partial C(x)/\partial x$ em cm^{-4}, D será dado por cm^2/s. A equação (3.19) relaciona o gradiente de concentração de impurezas ao seu fluxo, e é puramente fenomenológica, pois não nos dá nenhuma informação sobre o valor de D. Esta equação é a chamada de primeira lei de Fick.

É claro que a concentração não é apenas uma função da profundidade. Ela é também uma função do tempo (por isso usamos o símbolo de derivada parcial em (3.19)). Isso pode ser deduzido observando-se que imediatamente após o início da difusão a superfície terá atingido a concentração de equilíbrio, C_o, enquanto que a alguma profundidade x arbitrária, a concentração será zero. Já para um tempo t_1 arbitrário, teremos à mesma profundidade x uma concentração C_1, que é algum valor entre 0 e C_o. Para um tempo infinito o semicondutor teria atingido o equilíbrio, e a concentração seria uniforme pelo cristal e igual a C_o. A difusão é portanto um processo que ocorre em um sistema fora do equilíbrio.

Podemos obter facilmente uma equação que nos dá a taxa de variação de impurezas. Partimos de um elemento de volume de arestas Δx, Δy e Δz, como mostra a figura 3.9. A taxa de material que entra no cubo, Q_e, supondo fluxo apenas na direção x, é,

$$Q_e = J_x \, \Delta y \, \Delta z \qquad (3.20)$$

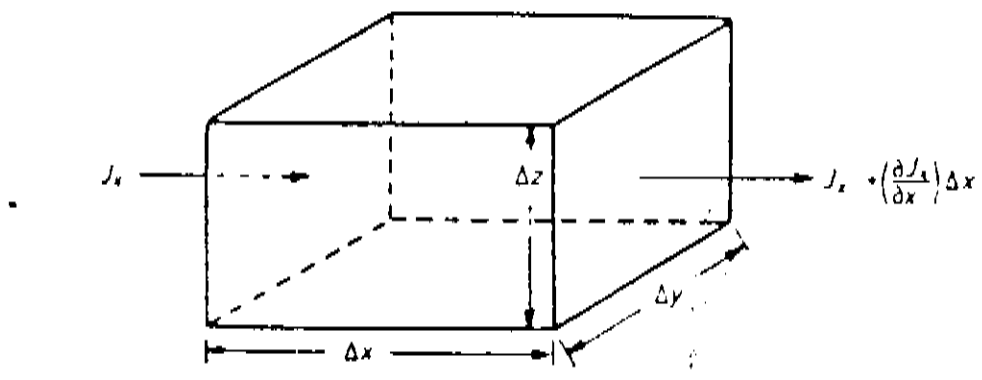

Figura 3.9 - Elemento de volume ΔxΔyΔz para cálculo dos fluxos de material que difunde.

a taxa que sai, Q_s, será,

$$Q_s = \Delta y\ \Delta z \left(J_x + \frac{\partial J_x}{\partial x} \Delta x \right) \qquad (3.21)$$

A taxa de acumulação de material no cubo, $\Delta x\ \Delta y\ \Delta z\ \partial C/\partial t$, será igual à taxa de material que entra no volume menos a taxa que sai,

$$\Delta x\ \Delta y\ \Delta z \frac{\partial C}{\partial t} = \Delta y\ \Delta z \left(J_x - J_x - \frac{\partial J_x}{\partial x} \Delta x \right) \qquad (3.22)$$

de onde tiramos,

$$\frac{\partial C}{\partial t} = - \frac{\partial J_x}{\partial x} \qquad (3.23)$$

Substituindo (3.19) em (3.23) obtemos,

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial C(x,t)}{\partial x} \right) \qquad (3.24)$$

Se D não for uma função de x podemos escrever,

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} \qquad (3.25)$$

Esta é a segunda lei de Fick, que relaciona a taxa de variação da concentração de impurezas com a sua derivada em relação à posição.

A equação (3.25) pode ser resolvida de tal forma que possamos obter C(x,t) para quaisquer valores de x ou t. Um caso típico nos processos de difusão em semicondutores é considerar a fase externa que contém a impureza com um volume muito grande. Dessa forma a fase externa é praticamente uma fonte infinita de impurezas para o semicondutor. Isso equivale a dizer que uma vez que a difusão comece, a superfície do semicondutor irá atingir uma concentração de equilíbrio, C_o, que não irá mudar, por mais longo que seja o tempo tomado pela difusão. As condições de contorno nesse caso para a solução de (3.25) são,

$$C(x,t) = \begin{cases} 0 & , \text{ p/ } x>0,\ t=0 \\ C_o & , \text{ p/ } x=0,\ t>0 \end{cases} \qquad (3.26)$$

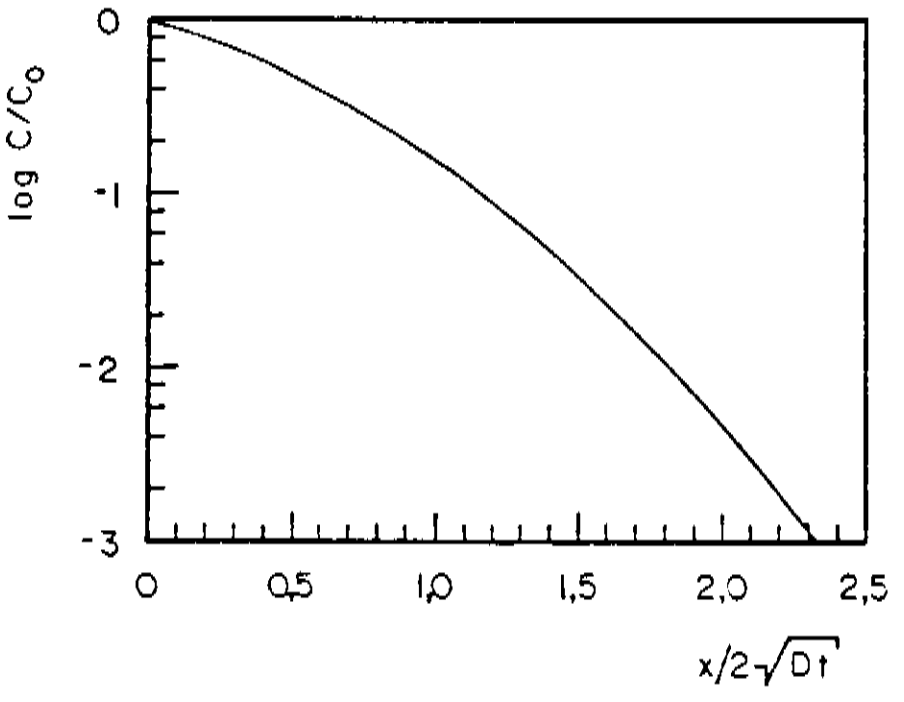

Figura 3.10 - Perfil de difusão (concentração x profundidade) obtido da solução de (3.25) com as condições de contorno (3.26).

A técnica usada na solução é através de transformadas de Laplace. A solução de (3.25) sujeita às condições de contorno acima é,

$$C(x,t) = C_o \, \mathrm{erfc}\left(\frac{x}{2\sqrt{D\,t}} \right) \qquad (3.27)$$

onde a função erfc(z), que é chamada de função erro complementar, é definida como,

$$\mathrm{erfc}(z) = \frac{1}{2\sqrt{\pi}} \int_z^{\infty} \exp\left(- \frac{z'^2}{2} \right) dz' \qquad (3.28)$$

Na figura 3.10 podemos ver um gráfico que mostra a concentração em função da profundidade (equação (3.27)), o que é normalmente chamado de perfil de difusão. Observe que nesse gráfico a maior concentração de impurezas é achada na superfície, o que é uma característica do processo de difusão.

3.6.1 - Mecanismos de Difusão em Semicondutores

Tendo visto alguns princípios da difusão passamos a estu-

dar o seu mecanismo físico em semicondutores, que nos interessa. Como vimos nos itens 3.2 e 3.3, um semicondutor é um cristal, e como tal é formado por um arranjo periódico de átomos, arranjo este que pode apresentar defeitos como vacâncias e átomos intersticiais. Um grande número de mecanismos físicos de difusão estão ligados diretamente aos defeitos pontuais, de onde vem a importância de estudá-los. Ao todo temos 11 diferentes mecanismos listados abaixo [10]:

1 - Troca
2 - Anel
3 - Intersticial
4 - Geração de intersticial
5 - "Crowdion"
6 - Vacância
7 - Divacância
8 - Relaxação
9 - Dislocação
10 - Contorno de grão
11 - Superficial

Dos mecanismos acima, seis são relacionados aos defeitos pontuais (mecanismos 3 a 8). Dois dos três últimos são relacionados a defeitos extensos, a difusão pelas dislocações e por contornos de grão.

Iremos mostrar que o coeficiente de difusão é normalmente proporcional à concentração de defeitos (quando a difusão realiza-se com a participação de um defeito). O número de defeitos extensos é pouco sensível à temperatura, variando muito pouco com uma mudança desta. Já a concentração de defeitos pontuais depende muito da temperatura. Esperamos dessa forma que em altas temperaturas a concentração de defeitos pontuais seja muito grande e tenhamos dessa forma difusão por estes como sendo o principal mecanismo, sendo o por defeitos extensos um mecanismo secundário e de pouca importância. Como estamos interessados nos mecanismos de alta temperatura (no GaAs a difusão de doadores só ocorre a altas temperaturas) não iremos estudar a difusão por dislocações e

contornos de grão, e nem a superficial que não resulta em uma camada de semicondutor dopado.

Começamos com o mecanismo de troca, que é representado na figura 3.11a. Neste primeiro mecanismo um átomo simplesmente difunde no cristal trocando de posição com outro átomo vizinho. Isso é pouco provável de ocorrer em cristais com uma grande densidade de empacotamento, já que o mecanismo exige que um átomo passe próximo ao outro com o qual ele troca de posição, o que requer uma rede com espaços relativamente abertos. Já no segundo mecanismo, o de anel, isso não é necessário. Neste caso três ou mais átomos trocam de posição, cada um ocupando a posição do seu vizinho seguinte no anel, como mostra a figura 3.11b. Embora esse mecanismo seja possível em cristais com alta densidade de empacotamento ele é pouco provável, uma vez que requer o movimento simultâneo de três ou mais átomos.

Os dois mecanismos acima são os que ocorreriam no corpo de cristais perfeitos. A introdução de defeitos pontuais em um cristal cria novos mecanismos de difusão. Temos por exemplo o terceiro mecanismo da lista, a difusão por átomos intersticiais, figura 3.12a. Neste caso o átomo que difunde tem um pequeno tamanho quando comparado aos átomos da rede. A difusão é realizada simplesmente com um átomo pulando de um interstício para outro. O pequeno tamanho do átomo irá facilitar esses pulos, mesmo em cristais altamente empacotados. Isso nos leva à conclusão de que a

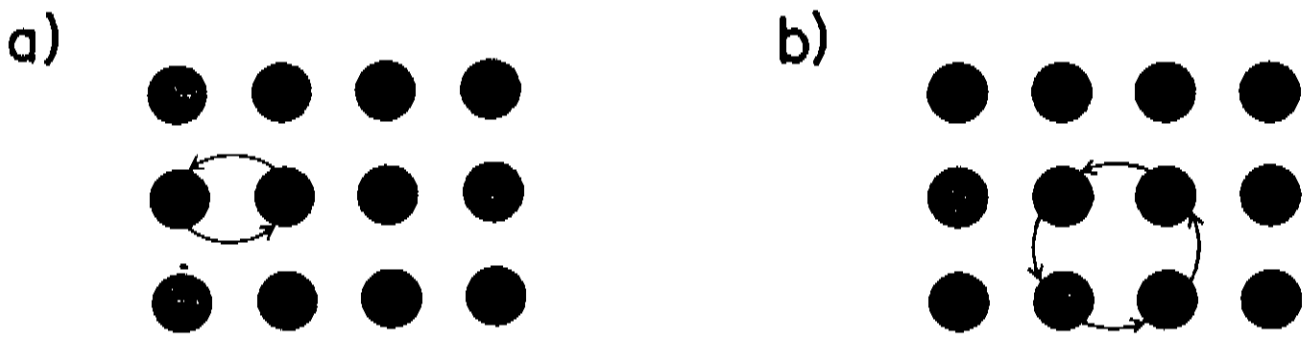

Figura 3.11 - Mecanismos de difusão em um cristal perfeito. a) Mecanismo de troca, e b) Mecanismo de troca em anel.

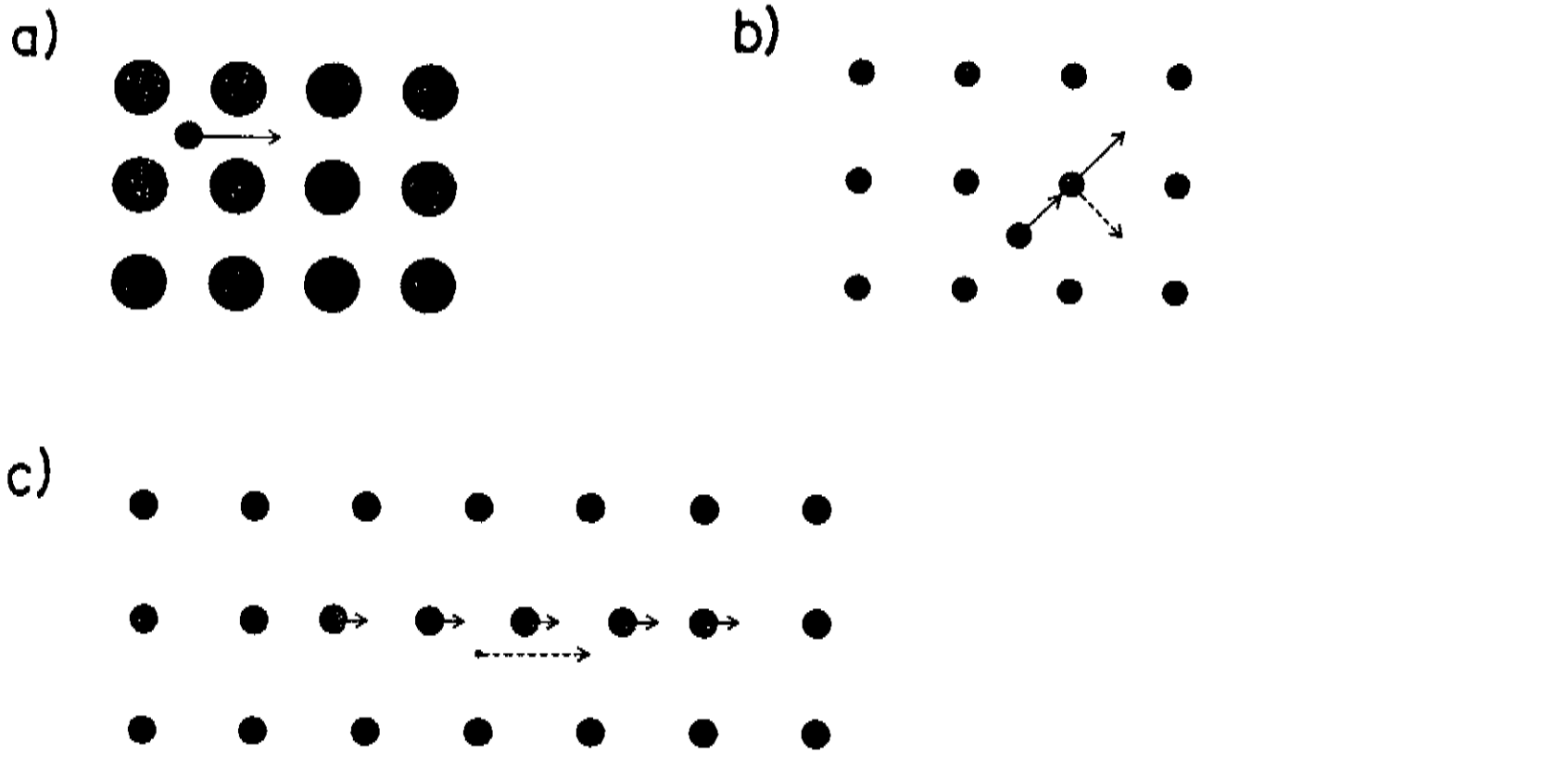

Figura 3.12 - Mecanismos de difusão intersticial. a) Intersticial simples, b) Geração de intersticiais, e c) "Crowdion".

energia empregada nesse processo (energia de ativação) deve ser menor do que nos dois processos anteriores.

Outra forma de difusão intersticial é o mecanismo de geração de intersticiais ou mecanismo intersticial indireto. Neste caso o átomo que difunde e os átomos da rede são aproximadamente iguais em diâmetro. A difusão ocorre com o átomo da impureza empurrando um átomo da rede de sua posição para um interstício vizinho. Após isso o átomo da impureza ocupa a posição deixada vaga pelo átomo da rede. Normalmente o deslocamento dos átomos é colinear, como mostrado na figura 3.12b.

O terceiro e último mecanismo da forma intersticial é o "crowdion". Neste tipo o átomo que está difundindo ocupa uma posição intersticial em uma "linha" de átomos ("crowdion"). A ocupação é feita de tal forma que os átomos dessa linha estarão ligeiramente mais próximos do seus vizinhos (na linha). Isso ocorrerá com até dez átomos na linha, deixando assim um espaço, no

centro da linha, para acomodar o átomo da impureza. A difusão irá ocorrer nesse mecanismo com o átomo saindo de sua posição central e passando para o interstício entre os dois átomos seguintes. Ao fazer isso ele move cada átomo do "crowdion" de uma pequena distância, como mostrado na figura 3.12c.

Os dois mecanismos seguintes são os devidos às vacâncias e divacâncias, que podem ser vistos nas figuras 3.13a e 3.13b, respectivamente. O mecanismo de difusão por vacâncias é simplesmente o pulo do átomo que difunde de sua posição para uma vacância vizinha. Isso não é nada mais do que um mecanismo de troca entre uma vacância e um átomo. Em muitos cristais podemos ter uma força de atração entre vacâncias, que seria responsável pela formação de pares de vacâncias ou divacâncias. A difusão neste caso ocorreria com o átomo pulando para uma vacância do par, o que equivale ao movimento da divacância e do átomo. Em semicondutores compostos a divacância pode ser usada como auxiliar em um mecanismo de troca, entre uma impureza em uma posição substitucional de um átomo V, por exemplo, com o átomo V vizinho. Neste caso temos uma vacância de um átomo III entre os dois sítios de troca, e é necessária a divacância para que os dois átomos possam passar um pelo outro realizando a troca e sem que haja uma grande distorção da rede.

O último mecanismo gerado por defeitos pontuais é o de relaxação. Este mecanismo ocorre quando vários átomos relaxam em direção a uma vacância, de tal forma que em uma região de raio de

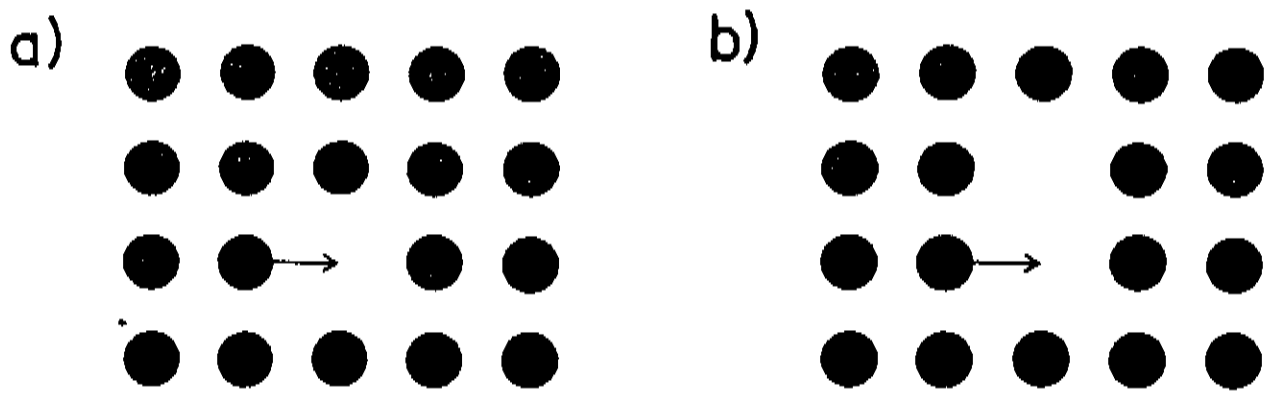

Figura 3.13 - Mecanismos de difusão por vacâncias. a) Mecanismo devido a uma vacância, e b) Mecanismo devido à divacâncias.

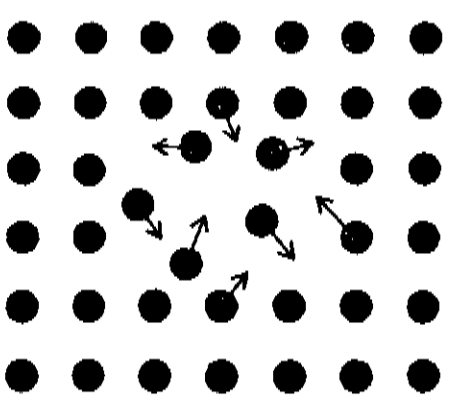

Figura 3.14 - Mecanismo de difusão por relaxação.

várias distâncias inter-atômicas os átomos estão fora de suas posições e irão difundir desordenadamente (figura 3.14), como as moléculas de um líquido. Devido a isso essa região é normalmente comparada com uma fusão local.

3.6.2 - Cálculo do Coeficiente de Difusão

Vimos nos itens anteriores quais são os defeitos presentes em um cristal e quais os mecanismos de difusão relacionados aos mesmos. Estudamos também as leis de Fick, que nos permitem a partir de um conhecimento do coeficiente de difusão estabelecer qual será o perfil de impurezas após a difusão. Vamos agora deduzir expressões para o coeficiente de difusão em função das propriedades microscópicas do semicondutor, tais quais distâncias e frequências de vibração atômicas. Vamos deduzir antes uma expressão para o coeficiente de difusão quando o fluxo de material que difunde é devido a vários mecanismos.

Vamos supor que uma impureza difunde em um semicondutor através de N mecanismos diferentes e simultâneos, cada um com um coeficiente de difusão D_i (i=1,2,...,N). O fluxo total de impurezas, J, será a soma de cada fluxo correspondente a cada um dos mecanismos, J_i. Usando a primeira lei de Fick podemos escrever,

$$J = \sum_{i=1}^{N} J_i = - \sum_{i=1}^{N} D_i \frac{\partial C_i}{\partial x} \qquad (3.29)$$

Podemos rescrever a equação (3.29) da seguinte forma,

$$J = - \left[\sum_{i=1}^{N} D_i \frac{\partial C_i}{\partial C} \right] \frac{\partial C}{\partial x} \qquad (3.30)$$

A equação acima tem a forma da primeira lei de Fick com o seguinte valor de D,

$$D = \sum_{i=1}^{N} D_i \frac{\partial C_i}{\partial C} \qquad (3.31)$$

Supondo que a fração de impurezas que difunde em cada modo é constante (C_i/C = constante), podemos simplificar (3.31),

$$D = \sum_{i=1}^{N} D_i \frac{C_i}{C} \qquad (3.32)$$

A equação acima nos dá uma relação entre o coeficiente de difusão medido em um experimento, D, e os coeficientes devidos aos diversos mecanismos atuantes, D_i.

Vamos agora deduzir uma expressão para o coeficiente de difusão (supondo apenas um modo de difusão) em termos das propriedades microscópicas do cristal. Começamos considerando uma rede FCC e dois planos (100) adjacentes (veja a figura 3.15). Chamando de Γ_{12} a frequência de pulos de átomos da impureza do

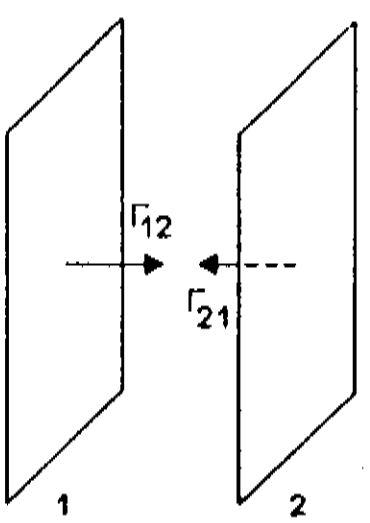

Figura 3.15 - Frequências de pulos de átomos entre planos vizinhos.

plano 1 para o plano 2, e Γ_{21} a do plano 2 para o plano 1, o fluxo de impurezas será,

$$J = n_1 \Gamma_{12} - n_2 \Gamma_{21} \qquad (3.33)$$

onde n_1 é a concentração de impurezas no plano 1 e n_2 no plano 2.

Vamos utilizar a suposição de que a frequência de pulos não depende da direção. Essa suposição forma a base da teoria do "passeio aleatório" ("random-walk"), que é usada nos estudos de difusão. Temos assim,

$$\Gamma_{12} = \Gamma_{21} = \Gamma \qquad (3.34)$$

$$J = \Gamma (n_1 - n_2) \qquad (3.35)$$

Chamando a distância entre planos vizinhos de λ, podemos relacionar as concentrações dos planos com as volumétricas, C,

$$n_1 - n_2 = \lambda (C_1 - C_2) \qquad (3.36)$$

de onde tiramos,

$$n_1 - n_2 = \lambda^2 \frac{\partial C}{\partial x} \qquad (3.37)$$

já que podemos escrever $C_1 - C_2 = -\lambda \, \partial C/\partial x$. Substituindo (3.37) em (3.36) achamos,

$$J = -\lambda^2 \Gamma \frac{\partial C}{\partial x} \qquad (3.38)$$

A equação acima é a primeira lei de Fick com o seguinte valor para o coeficiente de difusão, D,

$$D = \lambda^2 \Gamma \qquad (3.39)$$

Devemos lembrar que um átomo em uma rede FCC possui 12 vizinhos próximos, e que apenas 8 estão nos planos adjacentes, 4 no plano anterior e 4 no posterior. Podemos relacionar a frequência total de pulos, ν, com a de pulos entre os planos,

$$\Gamma = \frac{\nu}{3} \qquad (3.40)$$

Para a rede FCC a distância entre os planos (100) é metade do parâmetro de rede, a. Usando isso e a equação (3.40) achamos o coeficiente de difusão,

$$D = \frac{a^2 \nu}{12} \qquad (3.41)$$

A equação acima nos dá o coeficiente de difusão em termos de parâmetros que podem ser determinados.

Podemos relacionar ν com a frequência de vibração de um átomo da impureza, ω_o. Começamos supondo que o mecanismo de difusão é através de vacâncias. A probabilidade de acharmos uma vacância em um determinado sitio da rede, p, será dada pela concentração de vacâncias dividida pela de átomos da rede, que pode ser obtida de (3.4),

$$p = \frac{[V_{As}]}{[As_{As}]} = K_{10}\, P_{As_2}^{-1/2} = \exp\left(- \frac{\Delta g_V}{kT} \right) P_{As_2}^{-1/2} \qquad (3.42)$$

onde usamos uma expressão análoga a (3.2) para a reação de formação de uma vacância, supondo que ela seja de arsênio, e a espécie de arsênio na fase de vapor seja As_2. Na expressão acima também utilizamos (3.2) para substituir a constante de equilíbrio pela exponencial de Δg_V, que é a energia livre da reação de formação de uma vacância de arsênio.

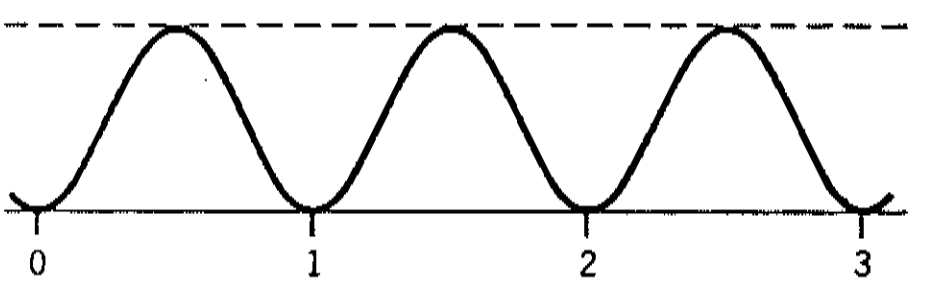

Figura 3.16 - Níveis de energia de uma impureza em um cristal. Os pontos de menor energia correspondem aos sítios da rede.

Para que um átomo troque de posição com uma vacância ele deve ultrapassar uma barreira de potencial, como mostra a figura 3.16 (os sítios normalmente ocupados são as posições energeticamente mais favoráveis). A frequência com que o átomo consegue ultrapassar essa barreira, ω, é dada por,

$$\omega = \omega_o \exp\left(-\frac{g_m}{kT}\right) \qquad (3.43)$$

onde g_m é a diferença entre a energia livre do átomo no topo da barreira e no seu sitio normal. A frequência de pulos para uma vacância vizinha, Ω, é igual ao produto da probabilidade de acharmos a vacância vezes a frequência com que o átomo ultrapassa a barreira,

$$\Omega = p\ \omega \qquad (3.44)$$

Por fim sendo z o número de vizinhos de um átomo, a frequência total de trocas será,

$$\nu = z\ \Omega = z\ \omega_o \exp\left(-\frac{g_m}{kT}\right) \exp\left(-\frac{\Delta g_V}{kT}\right) P_{As_2}^{-1/2} \qquad (3.45)$$

que substituindo na expressão para D, equação (3.41) resulta em,

$$D = \frac{a^2\ z\ \omega_o}{12} P_{As_2}^{-1/2} \exp\left(-\frac{g_m + \Delta g_V}{k\ T}\right) \qquad (3.46)$$

A equação acima nos dá um valor de D da forma,

$$D = D_o \exp\left(-\frac{Q}{kT}\right) \qquad (3.47)$$

com os seguintes valores de D_o e Q,

$$D_o = \frac{a^2\ z\ \omega_o}{12} P_{As_2}^{-1/2} \qquad (3.48a)$$

$$Q = g_m + \Delta g_V \qquad (3.48b)$$

a equação (3.47) tem a forma de arrehnius e é válida em um amplo intervalo de temperatura desde que mantenhamos a pressão de arsênio constante. Ela nos mostra que a dependência do coeficiente de difusão com a temperatura é exponencial, sendo Q a energia de ativação do processo de difusão, e D_o o termo pré-exponencial. Repare que o coeficiente de difusão depende da pressão de arsênio. Isso pode ser entendido uma vez que essa pressão controla a concentração de vacâncias, que é inversamente proporcional a uma potência da pressão. Dessa forma para maiores pressões teremos menores concentrações de vacâncias, o que reduz o movimento da impureza que difunde, já que o seu movimento requer a presença de uma vacância.

3.7 - Difusão de Enxofre em Arseneto de Gálio

Vamos apresentar neste item os diagramas de fase do Ga-As e Ga-As-S. Discutiremos os trabalhos anteriores ao nosso sobre a difusão de enxofre no arseneto de gálio, e apresentaremos os mecanismos de difusão propostos nesses trabalhos, seguido de uma pequena discussão sobre a validade desses modelos.

3.7.1 - Diagramas de Fase. Pressões de Vapor de Equilíbrio e Concentrações de Vacâncias no GaAs

Na figura 3.17 podemos ver o diagrama de fase do gálio e arsênio. Este foi determinado por Thurmond [11] a partir de dados obtidos por Hall [12] e por Köster e Thoma [13]. Na figura podemos ver que a região onde temos o GaAs sólido é bastante estreita, o que equivale a dizer que a concentração máxima de vacâncias no cristal deve ser relativamente pequena. Essa é uma característica dos semicondutores compostos do grupo III-V.

As concentrações das várias espécies gasosas são determinadas através de medidas de espectroscopia de massa, realizadas por Arthur [14]. Para as linhas a-a' (excesso de gálio) e b-b'

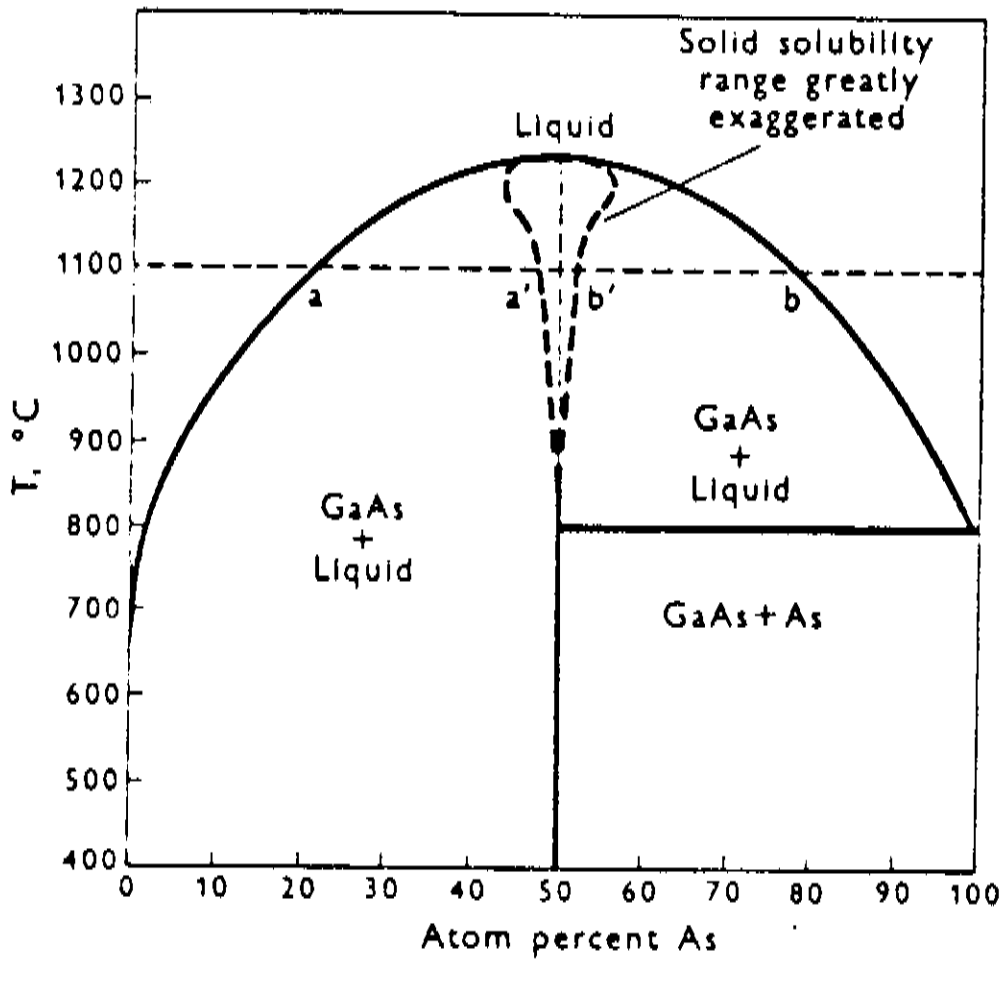

Figura 3.17 - Diagrama binário do GaAs [11].

(excesso de arsênio) as pressões das várias espécies podem ser vistas no gráfico da figura 3.18 (nessas regiões as pressões de

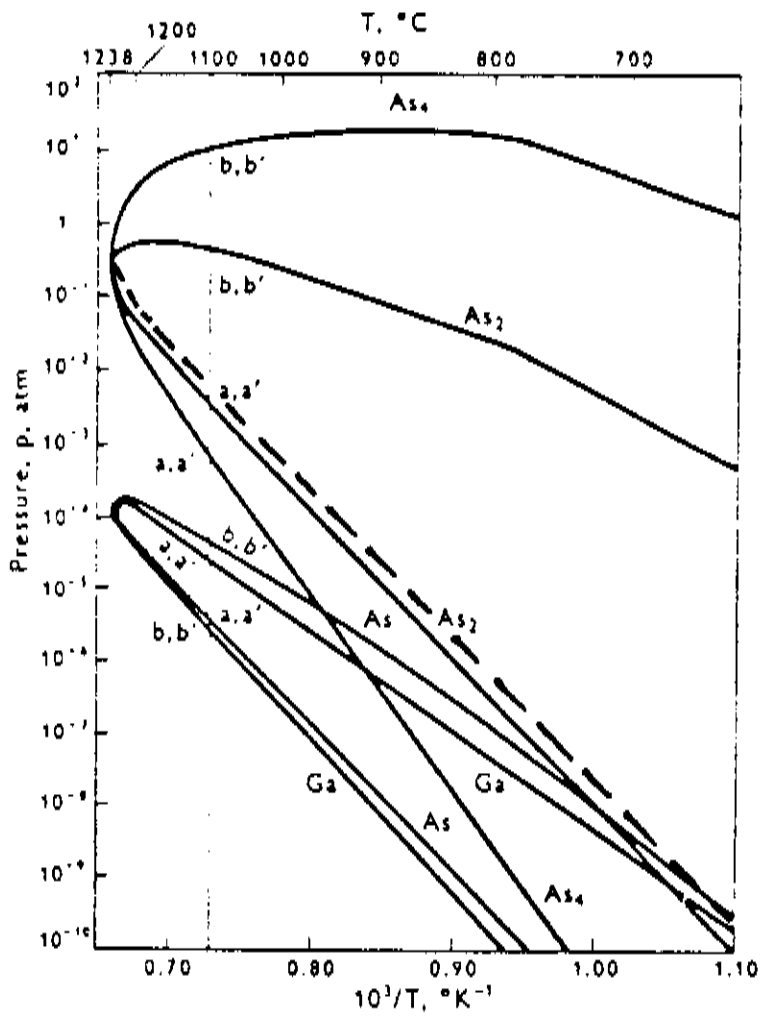

Figura 3.18 - Pressões de vapor das várias espécies de gálio e arsênio para o cristal estequiométrico (pressão de dissociação); para os pontos de formação de fases líquidas de gálio, a-a'; e para os de formação de fases líquidas de arsênio, b-b'.

vapor são constantes). Na mesma figura vemos também a pressão de dissociação das várias espécies [15]. Essa pressão é definida como a da espécie gasosa em equilíbrio com o cristal aquecido em vácuo sem a presença de fases líquidas. Observe que para pressões nos valores determinados pelas curvas a-a' e b-b' teremos a formação de fases líquidas, o que não é desejado durante a difusão. O gráfico da figura 3.18 é portanto útil para determinarmos as faixas de pressão de gálio ou arsênio que podem ser usadas durante a difusão (os valores entre as curvas a-a' e b-b').

Os limites da região de GaAs sólido no diagrama binário são determinados a partir de medidas de concentração de vacâncias no cristal em função da temperatura. Na figura 3.19 podemos ver um gráfico dessa concentração para um cristal em contato com fases líquidas ricas em gálio e arsênio. As curvas para a concentração de vacâncias de arsênio foi determinada por Potts e Pearson [16] a partir de medidas de mudança de parâmetro de rede dos cristais de GaAs aquecidos. Para as vacâncias de gálio é dado um valor arbitrário de concentração [9].

Por fim temos na figura 3.20 o diagrama de fase do

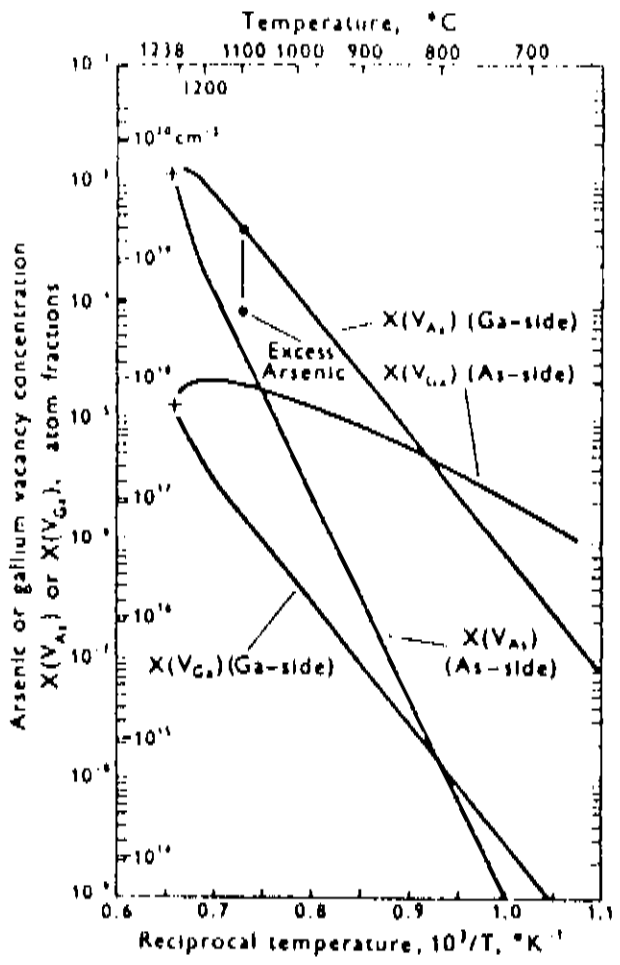

Figura 3.19 - Concentrações de vacâncias de gálio e arsênio em um cristal de GaAs em equilíbrio com uma fase líquida rica em gálio, a-a'; e rica em arsênio, b-b'.

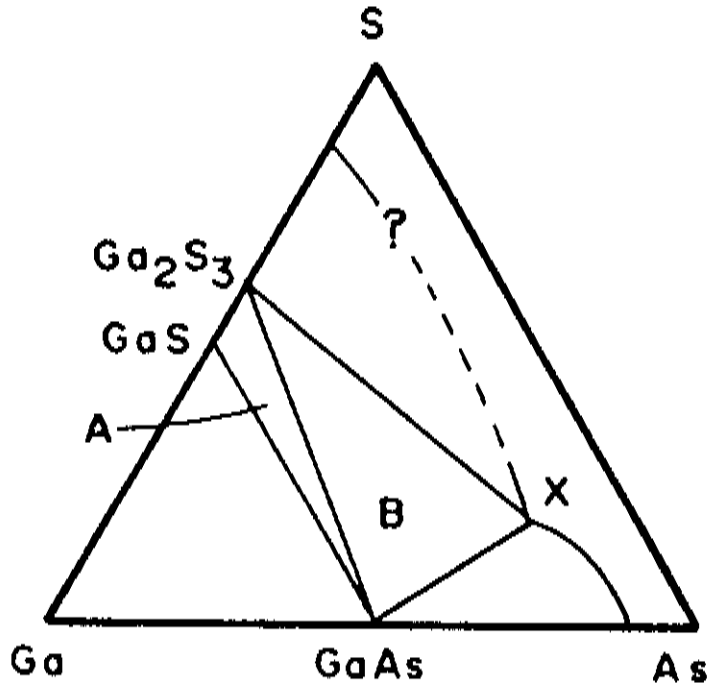

Figura 3.20 - Diagrama de fases do Ga-As-S [17].

Ga-As-S, como proposto por Matino [17]. Ele adaptou o diagrama determinado por Panish [18] para o Ga-As-Te, mudando as temperaturas de fusão dos vários compostos. Matino admite que não existem fases líquidas para temperaturas abaixo de 930°C, exceto um composto X, rico em arsênio (veja o diagrama de fases). As regiões A e B possuem cada uma três fases condensadas.

Todas essas considerações sobre os diagramas de fase são necessárias para que possamos entender o mecanismo de difusão do enxofre em arseneto de gálio, e analisar os resultados obtidos dos experimentos de difusão.

3.7.2 - Trabalhos sobre a Difusão de Enxofre em Arseneto de Gálio

Na literatura especializada podemos achar até a data de hoje cerca de dez trabalhos sobre a difusão de enxofre em arseneto de gálio. Os primeiros trabalhos datam de 1961, de Goldstein [19]. Ele fez experiências difundindo o enxofre em uma ampola fechada (veja o capitulo 4 onde comentamos sobre as técnicas de difusão) junto com pequenas quantidades de arsênio para evitar a sua perda pela superfície do cristal. As medidas do perfil de difusão foram feitas com enxofre radioativo. Os seus resultados mostram uma alta energia de ativação de 4,04eV. Goldstein observou que usando uma

densidade de enxofre no vapor maior que algumas $\mu g/cm^3$ ocorria a formação de uma camada de GaS na superfície do cristal. Entre todos os trabalhos posteriores, Goldstein obteve os menores coeficientes de difusão, o que foi atribuído a um processo de decapagem dos seus cristais de GaAs pelo vapor de enxofre. Algumas amostras de Goldstein perdiam até 5% de seu peso.

Como já vimos, a pressão de vapor de arsênio tem grande influência sobre a difusão. Um dos primeiros trabalhos onde isso foi notado é o de Vieland [20], que realizou várias experiências de difusão, com pressões de arsênio variadas. Seus resultados mostram que o uso de densidades menores que $3\mu g/cm^3$ de enxofre no vapor evitam a formação da camada de GaS sobre o cristal. Vieland obteve uma energia de ativação de 1,8eV, e coeficientes de difusão maiores que os de Goldstein, da mesma ordem que os relatados por outros autores posteriormente. Os seus resultados foram baseados em medidas de profundidade de junções PN.

O trabalho seguinte foi feito por Yeh [21] em 1964. Embora ele tenha usado a técnica de medidas de profundidade de junções PN, ele não apresenta nenhum resultado sobre coeficientes de difusão ou energia de ativação. A principal preocupação de Yeh foi demonstrar que cobrindo os cristais de GaAs com uma fina camada de SiO conseguia-se eliminar o problema de deposição de sulfetos sobre o cristal. Ele ainda observa que o SiO não é uma boa máscara contra a difusão, pois nem mesmo o uso de camadas grossas ($2\mu m$) conseguia eliminar a difusão. O próximo trabalho no ano de 1965 foi de Frieser [22] que demonstrou que o uso de fontes de enxofre compostas, como por exemplo Al_2S_3, evita a formação de sulfetos sobre o GaAs, uma vez que essas fontes compostas possuem baixas pressões de vapor. Frieser usa medidas em junções PN para obter uma energia de ativação de 1,86eV, e coeficientes de difusão maiores que os de Goldstein, porém ainda menores que os de outros autores. A causa desses resultados foi atribuída a uma decapagem da amostra pelo vapor, que embora não tenha provocado depósitos de sulfetos pode ainda ter sido suficientemente denso para decapar a amostra. Frieser comenta em seu trabalho que a máxima quantidade de enxofre eletricamente ativo é $4x10^{18}cm^{-3}$.

Kendall em 1968 é o autor do próximo trabalho [3]. Ele

usou enxofre radioativo e da mesma forma que Goldstein conseguiu determinar os perfis de difusão. A sua energia de ativação medida é de 1,63eV, em boa concordância com Vieland e Frieser. É interessante observar que até esse ponto os trabalhos realizados nos mostram que aparentemente existe uma concentração máxima de enxofre eltricamente ativo e que as pressões de vapor de arsênio possuem grande influência sobre a difusão, como já era esperado. Com relação ao mecanismo de difusão apenas Goldstein fez a suposição de que o enxofre difunde substitucionalmente através de um mecanismo de troca com vacâncias de arsênio.

O trabalho seguinte de Young e Pearson [23] foi bastante completo, oferecendo resultados suficientes para que fosse proposto um mecanismo de difusão. Neste trabalho os perfis de difusão foram determinados com o uso de enxofre radioativo, e as concentrações de elétrons por uma técnica de medida de reflexão de plasma com decapagem da amostra. Young e Pearson conseguiram mostrar claramente que existe uma máxima concentração de enxofre eletricamente ativo de $3x10^{18}cm^{-3}$, e que para alta concentração de enxofre, maior que $1x10^{18}cm^{-3}$, o perfil de carga já passa a ser diferente do de enxofre. A variação do coeficiente de difusão com a pressão de arsênio obtida das experiências de Young e Pearson leva-os a propor um modelo de difusão, que será descrito mais adiante. Eles demonstram ainda que a concentração de enxofre na superfície do cristal varia linearmente com a densidade de enxofre no vapor, o que não é previsto pela teoria. A energia de ativação obtida foi de 2,6eV, e os coeficientes de difusão medidos são muito próximos aos obtidos por Frieser.

O próximo trabalho, de Matino [17], mostra que a escolha de fontes que apresentem três fases condensadas durante a difusão irão resultar em um grau de liberdade zero para o sistema, e que com isso as pressões de vapor serão constantes e independentes da exata composição da fonte, resultando em difusões altamente reprodutíveis. Para a escolha da fonte Matino usou o diagrama de fases do Ga-As-S, que foi determinado partindo-se do diagrama do Ga-As-Te. Matino obteve o perfil de elétrons a partir de medidas de efeito Hall e resistividade da amostra. Ele obteve dois valores distintos para a energia de ativação, 4,35eV para fontes com alta

concentração de enxofre, e 2,95eV para fontes com baixa concentração.

Ao trabalho de Matino seguiu-se o de Tuck e Powell [24]. Eles usaram como fonte ^{35}S, que é um isótopo radioativo do enxofre. Isso permitiu a medida do perfil de enxofre, decapando a amostra e fazendo contagens radioativas. As temperaturas de difusão usadas por Tuck e Powell são de 1000°C e 1120°C. Eles realizaram várias experiências onde foi variada a concentração de enxofre na ampola. Os seus resultados mostram que a adição de arsênio à ampola faz com que a concentração de enxofre na superfície passe de uma potência 0,5 da quantidade de enxofre colocada na ampola para 1,5. Apenas o expoente 0,5 pode ser previsto na teoria, considerando-se que a maior parte do enxofre, para altas concentrações deste, seja não ionizada. O expoente 1,5 não pode ser previsto de forma alguma pela teoria. Por fim Tuck e Powell comentam que os perfis de difusão obtidos não são do tipo função erro complementar. O trabalho seguinte foi de Zahari e Tuck [25], que propuseram um modelo para a difusão de enxofre em GaAs que explica o perfil encontrado. Iremos discutir esse trabalho mais adiante.

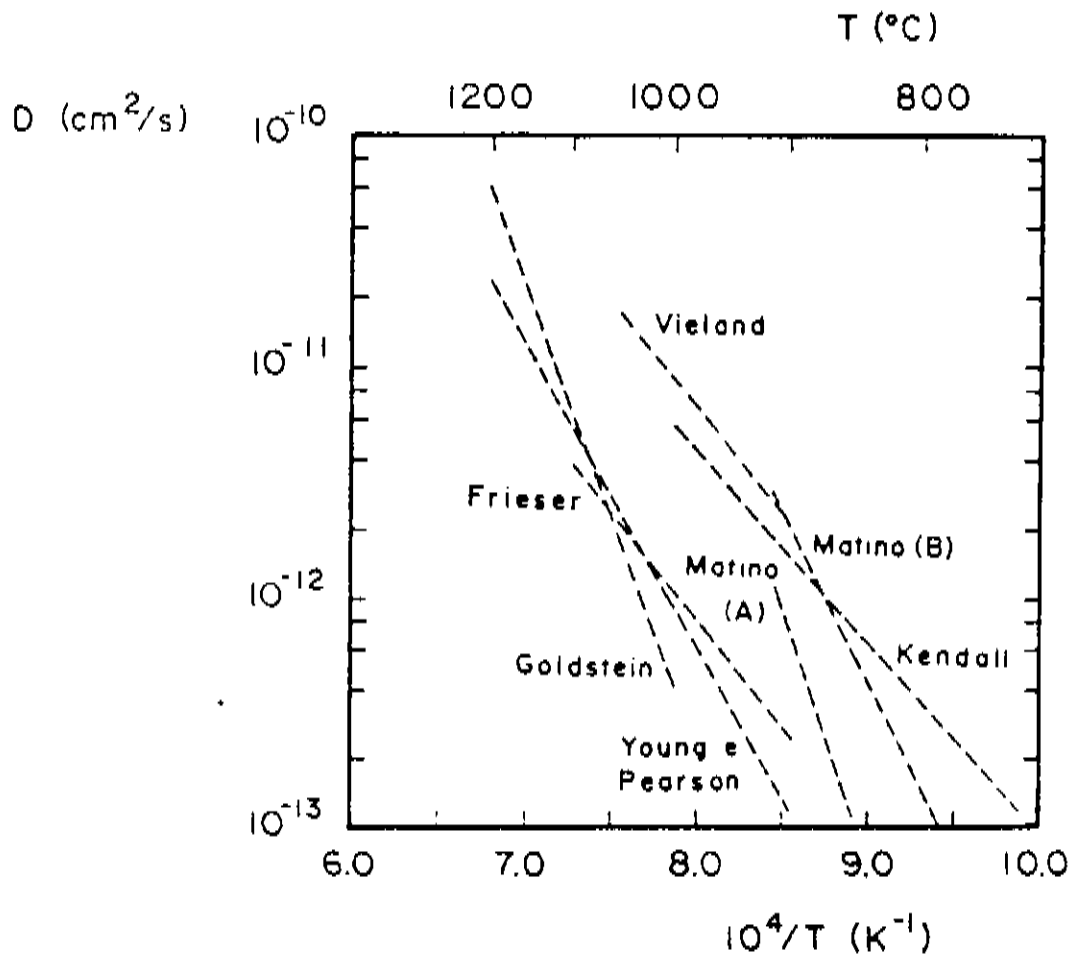

Figura 3.21 - Coeficientes de difusão de enxofre em GaAs em função da temperatura obtidos por vários autores.

Os últimos trabalhos foram de Prince et al. [26,27]. Eles desenvolveram um processo de difusão de enxofre em tubo aberto (não usaram uma ampola lacrada), onde pode-se usar amostras de até duas polegadas de diâmetro. A fonte de enxofre usada foi Ga_2S3, e a energia de ativação medida foi de 2,35eV. Prince foi o primeiro a aplicar as camadas na fabricação de transistores MESFET de GaAs. Na figura 3.21 vemos um gráfico que compara os coeficientes de difusão obtidos por vários autores.

3.7.3 - Mecanismo Físico de Difusão de Enxofre em GaAs

Em toda a literatura temos dois mecanismos propostos para a difusão de enxofre em GaAs, e dois modelos diferentes que serão vistos mais abaixo. Temos ainda dois modelos para a concentração de elétrons no cristal.

O primeiro mecanismo de difusão foi proposto por Goldstein [19]. Analisando os resultados obtidos da difusão de enxofre e selênio no GaAs ele conclui que o mecanismo provável de difusão é o de vacâncias, onde um átomo da impureza sai de sua posição substitucional e pula não para uma vacância vizinha, já que um átomo de arsênio é rodeado por quatro gálios, mas para uma segunda vizinha.

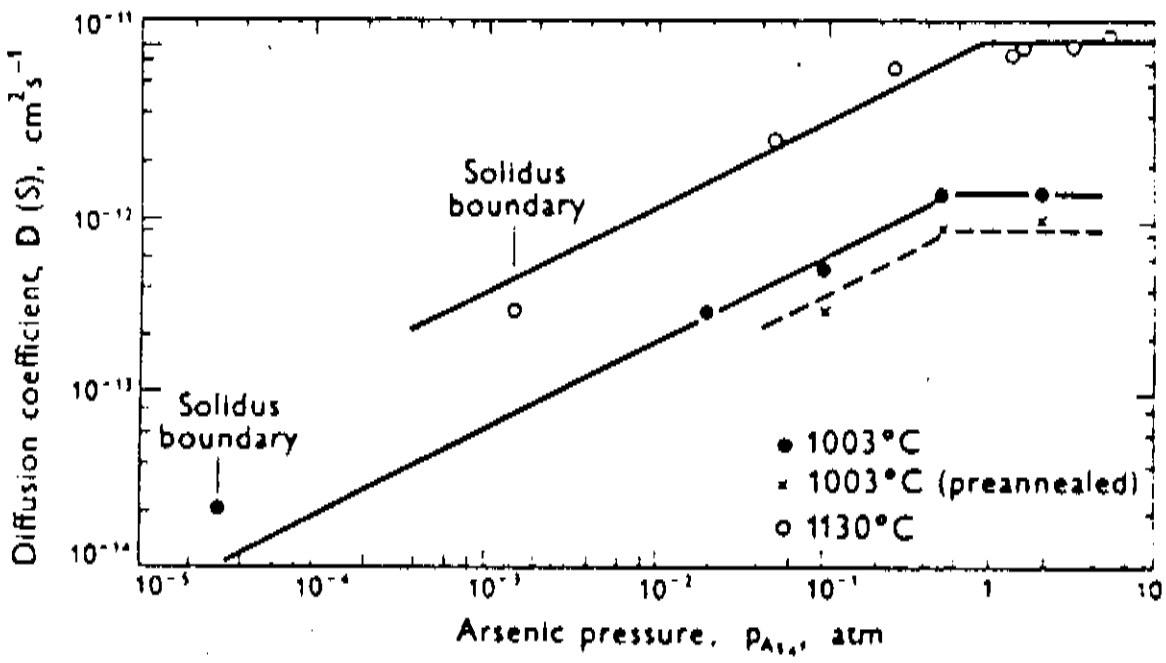

Figura 3.22 - Coeficiente de difusão de enxofre em GaAs em função da pressão parcial de arsênio [23].

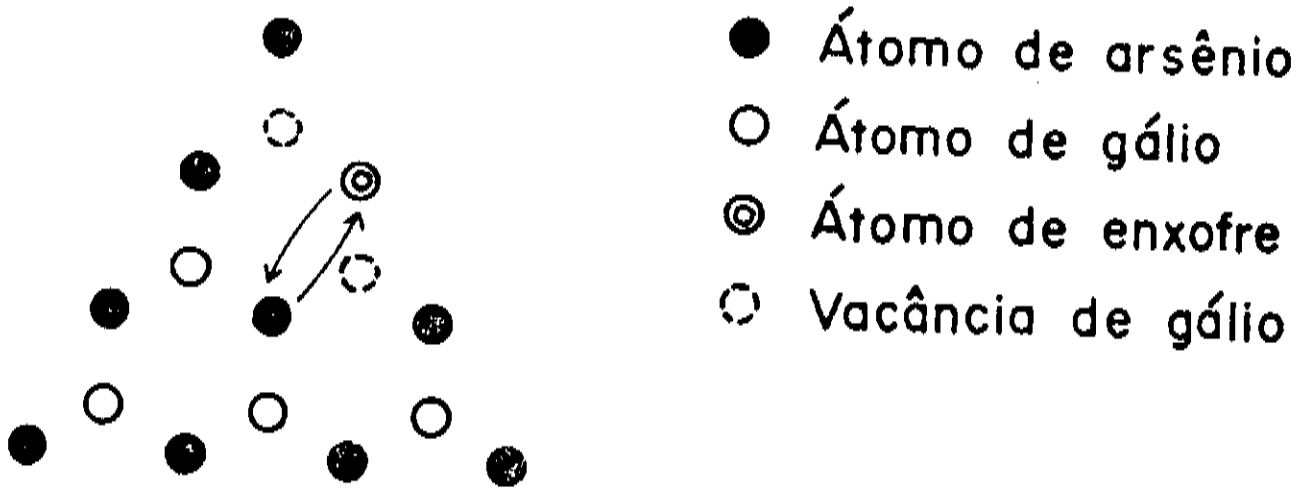

Figura 3.23 - Mecanismo de troca auxiliado por divacâncias para a difusão de enxofre em GaAs.

O segundo mecanismo e o primeiro modelo foi proposto Young e Pearson [23]. Eles observaram que o aumento da pressão de arsênio causa o aumento do coeficiente de difusão de enxofre, e que ocorre uma saturação do coeficiente para altas pressões. Na figura 3.22 podemos ver um gráfico que mostra essa dependência do coeficiente com a pressão. Para interpretar esse resultado Young e Pearson propõem um mecanismo de difusão de troca entre uma vacância de arsênio e um enxofre em um sitio substitucional de arsênio. Esse mecanismo de troca seria auxiliado por uma divacância de gálio entre o átomo de e a vacância, figura 3.23. Temos dessa forma duas espécies de enxofre, S_{As}, que seria imóvel, e um complexo, $V_{Ga}S_{As}V_{Ga}$, chamado de C_1, que seria móvel. A reação de mudança da espécie de enxofre é,

$$S_{As} + V_{Ga}V_{Ga} \rightarrow C_1 \qquad (R10)$$

Com a seguinte equação de equilibrio,

$$[S_{As}]\ [V_{Ga}V_{Ga}] = K_{10}\ [C_1] \qquad (3.49)$$

Supondo que a espécie de arsênio no vapor seja As_4, podemos escrever a concentração de divacâncias de gálio, usando (3.6), (3.9) e (3.7) como,

$$[V_{Ga}V_{Ga}] = K_6 \frac{K_3^{\,2}}{K_4^{\,2}} P_{As_4}^{\,1/2} = K_{11}\, P_{As_4}^{\,1/2} \qquad (3.50)$$

Por fim a concentração total de enxofre é,

$$[S] = [C_1] + [S_{As}] \qquad (3.51)$$

Resolvendo as três equações anteriores obtemos,

$$[C_1] = \left[\frac{K_{12}\, P_{As_4}^{\,1/2}}{1 + K_{12}\, P_{As_4}^{\,1/2}} \right] [S] \qquad (3.52)$$

onde K_{12} é dado por,

$$K_{12} = \frac{K_{11}}{K_{10}} \qquad (3.53)$$

Chamando de D_1 o coeficiente de difusão de C_1, obtemos usando (3.31) o seguinte coeficiente efetivo de difusão, D (lembre que o coeficiente de difusão do S_{As} é considerado como sendo zero),

$$D = D_1 \left[\frac{K_{12}\, P_{As_4}^{\,1/2}}{1 + K_{12}\, P_{As_4}^{\,1/2}} \right] \qquad (3.54)$$

A equação acima mostra o comportamento observado para o coeficiente de difusão. Para baixas pressões de arsênio o coeficiente deve aumentar com a raiz da pressão, o que é observado. A alguma pressão suficientemente alta o coeficiente deve atingir o valor D_1, permanecendo nesse valor, independente de um aumento ainda maior da pressão. Observe que este modelo é incompleto no sentido de que ele prevê um coeficiente de difusão constante para uma dada temperatura e pressão. Isso significa que o perfil de difusão deve ser do tipo função erro complementar, o que é observado para baixas concentrações de enxofre. Para altas concentrações o perfil assume uma forma complexa, que Young e Pearson atribuem a um decréscimo do coeficiente de difusão devido à formação de

complexos de enxofre em grande quantidade. Estes seriam $V_{Ga}S_3$ e precipitados de Ga_2S_3.

Temos por fim o modelo de Zahari e Tuck [25], que prevê exatamente o perfil de enxofre no cristal. Está implícito no seu trabalho que eles usam o mecanismo de troca de Goldstein. No modelo é resolvida numericamente a segunda lei de Fick (equação (3.25)). Isto é feito em um computador, onde os autores criaram basicamente dois "vetores" (o problema foi tratado unidimensionalmente), um que continha a concentração de enxofre, e o outro que continha a concentração de átomos de arsênio. Cada posição do vetor é relacionada a uma "fatia" do cristal que pode conter muitos planos atômicos. As fatias são espessas o suficiente de tal forma que não tenhamos uma quantidade excessivamente grande de pontos no vetor, sendo ainda praticamente constante a concentração de átomos (enxofre) em cada fatia. O cálculo é feito supondo-se que a quantidade de átomos de enxofre e de arsênio que pulam de uma fatia para outra é proporcional ao produto da sua concentração pela de vacâncias de arsênio. Dessa forma é feita uma simulação do fluxo de arsênio e enxofre no cristal, resultando no perfil. Zahari e Tuck usam o fato de que o enxofre difunde muito mais

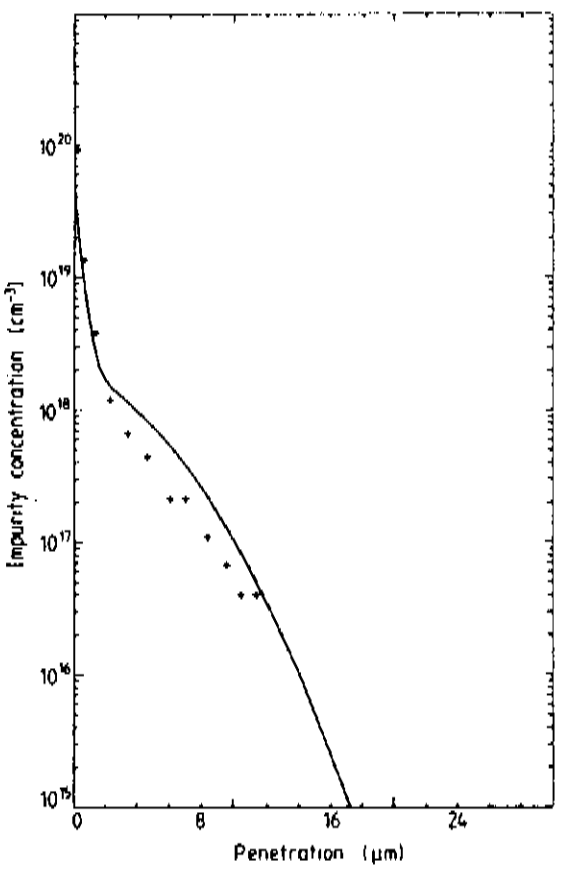

Figura 3.24 - Perfis teóricos (linha cheia) e pontos experimentais obtidos da difusão de enxofre em GaAs [24,25].

rápido do que o arsênio para obter o perfil medido experimentalmente. Segundo o seu modelo, na superfície teremos um equilíbrio entre as concentrações de vacâncias, enxofre e arsênio. Já uma camada abaixo da superfície não estará em equilíbrio, pois ela terá uma alta concentração de enxofre e baixa de vacâncias, já que o coeficiente de difusão dos átomos de arsênio é pequeno, e portanto eles não se movimentam rápido o suficiente para suprir a quantidade de vacâncias que é consumida pelos átomos de enxofre que chegam à camada. Essa queda na concentração de vacâncias deve diminuir o coeficiente de difusão, como é observado experimentalmente. Por fim em uma camada mais profunda devemos voltar a ter o equilíbrio, já que nesse caso a concentração de enxofre é pequena e consequentemente o consumo de vacâncias também o é. Na figura 3.24 vemos um gráfico que compara os perfis teóricos e experimentais.

Uma crítica ao trabalho de Zahari e Tuck, é que embora o coeficiente de difusão do arsênio seja muito pequeno, o de vacâncias é bem maior, sendo inclusive maior que o de enxofre. Isso é concluido por nós do trabalho de Casey Jr. [9], que relaciona o coeficiente de difusão de vacâncias, D_V, ao de arsênio, D_{As}, como,

$$D_{As} = 0,78\ D_V\ X_V \qquad (3.55)$$

Onde X_V é a concentração de vacâncias de arsênio, e 0,78 um fator de correlação [10]. Para uma temperatura de 700°C, Goldstein [19] prevê um valor de D_{As} de $10^{-17} cm^2/s$, e temos X_V de $5x10^{-7} cm^{-3}$, o que dá um valor de D_V da ordem de $10^{-11} cm^2/s$, o que é muito maior que o coeficiente de difusão de enxofre nessa temperatura, da ordem de $10^{-14} cm^2/s$. Portanto deve haver um equilíbrio de vacâncias no cristal, o que invalida o modelo de Zahari e Tuck.

Com relação ao fato de que para altas concentrações de enxofre a concentração de elétrons é bem menor, temos dois modelos que podem ser aplicados. O primeiro é a degenerescência da banda de condução [28]. O GaAs possui uma pequena massa efetiva para os elétrons. Isso pode ser visto pela grande curvatura da banda de condução (figura 3.2), que implica em uma pequena densidade

efetiva de estados, N_C, da ordem de $10^{17} cm^{-3}$ na temperatura ambiente. Dessa forma quando começamos a preencher os níveis de energia da banda com os elétrons dos átomos de enxofre, observamos que o nível sobe rapidamente, o que requer uma energia de ionização cada vez maior para podermos tirar um elétron do enxofre e passá-lo para a banda de condução, energia essa que não é disponível. Dessa forma uma alta concentração de enxofre implica em que a maioria deste seja não ionizada. A fração de impurezas ionizadas, N_D^+/N_D, será dada por [29],

$$\frac{N_D^+}{N_D} = 1 - \left[1 + \frac{1}{2} \exp\left(\frac{E_D - E_F}{k\ T} \right) \right]^{-1} \qquad (3.56)$$

Onde E_F é a energia do nível de Fermi, e E_D o nível de energia criado pela impureza. A concentração de elétrons, n, pode ser achada a partir do nível de Fermi de uma forma bastante simples, onde supomos que a banda é parabólica para altas concentrações de elétrons,

$$n = N_C \left(\frac{4}{\pi} \right)^{1/2} F_{1/2}\left(\frac{E_F - E_C}{k\ T} \right) \qquad (3.57)$$

onde $F_{1/2}$ é a integral de Fermi-Dirac [30]. Agora supondo que cada elétron da banda vem de um átomo de enxofre ionizado,

$$n = N_D^+ \qquad (3.58)$$

podemos resolver as equações acima, pelo menos numericamente, e obter o gráfico da figura 3.25, onde vemos a fração ionizada em função da concentração de doadores. Observe que para altas concentrações de enxofre, a maioria, cerca de 97% para N_D igual a $1x10^{20} cm^{-3}$, será não ionizada.

O segundo modelo para a diferença entre as concentrações de elétrons e de enxofre é o de Venkatasubramanian et al. [31]. Eles mostram que formam-se no GaAs dopado com enxofre complexos $V_{Ga}S_{As}$, que são aceitadores de nível profundo, o que provoca a compensação do cristal. O nível de compensação NA^-/ND^+ varia de 0,5 para uma concentração de elétrons de $1x10^{17} cm^{-3}$, até 0,91 para $3,8x10^{18} cm^{-3}$. Esse comportamento é semelhante ao da não ionização

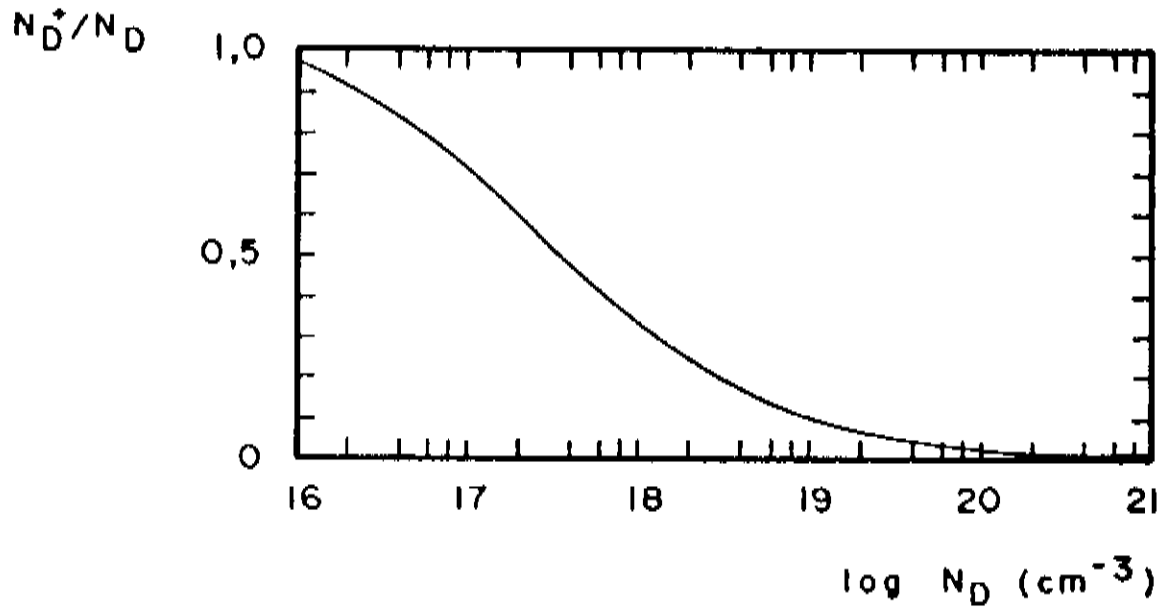

Figura 3.25 - Fração de impurezas doadoras ionizadas, N_D^+/N_D, no GaAs à temperatura ambiente.

devida à degenerescência do nível de Fermi.

Temos ainda que considerar as espécies existentes na fase de vapor para que possamos compreender totalmente a difusão. Casey Jr. [9] fez essa consideração e suas conclusões não são boas. Segundo ele infelizmente não conhecemos as espécies no vapor de um sistema Ga-As-S. As espécies principais seriam os vapores de As_2 e As_4 para um sistema contendo gálio e arsênio [32], e S_2, S_6 e S_8 para um vapor de enxofre [33]. Há fortes evidências de que a adição de gálio à fonte de difusão altera significativamente a concentração de espécies no vapor, tornando o Ga_2S como a espécie de enxofre predominante. Essas evidências são dadas pelo trabalho de Arthur [34], que aponta o Ga_2Te como dominante para um sistema Ga-Te. Temos ainda a observação de que para um sistema Ga-Ga_2O_3 a espécie principal é o Ga_2O [35].

A conclusão que tiramos é que não podemos analisar a incorporação de enxofre na superfície do GaAs, se não sabemos as espécies de enxofre no vapor, já que a reação de incorporação é dependente das espécies.

UNICAMP
BIBLIOTECA CENTRAL

3.8 - Conclusões

A difusão de impurezas em um semicondutor é um importante meio de obtermos camadas dopadas. Podemos através de um controle da temperatura e tempo de difusão variar a profundidade e concentração da camada obtida. Para um semicondutor binário as pressões de vapor dos seus componentes e da impureza terão grande influência no coeficiente de difusão e na concentração de impurezas na superfície da amostra. Mostramos nos itens anteriores que a incorporação de impurezas e sua difusão são intimamente relacionadas com os defeitos pontuais do semicondutor. Partindo de expressões para as concentrações de vacâncias e frequências de vibração atômicas obtivemos uma expressão para o coeficiente de difusão.

Com relação à difusão de enxofre em GaAs, concluimos que existe uma certa divergência quanto aos valores de energia de ativação do processo. Isso aparentemente é devido a diferentes condições experimentais, como por exemplo pressão de vapor de arsênio, composição da fonte de difusão, preparação da amostra, etc. O verdadeiro mecanismo físico de difusão ainda não é conhecido, e temos apenas dois modelos que explicam o comportamento da difusão, um deles, o de Zahari e Tuck, que aparentemente esta errado. Ainda é impossível apresentar qualquer modelo para a reação de incorporação de superfície, uma vez que não conhecemos as espécies gasosas no vapor do sistema Ga-As-S. São necessários estudos para a determinação do exato diagrama de fases do Ga-As-S e espécies gasosas presentes no vapor antes que seja descoberto o correto mecanismo para a difusão.

3.9 - Referências

[1] - N. W. Ashcroft e N. D. Mermin, *Solid State Physics*, 1ª edição, RINEHART AND WINSTON, New York, 1976.

[2] - J. R. Chelikowsky e M. L. Cohen, "Nonlocal Pseudopotential Calculations for the Electronic Structure of Eleven Diamond

and Zinc-Blende Semiconductors", *Physical Review B*, vol. 14, n. 2, p. 556 (1976).

[3] - D. L. Kendall, *Semiconductors and Semimetals*, vol.4, p. 163, R. K. Willardson e A. C. Beer editores, ACADEMIC PRESS, New York, 1968.

[4] - F. A. Kröger, *The Chemistry of Imperfect Crystals*, JONH WILEY AND SONS, New York, 1964, p. 618.

[5] - J. C. Wooley, C. M. Gillett e J. A. Evans, "Some Electrical and Optical Properties of InSb-In_2Te_3 Alloys", *J. Phys. Chem. Solids*, vol. 16, n. 1/2, p. 138 (1960).

[6] - B. Tuck, *Introduction to Diffusion in Semiconductors*, IEE Monograph Series 16, PETER PEREGRINUS LTD., England, 1974.

[7] - B. Tuck, *Atomic Diffusion in III-V Semiconductors*, 1ª edição, ADAM HILGER, Bristol, 1988.

[8] - D. R. F. West, *Ternary Equilibrium Diagrams*, 2ª edição, CHAPMAN AND HALL, New York, 1985.

[9] - H. C. Casey, Jr., *Atomic Diffusion in Semiconductors*, D. Shaw editor, PLENUM PRESS, New York, 1973.

[10] - J. R. Manning, *Diffusion Kinetics for Atoms in Cristals*, D. VAN NOSTRAND COMPANY, INC., New York, 1968.

[11] - C. D. Thurmond, "Phase Equilibria in the GaAs and the GaP Systems", *J. Phys. Chem. Solids*, vol. 26, n. 5, p. 785 (1965).

[12] - R. N. Hall, "Solubility of III-V Compound Semiconductors in Column III Liquids", *J. Electrochem. Soc.*, vol. 110, n. 5, p. 385 (1963).

[13] - W. Koster e B. Thoma, *Z. Metall.*, vol. 46, p. 291 (1955).

[14] - J. R. Arthur, "Vapor Pressures and Phase Equilibria in the Ga-As System", *J. Phys. Chem. Solids*, vol. 28, n. 11, p. 2257 (1967).

[15] - N. N. Sirota, *Semiconductors and Semimetals*, vol. 4, p. 35, R. K. Willardson e A. C. Beer editores, ACADEMIC PRESS, New York, 1968.

[16] - H. R. Potts e G. L. Pearson, "Annealing and Arsenic Overpressure Experiments on Defects in Gallium Arsenide", *J. Appl. Phys.*, vol. 37, n. 5, p. 2098 (1966).

[17] - H. Matino, "Reproducible Sulfur Diffusion into GaAs",

Solid-State Electronics, vol. 17, p. 35 (1974).

[18] - M. B. Panish, "The Ga-P-Te System", *J. Electrochem. Soc.*, vol. 114, n. 11, p. 1161 (1967).

[19] - B. Goldstein, "Diffusion in Compound Semiconductors", *Physical Review*, vol. 121, n. 5, p. 1305 (1961).

[20] - L. J. Vieland, "The Effect of Arsenic Pressure on Impurity Diffusion in Gallium Arsenide", *J. Phys. Chem. Solids*, vol. 21, p. 318 (1961).

[21] - T. H. Yeh, "Diffusion of Sulfur, Selenium, and Tellurium in Gallium Arsenide", *Journal of The Electrochemical Society*, vol. 111, n. 2, p. 253 (1964).

[22] - R. G. Frieser, "Donor Diffusion into GaAs from Group VI Compounds", *Journal of the Electrochemical Society*, vol. 112, n. 7, p. 697 (1965).

[23] - A. B. Y. Young e G. L. Pearson, "Diffusion of Sulfur in Gallium Phosphide and Gallium Arsenide", *J. Phys. Chem. Solids*, vol. 31, p. 517 (1970).

[24] - B. Tuck e R. G. Powell, "High-Temperature Diffusion of Sulfur in GaAs", *J. Phys. D : Appl. Phys.*, vol. 14, p. 1317 (1981).

[25] - M. D. Zahari e B. Tuck, "Effect of Vacancy Reduction on Diffusion in Semiconductors", *J. Phys. D : Appl. Phys.*, vol. 15, p. 1741 (1982).

[26] - M. Oren e F. C. Prince, "GaAs IC Fabrication Using a Novel Open-Tube Sulfur Diffusion Technique", *Elec. Lett.*, vol. 22, n. 4, p. 221 (1986).

[27] - F. C. Prince, M. Oren e M. Lam, "High Mobility GaAs Layers Obtained by Open Tube Sulfur Diffusion", *Appl. Phys. Lett.*, vol. 8, n. 8, p. 546 (1986).

[28] - C. J. Hwang e J. R. Brews, "Electron Activity Coeficients in Heavily Doped Semiconductor with Small Effective Mass", *J. Phys. Chem. Solids*, vol. 32, p. 837 (1971).

[29] - W. Shockley, *Electrons and Holes in Semiconductors*, D. VAN NOSTRAND, Princeton, N. J., 1950.

[30] - S. M. Sze, *Physics of Semiconductors Devices*, 2ª edição, JONH WILEY & SONS, New York, 1981.

[31] - R. Venkatasubramanian e S. K. Ghandhi, "Compensation

Mechanisms in n^+-GaAs Doped with Sulphur", *Journal of Crystal Growth*, vol. 97, p. 827 (1989).

[32] - J. Drowart e P. Goldfinger, *J. Chim. Phys.*, vol. 55, p. 721 (1958).

[33] - J. R. West, *Indust. and Eng. Chem.*, vol. 42, p. 713 (1950).

[34] - J. R. Arthur, comunicação privada.

[35] - S. Antkiw e V. H. Dibeler, "Mass Spectrum of Gallium Vapor", *J. Chem. Phys.*, vol. 21, n. 10, p. 1890 (1953).

CAPITULO IV

DIFUSÃO DE ENXOFRE EM GaAs : EXPERIMENTO

4.1 - Introdução

Neste capitulo passamos a descrever as experiências realizadas e os resultados obtidos, e a analizar esses resultados. Todas as experiências de difusão foram realizadas no forno rápido que foi descrito no capítulo 2. Com relação à difusão o nosso objetivo foi desenvolver um processo de curto tempo, o que chamamos de difusão por processamento térmico rápido, que pudesse ser aplicado em amostras de grande diâmetro, de até duas polegadas. O objetivo foi plenamente atingido, e nos itens seguintes deste capitulo iremos relatar quais são as condições que permitem uma difusão rápida e que tipo de resultado é obtido.

4.2 - Projeto do Porta-Amostra

Para a realização da difusão foi necessário projetar e construir um porta-amostra que recebesse a amostra e a fonte de difusão. Tínhamos duas opções, um porta-amostra do tipo ampola selada, ou um do tipo tubo aberto.

No primeiro tipo de porta-amostra, ampola selada, a amostra e a fonte de difusão são colocadas no interior de um tubo (figura 4.1), geralmente de quartzo de alta pureza, que é evacuado até uma pressão menor que 1×10^{-3} Torr, e depois selado. Dessa forma a amostra e a fonte de difusão ficam lacradas no interior de uma ampola em vácuo, o que garante um ambiente de alta pureza para o processo de difusão. Isso é muito importante, pois queremos ter o controle do tipo e quantidade de impurezas a serem introduzidas na amostra durante a difusão e necessitamos, portanto, de um ambiente de difusão bastante limpo. Outra vantagem da ampola

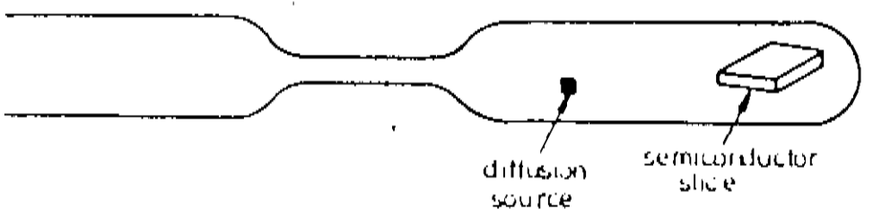

Figura 4.1 - Ampola para difusão em tubo fechado.

selada é a possibilidade de termos um bom controle da pressão do vapor das impurezas colocadas no interior da ampola e uma grande gama de variação dessas pressões. Isso é possivel já que temos a liberdade de variar a quantidade de material da fonte e o volume da ampola, que é bem conhecido. Essa maior gama de variação de pressões permite um estudo bastante aprofundado da difusão, o que é importante, pois pode-se conhecer melhor os mecanismos de difusão em cristais, o que não seria possivel se não tivéssemos a pressão como uma variável do processo. Como desvantagens a ampola selada não é aplicável a amostras grandes, ficando o seu uso restrito a amostras de no máximo dois centímetros de lado. Outra desvantagem é a necessidade de um equipamento de vácuo (bomba, conexões para vácuo, etc.) e um maçarico para a selagem da ampola, o que torna o processo não muito prático.

Outro tipo de porta amostra é o chamado tubo aberto (figura 4.2). Neste caso a amostra e a fonte de difusão estão colocadas próximas uma da outra no interior de um suporte que não é lacrado. A vantagem deste tipo de porta-amostra é que podemos trabalhar com amostras de grande diâmetro, da ordem de várias polegadas. Como exemplo temos o nosso porta-amostra que pode receber "wafers" de até duas polegadas de diâmetro. Outra vantagem é a grande praticidade do tubo aberto, já que não há necessidade de um sistema de vácuo, nem de selagem de ampola, o que facilita bastante o manuseio de amostras. As desvantagens do tubo aberto são devidas à falta de selagem. A primeira é a pequena gama de valores da pressão do vapor da fonte que podem ser utilizados

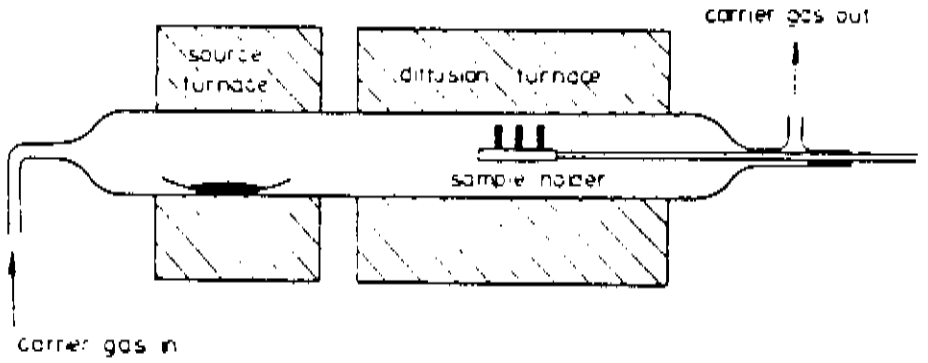

Figura 4.2 - Porta-amostra para difusão em tubo aberto.

durante a difusão. Esta geralmente é igual à pressão de vapor do material da fonte na temperatura de difusão e só pode ser mudada caso alteremos a composição da fonte ou mudemos a temperatura de difusão. Uma alteração da quantidade de material da fonte (sem alterar a sua composição) não muda a pressão, como ocorre com uma ampola lacrada. Existe um limite para a pressão da fonte, embora esta fique fixada na sua pressão de vapor. Este limite é que a pressão da fonte deve ser sempre menor que a atmosférica, já que geralmente o suporte é colocado em um forno com um fluxo de gás inerte à pressão atmosférica. Não podemos ter dessa forma a pressão da fonte maior que a atmosférica, pois isso implicaria na sublimação da fonte e a sua exaustão (todos os vapores da fonte seriam arrastados do interior do forno). A segunda desvantagem do tubo aberto é a maior susceptibildade à contaminação da amostra por impurezas existentes no gás inerte usado no forno, ou por impurezas deixadas no porta-amostra devido à sua manipulação.

Uma variante do tubo aberto onde a amostra não está próxima da fonte de difusão no mesmo suporte é o caso onde temos um fluxo de gás dopante no interior do forno. Neste caso a fonte de difusão pode ser sólida líquida ou gasosa. No caso de uma fonte sólida, esta seria colocada na entrada do forno onde é injetado um gás portador, que é inerte. A fonte irá perder material na forma de vapor que será carregado para as amostras que recebem a difusão. Podemos ter ainda uma fonte líquida, onde neste caso o gás portador é borbulhado no líquido antes de entrar no forno,

carregando uma parte do liquido na forma de vapor. Por fim podemos ter uma fonte na forma de gás, que seria simplesmente injetado na entrada do forno junto com o gás de arraste.

4.2.1 - Projeto do Suporte de Grafite

Para a realização das difusões rápidas optamos pelo sistema de tubo aberto. O motivo da escolha foi que o tubo aberto permite o uso de amostras de grande diâmetro, o que foi um dos requisitos iniciais para o desenvolvimento do processo. Entre o uso de um suporte e um fluxo gasoso de dopantes optamos pelo suporte. O motivo dessa opção é que as paredes do forno (tubo de quartzo) estão relativamente frias, o que provocaria a deposição do material dopante nas paredes caso usássemos um fluxo gasoso de dopantes (isso ocorre para uma fonte sólida, que é justamente o nosso caso, como iremos ver nos próximos itens). A deposição de dopantes na parede do forno é indesejável, pois implica em um consumo excessivo de dopantes e torna as paredes opacas, o que prejudica a transmissão de radiação das lâmpadas por estas. Outro motivo da escolha do suporte foi a maior facilidade que temos de aquecer a amostra e a fonte, pois podemos fazer o aquecimento indiretamente. Neste caso é o suporte quem recebe a radiação e aquece a amostra e a fonte que são colocadas no seu interior. Essa maior facilidade é devido ao fato de que a amostra ou a fonte podem ser transparentes a certos comprimentos de onda da radiação, o que prejudica o aquecimento ou o torna desigual entre a fonte e a amostra. Já o suporte pode ser feito de um material totalmente opaco à radiação das lâmpadas, o que resultaria em uma absorção mais uniforme da radiação e um consequente aquecimento maior e mais uniforme. Caso o suporte seja feito de um material bom condutor de calor isso irá melhorar a uniformidade de temperatura da amostra e da fonte o que é mais uma vantagem do uso do suporte.

Na figura 4.3 podemos ver um desenho do porta-amostra mostrando as posições da amostra e da fonte de difusão no seu interior. O porta-amostra é constituido por 3 discos e um anel. O material usado na sua fabricação foi um tipo de grafite de alta

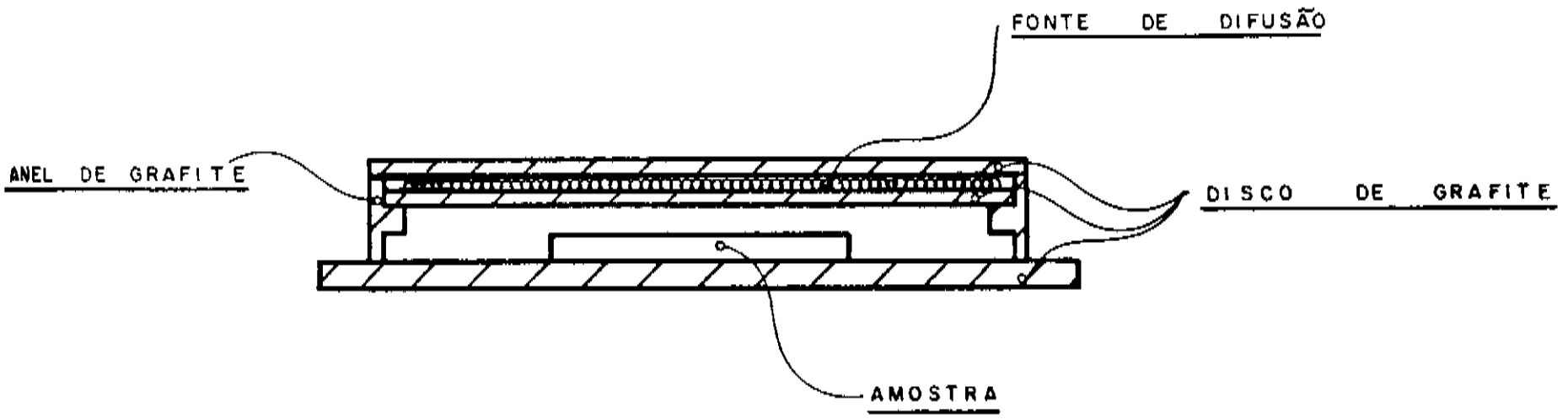

Figura 4.3 - Desenho em corte do porta-amostra de grafite mostrando as posições da fonte de difusão e da amostra.

pureza, o mesmo tipo usado na fabricação de "boats" de grafite para fornos de crescimento do tipo LPE (epitaxia em fase líquida). Este tipo de grafite possui uma alta condutividade térmica, o que é uma propriedade desejável para o material do porta-amostra. Na figura 4.4 está o desenho detalhado de cada peça do porta-amostra. Como podemos observar na figura 4.3 a amostra é colocada sobre o disco inferior, e é aquecida por ele. Sob a amostra é colocado outro disco que contém cerca de 200 furos de diâmetro 0,4mm, com espaçamento aproximado de 1,5mm um do outro. Sobre esse segundo disco é colocada a fonte de difusão. Os vapores da fonte atingem a amostra através dos furos existentes no disco, e como a fonte é distribuida uniformente sobre o disco, conseguimos garantir que o vapor atinge a amostra de maneira bastante uniforme. Por fim sobre a fonte é colocado o tercero disco que a recobre, e serve para aquece-lá, além de impedir uma grande perda de vapor da fonte para o interior do forno. O segundo disco e o terceiro são suportados por um anel, também de grafite, que é colocado sobre o primeiro disco. Esse anel mantém o segundo disco a uma distância de 0,5mm da amostra. Para que o suporte tenha a menor capacidade térmica, e apresente portanto uma grande taxa de aquecimento, ele deve ser feito com discos bem finos. No nosso projeto adotamos uma espessura de 0,5mm para os discos, o que garante uma baixa capacidade térmica e ainda proporciona uma boa rigidez mecânica aos discos.

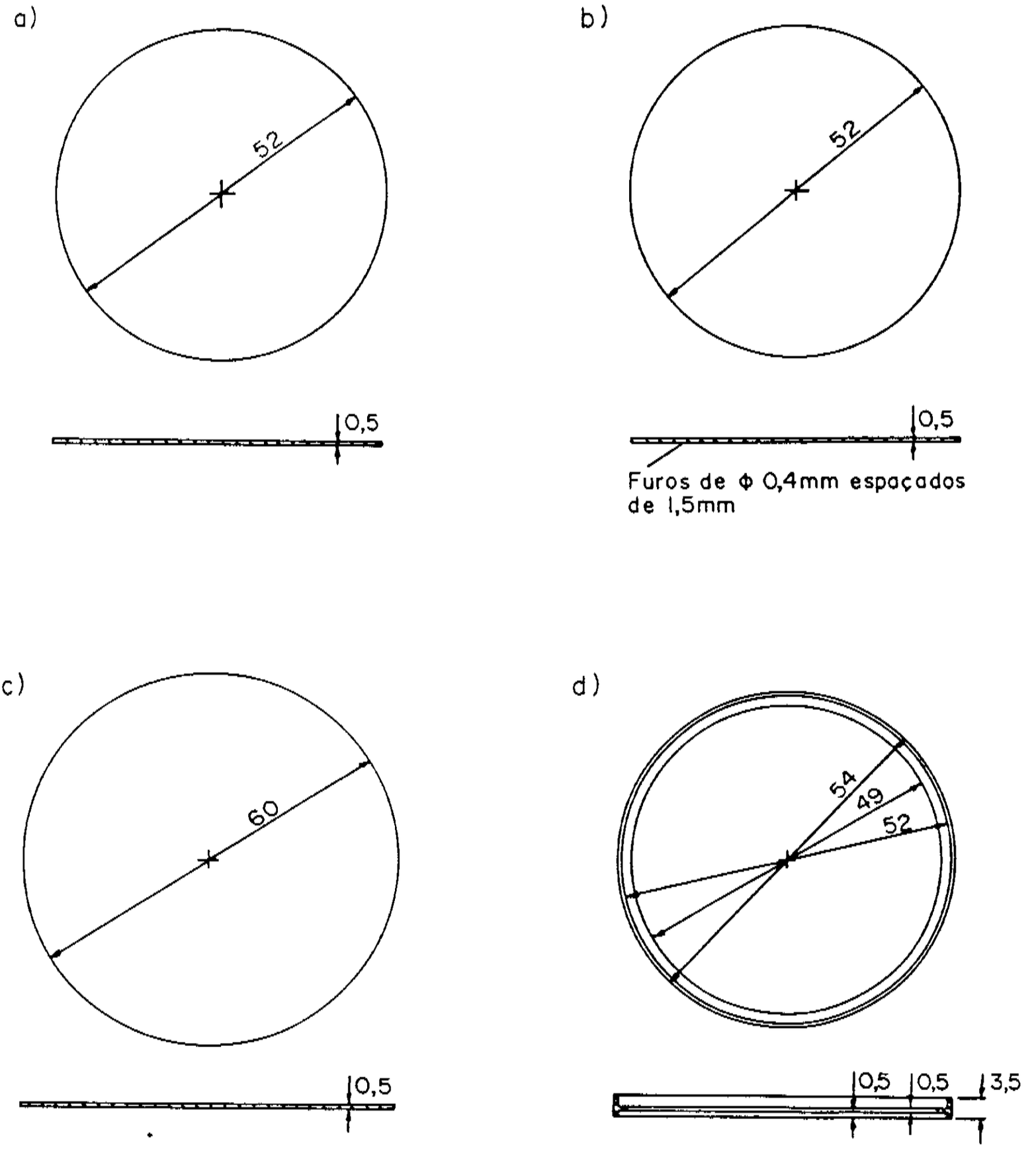

Figura 4.4 - Desenho detalhado das peças do porta-amostra. a) tampa superior, b) disco intermediário, c) disco de base e d) anel de suporte.

A fabricação dos discos e do anel foi feita partindo-se de um bloco de grafite de 4mm de espessura que foi preso em um suporte e torneado em um torno para usinagem de peças metálicas. Após o torneamento as peças foram limpas em tricloroetileno para eliminar alguma contaminação introduzida durante o processo de usinagem e enxaguadas em acetona e metanol e depois secas com um jato de N_2. Depois disso as peças foram aquecidas para uma completa secagem, primeiro a 100°C por 2 minutos, sendo então aquecidas a 400°C por mais 2 minutos e finalmente a 900°C por 1 minuto, isso em uma atmosfera de N_2 de alta pureza. Essa secagem a alta temperatura foi feita no forno rápido e garante que as peças estejam livres de contaminantes.

4.2.2 - Suporte de Quartzo

A colocação do porta-amostra de grafite no interior do tubo de quartzo apresenta um problema. O porta-amostra é redondo e deve ficar apoiado no tubo cilíndrico, o que resulta em um apóio instável, uma vez que todo o suporte fica sustentado por dois pontos (as extremidades do disco de grafite que ficam em contato com o tubo). Para evitar esse problema construímos um suporte de quartzo para o porta-amostra. O seu desenho pode ser visto na figura 4.5, e na figura 4.6 temos uma fotografia que mostra o porta-amostra colocado sobre o suporte. O motivo de usar quartzo para o material desse suporte é que este é transparente à radição e não perturba, portanto, a iluminação do porta-amostra de grafite, o que garante a sua uniformidade de aquecimento.

Como podemos observar nas figuras 4.5 e 4.6, o porta-amostra de grafite é sustentado pelas quatro pequenas pontas de quartzo do suporte. O contato entre as pontas e o disco de grafite é feito o menor possível, já que as pontas estão frias, pois o quartzo não é aquecido pela radiação, e esse contato das pontas frias com o grafite pode criar uma variação de temperatura na região do disco próxima ao ponto de contato. Observe ainda que os pontos de contato estão localizados na periferia do disco de base, e que este é mais largo que os outros discos, o que mantém as

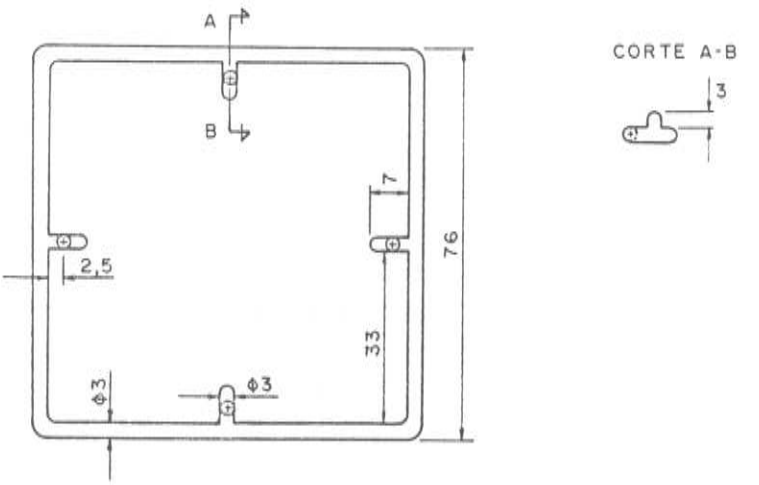

Figura 4.5 - Desenho do suporte de quartzo.

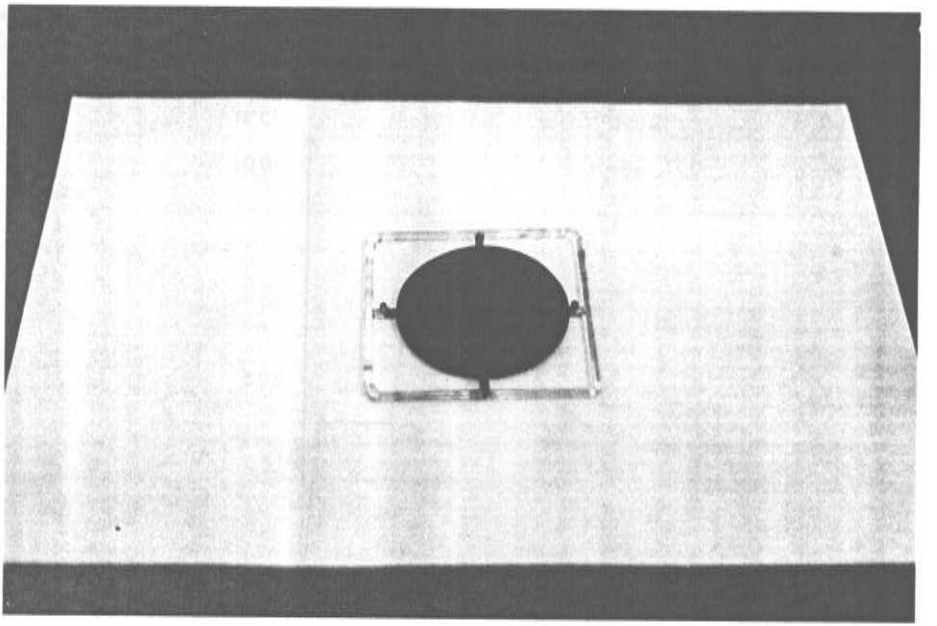

Figura 4.6 - Fotografia do porta-amostra colocado sobre o suporte de quartzo.

pontas afastadas da região onde é colocada a amostra. Esses cuidados são importantes para garantir a uniformidade de temperatura do suporte de grafite.

4.3 - Ciclo Térmico

Como foi comentado no capitulo 2, o sensor de temperatura do forno é um disco de grafite de 1,5mm de espessura, que é mantido permanentemente no interior do forno. O motivo de se usar um sensor ao invés de medir a temperatura diretamente no porta-amostra, é que isso simplifica a montagem mecânica do termopar, já que a sua colocação no porta-amostra não é prática, além de dificultar a manipulação deste último.

A espessura dos discos do porta-amostra sobrepostos e igual à do disco sensor de temperatura, e eles são feitos do mesmo material (grafite) o que deve à princípio garantir que ambos, o suporte e o sensor, tenham uma mesma variação de temperatura.

O aquecimento até a temperatura de difusão é feito em duas

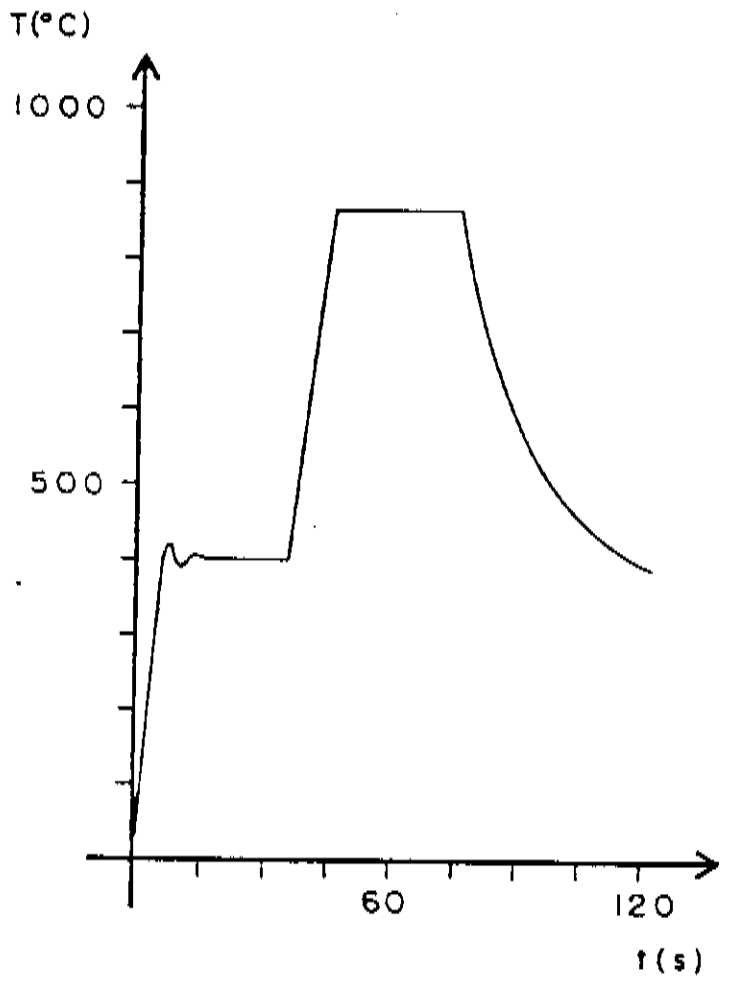

Figura 4.7 - Ciclo térmico usado nas difusões.

etapas, primeiro até 400°C e depois até a temperatura final. O pré-aquecimento em 400°C é necessário para estabelecer uma temperatura fixa para o início do aquecimento final. Dessa forma eliminamos uma possível dependência do ciclo final de aquecimento com a temperatura ambiente, já que estamos partindo sempre de 400°C e não da temperatura ambiente. Outra vantagem é que o porta-amostra irá sofrer um menor aquecimento até a temperatura final, o que ajuda a minimizar gradientes de temperatura que ocorram no porta-amostra durante o aquecimento. Na figura 4.7 podemos ver o ciclo térmico (temperatura em função do tempo) usado nas experiências de difusão. O pré-aquecimento é feito durante 30 segundos.

Algumas experiências de recozimento de amostras de GaSb foram feitas no forno rápido com um suporte de grafite semelhante ao descrito no item 4.2, porém de menor diâmetro (ϕ 25mm). A temperatura de fusão do GaSb é de 712°C. Foram realizados ciclos de aquecimento a temperaturas de 500°C, 550°C, 600°C e 650°C. Observamos que os aquecimentos a 600°C e 650°C produzem a fusão da amostra (tamanho de 3x3mm). Atribuimos a fusão a uma maior taxa de aquecimento do suporte onde é colocada a amostra, que se deve a uma menor inércia térmica do suporte. Dessa forma o suporte pode atingir temperaturas maiores do que o sensor, durante o ciclo de aquecimento, antes da estabilização da temperatura no sensor.

A experiência com o recozimento de cristais de GaSb nos forneceu um importante resultado que é a existência de uma diferença de temperatura, que pode ser muito grande, entre o sensor e o porta-amostra. Passamos então a realizar experiências para determinar qual é o comportamento térmico do porta-amostra. Essas experiências foram realizadas medindo-se simultaneamente a temperatura do sensor e do porta-amostra. A medida da temperatura do sensor foi feita simplesmente com o termômetro do forno e o termopar colocado no sensor. Já a medida de temperatura do porta-amostra foi feita colocando-se um termopar em seu interior em contato com o disco de base. Os fios do termopar saem do interior do suporte através de dois furos neste disco. A tampa frontal do forno teve que ser substituída por uma com um furo no seu centro para permitir a passagem dos fios. A medida de

temperatura foi feita com um termômetro ligado ao termopar.

As experiências foram realizadas aquecendo o forno partindo de uma temperatura entre 40°C e 60°C, com medida de temperatura idêntica por ambos os termômetros. O aquecimento foi realizado com potência total das lâmpadas. Experimentamos três espessuras de discos de base, 0,48mm, 0,58mm e 0,64mm. Para as duas primeiras medidas observamos um sobreaquecimento do porta-amostra, como é mostrado na figura 4.8, onde comparamos as temperaturas do sensor e do porta-amostra. Para o disco de 0,64mm observamos um sub-aquecimento de duração aproximada de 20 segundos. O sobreaquecimento do porta-amostra é caracterizado por um pico de temperatura ΔT, que tem uma duração máxima de 15 segundos. Após o pico a temperatura do porta-amostra cai, praticamente igualando-se à do sensor, com uma diferença máxima de 2°C. Essa experiência foi repetida para cada disco de base com 5 temperaturas finais diferentes, de 400°C a 800°C com intervalos de 100°C. As medidas de ΔT em função da temperatura final do sensor podem ser vistas no gráfico da figura 4.9. Para o disco de 0,58mm achamos os menores picos de temperatura. Observamos neste caso que o pico aumenta partindo da temperatura final de 400°C, atingindo o máximo em 500°C, quando então começa a cair. Um comportamento

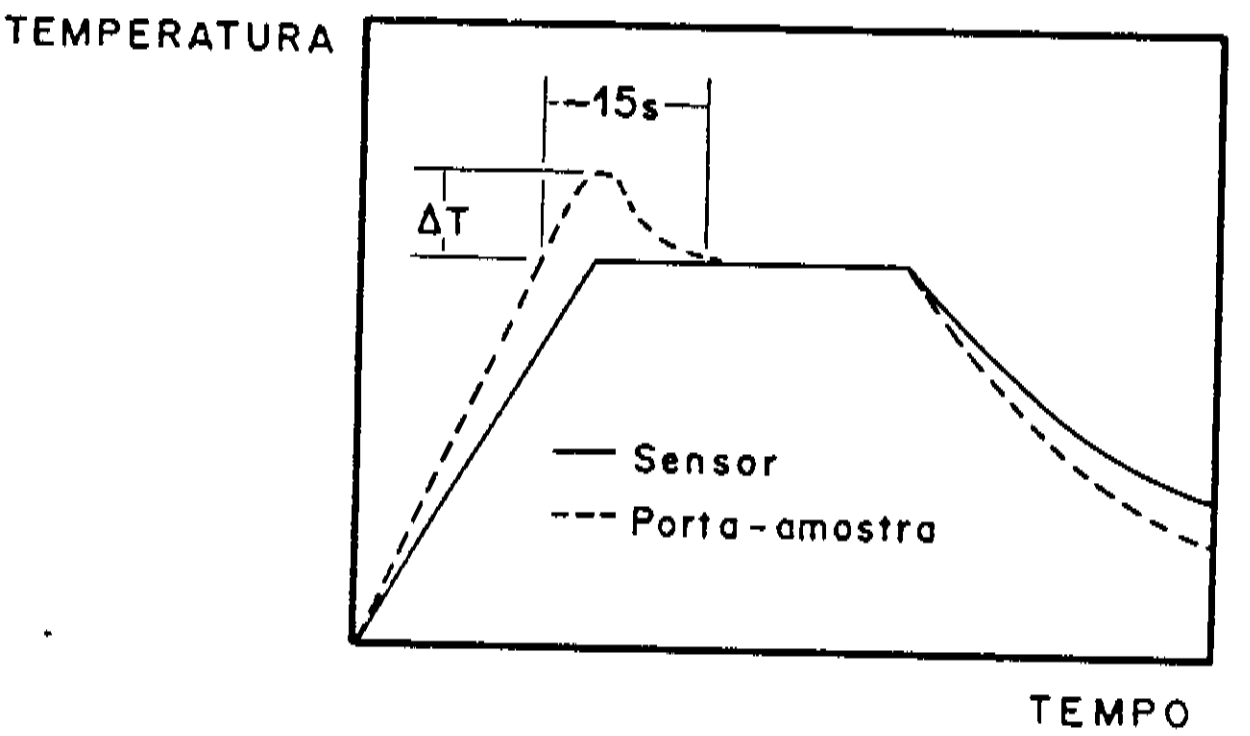

Figura 4.8 - Temperaturas no sensor (linha cheia) e no porta-amostra de grafite de uma polegada de diâmetro (linha pontilhada).

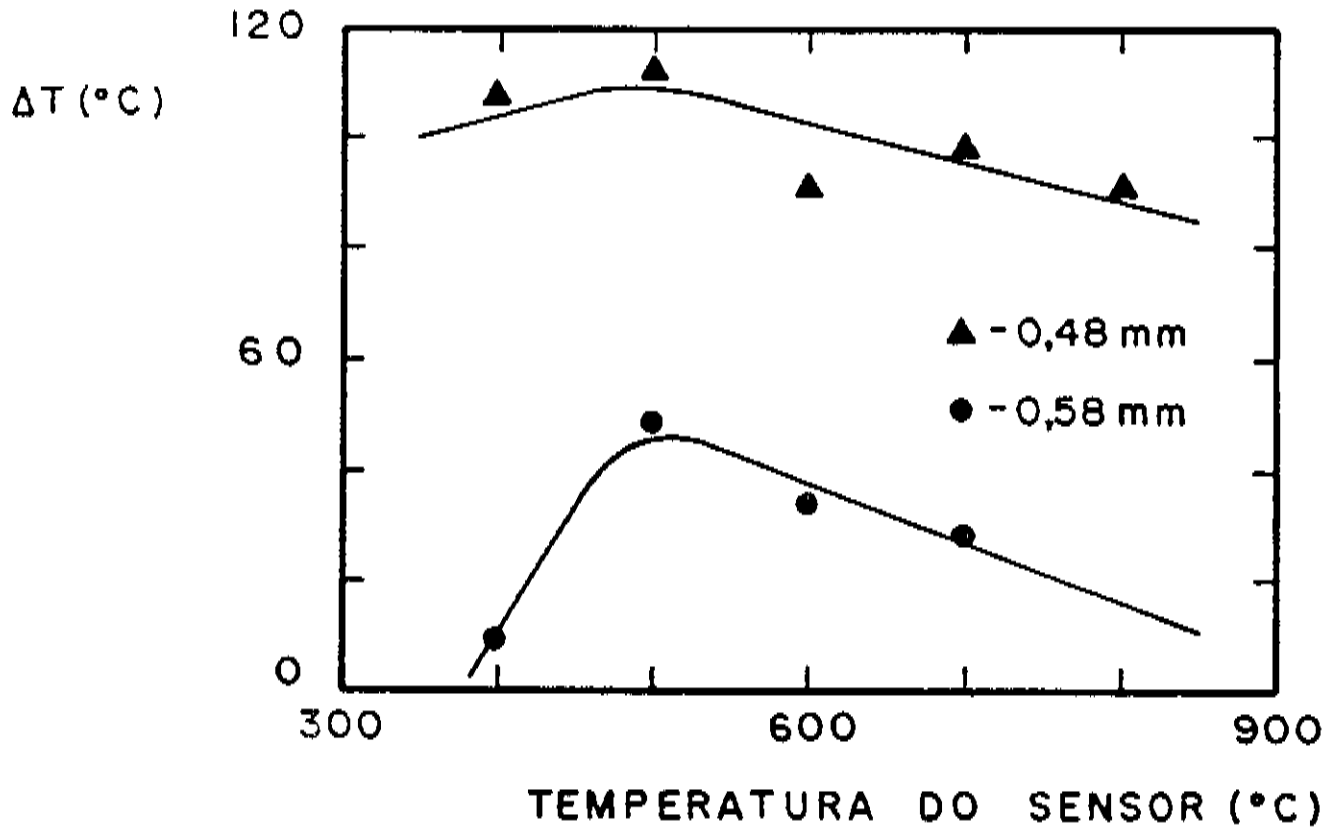

Figura 4.9 - Sobreaquecimento do porta-amostra, ΔT, em função da temperatura final do sensor para duas espessuras do disco de base.

semelhante porém menos acentuado é observado para o disco de 0,48mm, com a diferença de que os picos são em média três vezes maiores para este disco. O comportamento descrito é explicado observando-se que a partir de aproximadamente 500°C a taxa de aquecimento do forno começa a cair, chegando a zero em cerca de 950°C. Dessa forma, até 500°C esperamos que os picos cresçam, pois para uma temperatura mais alta o aquecimento deve demorar mais, o que aumenta as diferenças de temperatura entre o sensor e o porta-amostra. A partir de 500°C podemos concluir que a taxa de aquecimento do porta-amostra deve cair mais rapidamente, pois este está mais quente que o sensor. Isso fará com que as diferenças de temperatura diminuam, conforme a temperatura final seja cada vez maior.

Embora tenhamos observado esses picos de alta temperatura no porta-amostra, isso não prejudicou as nossas experiências de difusão, por dois motivos. Primeiro o nosso disco de base tem uma espessura de 0,54mm, próxima do valor do segundo disco, 0,58mm, e portanto devemos esperar picos menores que cerca de 30°C, já que como veremos adiante, as temperaturas de difusão são maiores que

830°C, o que deve minimizar os picos. O segundo motivo é que a duração dos picos é de aproximadamente 15s, o que é um tempo pequeno, quando comparado aos tempos de difusão, maiores que 30s. Algumas difusões foram feitas com 30s de tempo de duração, porém o uso de pequenos tempos, comparáveis à duração do pico de sobre-aquecimento, implica em altas temperaturas, maiores do que 850°C, onde praticamente não temos sobre-aquecimento do porta-amostra.

4.4 - Escolha e Preparação das Amostras

No nosso trabalho optamos pelo uso de substratos do tipo semi-isolante, com resistividade maior que $10^6 \Omega cm$. Usamos dois tipos de substratos, dopados com cromo e não dopados. Os substratos usados tiveram três procedências diferentes, o dopado com cromo foi obtido da Cominco, e os não dopados da Sumitomo e Wacker. Na tabela 4.1 podemos ver as características desses substratos, as suas resistividades, a mobilidade dos portadores (elétrons), a densidade superficial de pontos revelados por ataques químicos, EPD, que está relacionada com a densidade de dislocações no cristal, e a sua dopagem. O motivo desses substratos possuirem alta resistividade é a existência de níveis profundos no "gap" (níveis de energia no meio do "gap") que irão funcionar como armadilhas para os elétrons ou buracos que foram

Fornecedor	Cominco	Sumitomo	Wacker
Resistividade (Ωcm)	$>10^6$	1.4×10^8	3×10^7
Mobilidade (cm^2/Vs)	-	-	7300
EPD (cm^{-2})	-	26000	52000
Dopagem	Cr	Nenhuma	Nenhuma

TABELA 4.1 - Características dos "wafers" usados na difusão.

gerados por uma eventual dopagem não intencional do substrato. No caso do substrato dopado com cromo, este gera o nível profundo, enquanto que nos outros substratos o nível é gerado por defeitos intrinsecos ao cristal de GaAs. Todos os substratos usados nos experimentos de difusão foram de alta qualidade, próprios para trabalhos de fabricação de transistores MESFET. Escolhemos substratos semi-isolantes para serem utilizados como camada para isolação entre MESFET's, como veremos no capitulo 6.

Para as experiências de difusão usamos pequenos cristais de 1x1cm, que foram cortados (clivados) dos wafers. Para a fabricação de transistores usamos pedaços maiores de 2x2cm. Após o corte iniciamos a limpeza dos cristais. Esta é feita decapando-se uma camada de 1μm da superfície do cristal, que pode conter defeitos gerados pelo seu polimento. Após a decapagem obtemos uma superfície limpa e livre de defeitos. A decapagem é feita com uma solução de ácido sulfúrico, água oxigenada e água deionizada, $H_2SO_4:H_2O_2:H_2O$(DI), numa proporção 8:1:1, à temperatura ambiente, por 60 segundos. Após o ataque químico a amostra é enxaguada com bastante água deionizada e sêca com um jato de nitrogênio puro. A amostra é então levada para o forno e é feita a difusão. Podemos guardar o cristal após a limpeza, porém a sua exposição durante um longo tempo à atmosfera provocará a oxidação da superfície. Antes da difusão esse óxido deve ser removido, e neste caso usamos uma solução de ácido cloridrico e água deionizada, $HCl:H_2O$(DI), numa proporção de 1:3. O cristal é mantido nessa solução por 60 segundos, e após isso é enxaguado com bastante água DI e séco com um jato de nitrogênio. Todas as duas soluções ácidas mantém a morfologia da superfície do cristal inalterada, ou seja polida e brilhante.

4.5 - Escolha e Preparação da Fonte de Difusão

A fonte de difusão é a fase externa ao semicondutor que contém a impureza a difundir. Como já vimos, a correta e cuidadosa escolha da composição da fonte irá nos garantir uma difusão

reprodutivel e sem danos à superfície do cristal devidos à formação de novas fases durante o aquecimento deste. Obviamente a fonte deve conter a impureza a difundir, que no nosso caso é o enxofre. Ela deverá conter também arsênio, já que o GaAs apresenta o problema de evaporação incongruente para temperaturas maiores que cerca de 630°C. Essa evaporação incongruente é caracterizada pela perda de arsênio da superfície, criando uma camada rica em gálio, o que é indesejável durante a difusão. Para evitar isso devemos criar um vapor que contenha arsênio gasoso no ambiente onde se encontra a amostra a difundir. Esse vapor irá saturar a atmosfera ao redor do cristal com arsênio, fazendo com que a perda deste pela superfície do cristal entre em equilíbrio dinâmico e se mantenha uma concentração normal de arsênio na superfície.

No nosso trabalho usamos o critério de Matino [1] para a escolha da composição da fonte de difusão. Segundo matino, as fontes que apresentam três fases condensadas possuem um grau de liberdade zero, e consequentemente as pressões de vapor dos elementos que compõem a fonte são constantes e independentes da exata composição. Na figura 3.10 podemos ver o diagrama de fase do Ga-As-S para temperaturas menores que 962°C. As regiões A e B contém cada uma três fases condensadas e serão preferíveis para a escolha da composição da fonte. Usando o critério de Matino

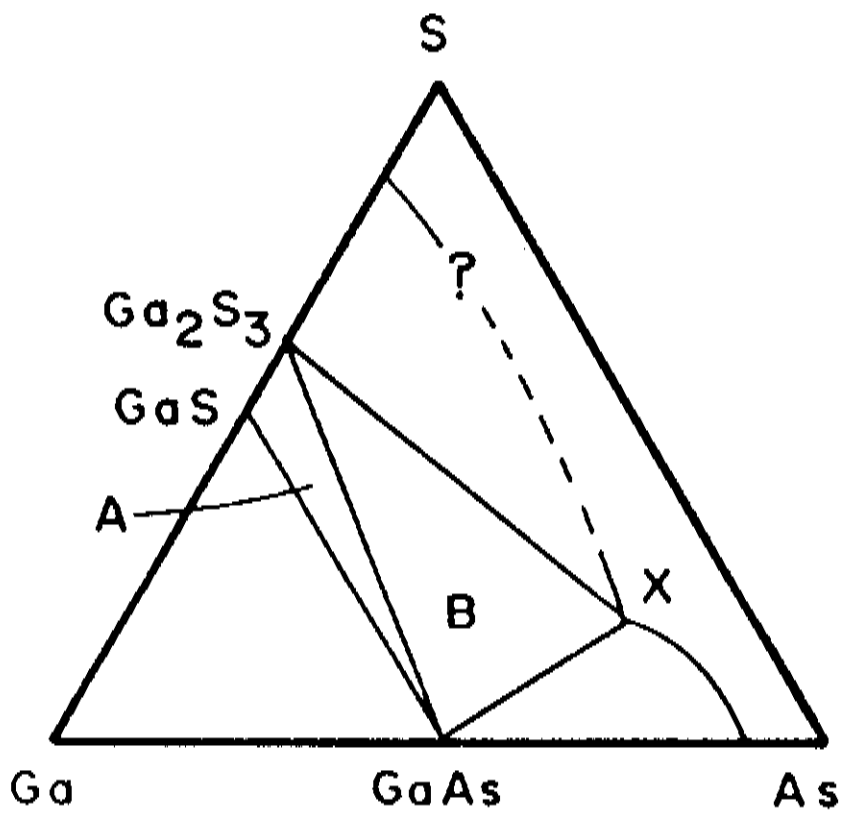

Figura 4.10 - Diagrama de fases do Ga-As-S [1].

escolhemos uma mistura de Ga_2S_3 e GaAs para a nossa fonte. O Ga_2S_3 é obviamente a fonte de enxofre, e o próprio GaAs é a de arsênio. O Ga_2S_3 usado possue alta pureza, 99.99%, e foi obtido da Alpha Products. Sua forma é de um pó branco, que é lacrado dentro de uma ampola. O GaAs usado na fonte foi obtido de lâminas poli-cristalinas de alta pureza (concentração de portadores menor que $5x10^{15} cm^{-3}$ para wafers obtidos a partir de crescimentos usando essas lâminas como fonte de GaAs), que foram moídas até atingirem a forma de um pó bastante fino. Antes de moer as lâminas estas foram limpas com a mesma solução usada na decapagem para limpeza das amostras a difundir. Para moer as lâminas usamos um pequeno pilão de cerâmica com um socador. O fato de usarmos o GaAs na fonte na forma de pó implica em termos uma área de superfície do pó muito maior que a da amostra. Dessa forma o pó irá perder muito mais arsênio do que a superfície da amostra, mantendo a integridade desta última. A pressão de vapor de arsênio no interior do porta-amostra será a de dissociação do cristal, supondo que a perda de arsênio para o interior do forno é pequena (o porta-amostra é do tipo tubo aberto).

Todas as composições usadas para a fonte de difusão localizaram-se na linha entre o Ga_2S_3 e o GaAs, a fronteira entre as regiões A e B. Após o aquecimento a fonte perde arsênio e enxofre, entrando na região A. As composições usadas são mostradas na tabela 4.2.

	Concentração (% de peso)	
Fonte	Ga_2S_3	GaAs
S1	1,0	99,0
S5	5,0	95,0
S12	12,5	87,5
S50	50,0	50,0

TABELA 4.2 - Composição das fontes de difusão.

A colocação da fonte de difusão no porta-amostra é bastante simples. Separamos cerca de 0,5g do pó resultante da mistura do Ga_2S_3 com o GaAs e o espalhamos sobre o disco de grafite furado, tomando o cuidado de coletar o pó que atrevessa os furos para colocá-lo de volta sobre o disco. Feito isso a fonte é aquecida a 800°C por 60s no forno rápido antes que a primeira difusão seja feita. Esse aquecimento elimina a umidade da fonte e os óxidos de gálio, arsênio e enxofre, que possuem todos um baixo ponto de evaporação. Caso a fonte fique guardada dentro do porta-amostra por muito tempo, é necessário que ela seja aquecida, da forma comentada acima, antes da difusão. Na tabela 4.3 podemos ver

Composto	TF (°C)	TE (°C)	Cor do Sólido
GaAs	1238	-	cinza
Ga_2O_3	1900	-	branco
Ga_2O	>660	sublima 550	marrom escuro
$Ga_2(SO_4)_3$	dissoc. 690	-	branco
GaS	965	-	c. amarelo
Ga_2S_3	1255	-	c. amarelo a. branco
Ga_2S	d. vac. 800	-	cinza escuro
As_2O_5	decomp. 315	-	branco
As_2O_3	sublima 193	457	vitreo
As_2S_2	307	565	vermelho-marrom
As_2S_5	sublima	decomp. 500	amarelo
As_2S_3	300	707	amarelo
SO_3	17	45	vitreo
SO_2	-73	-10	incolor (liq.)

abreviações : c. - cristal, a. - amorfo, d. vac. - dissocia em vácuo, dissoc. - dissocia, e decomp. - decompõe.

TABELA 4.3 - Propriedades físicas (temperaturas de fusão e evaporação) de materiais que podem existir na fonte de difusão.

algumas propriedades dos materiais que podem existir na fonte. Da tabela podemos verificar que apenas o Ga_2O_3 resiste a altas temperaturas. Os outros óxidos possuem temperatura de fusão menor que 800°C, e a maioria irá sublimar ou decompor quando aquecidos.

É interessante observar que a cor resultante da mistura de Ga_2S_3 com o GaAs variou de um cinza metálico de tom médio (cor característica do GaAs) até um cinza médio, quando tinhamos alta concentração de Ga_2S_3 na fonte. Para essas fontes de alta concentração notamos após a primeira difusão uma mudança de cor da fonte, de cinza médio para um vermelho ocre, que é devido à formação de Ga_2S na fonte, que é a terceira fase condensada da região A do diagrama de fases. Isso demonstra claramente a perda de enxofre pela fonte, devido à decomposição do Ga_2S_3 formando GaS mais algum composto de enxofre, provavelmente S_2, S_4, ou Ga_2S (no capitulo 3 comentamos que esta espécie pode estar presente em grande quantidade no vapor) na forma de gás, que irá deixar a

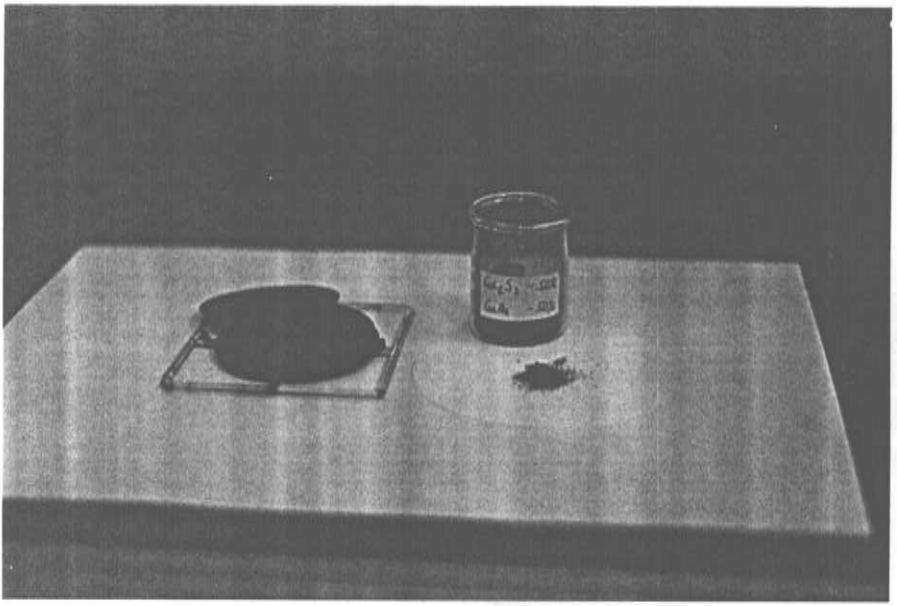

Figura 4.11 - Fonte de difusão no porta-amostra após o primeiro aquecimento. No becker ao lado vemos a fonte não usada.

fonte atingindo o GaAs. Compostos contendo arsênio e enxofre não podem estar presentes na fonte, já que todos eles possuem ponto de ebulição menor do que as temperaturas tipicamente usadas na difusão. Os únicos compostos de cor amarela que podem existir na fonte são o Ga_2S e Ga_2S_3. Na figura 4.11 podemos ver uma fotografia que mostra a fonte de difusão após o primeiro aquecimento e o pó resultante da mistura de GaAs com Ga_2S_3 em um becker.

4.6 - Determinação das Faixas de Temperaturas Utilizáveis

Nossas primeiras experiências de difusão não nos deram resultados bons. Os primeiros cristais difundidos apresentaram uma péssima morfologia de superfície, caracterizada pela formação de pequenos depósitos de algum composto químico na superfície,

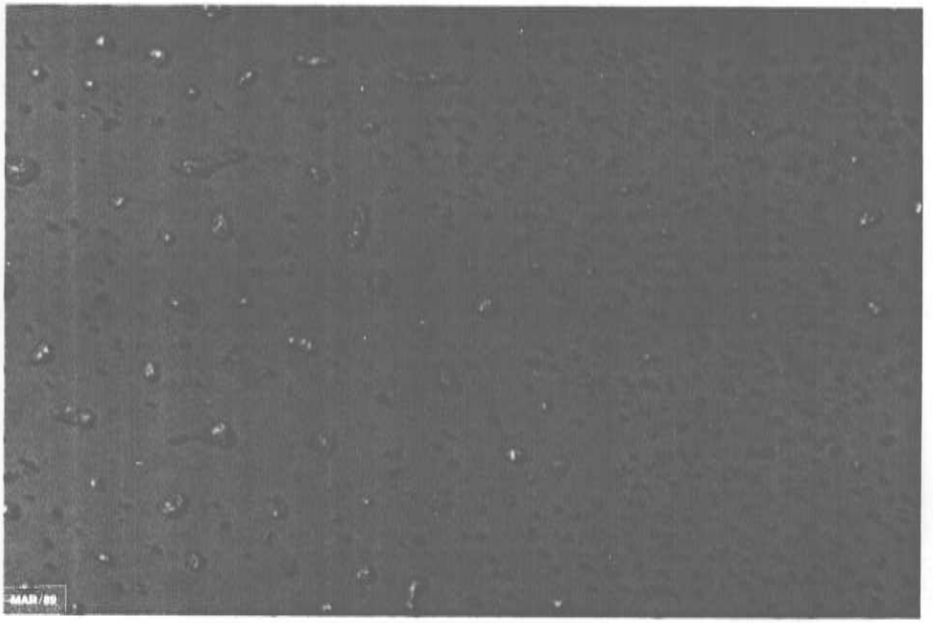

Figura 4.12 - Fotografia da superfície do cristal mostrando depósitos de sulfetos.

provavelmente um sulfeto, como apontado por Goldstein [2]. A morfologia característica nesse caso é uma superfície opaca, leitosa e sem brilho, que é resultante da deposição, que por sua vez é devida a uma alta concentração de enxofre na forma de vapor durante a difusão. Na figura 4.12 podemos ver uma fotografia mostrando os depósitos na superfície do GaAs.

Esse problema nos levou à realização de várias séries de experiências, numa tentativa de determinar quais seriam as faixas de temperaturas onde teríamos a difusão sem formação de depósitos na superfície do cristal. As experiências foram realizadas seguindo os procedimentos de limpeza e ciclos térmicos descritos nos itens 4.4 e 4.3, respectivamente. Em todos os ciclos térmicos usamos um pré-aquecimento de 400°C/30s, com tempo de difusão de 30s. Sempre antes do aquecimento fizemos uma purga da atmosfera no interior do forno com um alto fluxo de N_2 puro, cerca de 8l/min. durante 5 minutos. Isso é feito para retirar todo o oxigênio do interior do forno, garantindo uma atmosfera não oxidante.

Logo nos primeiros testes abandonamos as fontes de alta concentração, mais especificamente a de 50% e 5% de Ga_2S_3. O motivo foi a formação dos depósitos na superfície das amostras. Posteriormente voltamos a essas fontes, só que trabalhando com temperaturas bem inferiores às dos primeiros experimentos. A fonte de 12,5% foi testada posteriormente. Concentramos nossas atenções na fonte de 1% de Ga_2S_3. A escolha dessa fonte foi devida ao menor uso de sulfeto de gálio, pois tínhamos apenas uma ampola com 10g de Ga_2S_3, e era importante não gastar muito material nas primeiras experiências. Começamos os testes partindo de altas temperaturas, já que um processo de difusão em pequenos tempos iria requerer em compensação uma alto coeficiente de difusão, o que implica no uso de altas temperaturas.

Os resultados dos testes podem ser vistos na forma de um gráfico na figura 4.13, que mostra a faixa de temperaturas utilizáveis para a difusão, em função da composição da fonte (concentração de Ga_2S_3). Nesse gráfico observamos que existe um limite de temperatura superior e um inferior. O superior é devido à formação de sulfetos na superfície da amostra, como já dissemos. Já o limite inferior é devido ao fato de não observarmos uma

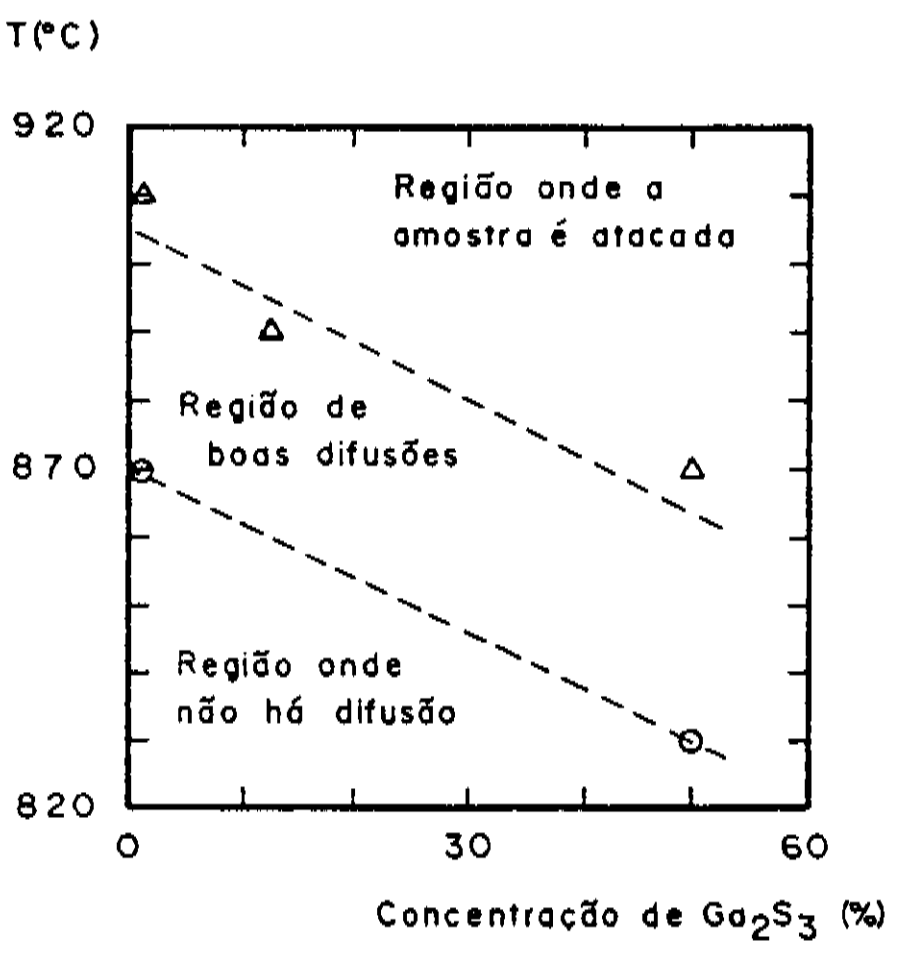

Figura 4.13 - Gráfico mostrando as faixas de temperaturas utilizáveis.

camada difundida para baixas temperaturas. Isso ocorre do fato de termos nesse caso um pequeno coeficiente de difusão, o que corresponde à formação de uma camada difundida extremamente rasa, que não é detectada. Logicamente, se mantermos a amostra por tempos muito longos nessa temperatura inferior obteremos uma camada apreciavel. O limite inferior de temperatura é uma função do tempo de difusão, e ele deve diminuir com o aumento do tempo, como podemos deduzir do que foi dito anteriormente. Já o limite superior é apenas uma função das propriedades físicas do material da fonte (pressão de vapor). Deve ficar bastante claro que os limites de temperatura mostrados no gráfico não são precisos. O gráfico apenas especifica que próximo aos limites mostrados começamos a ter problemas, ou há formação de depósitos de sulfetos, ou não há difusão.

Na figura 4.13 podemos observar um estranho efeito. As fontes com maior concentração de Ga_2S_3 possuem temperaturas de trabalho menores que as fontes menos concentradas. Isso aparentemente contradiz a regra de Matino, já que todas as fontes

estão na região A do diagrama, onde as pressões de vapor são fixas e constantes. Portanto devemos esperar que as faixas sejam iguais para quaisquer concentrações de sulfeto de gálio. A explicação dessa anormalidade vem do fato de usarmos um suporte do tipo tubo aberto, que não é um sistema fechado. Podemos ter transporte de vapor de enxofre do suporte para o interior do forno. Dessa forma se supormos uma taxa finita de evaporação de enxofre, proporcional à quantidade deste na fonte, podemos concluir que as fontes de maior concentração devem manter a pressão de vapor do enxofre mais próxima do equilíbrio. As fontes de menor concentração terão dessa forma menores pressões de vapor, e poderão ser usadas a temperaturas maiores, o que explica a anormalidade observada no gráfico.

4.7 - Caracterização das Amostras

Uma vez feita a difusão é necessário caracterizar a camada obtida. Essa caracterização irá nos dar como resultados [3] o perfil de difusão de impurezas, o perfil de cargas elétricas, que nem sempre é igual ao de impurezas, a resistividade da camada, qual o tipo de portador gerado pela impureza, e qual a mobilidade desses portadores. É claro que todas essas informações não são obtidas com um único método de caracterização, mas com uma combinação de vários métodos. Os resultados comentados acima são fornecidos através de técnicas básicas de caracterização, medidas de efeito Hall, resistividade por quatro pontas, C-V, C-V eletroquímico, SIMS (espectroscopia de massa de ion secundário), medidas de decaimento radioativo e medidas de profundidade de junções PN. Experiências mais sofisticadas irão obter o perfil de mobilidade, ou seja a mobilidade dos portadores em função da profundidade, as energias de ionização, ou até a posição da impureza na rede, substitucional ou intersticial, entre outros dados. Estamos interessados nas técnicas de caracterização mais básicas, razão pela qual elas serão apresentadas e discutidas abaixo.

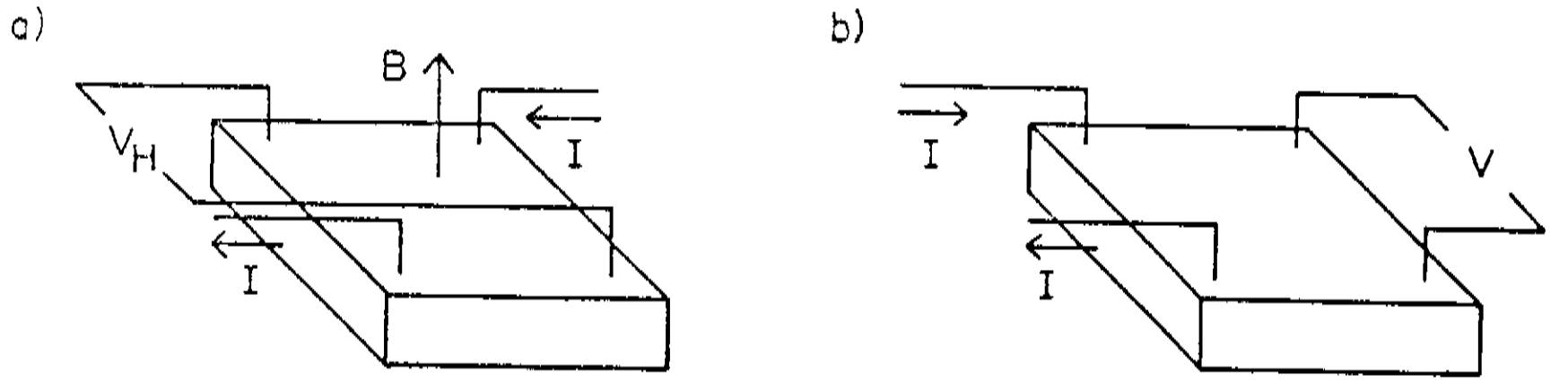

Figura 4.14 - a) Montagem da amostra para medidas de efeito Hall.
b) Montagem para medida de resistividade.

Entre as técnicas básicas de caracterização, as medidas de efeito Hall irão determinar o tipo e a mobilidade dos portadores, a resistividade da camada difundida, e até o perfil de portadores se fizermos as medidas com decapagem periódica da amostra. Da mesma forma a técnica de medida de resistividade por quatro pontas com a decapagem da amostra também determinará o perfil de portadores. Essa técnica é menos completa do que o efeito Hall, já que ela apenas nos fornece a resistividade. As técnicas de C-V e C-V eletroquímico irão exclusivamente determinar o perfil de portadores na amostra, sendo bastante úteis e práticas, já que podemos montar um sistema completo para a medida (existem sistemas comerciais de C-V eletroquímico). Por fim as medidas de decaimento radioativo e SIMS nos fornecem o perfil de impurezas difundidas. Todas a técnicas acima são destrutivas (danificam a camada a ser medida), com exceção do C-V que pode ou não ser destrutiva, dependendo de como é preparada a amostra.

O efeito Hall é uma técnica bastante simples, que consiste na medida da força eletromotriz, V_H, gerada pelo movimento de cargas na camada difundida, movimento esse efetuado na presença de um campo magnético, B, normal à superfície da amostra, como mostra a figura 4.14a. A força eletromotriz é gerada pelo acúmulo de cargas nas laterais da amostra, como mostrado na figura. Esse acúmulo é resultante da força criada pelo campo magnético sobre os

portadores de carga. Logicamente esse acúmulo não é infinito, já que ele cria um campo elétrico transversal que irá anular a força magnética sobre as cargas. Igualando as expressões para a força magnética sobre os portadores e para a força elétrica criada pelo acúmulo lateral de carga obtemos uma expressão para V_H,

$$\frac{V_H}{L} = v_x B \qquad (4.1)$$

onde v_x é a velocidade do portador, e V_H/L o campo elétrico perpendicular ao fluxo de corrente. Observe que V_H será positivo se os portadores forem elétrons, e negativo se lacunas, já que o fluxo de elétrons é na direção do campo F_x (sua velocidade é positiva). O sinal de V_H medido permite, portanto, a determinação do tipo de portador. Lembrando que a velocidade dos portadores é proporcional ao campo, sendo a constante de proporcionalidade a mobilidade do portador, μ, e lembrando a lei de Ohm, temos respectivamente,

$$v_x = \mu F_x \qquad (4.2a)$$

$$J_x = \sigma F_x \qquad (4.2b)$$

onde σ na lei de Ohm é a condutividade elétrica da camada. Podemos facilmente determinar a relação média entre a resistividade e a altura da camada difundida. Isso é feito da forma mostrada na figura 4.14b, onde injetamos corrente em uma das laterais da amostra e medimos a tensão na outra lateral, V_1. A relação entre a resistividade e a espessura da camada, h, será,

$$\frac{\rho}{h} = \frac{V_1}{I} f \qquad (4.3)$$

onde f é um fator de correção que depende da geometria da amostra, como demonstrado na referência 4. A razão ρ/h é normalmente chamada de resistência de folha ou de placa de amostra.

Substituindo σ por $1/\rho$ em (4.2b), e o valor de F_x dessa expressão em (4.2a), achamos,

$$v_x = \mu J_x \rho \qquad (4.4)$$

que substituindo em (4.1) achamos,

$$V_H = \mu J_x \rho B L \qquad (4.5)$$

multiplicando o lado direito de (4.5) por h/h verificamos prontamente o termo J_x L h como I_x, e ρ/h como a resistência de folha (que é um valor conhecido). Obtemos dessa forma,

$$\frac{V_H}{I_x} = \mu B \frac{\rho}{h} \qquad (4.6)$$

Da equação (4.6) podemos determinar o valor de μ, já que ele é o único desconhecido na expressão, enquanto que os outros termos podem ser medidos na experiência. Na equação acima μ é a chamada mobilidade Hall e seu valor só será diferente da mobilidade real para pequenos campos magnéticos aplicados à amostra.

A resistividade da amostra será dada por,

$$\rho = \frac{1}{q \mu N} \qquad (4.7)$$

onde N é a concentração de cargas, e q a carga do elétron. Se multiplicarmos ambos os lados de (4.7) por 1/h achamos,

$$\frac{\rho}{h} = \frac{1}{q \mu N h} \qquad (4.8)$$

Conhecemos da expressão (4.8) todas as variáveis, com excessão de N h, que é a concentração de portadores por unidade de área de superfície do cristal, que pode ser determinada.

Outra técnica que é usada para a medida de resistividade é o sistema de quatro pontas, que é constituido por uma fonte de corrente, um voltímetro e quatro agulhas igualmente espaçadas que fazem contato elétrico com a superfície do substrato. O sistema é montado conforme a figura 4.15. Nessa configuração injetamos corrente na amostra pelas agulhas exteriores e medimos a tensão nas agulhas interiores. Dessa forma uma eventual resistência de contato entre a amostra e as agulhas não interferirá nas medidas.

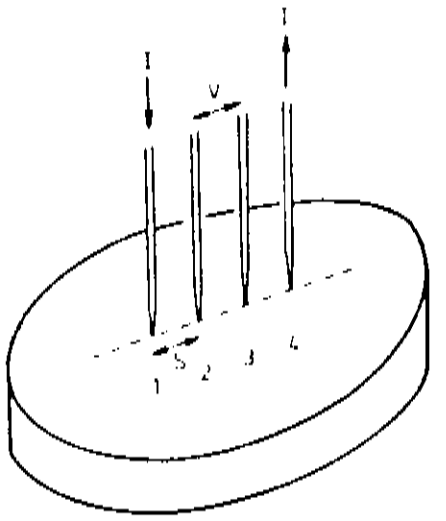

Figura 4.15 - Montagem para medida de resistividade por quatro pontas.

A resistividade será dada por,

$$\rho = f \frac{V}{I} \qquad (4.9)$$

onde f é um fator de correção que depende da geometria da amostra e pode ser achado em tabelas [5].

A técnica de C-V é bastante simples e muito utilizada para determinar o perfil de portadores em uma camada dopada. Ela consiste basicamente na medida de capacitância de uma junção formada pelo semicondutor mais um contato externo do tipo Schottky. Esse contato irá criar uma região de depleção no semicondutor, quando polarizado adequadamente, região esta que é caracterizada pela ausência de cargas livres. Neste caso temos um capacitor, onde o dielétrico é a região de depleção, e as placas são o contato Schottky e a região neutra não depletada do semicondutor. A capacitância por unidade de área, C_a, é dada simplesmente por,

$$C_a = \frac{\varepsilon}{W} \qquad (4.10)$$

onde W é a profundidade da interface entre a região neutra e a de depleção, e ε é a constante dielétrica do semicondutor. De uma medida da capacitância podemos achar a profundidade da interface. Outra relação importante é,

$$\frac{d(1/C^2)}{dV} = \frac{2}{q\ \varepsilon\ N} \qquad (4.11)$$

que nos mostra que a derivada do inverso do quadrado da capacitância em relação à tensão aplicada ao contato Schottky, V, é proporcional ao inverso da concentração de carga, N. Fazendo medidas da capacitância e da derivada acima podemos para um dado valor de V achar a profundidade da interface e a concentração de carga nesse ponto, o que é resulta no perfil de cargas na camada. É importante observar que a capacitância da região de depleção é insensível a variações de dopagem que ocorram dentro de uma distância da interface da ordem de um comprimento de Debye, L_D, definido como,

$$L_D = \left(\frac{\varepsilon\ k\ T}{q^2\ N} \right)^{1/2} \qquad (4.12)$$

onde k é a constante de Boltzmann e T a temperatura. Observe ainda que para uma dopagem da ordem de $10^{18} cm^{-3}$ no GaAs L_D vale apenas 40Å, enquanto que para $10^{15} cm^{-3}$ ele vale 0,2μm. Disso podemos concluir que a precisão das medidas de C-V está limitada a camadas de média e alta concentração de dopantes.

Outra técnica de medida do perfil de cargas é o C-V eletroquímico. Esta técnica elimina um inconveniente que o C-V apresenta, que é a diminuição da capacitância para um aumento na tensão do contato Schottky com o consequente afastamento da interface. Essa diminuição irá limitar a aplicação da técnica de medida, já que no sistema temos capacitâncias parasitas devidas a cabos de ligação, entrada de aparelhos, conectores, etc., que devem ser sempre menores do que a capacitância do contato Schottky. No C-V eletroquímico a amostra sofre um processo de decapagem eletroquímica entre as medidas de capacitância, que irá retirar parte das camadas já medidas, mantendo a interface sempre próxima do contato Schottky, o que garante uma alta relação capacitância de depleção/capacitância parasita.

Entre as técnicas de medida do perfil de impurezas, o SIMS é a mais prática. Essa técnica consiste na análise de produtos de "sputtering" obtidos da amostra. Essa análise é feita por um

espectrômetro de massa que é ajustado para coletar e contar átomos do material que difundiu ou de algum composto entre este e os átomos do substrato. Essa contagem será proporcional à concentração de átomos difundidos. Para que o espectrômetro de massa funcione adequadamente, ele deve ser operado em um ambiente com alto vácuo. Da mesma forma a amostra e o sistema de decapagem por "sputtering" devem ser colocados em alto vácuo, para que gases residuais não interfiram nas reações de formação de produtos da amostra, nem na contagem de átomos pelo espectrômetro. A determinação do perfil de difusão é feita decapando-se continuamente a amostra e fazendo-se contagens de átomos a intervalos de tempo periódicos. Obtemos do SIMS um gráfico do número de contagens em função do tempo de "sputtering". O perfil real é obtido fazendo a mesma análise em um amostra padrão, onde conhecemos a concentração de dopantes e a profundidade da camada, nas mesmas condições da análise feita na amostra desconhecida. Dessa forma conseguimos relacionar o número de contagens à concentração e o tempo de "sputtering" à profundidade.

Embora o SIMS seja bastante prático e de certa forma uma técnica até simples, ele é um equipamento de custo muito elevado (da ordem de meio milhão de dólares), o que limita a sua aquisição a grandes centros de pesquisa, e torna o acesso a esse equipamento bastante limitado.

Temos por fim a técnica de difusão de material radioativo para determinar o perfil de impurezas. Como o próprio nome já diz essa técnica consiste na difusão de algum isótopo radioativo do material que desejamos difundir, isótopo este que sofre decaimento radioativo emitindo partículas beta ou gama. A determinação do perfil é feita decapando-se a amostra e fazendo contagens de emissão de partículas entre as decapagens. Embora essa técnica não necessite um equipamento caro como SIMS (aqui basta um detetor/contador de partículas), ela não é muito prática, já que precisamos usar materiais radioativos, o que requer muito cuidado na sua manipulação.

4.8 - Resultados Obtidos

Dos métodos de caracterização descritos anteriormente usamos o efeito Hall, o C-V eletroquímico, e uma amostra foi analisada por SIMS. Não usamos medidas de quatro pontas, já que a caracterização por efeito Hall nos fornece a resistência de folha da amostra e a mobilidade média, o que a torna mais interessante e útil. Medidas de perfil por enxofre radioativo não foram realizadas, pois não dispúnhamos do material, e mesmo que dispuséssemos, o uso de material radioativo em tubo aberto é desaconselhável, pois a perda de material na fase vapor è muito grande, o que provoca uma grande contaminação do ambiente da experiência.

Na tabela 4.4 podemos ver os resultados das medidas de efeito hall para as amostras difundidas com a fonte de 1% de Ga_2S_3. As medidas foram feitas com um campo magnético de 0,5T e correntes de 100μA aplicadas às amostras. Na tabela temos o número da amostra, a temperatura e tempo de difusão, o wafer usado, o número da fonte e a quantidade de vezes que ela foi usada (por exemplo, para a amostra DR18 usou-se pela oitava vez a fonte número um), a mobilidade, concentração de portadores por unidade de área e a resistência de folha da amostra. Observe que experimentamos três fornecedores de wafers, e que todos eles apresentaram bons resultados, o que vemos pelas mobilidades das amostras. Para a Sumitomo temos a amostra DR19 com um valor de mobilidade de 2050cm^2/Vs, para a Cominco temos DR15 com 2200cm^2/Vs, e para a Wacker DR27 com 2030cm^2/Vs. Obtivemos também baixos valores de mobilidade, para as amostras DR11, DR14 e DR22, porém especificamente nessas três amostras observamos a formação de sulfetos nas suas superfícies, o que indica uma alta concentração de enxofre na forma de vapor durante a difusão, que pode ter atacado a amostra provocando defeitos na sua superfície, ou mesmo a precipitação de enxofre em camadas abaixo da superfície. Esses dois fenômenos provocam a diminuição da mobilidade, já que criam defeitos no cristal. Se desconsiderarmos essas amostras vemos que existe uma correlação (não muito forte) entre a temperatura de difusão e a mobilidade. Em geral as

Amostra #	Temper. °C	Tempo s	Wafer	Fonte #	Fonte vez	μ cm^2/Vs	Concen. $x10^{12} cm^{-2}$	Res. Ω/□
DR11	890	30	S	1	1	900	5,44	1280
DR14	890	30	S	1	4	490	15,80	810
DR15	870	30	C	1	5	2200	0,65	4350
DR16	900	30	C	1	6	1700	2,16	1700
DR18	900	15	S	1	8	1450	1,11	3900
DR19	870	60	S	1	9	2050	3,26	930
DR20	890	30	C	1	10	1530	3,93	1040
DR21	890	60	C	1	11	2020	4,79	650
DR22	900	30	S	1	12	740	6,98	1220
DR25	900	30	W	2	1	1700	12,10	300
DR26	900	30	W	2	2	1870	5,93	560
DR27	900	30	W	2	3	2030	3,80	810
DR28	900	30	W	2	4	1790	7,74	450

TABELA 4.4 - Resultados das medidas de efeito hall para amostras difundidas com a fonte de 1% de Ga_2S_3. Na tabela C - Cominco, S - Sumitomo e W - Wacker.

amostras de mais baixa temperatura apresentam mobilidades maiores. Temos dessa forma para 870°C valores de 2050 a 2200cm^2/Vs, a 890°C de 1530 a 2020cm^2/Vs e a 900°C de 1450 a 2030cm^2/Vs. Com relação ao tempo de difusão não podemos concluir se há alguma correlação, pois a maioria das difusões foi feita a 30 segundos, só duas foram feitas a 60 e uma a 15 segundos. Com relação à concentração de portadores por unidade de área devemos tomar um cuidado de só comparar resultados de difusões realizadas próximas na sequência de uso da fonte, pois podemos ter um efeito de perda de enxofre por ela, com a consequente saída da composição da região A do diagrama de fases, o que provoca uma menor repetibilidade dos

Amostra #	Temper. °C	Tempo s	Wafer	Fonte #	Fonte vez	μ cm^2/Vs	Concen. $x10^{12}cm^{-2}$	Res. Ω/□
DR29	875	30	W	3	1	2600	6,03	400
DR30	870	30	W	3	2	2170	12,20	240
DR31	890	30	W	3	3	1710	16,70	220
DR32	870	30	W	3	4	1560	5,03	800

TABELA 4.5 - Resultados das medidas de efeito hall para amostras difundidas com a fonte de 12,5% de Ga_2S_3. Na tabela W - Wacker.

resultados. Observe que a concentração aumenta com a temperatura como podemos ver das amostras DR15 e DR16, DR19 e DR21, e DR20 e DR22. A concentração também aumenta com o tempo de difusão, o que concluimos das amostras DR15 e DR19, DR16 e DR18, e DR20 e DR21. A resistência de folha é inversamente proporcional à mobilidade e à concentração, e portanto a sua dependência com a temperatura e tempo deve ser mais complexa.

Na tabela 4.5 podemos ver os resultados de medidas hall em amostras difundidas com a fonte de 12,5%. Nesse caso só usamos wafers da Wacker. Não podemos observar nenhuma correlação da mobilidade e da concentração de portadores com a temperatura e tempo de difusão devido ao pequeno número de dados. Não realizamos muitas experiências com a fonte de 12,5%, pois percebemos que era possivel aumentar ainda mais a concentração de enxofre na fonte, o que tornaria o processo mais repetitivel, já que a fonte perderia menos enxofre, relativamente à concentração total deste, entre as difusões. Observando a tabela vemos que a mobilidade é da mesma ordem que a obtida com a fonte de 1%, porém as concentrações de portadores são maiores, entre $5x10^{12}cm^{-2}$ e $1,7x10^{13}cm^{-2}$, e as resistências de folha menores, entre 220Ω/□ e 800Ω/□, enquanto que para a fonte de 1% ela sempre foi maior que 300Ω/□, atingindo até 4kΩ/□.

Temos por fim as medidas de efeito hall das amostras di-

Amostra #	Temper. °C	Tempo s	Wafer	Fonte #	Fonte vez	μ cm^2/Vs	Concen. $x10^{12} cm^{-2}$	Res. Ω/□
DR35	870	30	W	5	1	1800	10,60	320
DR36	870	30	W	5	2	1790	11,60	300
DR37	870	30	W	5	3	2000	15,80	200
DR38	830	90	W	5	4	1770	5,85	600
DR39	860	90	W	5	5	1690	14,70	250
DR40	850	90	W	5	6	1670	3,77	990
DR41	830	90	W	5	7	2250	2,32	1190
DR42	860	90	W	5	8	1730	11,30	320
DR43	850	60	W	5	9	1450	7,78	550
DR46	860	90	W	5	12	1890	22,00	150
DR47	860	90	W	5	13	1600	16,70	240
DR48	860	90	W	5	14	1800	19,30	180
DR49	860	90	W	5	15	1300	3,63	-
DR50	860	90	W	5	16	770	2,62	-
DR52	860	90	W	5	18	1000	7,60	-
DR55a	860	90	W	6	3	1960	12,30	260
DR55b	860	90	W	6	3	2100	8,10	-
X1	860	90	W	6	-	2100	9,67	300

TABELA 4.6 - Resultados das medidas de efeito hall para amostras difundidas com a fonte de 50% de Ga_2S_3. Na tabela W - Wacker.

fundidas com a fonte de 50% de Ga_2S_3. Na tabela 4.6 podemos ver os resultados. Novamente só usamos wafers da Wacker. Repare que o uso dessa fonte resulta em amostras com alta concentração (para os tempos de difusão usados), ao redor de $1x10^{13} cm^{-2}$. Apenas quatro

amostras apresentam concentrações entre 2,2 e $3,8x10^{12}cm^{-2}$. Observe que com essa fonte usamos tempos longos de difusão, de 90 segundos. Isso foi necessário devido á baixa temperatura de difusão, que foi compensada por um tempo maior. Aparentemente a mobilidade não possui nenhuma correlação com a temperatura, o que não foi observado para a fonte de 1%. Para a fonte de 50% obtivemos mobilidades entre $1450cm^2/Vs$ e $2250cm^2/Vs$. Estamos desconsiderando as amostras DR49, DR50 e DR52, que apresentaram mobilidades abaixo da média. Para essas três amostras a difusão foi feita após a fonte ter ficado muito tempo guardada. Isso pode ter provocado a oxidação de parte da fonte, e a consequente difusão de oxigênio nas amostras. Este pode formar precipitados na forma de aglomerados de oxigênio ("clusters") que seriam responsáveis pela baixa mobilidade. Temos fortes razões para acreditar nisso, como veremos dos resultados da análise por SIMS. Esses resultados nos mostraram que era necessário um tratamento térmico na fonte antes de seu uso na difusão (para eliminar óxidos e umidade da fonte). Até antes da amostra DR52 as difusões eram realizadas sem esse tratamento. Isso não prejudicou as experiências anteriores, que foram sempre feitas com intervalos de tempo pequenos entre si, e sempre guardávamos a fonte em uma atmosfera seca e inerte. Observe na tabela 4.6 que a fonte teve um uso maior, 18 vezes contra 12 vezes para a fonte de 1%. Com relação á repetibilidade observamos um efeito de queda na concentração de portadores. Isso para as sequências de amostras DR37, DR38 e DR39; e DR40, DR41 e DR42. Cada sequência de amostras foi feita a 850°C, 830°C e 860°C. Observe que de uma sequência para a outra a concentração de portadores cai. Isso pode ser devido a uma queda na pressão de vapor de enxofre, resultante da cristalização do Ga_2S_3 ou à formação de GaS. Neste caso estamos supondo que o Ga_2S_3 decompõem em GaS, Ga e S_2 (esta seria uma reação possível), e que a decomposição inicial é maior, causando altas pressões de enxofre. Embora haja esse efeito de queda na concentração, podemos observar que as difusões a 860°C são bastante reprodutíveis, como podemos ver das amostras DR39, DR42, DR46, DR47, DR48, DR55a, DR55b e X1, onde a mobilidade variou de $1600cm^2/Vs$ a $2100cm^2/Vs$, e a concentração de $8,1x10^{12}cm^{-2}$ a

$2,2x10^{13}cm^{-2}$. Com essa fonte obtivemos baixos valores de resistência de folha, de 150Ω/□ até 1190Ω/□.

Em quase todas as amostras foram feitas medidas de perfil de carga com o Polaron. Nas figuras 4.16 a 4.19 podemos ver os resultados típicos. Na figura 4.16 está o perfil da amostra DR11, típico para a fonte de 1% de Ga_2S_3. Observe que o perfil é aparentemente do tipo função erro complementar, e que a concentração de superfície é da ordem de 2 a $3x10^{18}cm^{-3}$, o que corresponde a uma camada N^+. A longa cauda do perfil é atribuída a um erro sistemático de medida do Polaron. Os pontos dessa cauda representam medidas de baixa concentração de dopantes, o que equivale a uma largura da região de depleção muito grande no semicondutor, e uma capacitância baixa. Esses pontos correspondem ao início da região não difundida do substrato. Desconsiderando a cauda vemos que a profundidade da camada está entre $0,2\mu m$ e $0,25\mu m$, o que é bastante raso. Já na figura 4.17 vemos outro perfil obtido com a fonte de 1%, só que neste caso a amostra após a difusão apresentou uma alta resistência de folha. Como era de se esperar temos uma camada bastante fina, de menos de $0,15\mu m$. Observe que o Polaron teve dificuldades pararealizar a medida, como podemos ver no gráfico, onde temos apenas três pontos, entre $10^{16}cm^{-3}$ e $10^{18}cm^{-3}$. O quarto ponto caiu sobre o substrato não difundido, o que causou a longa cauda no perfil. Na figura 4.18 está o perfil obtido na amostra DR32 com a fonte de 12,5%. Observe que a concentraçãao de superfície deve estar entre $1x10^{18}cm^{-3}$ e $2x10^{18}cm^{-3}$. Neste caso temos uma camada de profundidade entre $0,3\mu m$ e $0,35\mu m$. Por fim na figura 4.19 está o perfil para uma amostra difundida com a fonte de 50%. Temos uma concentração de superfície na mesma ordem da obtida na amostra com a fonte de 12,5%, porém agora a profundidade da camada é cerca de $0,5\mu m$, o que resulta na menor resistência de folha para as amostras com a fonte de 50%. Como temos temperaturas de difusão aproximadamente iguais para as fontes de 12,5% e 50%, e já que a mudança da composição da fonte, no nosso caso, não deve mudar as pressões de vapor de enxofre e arsênio, podemos comparar os perfis DR32 e DR46. As concentrações de superfície devem ser iguais, como observado, mas a profundidade da camada deve ser maior para DR46,

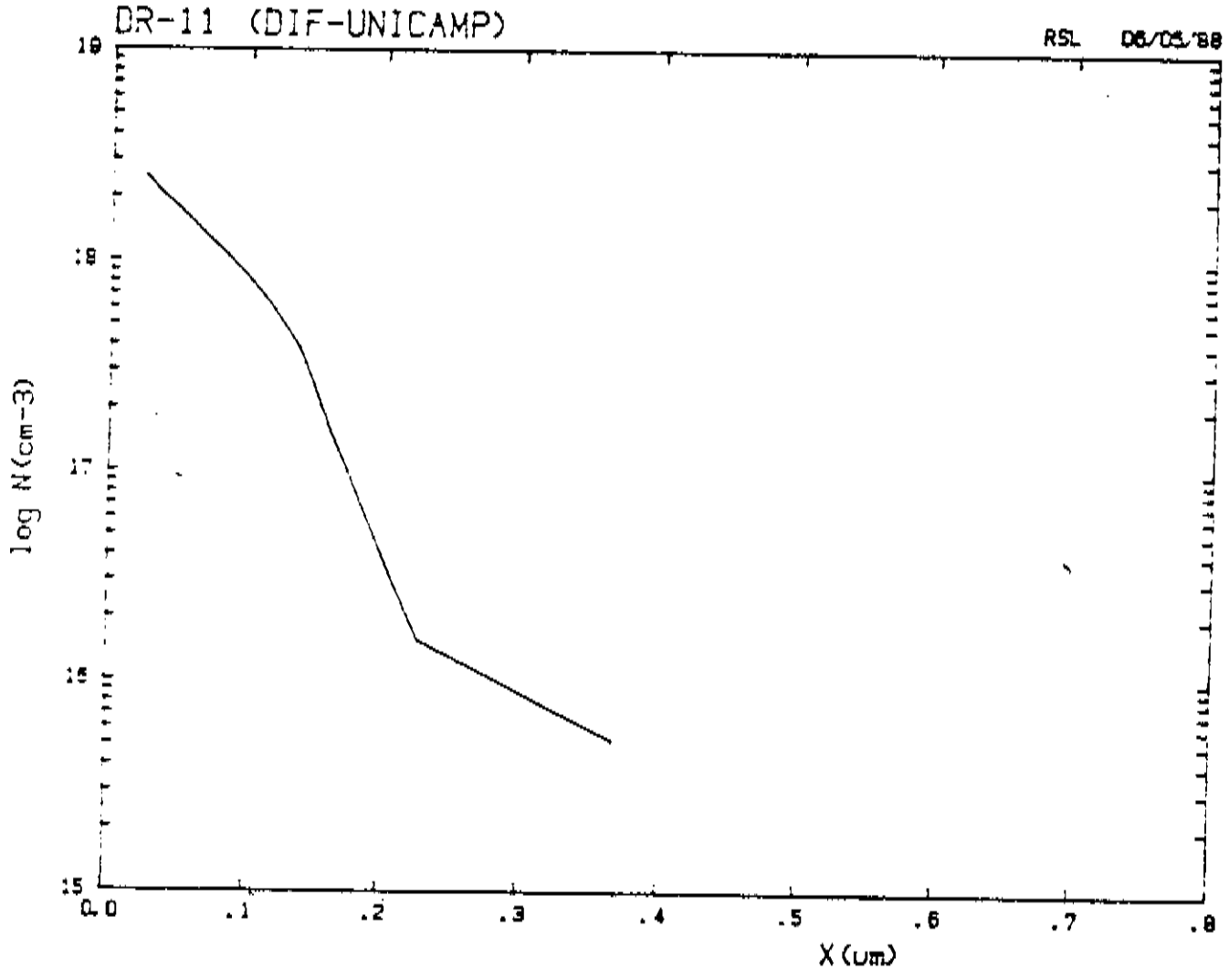

Figura 4.16 - Perfil de carga obtido pelo Polaron para a amostra DR11, difundida com a fonte de 1%, a 890°C-30s.

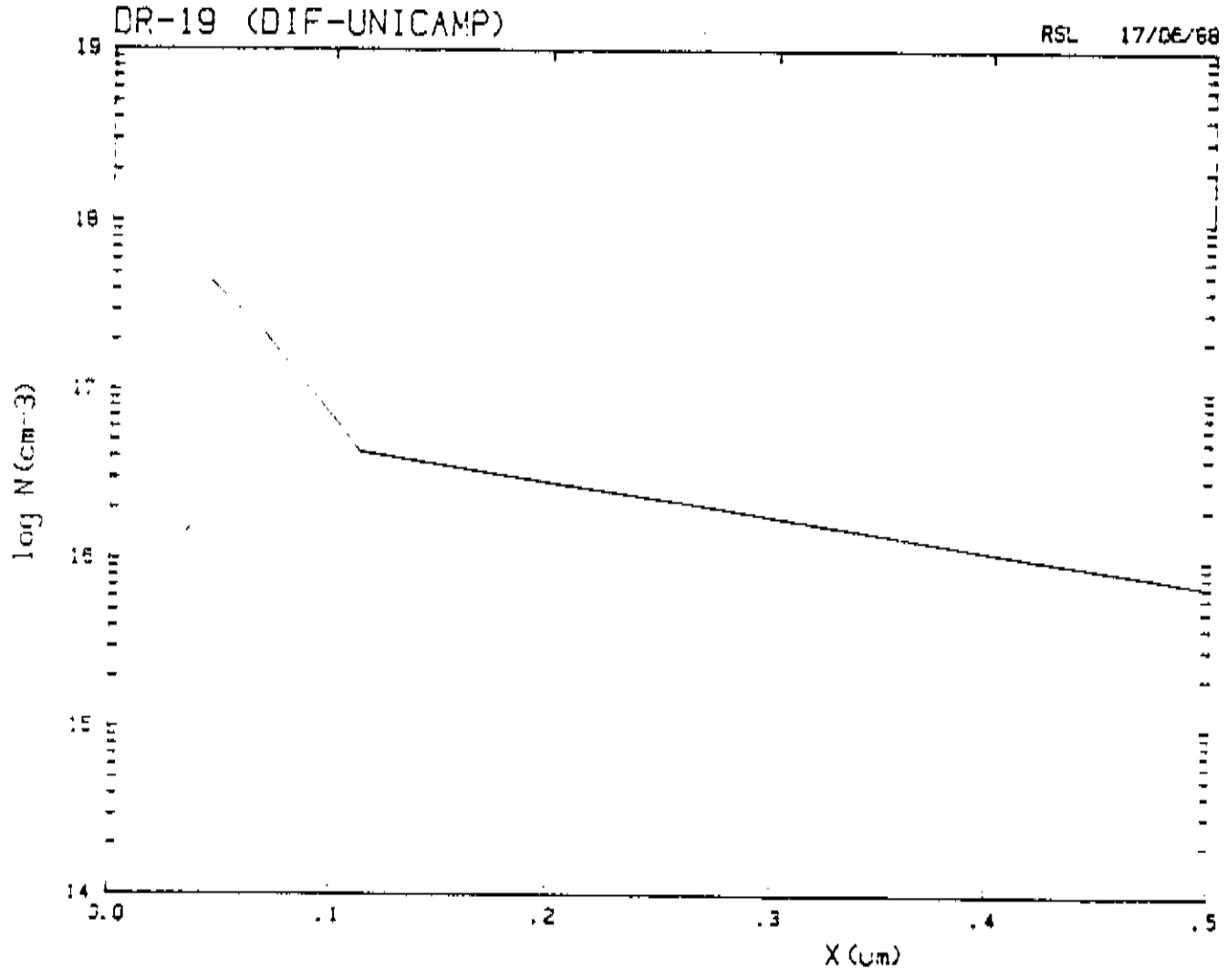

Figura 4.17 - Perfil de carga obtido pelo Polaron para a amostra DR19, difundida com a fonte de 1%, a 870°C-60s.

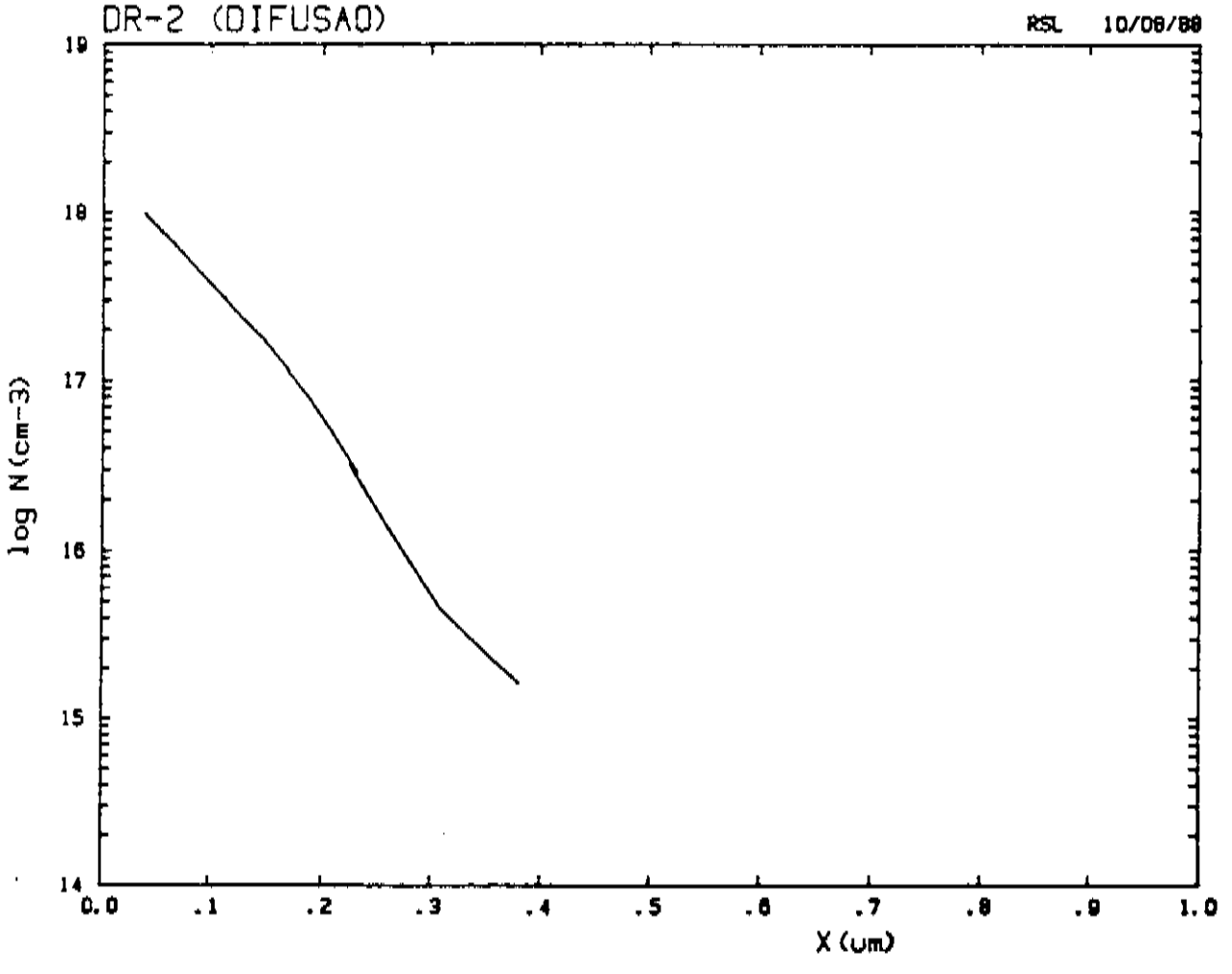

Figura 4.18 - Perfil de carga obtido pelo Polaron para a amostra DR32, difundida com a fonte de 12,5%, a 870°C-30s.

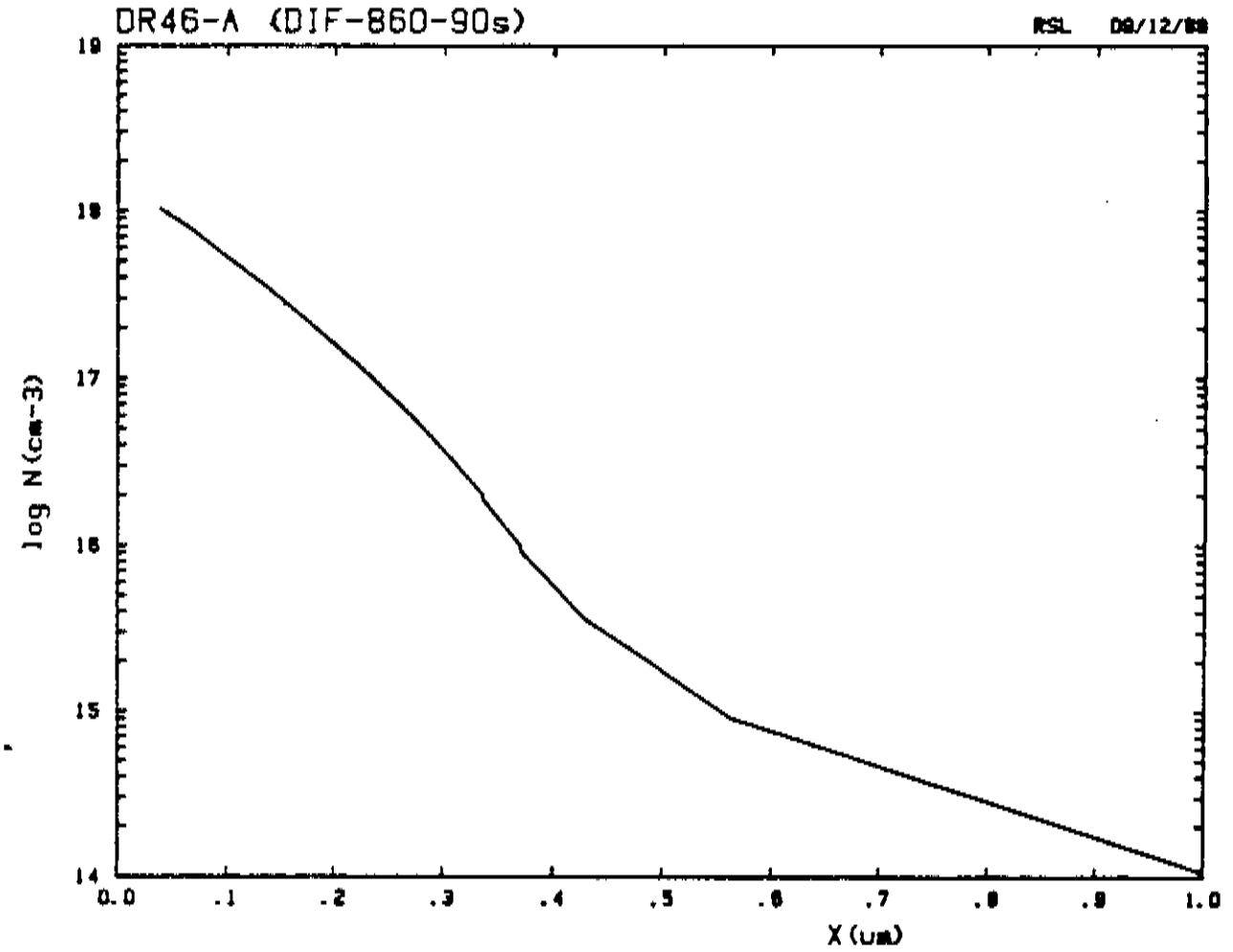

Figura 4.19 - Perfil de carga obtido pelo Polaron para a amostra DR46, difundida com a fonte de 50%, a 860°C-90s.

devido ao maior tempo de difusão, o que também foi observado. Em resumo as análises do Polaron sempre nos deram camadas de profundidade variável entre 0,1μm e 0,5μm, com concentrações de superfície da ordem de $10^{18}cm^{-3}$.

As últimas análises feitas foram as de SIMS na amostra DR52. O objetivo dessas análises era qualificar três equipamentos de SIMS. Essa qualificação foi feita por um dos professores do nosso grupo, que viajou para a Europa para realizar esse trabalho. A qualificação consistia em observar o perfil de enxofre e verificar se havia alguma interferência da molécula de O_2, que tem a mesma massa que o enxofre, sobre esse perfil. Em nenhum dos equipamentos observou-se tal interferência. A análise que temos são apenas gráficos do número de contagens por tempo de "sputtering", o que nos dá pouca informação sobre o perfil real. Na figura 4.20 vemos várias curvas, para o enxofre, gálio, GaO, AsS e GaS. Observe em todas as curvas um ruído de fundo no número de contagens, o que é uma característica dos perfis de SIMS. A análise do Gálio é realizada para que se tenha uma referência, já que a sua concentração no cristal é constante. Escolhemos dessas curvas duas, as nas quais temos uma maior faixa dinâmica do número de contagens (relação entre a máxima contagem e a de fundo). Os elementos escolhidos foram o enxofre com um potencial de 40V aplicado a uma grade de filtro do espectrômetro (S 32 40), e o GaS. Na figura 4.21 temos o perfil do enxofre separado dos demais, sem e com a subtração do ruído de fundo. Essa subtração do ruído aumenta a faixa dinâmica do sinal, o que é interessante. Na figura 4.22 temos os mesmos gráficos, só que para o GaS. Observe que o perfil de enxofre que obteríamos desses gráficos não é do tipo função erro complementar. Observe na figura 4.20 que existe uma difusão de oxigênio da superfície do cristal para o seu interior. É importante lembrar que essa amostra apresentou uma baixa mobilidade. Na figura 4.23 temos o número de contagens de enxofre obtido pelo SIMS da Cameca. Esse equipamento possui um filtro de massa no espectrômetro com grande resolução. Repare nesse gráfico que o perfil não é do tipo erro complementar, e podemos observar duas frentes de difusão, uma com alta concentração e coeficiente de difusão baixo, e outra (na região mais profunda da camada

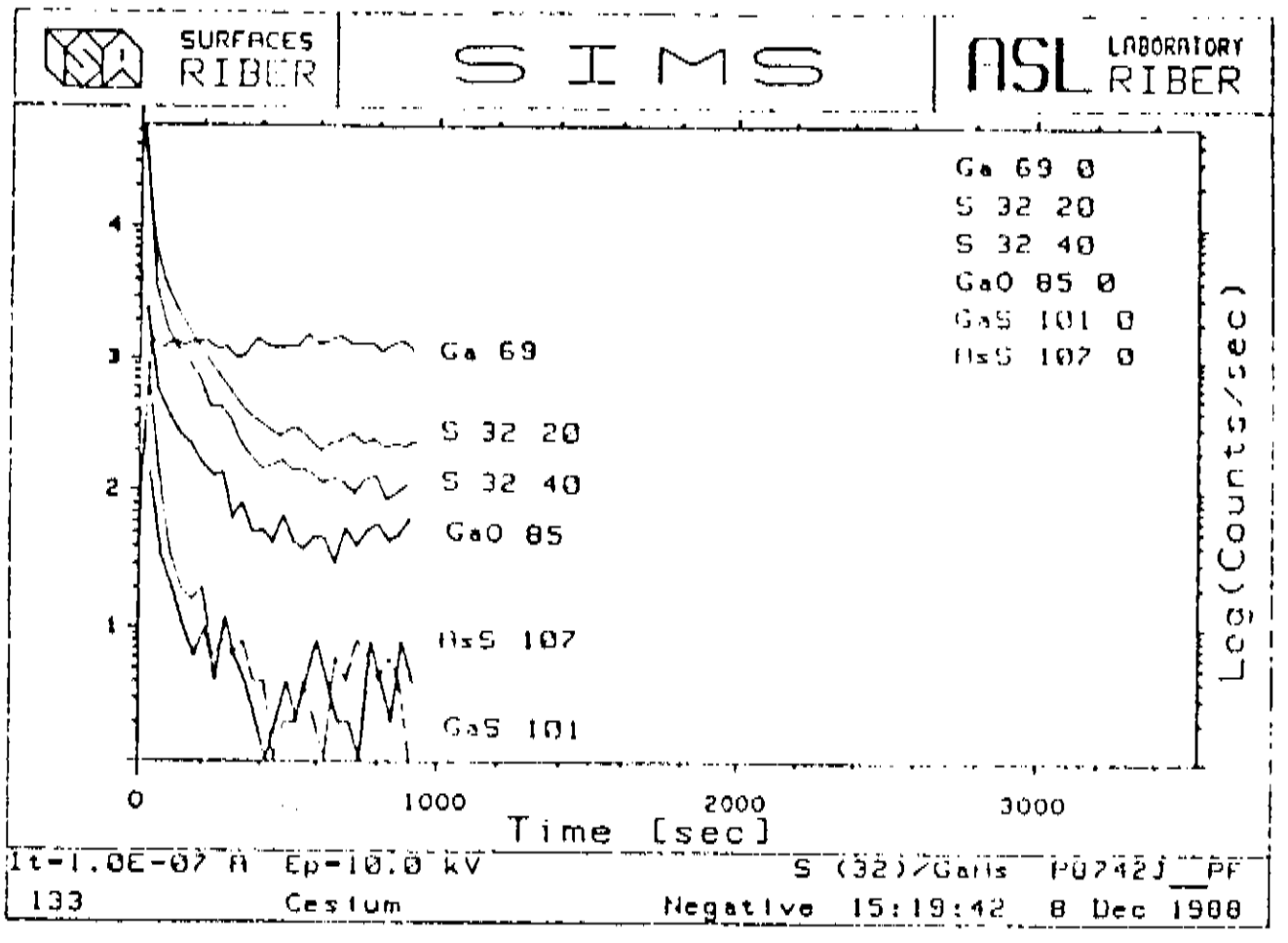

Figura 4.20 - Número de contagens de vários elementos e compostos em função do tempo de "sputtering" obtido por análise de SIMS na amostra DR52. O equipamento usado é de fabricação da Riber.

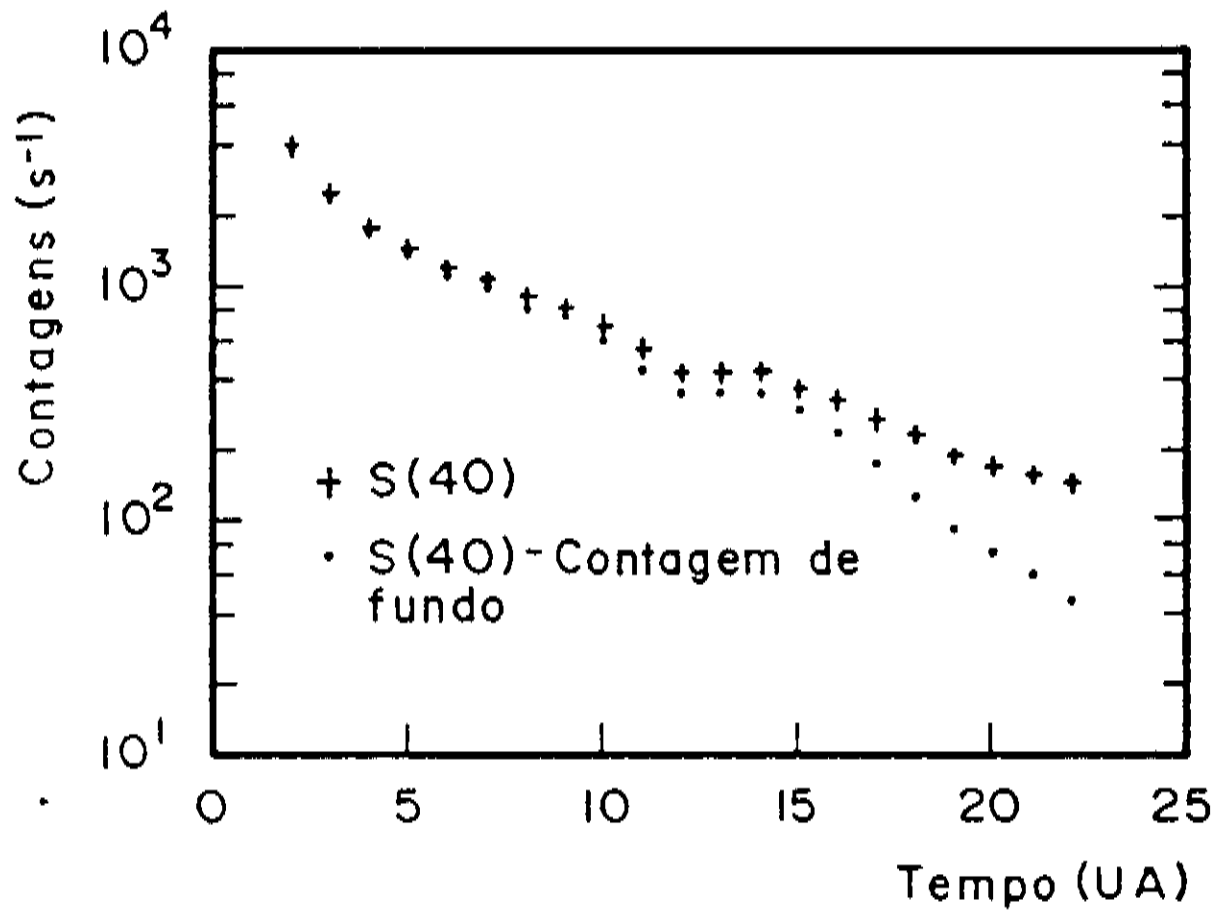

Figura 4.21 - Perfil de S (com potencial de 40V) separado do gráfico da figura 4.20. Temos no mesmo gráfico o perfil subtraído do ruído de fundo.

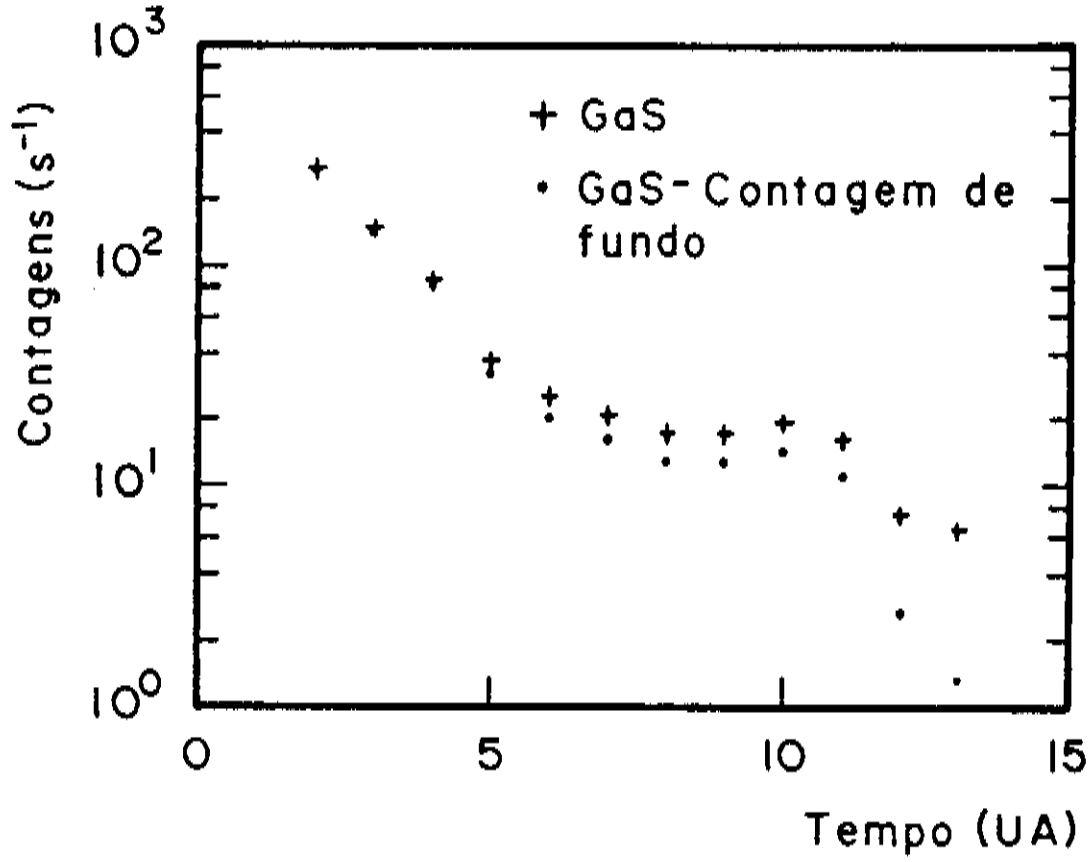

Figura 4.22 - Perfil de GaS separado do gráfico da figura 4.20. Temos no mesmo gráfico o perfil subtraído do ruído de fundo.

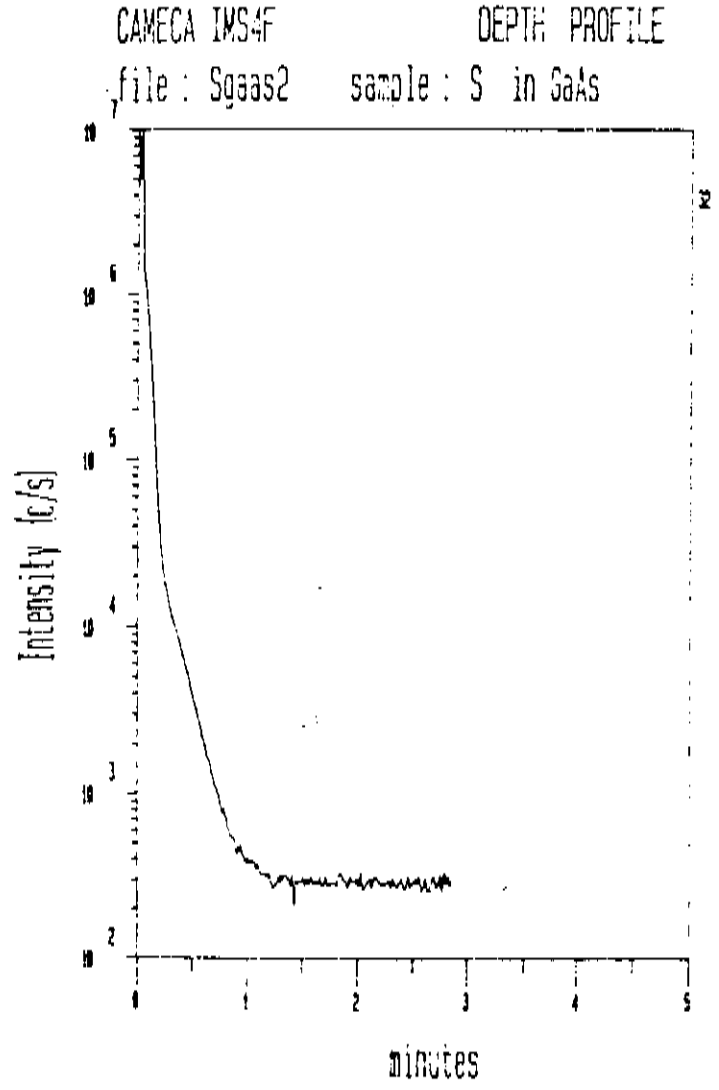

Figura 4.23 - Número de contagens de enxofre em função do tempo de "sputtering" obtido por análise de SIMS. O equipamento usado é de fabricação da Cameca.

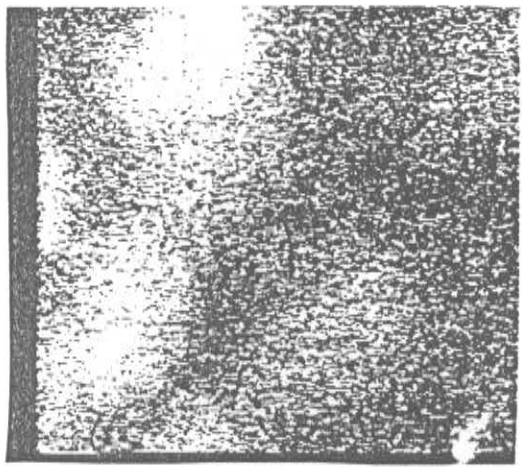

Figura 4.24 - Fotografia de aglomerados de oxigênio observados por SIMS. O equipamento usado é de fabricação da Perkin Elmer. Na foto a região analisada tem 300μm de largura.

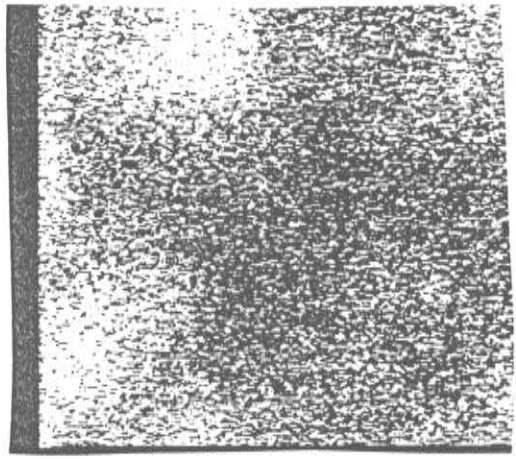

Figura 4.25 - Fotografia de aglomerados de enxofre observados por SIMS. O equipamento usado é de fabricação da Perkin Elmer. Na foto a região analisada tem 300μm de largura.

difundida) com baixa concentração e alto coeficiente de difusão. Temos por fim as fotografias das figuras 4.24 e 4.25. Elas foram tiradas por SIMS e indicam quais são as áreas dos cristais que estão produzindo as espécies químicas analisadas. Dessa forma na figura as regiões brancas são áreas de onde obtemos uma alta contagem de espécies, enquanto que nas regiões escuras a contagem é baixa. Observe na figura 4.24, obtida da análise de oxigênio, que existem aglomerados deste no cristal. Da mesma forma na figura 4.25 observamos aglomerados de enxofre. Essas duas fotografias foram feitas no equipamento de fabricação da Perkin Elmer.

4.9 - Análise dos Resultados

Podemos obter a partir do perfil de portadores o coeficiente de difusão, supondo que cada átomo de enxofre doe um elétron à rede. Isso não é verdade, mas como uma primeira aproximação obter o coeficiente dessa forma já é bastante importante, pois nos dá idéia da sua ordem de grandeza. Para o cálculo do coeficiente desenvolvemos o programa COEF_DIF que o calcula passando uma função erro complementar por dois pontos dados do perfil. A função erro complementar e sua inversa são calculadas numericamente, o que torna o programa bastante lento. Usamos o programa na análise do perfil de DR46, e os resultados são vistos na tabela 4.7. Escolhemos os pontos fora da região de cauda do perfil. Um ponto foi fixado em $2{,}1\times10^{16}cm^{-3}$ e o outro variado de $9{,}4\times10^{16}cm^{-3}$ até $8{,}1\times10^{17}cm^{-3}$. Observamos um aumento do coeficiente de difusão de cerca de 10% quando mudamos um ponto de $9{,}4\times10^{16}cm^{-3}$ para $3{,}1\times10^{17}cm^{-3}$. Acima dessa concentração o coeficiente é praticamente constante. A variação correspondente na concentração de superfície (C_o) é uma queda seguida de um patamar constante. Esse resultado nos sugeriu que estávamos tendo ligeiras variações no coeficiente de difusão. Isso nos levou a aplicar a técnica de Boltzman-Matano [3] para obter o coeficiente de difusão em função da concentração de cargas.

Na técnica de Boltzman-Matano partimos da segunda lei de

Ponto #1			
Concentração $\times 10^{16} cm^{-3}$	Profundidade μm	Conc. de Sup. $\times 10^{18} cm^{-3}$	Coef. de Dif. $\times 10^{-13} cm^2/s$
9,4	0,233	1,647	8,33
31,0	0,145	1,198	9,16
53,0	0,100	1,219	9,11
81,0	0,060	1,271	8,99

TABELA 4.7 - Coeficientes de difusão e concentrações de superfície calculados a partir de dois pontos do perfil. Na tabela vemos os vários valores do primeiro ponto. O segundo é fixado em $2,1 \times 10^{16} cm^{-3}$, com uma profundidade de $0,305 \mu m$.

Fick,

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x}\left(D(C) \frac{\partial C}{\partial x} \right) \qquad (4.13)$$

onde fazemos a seguinte substituição,

$$\eta = \frac{x}{t^{1/2}} \qquad (4.14)$$

obtendo,

$$\frac{\partial C}{\partial t} = - \frac{x}{2\, t^{3/2}} \frac{\partial C}{\partial \eta} \qquad (4.15a)$$

$$\frac{\partial C}{\partial x} = \frac{1}{t^{1/2}} \frac{\partial C}{\partial \eta} \qquad (4.15b)$$

que substituindo em 4.13 resulta em,

$$- \frac{1}{2\, \eta} \frac{\partial C}{\partial \eta} = \frac{\partial}{\partial \eta}\left(D \frac{\partial C}{\partial \eta} \right) \qquad (4.16)$$

a equação (4.16) pode ser integrada em C de 0 a C_1, e em dη, resultando em,

$$D(C_1) = -\frac{1}{2t}\left.\frac{dx}{dC}\right|_{C=C1}\int_0^{C_1} x\,dC \qquad (4.17)$$

Obtemos um expressão que nos dá o coeficiente de difusão em função da concentração desde que seja conhecido o perfil de difusão, para que possamos calcular a integral e a derivada em (4.17). Para a realização desses calculos desenvolvemos o programa ANALISBM que a partir do perfil nos dá o coeficiente calculado de duas maneiras diferentes, usando os valores dos pontos da curva (resulta nos valores Dponto no programa, veja o apêndice), ou uma exponencial interpolada por 4 pontos consecutivos da curva, onde nesse caso obtemos uma variação mais suave do coeficiente com a concentração. Na figura 4.26 podemos ver um resultado típico. Nessa curva está o coeficiente de difusão em função da concentração para a amostra DR46. Observe a existência de três regiões diferentes na curva. A

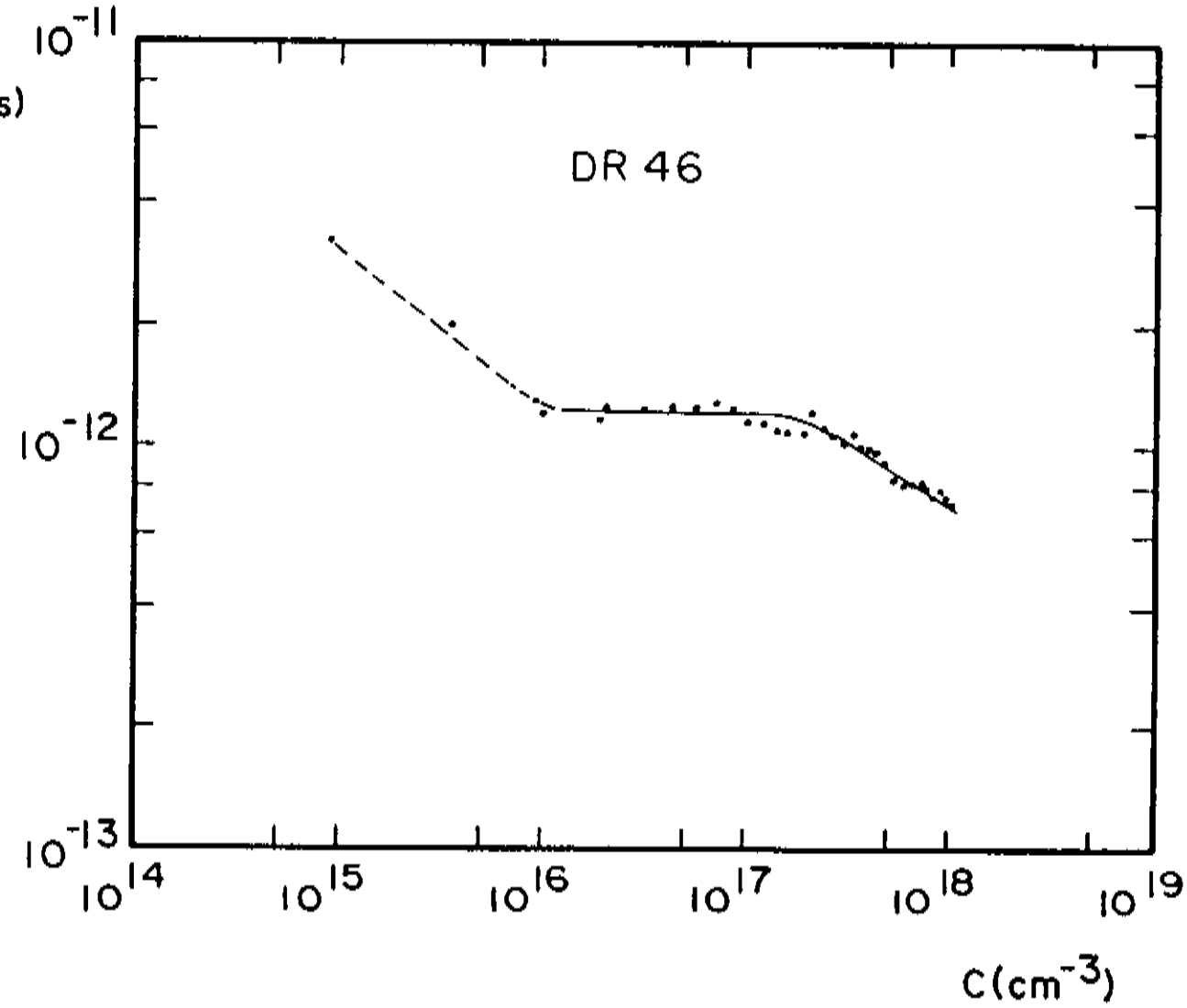

Figura 4.26 - Coeficiente de difusão em função da concentração obtido pela técnica de análise de Boltzman-Matano aplicada na amostra DR46.

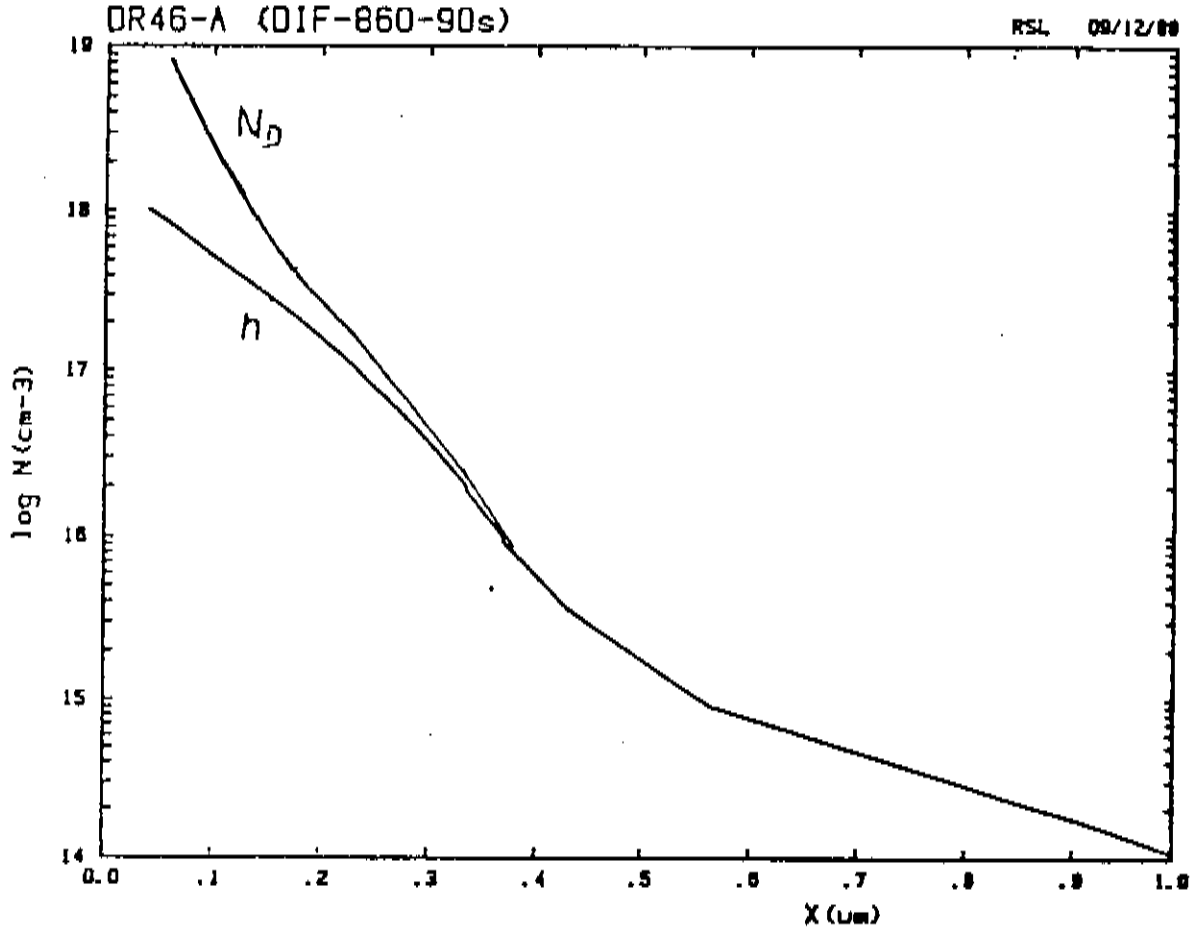

Figura 4.27 - Perfis de carga e de enxofre. Este último é obtido a partir do de carga (veja o texto).

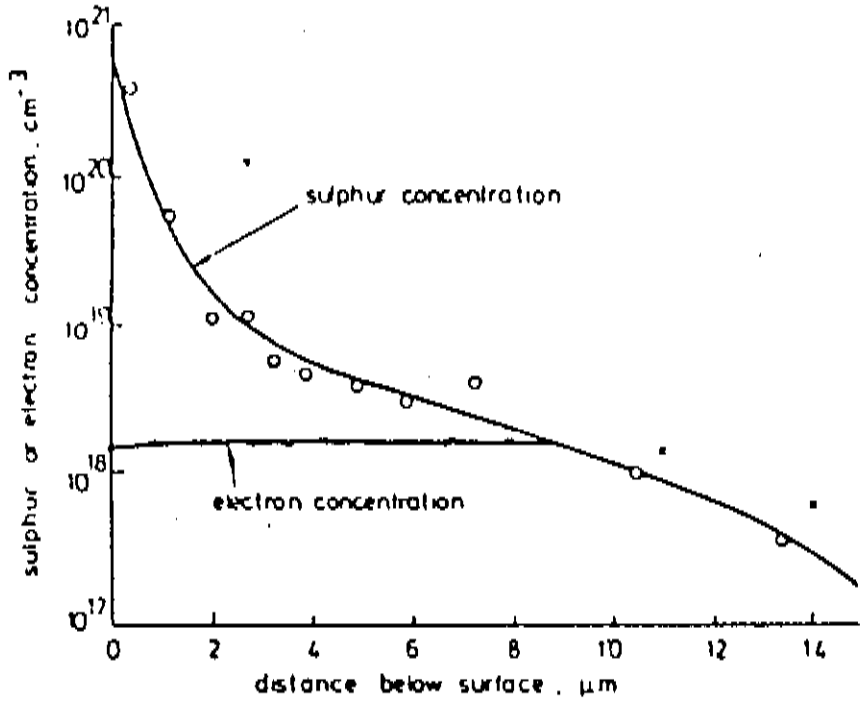

Figura 4.28 - Perfis de carga e enxofre obtidos por Young e Pearson.

primeira e a de baixa concentração onde o coeficiente tem uma rápida queda. Essa região é onde o Polaron apresenta uma extensa região de depleção, produzindo uma grande incerteza no perfil.

Devido a isso estamos considerando que há uma grande incerteza no coeficiente. A região seguinte, para médias concentrações, apresenta um coeficiente constante, como podemos ver no grafico. Essa região segue a função erro complementar. A partir de aproximadamente 3 a $5x10^{17}cm^{-3}$ temos o inicio da terceira região, onde o coeficiente cai cerca de meia ordem de grandeza do seu valor no patamar. Isso mostra claramente que o perfil não segue a função erro complementar. A explicação pode ser achada no trabalho de Young e Pearson [6]. Para altas concentrações de enxofre, maiores do que cerca de $1x10^{18}cm^{-3}$, eles acharam que não é todo o enxofre que doa um elétron ao cristal. Podemos atribuir o início da queda do coeficiente a esse efeito. Teríamos dessa forma para altas concentrações uma porcentagem de enxofre eletricamente ativo menor do que 100%, sendo que conforme aumentamos a sua concentração essa porcentagem cai ainda mais. Isso foi mostrado no capitulo 3. Isso faz com que o coeficiente medido a partir do perfil de cargas diminua, já que ele é proporcional à área sob o perfil limitada pela curva $C=C_1$ (essa área é dada pela integral na equação (4.17)).

Da discussão acima podemos concluir que o coeficiente de difusão obtido a partir do Polaron é igual ao de enxofre para concentrações de elétrons entre $10^{16}cm^{-3}$ e $10^{17}cm^{-3}$.

Supondo que a diminuição da ionização de impurezas para altas concentrações é o mecanismo responsável pela diminuição de carga (o outro mecanismo é a compensação gerada pela associação $V_{Ga}S_{As}$ como proposta por Venkatasubramanian [7], que possui um comportamento semelhante ao da diminuição da ionização), podemos calcular o perfil de enxofre a partir do perfil elétrico. Para isso usamos os resultados do capitulo 3, onde temos um gráfico, figura 3.28, que nos dá a fração de impurezas ionizadas em função da sua concentração. Na figura 4.27 podemos comparar ambos os perfis. Comparando-os com os obtidos por Young e Pearson, figura 4.28, vemos que a semelhança é muito grande. Comparando o perfil de enxofre calculado ao medido por SIMS, figura 4.23, vemos que a semelhança é grande também. Temos por fim a análise de Boltzman-Matano aplicada à amostra DR52 para o perfil de enxofre obtido por SIMS (figura 4.21). Observe que há uma queda no

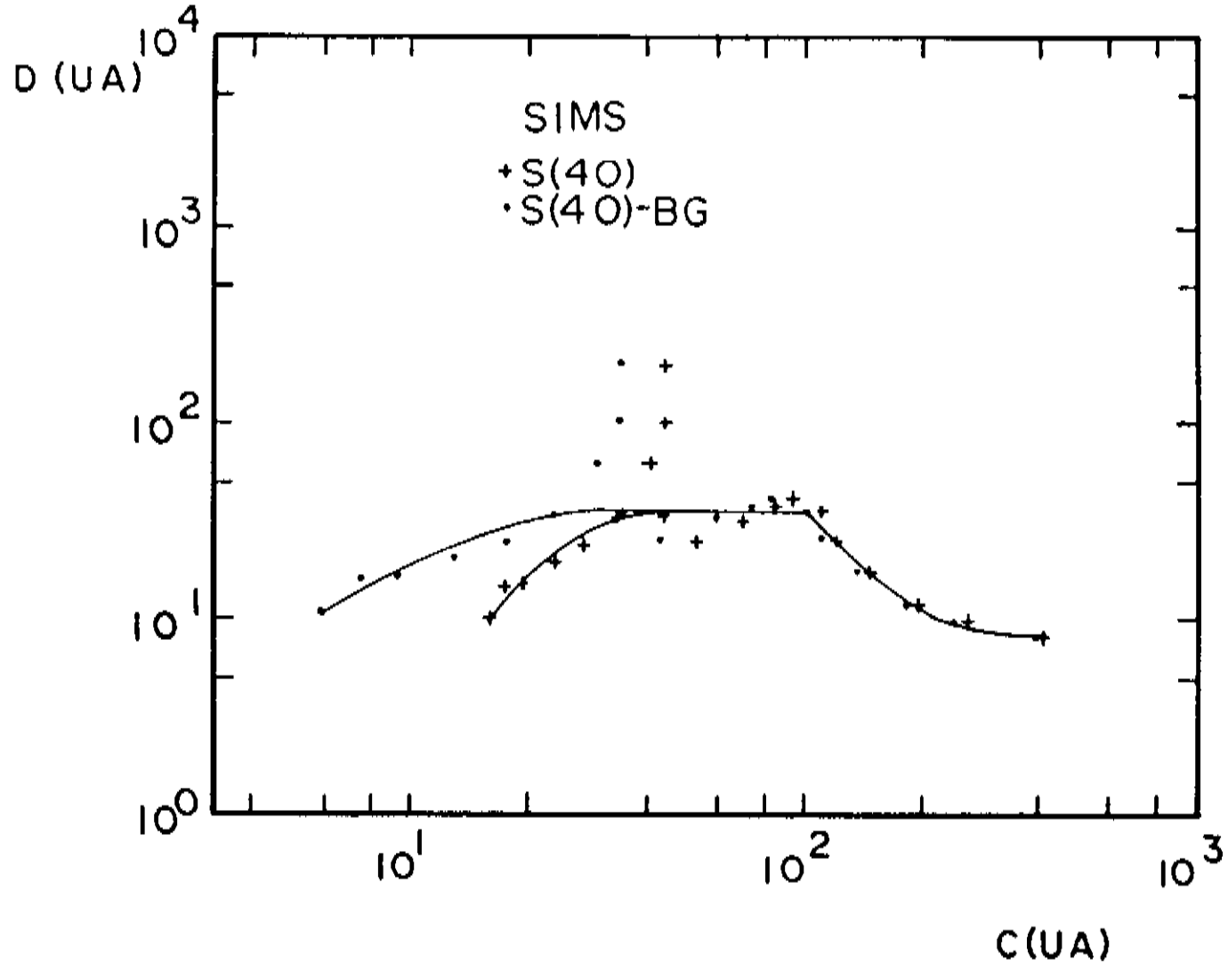

Figura 4.29 - Análise de Boltzman-Matano do perfil de S obtido por SIMS (figura 4.21).

coeficiente de difusão para altas concentrações de enxofre, de quase uma ordem de grandeza, como foi observado por Young e Pearson, e por Tuck e Powell [8].

Na figura 4.30 vemos um gráfico do coeficiente de difusão em função da temperatura obtido com o uso da fonte de 50% de Ga_2S_3. Esse coeficiente foi obtido da análise de Boltzman-Matano, nos perfis do Polaron, e é o valor médio da região de coeficiente constante, onde as concentrações de elétrons e de enxofre devem ser iguais. Obtivemos uma energia de ativação de 4,46eV, com um termo pré-exponencial (D_o na expressão de Arrenhius) de $5,44 \times 10^7$. Matino obtém usando uma fonte na mesma região do diagrama de fases uma energia de 4,35eV e D_o de $3,78 \times 10^6$, em ótima concordância com os nossos resultados. A repetibilidade da medida do coeficiente de difusão das amostras difundidas a 860°C-90s foi excelente. Praticamente tivemos quatro amostras que apresentaram o mesmo valor de coeficiente.

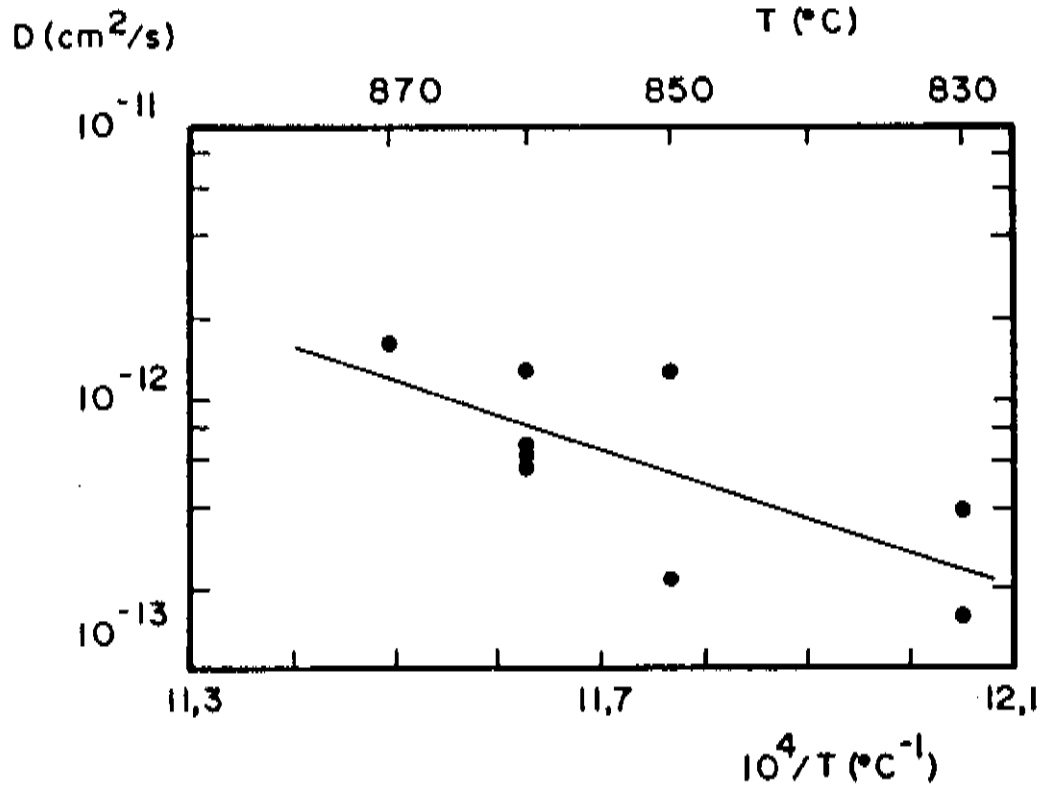

Figura 4.30 - Coeficiente de difusão de enxofre em função da temperatura para a fonte de 50% de Ga_2S_3.

Podemos comparar as mobilidades obtidas em nosso experimento com os valores relatados por Prince [9]. Ele obteve mobilidades médias de 4300cm^2/Vs para difusões de enxofre em GaAs por tubo aberto em fornos convencionais. O nosso valor de mobilidade é menor, pois temos camadas de maior concentração do que as obtidas por Prince. No seu trabalho a concentração de superfície era ao redor de 3 a $5x10^{17}cm^{-3}$, enquanto que nós obtivemos entre 1 e $3x10^{18}cm^{-3}$. Na referência 5 vemos que para essas concentrações o valor esperado de mobilidade é ao redor de 2000cm^2/Vs. Na média nossas amostras estão com um valor de 1800cm^2/Vs, o que é bem próximo do valor esperado. A amostra DR52 apresentou uma baixa mobilidade, que pode ser atribuida à formação dos "clusters" de oxigénio e enxofre. Infelizmente não temos analises de SIMS em amostras que apresentaram uma alta mobilidade, para verificar se a sua queda está realmente relacionada à formação desses "clusters".

4.10 - Conclusões

Descrevemos neste capitulo o projeto de um porta-amostra para uso em difusões de enxofre em GaAs em tubo aberto por processamento térmico rápido. Descrevemos também a preparação das amostras a difundir e da fonte de difusão. A escolha da composição da fonte seguiu o critério de Matino. Usamos fontes de várias concentrações de enxofre nas experiências, e determinamos quais são as faixas de temperatura em que cada fonte pode ser usada.

Das experiências de difusão realizadas podemos concluir que para a fonte de 1% de concentração de Ga_2S_3, a mobilidade das camadas obtidas está relacionada à temperatura de difusão, sendo maior para menores temperaturas. Essa fonte produziu amostras de alta concentração de elétrons de superfície, ao redor de $3x10^{18}cm^{-3}$. Como podemos ver na referência 7, essa concentração de elétrons implica em um concentração de enxofre maior que $10^{19}cm^{-3}$, o que pode criar a precipitação de parte do enxofre na forma Ga_2S_3, e a formação de defeitos complexos do tipo $V_{Ga}S_{As}$. Com o aumento da temperatura a concentração de enxofre deve aumentar (o que pode ser observado na tabela 4.4), aumentando a concentração de elétrons, e a de defeitos, o que resulta na queda da mobilidade. Como era esperado, a concentração de cargas na camada aumenta com o tempo e temperatura de difusão. Já para as fontes de maior concentração de enxofre observamos concentrações de superfície (de elétrons) ao redor de $1x10^{18}cm^{-3}$, com uma mobilidade média de $1800cm^2/Vs$, que está adequada para a camada obtida. A profundidade das camadas variou de $0,1\mu m$ a $0,5\mu m$, dependendo da temperatura e tempo de difusão. A energia de ativação para a fonte de 50% foi de 4,46eV, em excelente concordância com o trabalho de Matino. Dos perfis de SIMS concluimos que existe uma diferença entre a quantidade total de enxofre na camada e a concentração de elétrons. Embora o perfil de carga siga aproximadamente uma função erro complementar, o de enxofre não a segue, sendo mais complexo. Em resumo, desenvolvemos um processo de difusão de enxofre por processamento térmico rápido que resulta em finas camadas do tipo N^+ de boa qualidade.

4.11 - REFERÊNCIAS

[1] - H. Matino, "Reproducible Sulfur Diffusion into GaAs", *Solid-State Electronics*, vol. 17, p. 35 (1974).

[2] - B. Goldstein, "Diffusion in Compound Semiconductors", *Physical Review*, vol. 121, n. 5, p. 1305 (1961).

[3] - B. Tuck, *Introduction to Diffusion in Semiconductors*, IEE Monograph Series 16, PETER PEREGRINUS LTD., England, 1974.

[4] - L. J. Van der Pauw, "A Method of Measuring Specific Resistivity and Hall Effect of Discs of Arbitrary Shape", *Philips Research Report*, n. 13 (1958).

[5] - S. M. Sze, *Physics of Semiconductors Devices*, 2ª edição, JONH WILEY & SONS, New York, 1981.

[6] - A. B. Y. Young e G. L. Pearson, "Diffusion of Sulfur in Gallium Phosphide and Gallium Arsenide", *J. Phys. Chem. Solids*, vol. 31, p. 517 (1970).

[7] - R. Venkatasubramanian e S. K. Ghandhi, "Compensation Mechanisms in n^+-GaAs Doped with Sulphur", *Journal of Crystal Growth*, vol. 97, p. 827 (1989).

[8] - B. Tuck e R. G. Powell, "High-Temperature Diffusion of Sulfur in GaAs", *J. Phys. D : Appl. Phys.*, vol. 14, p. 1317 (1981).

[9] - F. C. Prince, M. Oren e M. Lam, "High Mobility GaAs Layers Obtained by Open Tube Sulfur Diffusion", *Appl. Phys. Lett.*, vol. 8, n. 8, p. 546 (1986).

CAPITULO V

TRANSISTORES DE EFEITO DE CAMPO DE GaAs

5.1 - INTRODUÇÃO

O transistor de efeito de campo de junção (JFET) foi proposto e analisado por Schockley no início da década de 50 [1]. Os transistores de efeito de campo (FET) são basicamente dispositivos semicondutores de três terminais, onde pode-se controlar a resistência entre dois dos terminais através de uma tensão aplicada ao terceiro terminal. O FET é praticamente um resistor controlado por tensão.

No FET o fluxo de corrente pelo transistor é formado por apenas um tipo de portador (lacunas ou elétrons). Devido a esse fato o FET é chamado de transistor unipolar, em comparação com o transistor bipolar, onde temos ambos os tipos de portadores de corrente.

A existência de apenas um tipo de portador no FET faz com que ele opere em velocidades e frequências maiores que as dos transistores bipolares, pois não temos efeitos de armazenamento de portadores minoritários.

Outra caracteristica interessante nos FET's é o coeficiente de temperatura negativo do transistor (diminuição da corrente pelo transistor com o aumento de temperatura), o que o deixa termicamente estável, eliminando possiveis avalanches térmicas na sua operação e tornando mais fácil a ligação de FET's em paralelo. Temos ainda uma alta impedância de entrada no FET, o que o torna atrativo para uso em sistemas de micro-ondas.

Fóra a aplicação em sistemas de micro-ondas, os FET's ainda são usados em circuitos digitais, circuitos de chaveamento e amplificadores de alta impedância e desempenho. Hoje em dia uma grande parte dos circuitos integrados digitais é feita com transistores FET (MOSFET's de Si).

Existem vários tipos de transistores FET [2]. O tipo pro-

posto por Schockley era o de junção, também conhecido por JFET. Nesse transistor o terminal de entrada (terminal de "gate") é formado por uma junção P-N. Outro tipo de FET, que é mais usado em GaAs, é o MESFET ("MEtal Semiconductor Field Effect Transistor"). Nesse tipo de FET ao invés de uma junção P-N temos uma barreira Schottky. As vantagens do uso da barreira são a maior facilidade de sua formação para FET's de GaAs (formação em baixa temperatura, ao contrário da junção, que é em alta temperatura), menor resistência e portanto menor queda de tensão ao longo do "gate", e melhor dissipação de calor para dispositivos de potência. São essas as vantagens que tornam grande o uso de MESFET's ao invés de JFET's de GaAs.

A família dos transistores FET's fica completa com os MOSFET's ("Metal Oxide Semiconductor FET") e MISFET's ("Metal Insulator Semiconductor FET"). Esses transistores funcionam de forma análoga aos JFET's e MESFET's, porém fogem ao interesse deste trabalho, não sendo portanto estudados.

Vamos ver nos próximos itens o modelo de transistores com canal parcialmente saturado, com dopagem não uniforme de canal. Este modelo é aplicável aos transistores que nós fabricamos a partir dos cristais difundidos. Fora ele ainda existem basicamente dois outros modelos, o de transistores de canal saturado [3,4] e o de mobilidade constante [1,4], que são aplicados apenas em transistores com "gates" muito curtos ou altas tensões de "pinch-off", e em transistores com "gate" longo, respectivamente, o que não é o caso dos transistores que nós fabricamos.

A importância de estudarmos os modelos físicos para os transistores vem do fato de podermos prever todo o seu comportamento em corrente contínua ou alternada, apenas usando o modelo. A partir dessas previsões podemos projetar circuitos integrados com grande segurança de que o seu funcionamento será correto.

5.2 - Aspectos Gerais

Os transistores de efeito de campo [2] são basicamente dispositivos de três terminais. Dois deles são os principais, e são chamados de dreno e fonte. O terceiro terminal controla a resistência elétrica entre os terminais de dreno e fonte, e é chamado de porta ou "gate". Normalmente uma corrente elétrica flui do terminal de fonte para o terminal de dreno. O valor desse fluxo de corrente é controlado pelo potencial elétrico aplicado ao terminal de "gate". Podemos dizer de maneira equivalente que estamos controlando a resistência entre os terminais de fonte e dreno a partir da tensão de "gate".

Na figura 5.1 vemos a constituição física de um transistor FET. Ele é formado basicamente por um cristal semicondutor tipo N

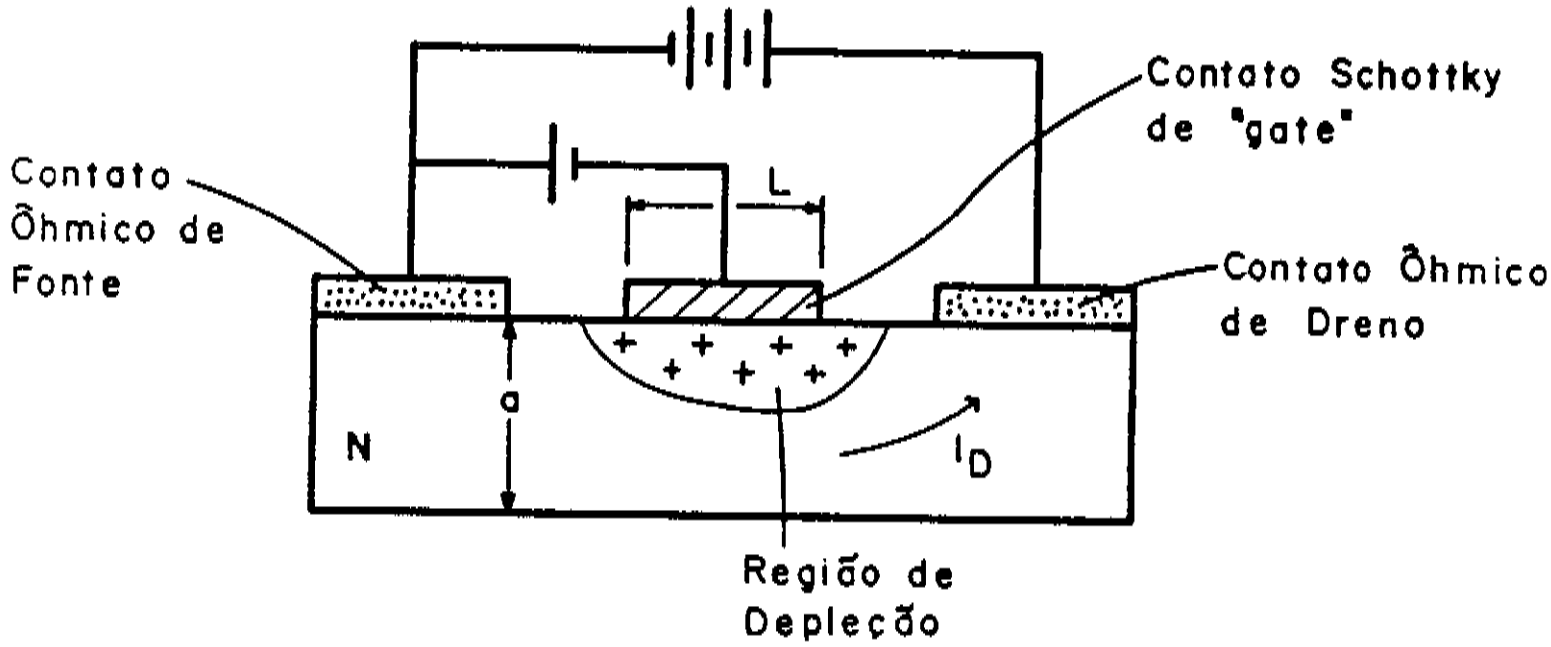

Figura 5.1 - Constituição física de um FET, com a identificação dos terminais e a polarização aplicada (FET de canal N).

e pelos eletrodos de "gate", dreno e fonte. A região do cristal compreendida entre os eletrodos, e por onde flui a corrente pelo transistor, é chamada de canal. No caso da figura 5.1 diz-se que o transistor possui canal tipo N, pois obviamente o canal é formado

por um semicondutor tipo N. Podemos ter FET's de canal P, porém em GaAs isso não é usual, devido às pobres características dos transistores de canal P (a baixa mobilidade das lacunas no GaAs degrada as características do transistor).

Os contatos dos terminais de dreno e fonte são do tipo ôhmico, o que produz uma baixa resistência elétrica entre a metalização de contato e o cristal semicondutor, não influenciando portanto, pelo menos a princípio, no fluxo de corrente pelo canal. Já no terminal de "gate" podemos ter dois tipos de contatos. O primeiro seria um contato ôhmico para o "gate" tipo P no caso de um JFET, e o segundo seria um contato do tipo barreira Schottky, no caso de um MESFET.

Em ambos os tipos de transistor (JFET e MESFET) temos uma barreira de potencial entre o canal e o terminal de "gate", barreira esta que cria uma região de depleção que se extende para o interior do canal. No transistor JFET a região de depleção é criada pela polarização reversa da junção P-N existente entre o "gate" e o canal do transistor. Já para um MESFET ela é criada pela polarização reversa da barreira Schottky existente entre o metal de "gate" e o canal.

Na figura 5.1 podemos identificar algumas das dimensões características do transistor. As principais são o comprimento do "gate", L, que para dispositivos típicos em circuitos integrados está entre $2\mu m$ e $0,5\mu m$, e a largura do canal, W, que varia desde $5\mu m$ até $200\mu m$, podendo chegar até a ordem de milímetros para MESFET's de micro-ondas. Temos a terceira dimensão característica que é a profundidade do canal, a, que fica entre $0,1\mu m$ e $0,5\mu m$.

O princípio de operação físico do dispositivo é bastante simples. Sabemos que a região de depleção é caracterizada pela existência de cargas elétricas fixas, que não contribuem, portanto, ao fluxo de corrente pelo canal (a magnitude do fluxo de corrente é dada basicamente pela quantidade de cargas livres existentes no canal). Variando a polarização reversa do terminal de "gate" conseguimos controlar a espessura da região de depleção e consequentemente a quantidade de cargas livres no canal. Dessa forma para pequenas tensões de "gate" (em módulo) temos pequenas regiões de depleção, o que significa que existirá uma grande

quantidade de cargas livres no canal, produzindo um alto fluxo de corrente pelo transistor. Com o aumento da tensão de "gate" a região de depleção irá se extender para o interior do canal, diminuindo a quantidade de cargas livres, e portanto diminuindo a corrente.

Podemos analisar o funcionamento do FET de outra maneira fazendo uma analogia entre o seu canal e um resistor. Sabemos que um resistor tem o seu valor inversamente proporcional à área da seção transversal por onde flui a corrente. O FET funciona como um resistor de seção variável. Assim, para pequenas tensões de "gate" temos pequenas regiões de depleção, e portanto maiores seções do canal, e consequentemente menores resistências de canal, o que equivale a um maior fluxo de corrente pelo canal.

Na figura 5.2 vemos as curvas típicas de polarização de um transistor FET. Essas curvas nos mostram o comportamento da corrente de canal ou corrente de dreno (I_D) em função da tensão dreno-fonte ou simplesmente tensão de dreno (V_D) para vários valores da tensão de "gate" (V_G). Podemos observar a existência de três regiões características nas curvas. A primeira região, a linear, ocorre para pequenos valores da tensão de dreno, e como o próprio nome já diz, essa é a região de operação linear do dispositivo, onde a corrente de dreno varia linearmente com a

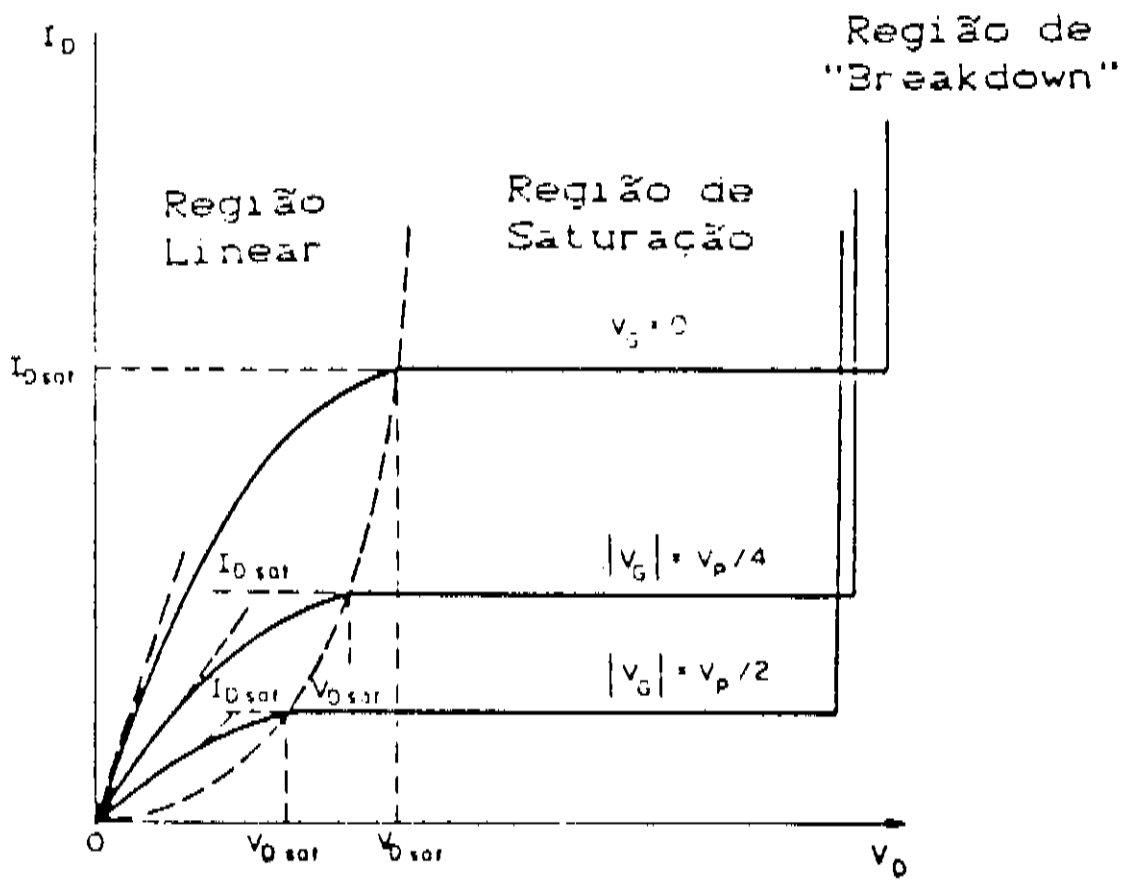

Figura 5.2 - Curvas típicas, I_D x V_D, para vários valores de V_G, de um FET de canal N.

tensão de dreno. Já para valores intermediários da tensão de dreno temos a região de saturação, onde a corrente de dreno não varia. Os valores da corrente e tensão no ponto onde é atingida a saturação são chamados de corrente e tensão de saturação (I_{Dsat} e V_{Dsat}, respectivamente). A região de saturação é usada para a amplificação de sinais pelo transistor. A última região é a de avalanche, onde a alta tensão de dreno provoca o fenômeno de multiplicação por avalanche, o que gera um aumento descontrolado da corrente de dreno, que pode levar o transistor à destruição.

Observando as curvas vemos que o aumento da tensão de "gate" (em módulo) produz a diminuição da corrente pelo transistor, como resultado da maior resistência de canal. Observamos ainda que a tensão de dreno de saturação diminue com o aumento da tensão de "gate", o que sera provado mais tarde.

5.3 - Característica Velocidade x Campo dos Eletrons no GaAs

Antes de passarmos ao estudo do modelo para MESFET com canal parcialmente saturado iremos estudar a curva de velocidade de arraste ("drift") em função do campo elétrico para os eletrons no GaAs. O interesse de se estudar essa curva é que os vários modelos existentes para transistores FET dependem da característica velocidade x campo dos elétrons [2,4] (FET de canal N), sendo que cada modelo irá dar ênfase a certa característica da curva, tendo aplicabilidade em condições bem definidas.

Na figura 5.3 vemos a curva de velocidade de arraste dos elétrons num cristal de GaAs em função do campo elétrico aplicado ao cristal [2] . É interessante observar as características dessa curva. Com o aumento do campo a partir do valor zero, os elétrons vão ganhàndo velocidade proporcionalmente ao campo, o que significa que nesse trecho da curva (baixo campo) a mobilidade, μ ($\mu = \partial V/\partial E$) é constante. Para um campo de aproximadamente 3×10^{3}V/cm observamos que os elétrons atingem um pico de velocidade (2×10^{7}cm/s) e observamos ainda que com o aumento do campo a sua velocidade começa a cair até atingir um valor de saturação entre 6

e 8×10^6 cm/s. A razão da queda de velocidade é que para grandes valores do campo há um espalhamento dos elétrons para um vale superior (veja o diagrama de bandas, figura 5.4), com energia 0,31eV maior, onde os elétrons possuem maior massa efetiva ($0,55m_e$ contra $0,067m_e$ no vale inferior, onde m_e é a massa do elétron) e

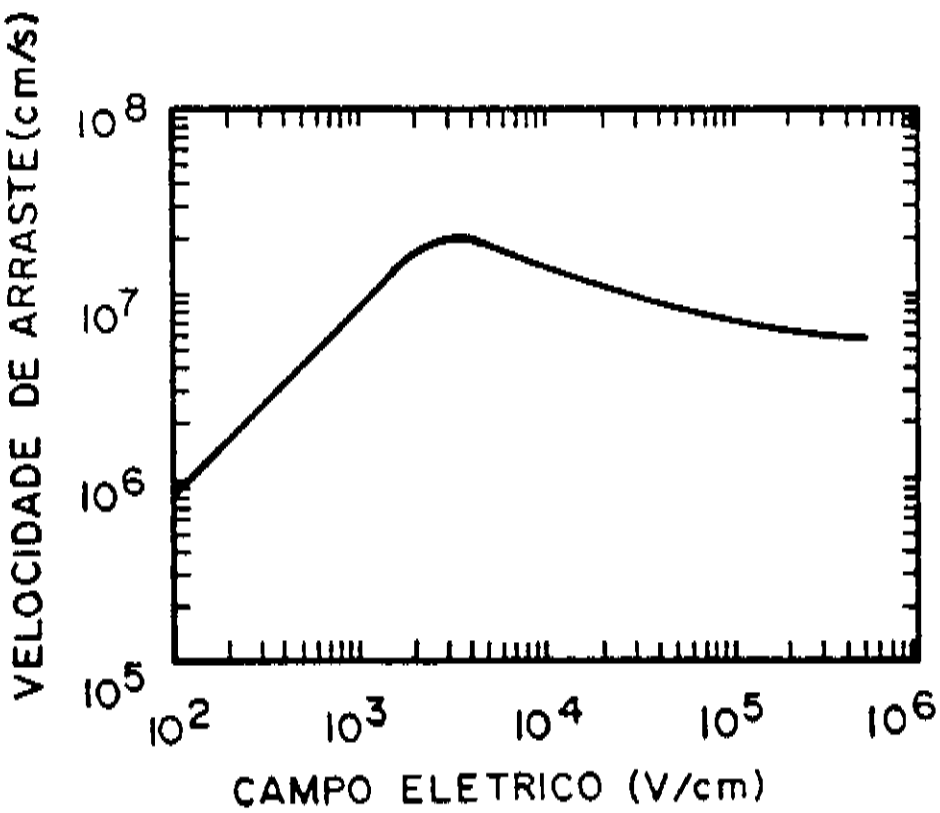

Figura 5.3 - Velocidade de arraste dos elétrons no GaAs em função do campo elétrico aplicado (o gráfico é obtido através de medidas em amostras de alta pureza à temperatura ambiente [2]).

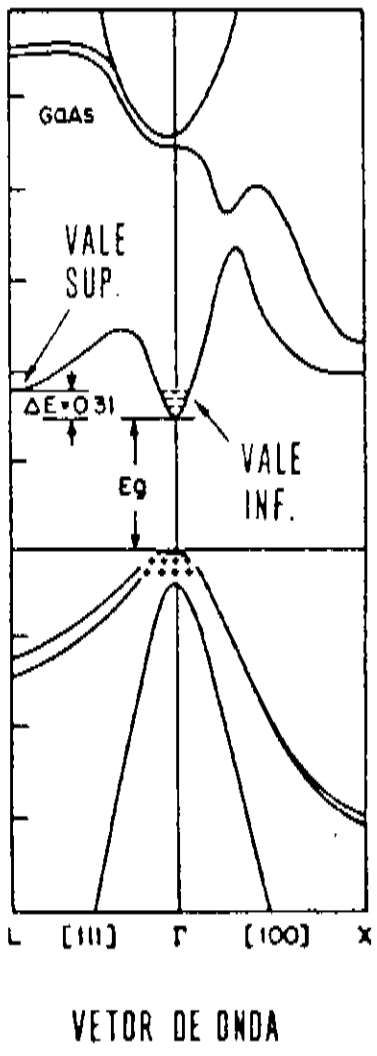

Figura 5.4 - Diagrama de bandas para o GaAs.

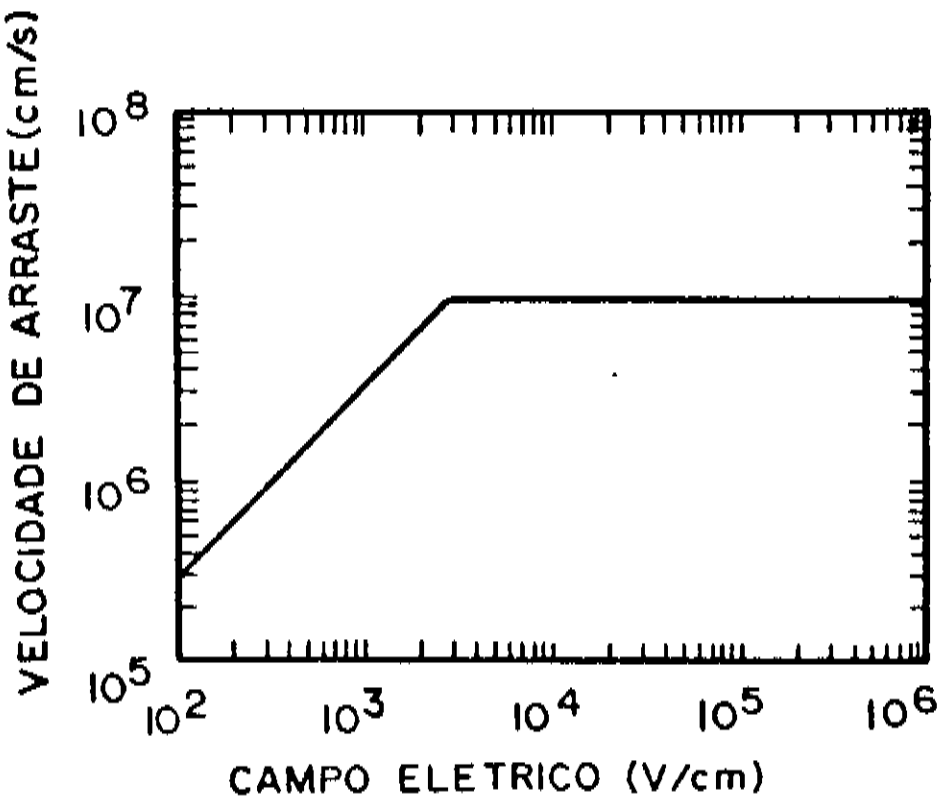

Figura 5.5 - Modelo da curva de velocidade de arraste em função do campo elétrico aplicado para elétrons no GaAs.

menor mobilidade, o que produz uma menor velocidade de arraste, sendo a causa da queda (mobilidade diferencial negativa).

Na figura 5.5 vemos o modelo adotado para a velocidade dos elétrons. Ele é uma simplificação da curva da figura 5.3, já que não apresenta o pico de velocidade para campos de 3×10^3V/cm. A sua descrição matemática também é bem simples, o que e adequado para a dedução das equações de corrente do transistor. A velocidade dos elétrons, v, é dada por,

$$v = \begin{cases} \mu F & , \quad F<F_s \\ v_s & , \quad F\geq F_s \end{cases} \qquad (5.1)$$

onde F é o campo elétrico aplicado, e v_s e F_s são respectivamente a velocidade de saturação e o campo que provoca essa saturação.

5.4 - Saturação de Corrente no MESFET de GaAs

Antes de estudarmos o modelo do canal parcialmente saturado vamos fazer uma consideração sobre a saturação de corrente no

transistor. Isso e importante, pois é a forma como a corrente satura que irá determinar qual é o modelo apropriado para o transistor.

Começamos observando o canal do transistor polarizado na região linear. Na figura 5.6 podemos ver um corte longitudinal do canal polarizado. Podemos mostrar [2] que a altura da região de depleção, y_d, está relacionada à sua queda de tensão, $V(y_d)$, por,

$$V(y_d) = \frac{1}{\varepsilon} \int_0^{y_d} y\, \rho(y)\, dy \qquad (5.2)$$

onde ε e a permissividade elétrica do GaAs, $1,16 \times 10^{-10}$ F/m, e $\rho(y)$ e o perfil de carga fixa devido aos doadores ionizados no canal depletado. A tensão $V(y_d)$ será a soma da tensão de "gate" aplicada, V_G, a tensão de potencial intrínseco de barreira, V_{bi}, e a tensão no canal neutro (não depletado), $V(x)$,

$$V(y_d) = - V_G + V_{bi} + V(x) \qquad (5.3)$$

Repare que V_G é negativa para MESFET's de canal N, e que existe uma queda de tensão no canal neutro, $V(x)$, devida à polarização de dreno. A tensão no canal possui valores fixos nas extremidades da fonte, x=0, e dreno, x=L,

$$V(x) = \begin{cases} 0 & , \quad x = 0 \\ V_D & , \quad x = L \end{cases} \qquad (5.4)$$

já que a fonte está no potencial zero, e o dreno em V_D. Ao longo

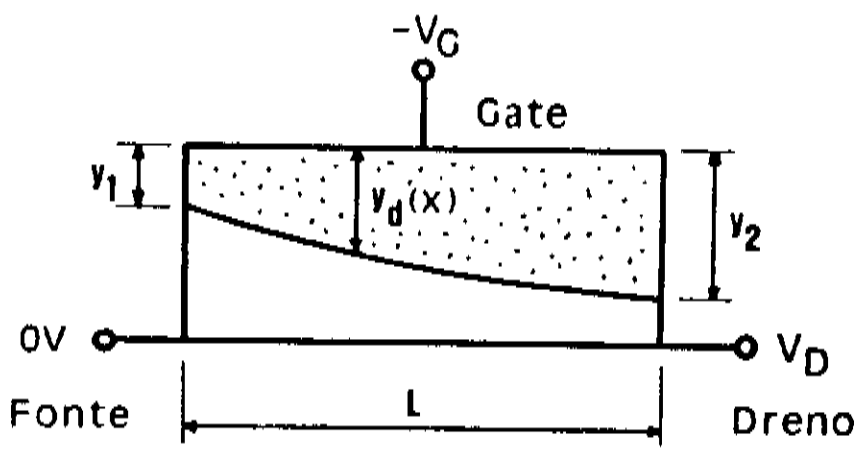

Figura 5.6 - Corte longitudinal do canal de um MESFET polarizado na região linear.

do canal o potencial deve aumentar, caso caminhemos em direção ao dreno. Da mesma forma a queda de tensão na região depletada deve aumentar também, quando vamos em direção ao dreno. Isso corresponde por sua vez a um aumento da altura da região depletada, como podemos concluir da equação (5.2) (y $\rho(y)$ tem sempre o mesmo sinal, e como $V(y_d)$ aumentou, o limite de integração superior deve aumentar). Isso explica a diminuição da altura do canal neutro na figura 5.6. Em direção ao dreno o potencial aumenta e depleta mais o canal.

A corrente pelo canal, I_D, sera dada pelo produto da quantidade de carga existente na seção condutora do canal, pela velocidade dos elétrons que atravessam essa seção,

$$I_D = \mu \frac{dV(x)}{dx} W \int_{y_d(x)}^{a} \rho(y)\, dy \qquad (5.5)$$

onde $dV(x)/dx$ é o campo elétrico no ponto x do canal, W a largura do canal, e a integral representa a densidade linear (sentido da largura) de carga livre no canal. Repare que a altura da região de depleção depende da posição dentro do canal (direção x), o que justifica o uso de $y_d(x)$ como limite de integração. Estamos supondo que a velocidade dos elétrons é simplesmente o produto da mobilidade pelo campo (lembre que o transistor está polarizado na região linear, e temos portanto campos sempre inferiores a F_s no canal neutro). Podemos reescrever (5.5) usando a integral de carga,

$$Q(A) = \int_{0}^{A} \rho(y)\, dy \qquad (5.6)$$

de onde obtemos,

$$I_D = \mu \frac{dV(x)}{dx} W \left[Q(a) - Q(y_d) \right] \qquad (5.7)$$

Usando o fato de que o canal fecha em direção ao dreno, e que a corrente pelo canal deve ser continua, de (5.5) observamos que fixando o valor da mobilidade, a integral de carga deve diminuir quando vamos em direção ao dreno, e que essa diminuição deve ser compensada por um aumento no campo elétrico no canal

neutro. Podemos concluir disso que encontramos o máximo campo elétrico na região de canal neutro sob a extremidade do "gate", no lado do dreno.

Logicamente a saturação de corrente no MESFET é provocada por uma saturação de velocidade dos portadores no interior do canal (eletrons para um canal tipo N). Vimos no item 5.3 que essa saturação de velocidade ocorre para um alto valor de campo elétrico, que, como explicado acima, so e encontrado sob o "gate" no lado do dreno. A conclusão que tiramos e que sera nesse ponto que a corrente atinge primeiro a saturação.

O campo elétrico no interior do canal é uma função do comprimento deste último. Isso pode ser deduzido observando-se que para uma tensão de dreno fixa, o campo será maior para um transistor de "gate" mais curto. Dessa forma um transistor de "gate" mais longo precisara de uma maior tensão de dreno do que um transistor de "gate" curto, para que seja atingido um campo de intensidade F_s que provoca a saturação de corrente.

Podemos achar a distribuição de potencial no canal neutro integrando (5.5). Para fazer isso devemos primeiro substituir dV(x) por $dy_d(dV/dy_d)$. A derivada pode ser achada de (5.2), como,

$$\frac{dV(y_d)}{dy_d} = \frac{y_d\ \rho(y_d)}{\varepsilon} \qquad (5.8)$$

Substituindo (5.8) em (5.5) e integrando em dx de zero, inicio do canal, até uma posição x' arbitrária, e em dy_d de y_1 ate y_x. Realizando essa operação e substituindo x' por x, obtemos,

$$I_D = \frac{\mu\ W}{x\ \varepsilon} \int_{y_1}^{y_x} \left(Q(a) - Q(y_d)\right)\ y_d\ \rho(y_d)\ dy_d \qquad (5.9)$$

onde y_1 e y_x são respectivamente as alturas das regiões de depleção na extremidade do "gate", lados da fonte, e no ponto x.

Pódemos resolver (5.9), partindo de uma escolha de valores para x, I_D, e V_G. Consideramos já fixados os valores de V_{bi}, ε, e μ, e calculado o valor de y_1, partindo de (5.2) e (5.3). Podemos assim calcular a integral do lado direito de (5.9), e determinar o valor de y_x. Logicamente y_d e V(x) estão relacionados por (5.2) e (5.3). Com o valor de y_x achado de (5.9), substituimos

em y_d de (5.2) e achaamos o $V(y_d)$ equivalente, que substituido em (5.3) nos dá V(x). Podemos obter dessa forma a distribuição de potencial para qualquer valor da corrente de canal.

O campo elétrico é achado numericamente calculando-se o potencial em um ponto x, V(x), e em um outro ponto afastado de Δx do primeiro, $V(x+\Delta x)$, onde Δx é um valor pequeno. Feito isso o campo e simplesmente,

$$F(x) = - \frac{V(x+\Delta x) - V(x)}{\Delta x} \qquad (5.10)$$

Estamos interessados em saber qual é a tensão de dreno que ira provocar a saturação na extremidade do "gate", lado do dreno, já que neste ponto temos o maior campo elétrico. Para isso resolvemos as equações acima, com a seguinte condição,

$$F(L) = F_s \qquad (5.11)$$

Isso pode ser feito numericamente, e para o caso extremo de um perfil de dopagem do tipo função delta, e para dopagem uniforme,

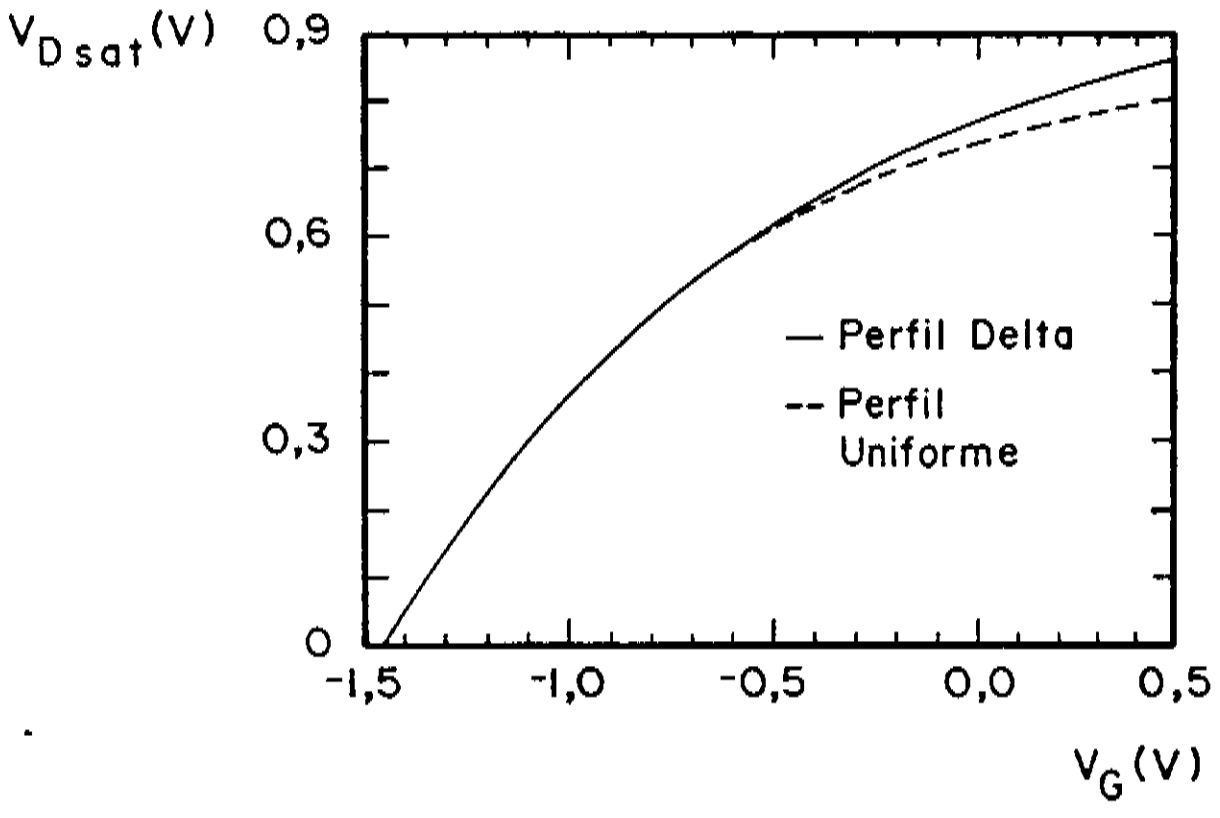

Figura 5.7 - Tensão de saturação, V_{Dsat}, em função de V_G para os casos de dopagem tipo função delta, q N_D $\delta(x-a)$, linha cheia, e dopagem uniforme, linha pontilhada. Em ambos os casos $V_{Dsat} < V_{po}$, e $\alpha = 0.55$ (veja o texto).

vemos na figura 5.7 a tensão de dreno de saturação, V_{Dsat}, em função da tensão de "gate". Repare que o perfil possui pequena influência na tensão de saturação.

Para continuar a análise da saturação devemos introduzir neste ponto a definição da tensão de "pinch-off", V_{po}. Ela é dada simplesmente por,

$$V_{po} = V(a) = \frac{1}{\varepsilon} \int_0^a y\, \rho(y)\, dy \qquad (5.12)$$

e seu significado é a queda de tensão no canal totalmente depletado (queda de tensão na região depletada, sentido da altura).

Podem acontecer dois fenômenos quando a corrente satura no lado do dreno. Caso o "gate" seja muito longo, precisaremos de uma tensão de dreno muito elevada para criar o campo de saturação, o que pode provocar o fechamento do canal no lado do dreno. Nesse caso temos,

$$V_{Dsat} > V_{po} - V_{bi} + V_G \qquad (5.13)$$

Os transistores que apresentam esse fenômeno seguirão o modelo de mobilidade constante ou modelo de Schokley [1,2,4]. O segundo caso é quando o canal é muito curto, onde para uma pequena tensão de dreno já temos o campo de saturação. Isso implica no não fechamento do canal no lado do dreno. Neste caso os transistores seguirão o modelo de canal parcialmente saturado ou o modelo de canal saturado, caso um aumento de tensão de dreno faça com que a região de saturação se extenda para o meio do canal ou caso ela preencha todo o canal. Este último só ocorre para transistores de "gate" muito curto, abaixo de 1μm.

Deve ficar bastante claro que quando nos referimos acima a um fechamento do canal estamos querendo dizer que a largura do canal neutro na extremidade do dreno sera muito pequena quando comparada as dimensões do MESFET. Em momento algum o canal pode fechar no lado do dreno, pois isso implica em um campo eletrico infinito nessa região, o que é fisicamente impossivel.

É claro que o valor da tensão de "pinch-off" irá influen-

ciar o fechamento ou não do canal do transistor (lado do dreno) quando este atinge a saturação. Concluimos isso observando que a tensão de dreno de saturação não depende da tensão de "pinch-off", ela é fixa e apenas uma função do perfil de dopagem, tensão de "gate", campo de saturação e comprimento de "gate". Assim para um dado conjunto dos parâmetros acima, um transistor de baixo V_{po} terá o canal fechado (lado do dreno), enquanto que um outro de alto V_{po} terá o canal aberto. Já o campo de saturação influencia de maneira inversa no fechamento do canal. Altos valores de F_s irão requerer altas tensões de dreno, o que pode provocar o fechamento. Na referência 5 é proposta uma variável adimensional, α, que irá nos indicar o que ocorre na saturação de corrente no MESFET. Essa variável é definida como,

$$\alpha = \frac{F_s L}{V_{po}} \qquad (5.14)$$

sendo que valores de α muito maiores que um correspondem ao fechamento do canal na saturação (modelo de Schokley), valores muito menores correspondem ao não fechamento com saturação total do canal (modelo do canal saturado) e valores ao redor de um ao não fechamento com saturação parcial do canal (modelo do canal parcialmente saturado). Na mesma referência são obtidas expressões para a distribuição de tensão e campo para um transistor com canal uniformemente dopado. Estas expressões são resolvidas numericamente e obtem-se a tensão de saturação em função de α e V_G. É sugerida uma expressão para interpolação dos resultados numéricos,

$$V_{Dsat} = \frac{F_s L \left(V_{po} - V_{bi} + V_G \right)}{F_s L + V_{po} - V_{bi} + V_G} \qquad (5.15)$$

Na figura 5.8 podemos ver um gráfico da tensão de saturação normalizada, u_s, em função de α para vários valores da tensão de "gate" normalizada, u_G. Essas tensões são definidas como,

$$u_s = \frac{V_{Dsat}}{V_{po}} \qquad (5.16a)$$

$$u_G = \frac{V_{b1} - V_G}{V_{po}} \qquad (5.16b)$$

Na figura 5.8 a linha cheia é obtida pela solução numérica das expressões obtidas na referência 5, enquanto que a linha tracejada é obtida da fórmula de interpolação (5.15). No mesmo gráfico vemos a solução numérica das equações (5.2), (5.3), (5.9) e (5.11) para o caso do perfil tipo função delta. Novamente repare que o perfil tem pouca influência na saturação de corrente.

A corrente no ponto onde ocorre a saturação é obtida tomando-se o produto da velocidade dos elétrons no ponto onde ocorre a saturação (extremidade do "gate") pela quantidade de carga existente no canal neutro,

$$I_{Dsat} = v_s \, W \int_{y_s}^{a} \rho(y) \, dy \qquad (5.17a)$$

$$I_{Dsat} = v_s \, W \, [\, Q(a) - Q(y_s) \,] \qquad (5.17b)$$

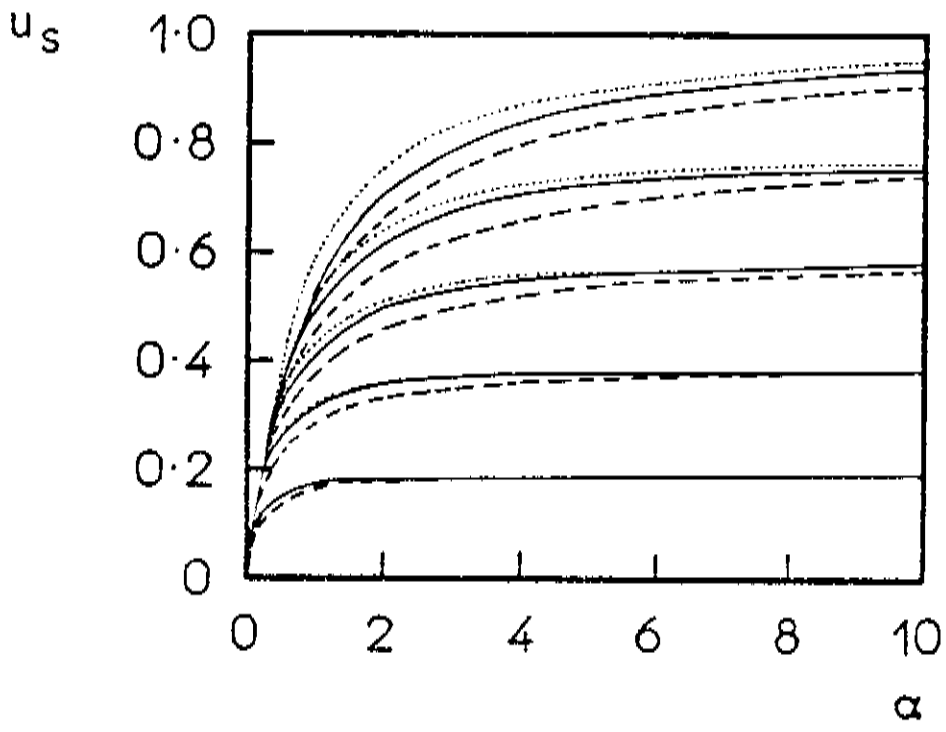

Figura 5.8 - Tensão de saturação adimensional em função do parâmetro α e u_G. No gráfico temos a solução numérica das equações obtidas na referência 5 (canal com dopagem uniforme), em linhas cheias; a solução da fórmula de interpolação (equação (5.15)), em linhas tracejadas; e pontilhada a solução de (5.11) e (5.9) para um perfil do tipo função delta (veja o texto).

onde y_s é a altura da região de depleção na borda do "gate", lado do dreno, quando ocorre a saturação de corrente. Lembre que nesse ponto os elétrons possuem velocidade de saturação, v_s. Está implícito em (5.17a) e (5.17b) que estamos supondo que a velocidade de saturação dos elétrons não varia com o perfil de dopagem. Os valores de y_s, V_{Dsat} e V_G estão relacionados por,

$$V_{Dsat} + V_{bi} - V_G = \frac{1}{\varepsilon} \int_0^{y_s} y\, \rho(y)\, dy \qquad (5.18)$$

5.5 - Modelo do Canal Parcialmente Saturado com Perfil Não Uniforme de Dopagem

Passamos agora a estudar o modelo de transistores com canal parcialmente saturado com perfil não uniforme de dopagem [2,4,6,7]. Este modelo irá ser aplicado para transistores com o parâmetro α próximos de um. Isso corresponde a transistores de 2μm de comprimento de "gate" com tensão de "pinch-off" menor que 4V. Faremos abaixo uma introdução ao modelo, seguida da dedução das correntes de dreno na região linear e na saturada. As tensões e correntes de saturação podem ser achadas partindo-se das equações deduzidas no item anterior.

Na figura 5.9 podemos ver um corte longitudinal do canal do transistor mostrando o que acontece quando ele é polarizado. Começando pela figura 5.9a, vemos o transistor com uma polarização de "gate" aplicada e uma pequena tensão de dreno. Nesse caso o transistor está na região linear, e a fraca polarização de dreno faz a interface entre a região de depleção e o canal neutro ser paralela ao "gate". Aumentando o potencial de dreno, figura 5.9b, a interface deixa de ser paralela, e no lado do dreno o canal é mais fechado, devido a depleção extra gerada pelo potencial de dreno. Aumentando a tensão chegamos à saturação, figura 5.9c. Nesse ponto o campo elétrico no lado do dreno é muito alto e provoca a saturação de velocidade dos elétrons. Aumentando ainda mais a tensão, figura 5.9d, observamos que o campo no interior do

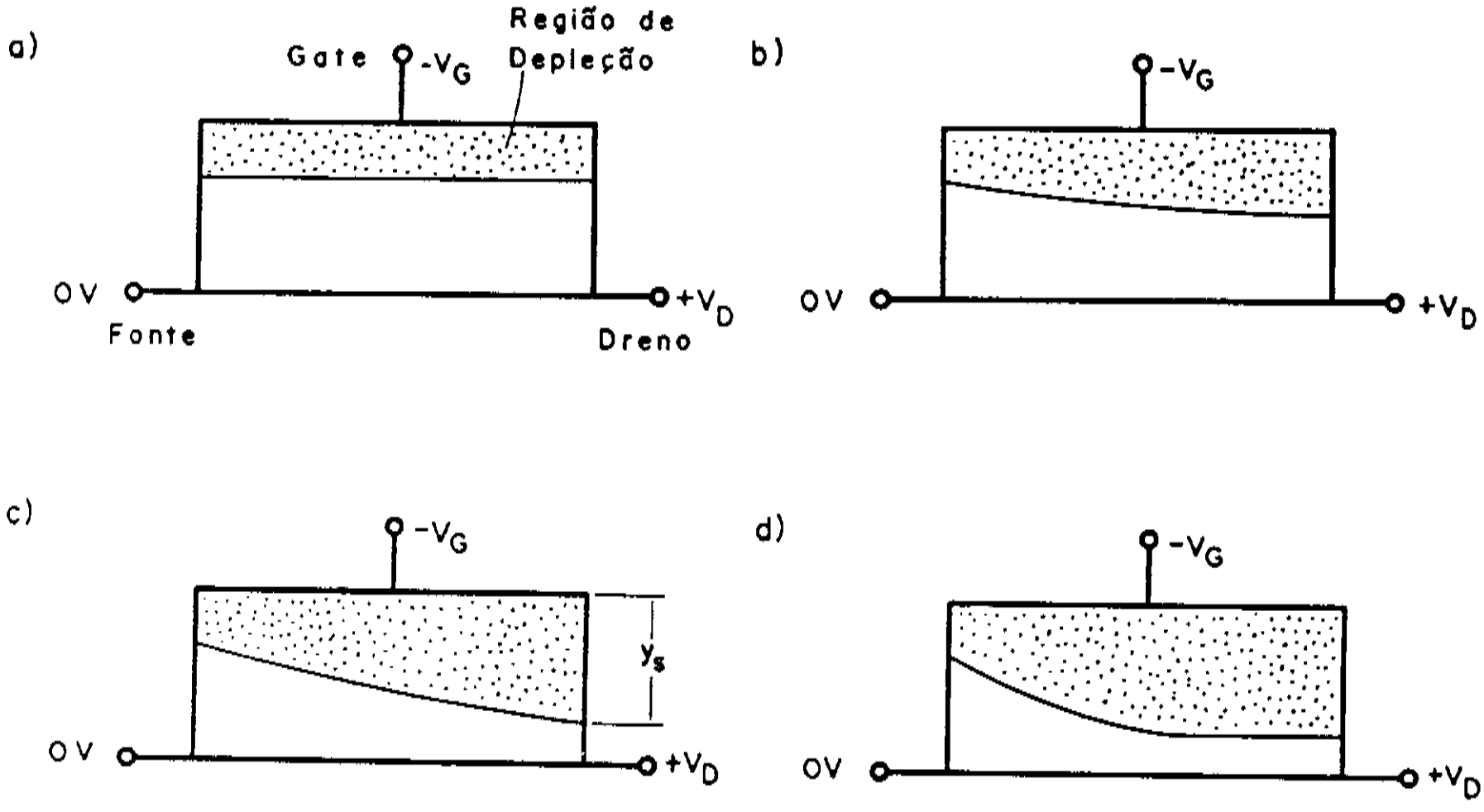

Figura 5.9 - Corte longitudinal do canal de um MESFET mostrando as larguras da região de depleção em, a) polarização na região linear, com baixa tensão de dreno. b) polarização na região linear com alta tensão de dreno. c) polarização no ponto de saturação e d) polarização na região de saturação.

canal neutro aumenta ainda mais, criando uma região de saturação que se estende para o interior do canal. Observe que na saturação ainda temos uma porção relativamente grande de canal não saturado. Na região de saturação formada os elétrons "viajam" pelo canal com a velocidade de saturação, embora o campo nessa região possa ser muito maior que F_s.

Para que cortemos o transistor, ou seja não tenhamos corrente de dreno, é necessario que o canal no lado da fonte esteja totalmente fechado. Como nesse ponto o potencial no canal neutro é zero, podemos definir uma tensão de limiar, V_{th}, que aplicada ao "gate" zera a corrente de dreno,

$$V_{th} = V_G(I_D=0) = - V_{po} + V_{bi} \qquad (5.19)$$

onde usamos (5.3) com $V(x) = 0$, e $V(y_d) = V_{po}$.

Um importante parâmetro do transistor é a sua transcondutância, g_m, que indica o quanto a corrente de dreno varia com a tensão de "gate". A transcondutância está ligada diretamente ao ganho de amplificação do transistor, de onde vem a importância deste parâmetro, que é dado por,

$$g_m = \frac{\partial I_D}{\partial V_G} \Bigg|_{V_D = \text{const.}} \qquad (5.20)$$

Podemos definir uma transcondutância na região de saturação,

$$g_{msat} = \frac{\partial I_{DSat}}{\partial V_G} \Bigg|_{V_D = \text{const.}} \qquad (5.21)$$

5.5.1 - Região Linear

Podemos obter facilmente a expressão para a corrente de dreno na região linear [6]. Isso é feito substituindo a equação (5.8) em (5.5), e fazendo a integração em dx de zero a L, e em dy_d de y_1 a y_2, de onde obtemos,

$$I_D = \frac{\mu W}{L \varepsilon} \int_{y_1}^{y_2} y_d \, \rho(y_d) \, dy_d \int_{y_d}^{a} \rho(y) \, dy \qquad (5.22)$$

ou usando a integral de carga,

$$I_D = \frac{\mu W}{L \varepsilon} \int_{y_1}^{y_2} \left(Q(a) - Q(y_d)\right) y_d \, \rho(y_d) \, dy_d \qquad (5.23)$$

onde y_1 e y_2 são respectivamente as alturas da região de depleção no lado da fonte e no lado do dreno, e são obtidas de,

$$- V_G + V_{bi} = \frac{1}{\varepsilon} \int_{0}^{y_1} y \, \rho(y) \, dy \qquad (5.24a)$$

$$- V_G + V_{bi} + V_D = \frac{1}{\varepsilon} \int_0^{y_2} y\ \rho(y)\ dy \qquad (5.24b)$$

A partir das expressões (5.22) e (5.23) podemos achar a transcondutância do transistor na região linear.

$$g_m = \frac{\partial I_D}{\partial y_1} \frac{\partial y_1}{\partial V_G} + \frac{\partial I_D}{\partial y_2} \frac{\partial y_2}{\partial V_G} \qquad (5.25a)$$

$$g_m = \frac{\mu W}{L} \left[Q(y_2) - Q(y_1) \right] \qquad (5.25b)$$

Obviamente estão implícitas algumas suposições nas deduções das equações, que são a mobilidade constante, região de depleção abrupta, e a chamada aproximação de canal gradual, que significa que a largura da região de depleção varia lentamente com a variação da posição sob o "gate" (sentido de comprimento do canal). Supomos ainda que a mobilidade não varia com a profundidade.

5.5.2 - Região de Saturação

No modelo do canal parcialmente saturado admitimos que a criação da região de saturação ocorre para tensões de dreno além da tensão de saturação. Veremos que um aumento na tensão de dreno provoca um aumento do comprimento da região saturada, que neste caso se extenderá da borda do "gate" para o seu interior. A esse efeito é dado o nome de ' modulação do comprimento do "gate" '.

Na figura 5.10 vemos um esquema mostrando as varias regiões do FET e tensões em cada região. A região I é a ôhmica, que tem comprimento L_1 e queda de potencial V_I. A região II é a de saturação, cujo comprimento é L_S, e queda de potencial V_{II}. Observe na figura que o alto campo elétrico no canal saturado na extremidade do "gate", lado do dreno, provoca a expansão da região de saturação para além do limite do "gate", em direção ao dreno. Essa é a terceira região, de saturação entre o espaçamento de "gate"-dreno, cuja queda de potencial é V_{III}. Logicamente a soma

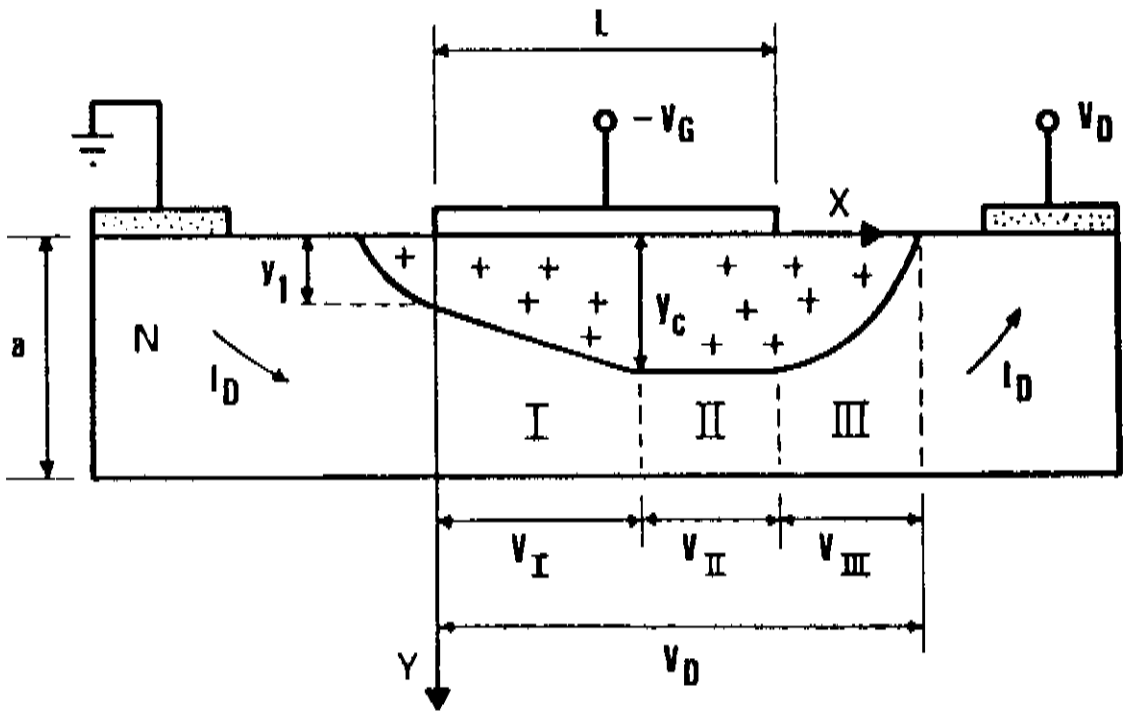

Figura 5.10 - Vista em corte do canal parcialmente saturado de um MESFET, mostrando as varias regiões do canal. As regiões I, II e III são a ôhmica, a de saturação sob o "gate" e a de saturação entre o "gate" e o dreno, respectivamente.

dos comprimentos das regiões I e II deve ser igual ao comprimento do "gate",

$$L = L_1 + L_S \qquad (5.26)$$

A tensão de dreno é a soma das tensões de cada uma das três regiões,

$$V_D = V_I + V_{II} + V_{III} \qquad (5.27)$$

A relação entre as tensões V_{II} e V_{III} sera dada pelo parâmetro K_D, que depende das características físicas do FET (espaçamento entre dreno e "gate", perfil de dopagem do canal, etc.).

$$K_D = \frac{V_{II}}{V_{II} + V_{III}} \qquad (5.28)$$

Na prática os valores de K_D são baixos, o que indica que é grande a queda de potencial na região III.

Na figura 5.10 podemos observar a largura das regiões de

depleção, na borda do "gate", lado da fonte, y_1, e na "interface" entre as regiões I e II, y_c. Os potenciais nesses dois pontos serão dadas pela expressão para o comprimento de uma barreira de potencial (equação 5.2), e são, respectivamente,

$$V_1 = \frac{1}{\varepsilon} \int_0^{y_1} y\ \rho(y)\ dy \qquad (5.29a)$$

$$V_c = \frac{1}{\varepsilon} \int_0^{y_c} y\ \rho(y)\ dy \qquad (5.29b)$$

A queda de tensão na região I será a diferença entre essas duas tensões.

$$V_I = V_c - V_1 \qquad (5.30a)$$

$$V_I = \frac{1}{\varepsilon} \int_{y_1}^{y_c} y\ \rho(y)\ dy \qquad (5.30b)$$

Podemos achar a corrente de canal na região ôhmica, I_{DI}, atraves da equação (5.5), onde integramos em x de 0 a L_1, e em y_d de y_1 a y_c. Ficamos assim com,

$$I_{DI} = \frac{\mu\ W}{L_1\ \varepsilon} \int_{y_1}^{y_c} \left(Q(a) - Q(y_d)\right)\ y_d\ \rho(y_d)\ dy_d \qquad (5.31)$$

Na região saturada, II, a corrente, I_{DII}, deve ser a mesma que na região I, para que haja a conservação de cargas eletricas,

$$I_{DI} = I_{DII} \qquad (5.32)$$

onde I_{DII} e dada por,

$$I_{DII} = v_s\ W \int_{y_c}^{a} \rho(y)\ dy \qquad (5.33)$$

ou usando a integral de carga, equação (5.6),

$$I_{DII} = v_s\ W\ \left[\ Q(a) - Q(y_c)\ \right] \qquad (5.34)$$

A queda de potencial na região II é achada resolvendo-se a equação de Laplace nessa região por separação de variáveis e

mantendo o primeiro termo da expansão da solução em serie de Fourier [7],

$$V_{II} = \frac{2\ a\ F_s}{\pi}\ \text{senh}\left(\frac{\pi\ L_S}{2\ a}\right) \qquad (5.35)$$

onde L_s é o comprimento da região saturada e F_s é o valor de campo elétrico que provoca a saturação da velocidade dos eletrons.

A queda de potencial no canal do transistor, é portanto,

$$V_D = \frac{1}{\varepsilon}\int_{y_1}^{y_c} y\ \rho(y)\ dy \ - \frac{1}{K_D}\ \frac{2\ a\ F_s}{\pi}\ \text{senh}\left(\frac{\pi\ L_S}{2\ a}\right) \qquad (5.36)$$

O conjunto de equações acima pode ser resolvido, pelo menos numericamente, de tal forma a obter-se valores para as diversas icognitas, resultando nas suas curvas caracteristicas $I_D \times V_D$.

Agora vamos supor que polarizamos o transistor no início da região de saturação, e que comecemos a aumentar a tensão de dreno. Esse aumento irá provocar um consequente aumento na tensão da região II, o que por sua vez faz com que L_S aumente (efeito de modulação do comprimento do "gate"). Porém o aumento correspondente em L_S não é proporcional ao aumento em V_{II} (canal não ôhmico), o que pode ser visto na expressão (5.35), onde V_{II} é proporcional ao seno hiperbólico contido nesta expressão. Este tem uma rápida variação com L_S. Dessa forma o aumento da tensão de dreno provoca um ligeiro aumento em L_S e a correspondente diminuição em L_1. Essa diminuição em L_1 é acompanhada de um diminuição em V_I para que o campo no interior da região I fique abaixo de F_s. Toda a tensão de dreno em excesso a de saturação irá cair sobre as regiões II e III, ficando a região I com uma tensão próxima a de saturação. Isso ocorre mesmo para tensões de dreno muito maiores que a de saturação. A pequena diminuição de V_I corresponderá a diminuição de y_c, e o consequente aumento da corrente pelo canal (equação (5.33)). Esse aumento de corrente implica em um aumento da trascondutância. Na figura 5.11 temos as curvas de um MESFET típico, obtidas pela solução das equações (5.26) a (5.36).

Podemos usar uma simples expressão para a corrente de dreno em função da tensão de "gate", como proposto na referência 5,

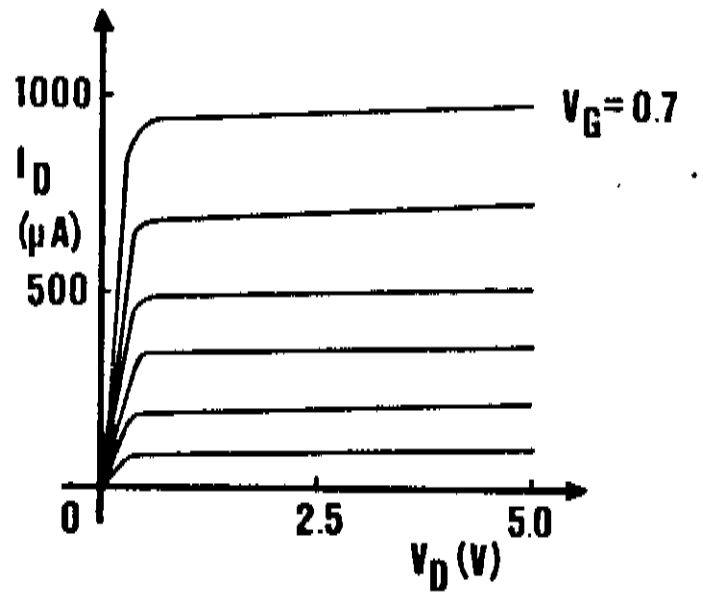

Figura 5.11 - Curvas caracteristicas de um MESFET de canal parcialmente saturado, obtidas pela solução numerica das equações (5.26) a (5.36). O MESFET possui V_{bi}=0,7V, a=0,15μm, W=5μm, L=2μm, v_s=1,2x10^{7}cm/s, F_s=3x10^{3}V/cm e μ=3000cm^2/Vs.

$$I_D = \beta \left(V_G - V_{th}\right)^2 \qquad (5.37)$$

O aumento da transcondutância e da corrente de dreno podem ser levados em consideração se escolhermos β como [5],

$$\beta = \frac{2\,\varepsilon\,\mu\,v_s\,W}{a\left(\mu\,V_{po} + 3\,v_s\left(L - L_S\right)\right)} \qquad (5.38)$$

O fato da expressão (5.37) ser valida faz com que o modelo do canal parcialmente saturado seja tambem chamado de 'modelo da lei do quadrado'. A transcondutância neste caso será,

$$g_{msat} = 2\,\beta\left(V_G - V_T\right) \qquad (5.39)$$

onde β é dado por (5.38). Observamos de (5.38) e (5.39) que a transcondutância irá depender de μ e v_s. Isto mostra que a saturação da velocidade dos elétrons já passa terá influência nas caracteristicas do transistor (além de influenciar a tensão e corrente de saturação influencia também a transcondutância).

O campo elétrico na região II, F(x'), onde x' é medido a

partir do início da região saturada,é obtido diferenciando-se a expressão para a distribuição de potencial nessa região [7].

$$F(x') = -F_s \cosh\left(\frac{\pi x'}{2a}\right) \qquad (5.40)$$

Da equação acima observamos que o campo elétrico irá aumentar, no interior da região II, além do seu valor de saturação. No nosso modelo para a velocidade dos elétrons desprezamos o pico de velocidade destes, ignorando consequentemente a mobilidade diferencial negativa, que ocorre para campos maiores que o de pico (figura 5.3).

Se levarmos em conta a mobilidade diferencial negativa, veremos que o aumento do campo no interior da região II irá provocar a diminuição da velocidade dos elétrons. Já que a corrente é contínua no interior do canal, esse efeito irá resultar na acumulação de cargas na região II.

Na região III o campo elétrico irá decrescer, partindo de seu valor máximo, $F(L_s)$ (início da região III), até atingir quase zero, no fim da região. Esse decréscimo no campo irá provocar a aceleração dos elétrons no interior da região III, criando uma região com acúmulo de cargas positivas, já que deve haver a conservação de carga nessa região. A esse efeito de criação de

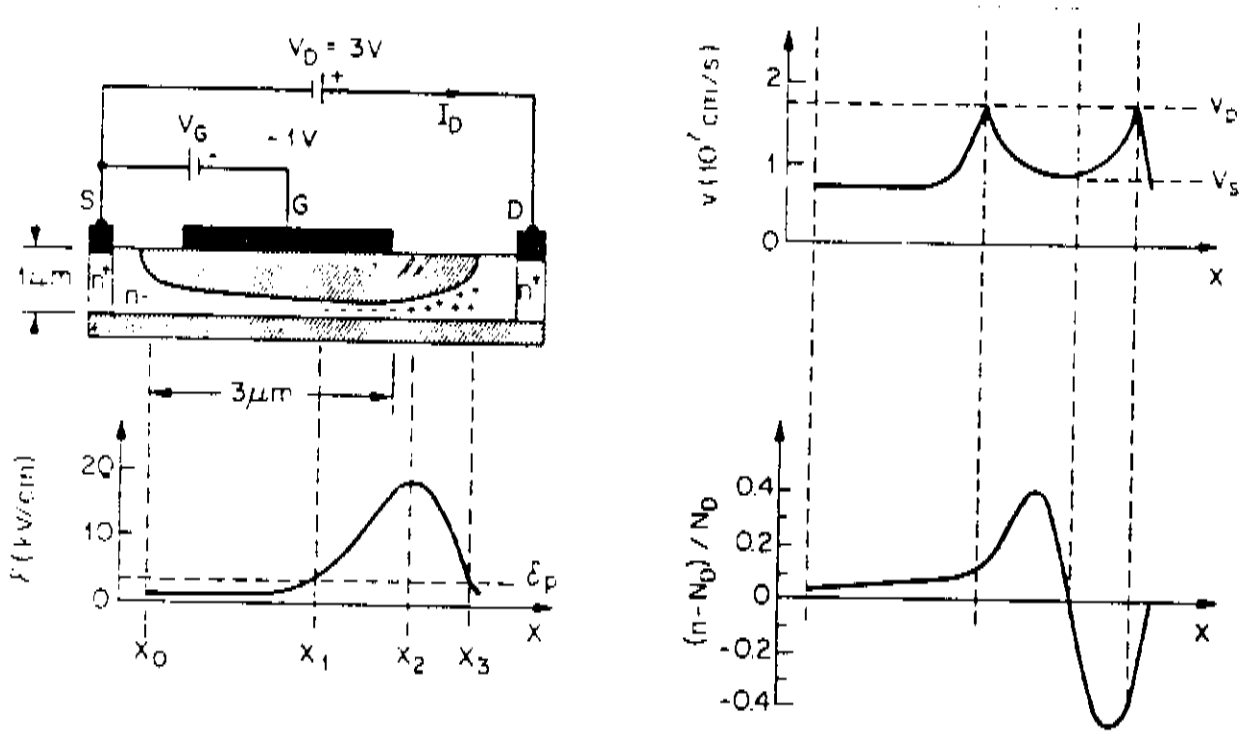

Figura 5.12 - Formação do domínio Gunn no transistor MESFET.

duas regiões carregadas com cargas de polaridade diferente é dado o nome de formação de dominio Gunn. Para MESFET's não auto-alinhados a saturação da corrente pelo canal (devida a saturação da velocidade dos elétrons) é muitas vezes associada à formação do dominio. Na figura 5.12 podemos ver um esquema do que acontece no interior do canal do transistor durante a formação do dominio Gunn.

5.6 - Conclusões

Mostramos nos itens anteriores o modelo de transistores MESFET com canal parcialmente saturado. Obtivemos expressões que podem ser resolvidas numericamente resultando nas curvas $I_D \times V_D$ para valores quaisquer de V_G. A saturação de corrente foi tratada à parte, já que o modo como ela ocorre determina qual será o modelo mais apropriado para o transistor.

No próximo capitulo vamos descrever as etapas de fabricação dos transistores e os resultados obtidos da sua caracterização.

5.7 - Referências

[1] - W. Schokley, "An Unipolar Field-Effect-Transistor", *Proc. IRE*, vol. 40, p. 1365 (1952).

[2] - S. M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, New York, 1981.

[3] - M. S. Shur, "Analytical Model of GaAs MESFET's", *IEEE Trans. Elec. Dev.*, vol. ED-25, n. 6, p. 612 (1978).

[4] - M. S. Shur, *GaAs Devices and Circuits*, Plenum Press, New York, 1985.

[5] - M. S. Shur, "Low Field Mobility, Effective Saturation Velocity and Performance of Submicron GaAs MESFET's", *Elec. Lett.*, vol. 18, n. 21, p. 909 (1982).

[6] - M. S. Shur e L. F. Eastman, "I-V Characteristics of GaAs MESFET with Nonuniform Doping Profile", *IEEE Trans. Elec. Dev.*, vol. ED-27, n. 2, p. 455 (1980).

[7] - A. B. Grebene e S. K. Ghandhi, "General Theory for Pinched Operation of the Junction-Gate FET", *Sol. State Elec.*, vol. 12, p. 573 (1969).

CAPITULO VI

FABRICAÇÃO E TESTES DOS TRANSISTORES

6.1 - INTRODUÇÃO

No capitulo anterior estudamos o funcionamento e constituição física dos MESFET's de GaAs. A principal caracteristica dos transistores estudados era o uso de um cristal dopado, tipo N, na formação do canal. Praticamente essa dopagem de canal pode ser obtida por varias formas, como por exemplo crescimento epitaxial, implante iônico ou difusão térmica.

Uma alternativa menos convencional para a dopagem do canal é a difusão por RTP de enxofre em GaAs, que foi vista no capitulo 4. Com relação ao implante iônico a difusão possui duas vantagens na formação do canal, que são a alta concentração de impurezas na superfície do cristal e o perfil de dopagem mais abrupto, o que resulta em transistores de menor resistência parasita de dreno e fonte e maior trancondutância.

Neste capitulo iremos descrever todo o processamento dos cristais difundidos com o objetivo de fabricar os transistores MESFET e apresentar os resultados obtidos da caracterização dos transistores produzidos.

6.2 - CARACTERISTICAS FÍSICAS DOS TRANSISTORES FABRICADOS

Os cristais uma vez difundidos e devidamente processados resultaram em dois tipos diferentes de transistores MESFET. Eles são transistores de depleção com 5μm de largura de canal e comprimento de "gate" de 2μm e 3μm. Em todos os dois tipos de transistor o espaçamento tipico entre o "gate" e o dreno ou fonte é de 1,5μm. Na figura 6.1 vemos um diagrama mostrando as dimensões tipicas dos transistores fabricados.

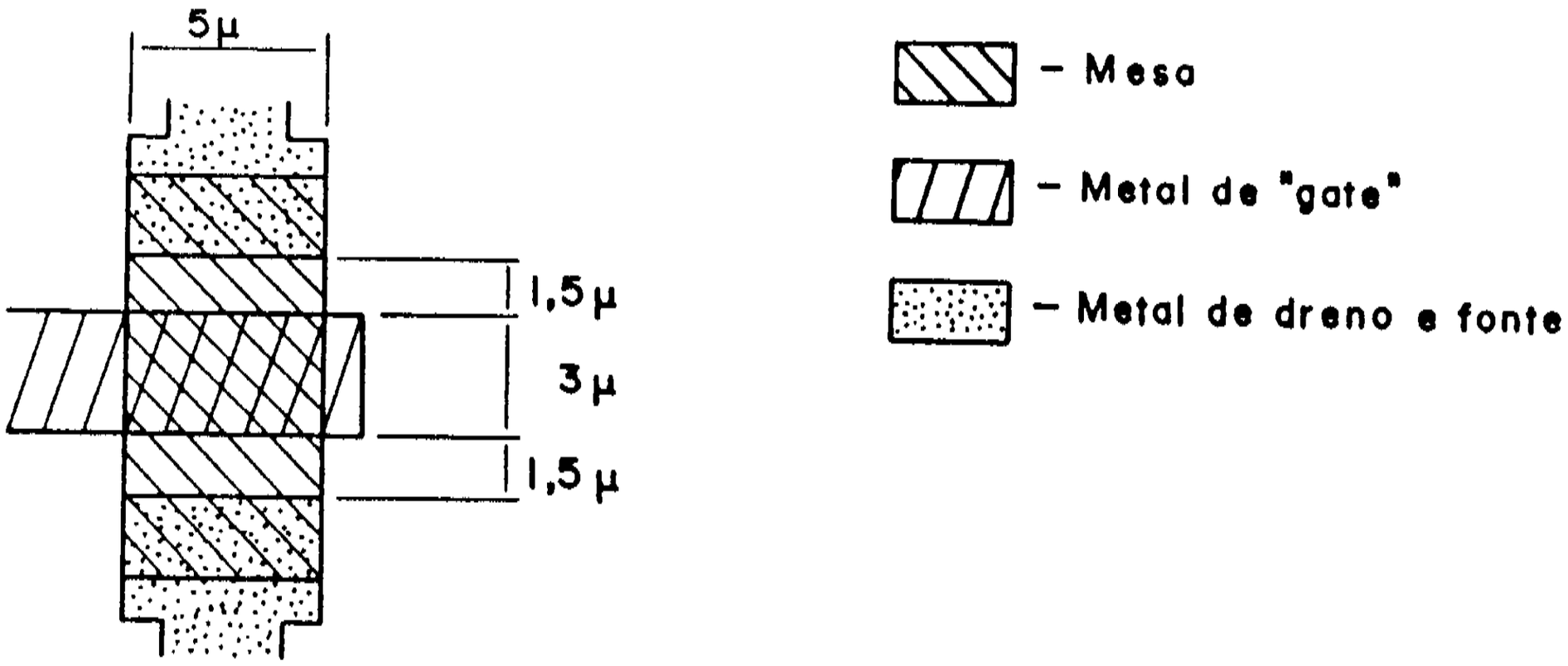

Figura 6.1 - Dimensões típicas dos transistores fabricados.

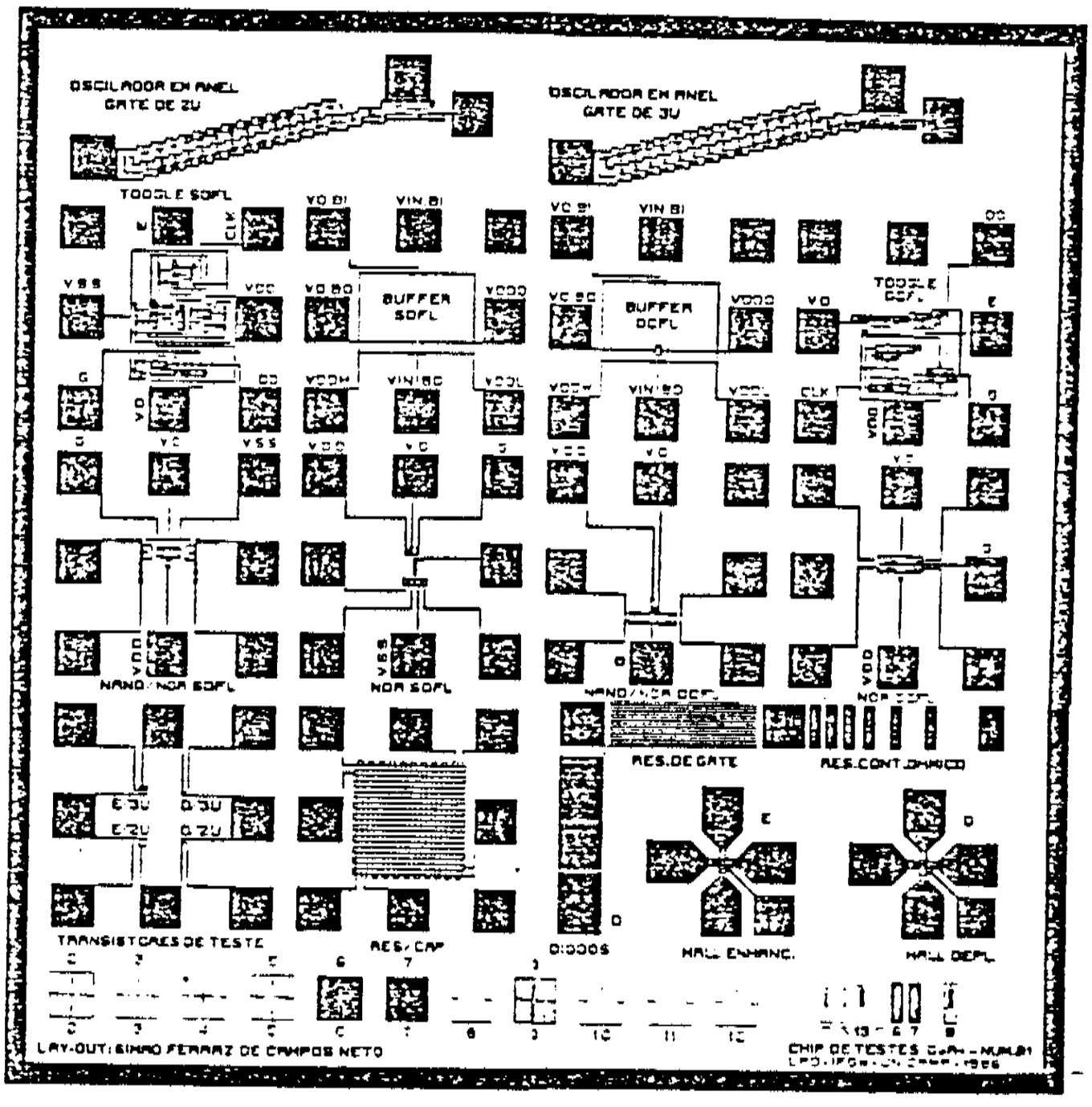

Figura 6.2 - Desenho do "chip" de testes.

Os transistores foram feitos a partir de um conjunto de máscaras para a fabricação de circuitos digitais para testes e caracterização dos processos de fabricação dos transistores. Esse conjunto de máscaras foi projetado no nosso grupo e a fabricação das mesmas foi realizada por uma firma norte-americana especializada (Advance Reproductions Corp.). Na figura 6.2 podemos

ver um desenho do "chip" de testes (superposição de varias mascaras). Os transistores fabricados estão localizados no canto inferior esquerdo do "chip", logo acima da inscrição "TRANSISTORES DE TESTE".

6.2.1 - Escolha dos Cristais Difundidos

No capitulo 4 vimos que o processo de difusão rápida produz cristais dopados com uma razoavel gama de concentração de placa (desde $5\times10^{11}cm^{-2}$ a $2\times10^{13}cm^{-2}$). O critério usado na escolha dos cristais a serem processados foi principalmente a sua resistência de placa. Um critério secundário foi a mobilidade media dos elétrons no cristal, que para as difusões realizadas deve ser sempre maior que $1600cm^2/Vs$, sendo rejeitados os cristais com mobilidade inferior a esse valor.

Os cristais escolhidos foram os de resistência de placa inferior a 400Ω/□, o que para as mobilidades tipicamente obtidas, entre 1600 e $2200cm^2/Vs$, corresponde a concentrações de placa maiores que $8\times10^{12}cm^{-2}$. A escolha desses cristais relativamente bem dopados foi feita com o proposito de produzir transistores com baixas resistências parasitas de dreno e fonte. Os cristais escolhidos possuiam boa morfologia de supefície, caracterizada pela ausência de defeitos na superfície do cristal apos a difusão.

É importante observar que os cristais utilizados tinham concentração superficial da ordem de $1\times10^{18}cm^{-3}$. A profundidade da camada difundida foi um fator sem importância, pois como veremos mais adiante a colocação do "gate" e feita após um ataque quimico que ajusta a profundidade do canal para um valor mais adequado ("gate recessed"). Por fim observamos que não foi tomado nenhum cuidado com relação a composição da fonte de difusão, temperatura ou tempo de difusão usados na obtenção dos cristais difundidos, já que o nosso objetivo era demonstrar a possibilidade de aplicação da difusão rápida na fabricação dos transistores MESFET, e não algum estudo mais sistemático sobre a influência dos parâmetros da difusão nas caracteristicas dos transistores.

6.3 - Tecnologia Empregada na Fabricação

A tecnologia empregada foi a de "mesas" na fabricação dos transistores. Nessa tecnologia a isolação entre os vários transistores em um mesmo cristal e feita pelo proprio substrato (cristal), que no nosso caso e semi-isolante. A difusão de enxofre e feita em toda a area do cristal ("broad area"), isso leva a necessidade de uma corrosão quimica (do cristal) para que seja retirada a camada difundida (condutora) onde ela não e necessaria (areas onde não existem os transistores). Essa corrosão cria estruturas em relevo na superficie do cristal, ou "mesas", de onde vem o nome dado a tecnologia.

6.3.1 - Etapas de Processo

Damos o nome de etapas de processo às varias fases do processamento dos cristais (tais quais a difusão, metalizações, corrosões quimicas, entre outras).

A fabricação dos transistores consiste ao todo de seis etapas diferentes de processo. A primeira que já vimos foi a difusão de enxofre no cristal para a formação da região condutora do canal. As cinco etapas seguintes são a formação de marcas de alinhamento no cristal, a formação das "mesas", a metalização de contato ôhmico de dreno e fonte, a metalização de "gate" para os transistores de depleção e por fim a metalização das areas de contato ("pads") para a conexão dos terminais externos dos transistores.

As metalizações e formação das "mesas" e marcas de alinhamento são feitas em areas bem definidas do cristal, ao contrario da difusão, que e feita em todo o cristal. A definição das areas a receberem a metalização ou a serem corroidas em cada etapa de processo e feita atraves de fotogravação com o uso de máscaras (citadas no item 6.2). Para cada etapa de processo temos uma fotogravação diferente o que implica no uso de uma mascara diferente (a fabricação dos transistores compreende o uso de cinco mascaras diferentes).

O tipo de fotogravação usada é a de impressão por contato. Neste tipo de fotogravação a máscara é colocada em contato com o cristal, e a sua imagem é transferida o mais fielmente possível para o cristal.

Nas várias etapas de processo usamos dois tipos diferentes de fotogravação. O primeiro tipo é o normal, que é usado para as corrosões químicas. O segundo tipo é o chamado fotogravação para "lift-off", e é usado para as metalizações. A seguir descrevemos os dois processos.

6.3.2. - FOTOGRAVAÇÃO NORMAL

A fotogravação é basicamente um processo fotográfico onde o cristal que vai ser fotogravado recebe uma fina camada de resina fotográfica ("fotoresist"), que é sensibilizada com a imagem da máscara de fotogravação e revelada a seguir.

A fotogravação começa com o espalhamento do "fotoresist" sobre o cristal. Este é feito com o uso de uma centrífuga chamada "spiner". Ela é basicamente um motor de alta rotação com um pequeno sistema de vácuo para prender o cristal ao eixo do motor.

Uma vez estando preso o cristal ao eixo do "spiner", são pingadas algumas gotas de "fotoresist" sobre o cristal e é ligado o motor do "spiner", que faz o cristal girar. Isso faz com que o "fotoresist" seja espalhado sobre o cristal, criando uma fina camada de aproximadamente 1.5μm de espessura, bem uniforme em todo o cristal.

O "fotoresist" usado em todos os processos de fotogravação foi o AZ 1350J de fabricação da Shipley. A velocidade e tempo de rotação usados no "spiner" são respectivamente 7000 rpm e 20 segundos.

Após o espalhamento do "fotoresist" este é curado ("pre-bake") em uma estufa. A cura é feita com o objetivo de secar o "fotoresist". Para a fotogravação normal o tempo de cura é de 20 minutos a uma temperatura de 90°C.

Uma vez espalhado o "fotoresist" sobre o cristal e feita a sua cura, este é exposto com a imagem da máscara. A exposição é

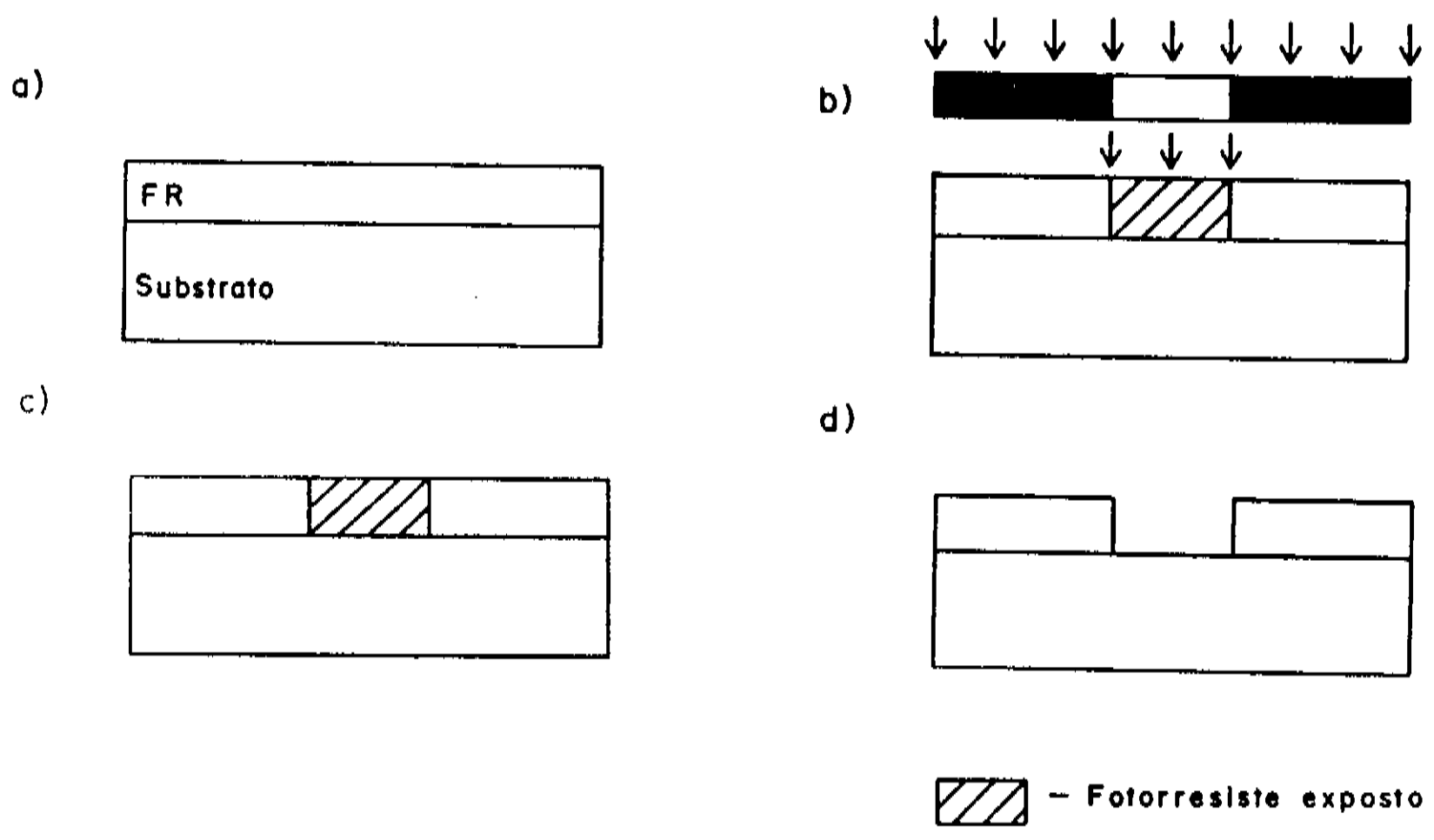

Figura 6.3 – Sequência de fotogravação normal. a) cristal coberto com "fotoresist", b) exposição do "fotoresist", c) áreas sensibilizadas pela exposição, e d) "fotoresist" revelado.

feita com luz ultra-violeta (o "fotoresist" é sensível a luz ultra-violeta) em uma fotoalinhadora da marca "Photo-Litographyc Products". Para a fotogravação normal o tempo de exposição é de 8 segundos.

Após a exposição é feita a revelação do "fotoresist". O revelador usado é o MF 312 da Shipley, diluído na proporção de 1:0,75 partes (em volume) de água. O tempo de revelação é de 12 a 15 segundos, dependendo da temperatura do revelador (temperatura ambiente). Durante a revelação o revelador irá dissolver todo o "fotoresist" que foi sensibilizado (exposto), deixando sobre o cristal apenas o "fotoresist" não sensibilizado. Dessa forma conseguimos áreas expostas na superfície do cristal, podendo ser feito, por exemplo, um ataque químico apenas nessas áreas, já que o "fotoresist" é resistente à ácidos.

Após a revelação os cristais são enxaguados em água deio-

nizada (DI) para retirar os resíduos do revelador, e secos com um jato de nitrogênio. Caso a revelação não tenha sido completa, podemos voltar o cristal para o revelador por mais alguns segundos.

Terminada a revelação do cristal este é colocado de novo na estufa ("pós-bake"), por mais 20 minutos a 90°C. Essa nova cura é feita com o objetivo de aumentar a aderência do "fotoresist" ao cristal e de torná-lo mais resistente à acidos.

Na figura 6.3 vemos a sequência típica de fotogravação do cristal. Na figura 6.3a vemos o cristal coberto com o "fotoresist" já seco. Na figura seguinte, 6.3b, o cristal está colocado em contato com a máscara e está sendo exposto (repare que a máscara possui áreas opacas que impedem a passagem da luz ultra-violeta). Na figura 6.3c vemos a área do "fotoresist" sensibilizada pela luz ultra-violeta, e finalmente na figura 6.3d o "fotoresist" já revelado, mostrando uma pequena parte da superfície do cristal exposta.

6.3.3 - Fotogravação para "Lift-Off"

A fotogravação para "lift-off" é usada para as metalizações do cristal. Existem várias diferenças entre a fotogravação normal e a para "lift-off". A primeira é que o "pré-bake" é feito a uma temperatura menor, 70°C, por 20 minutos. Neste processo não é feito o "pós-bake", e a exposição é feita com 13 segundos, ao invés de 8 segundos. A revelação é mantida inalterada, porém antes dela o cristal é mergulhado em clorobenzeno por 15 minutos. É usado o mesmo "fotoresist" (AZ 1350J) e o seu espalhamento é igual ao do processo normal.

O uso do clorobenzeno após a exposição endurece a camada superficial do "fotoresist", como mostram as figuras 6.4a e 6.4b. O "fotoresist" endurecido tem a sua revelação (dissolução) mais lenta. Dessa forma, durante a revelação, quando o revelador atinge a camada não endurecida ele passa a dissolver mais rapidamente o "fotoresist", inclusive dissolvendo lateralmente a camada, o que produz o perfil mostrado na figura 6.4c.

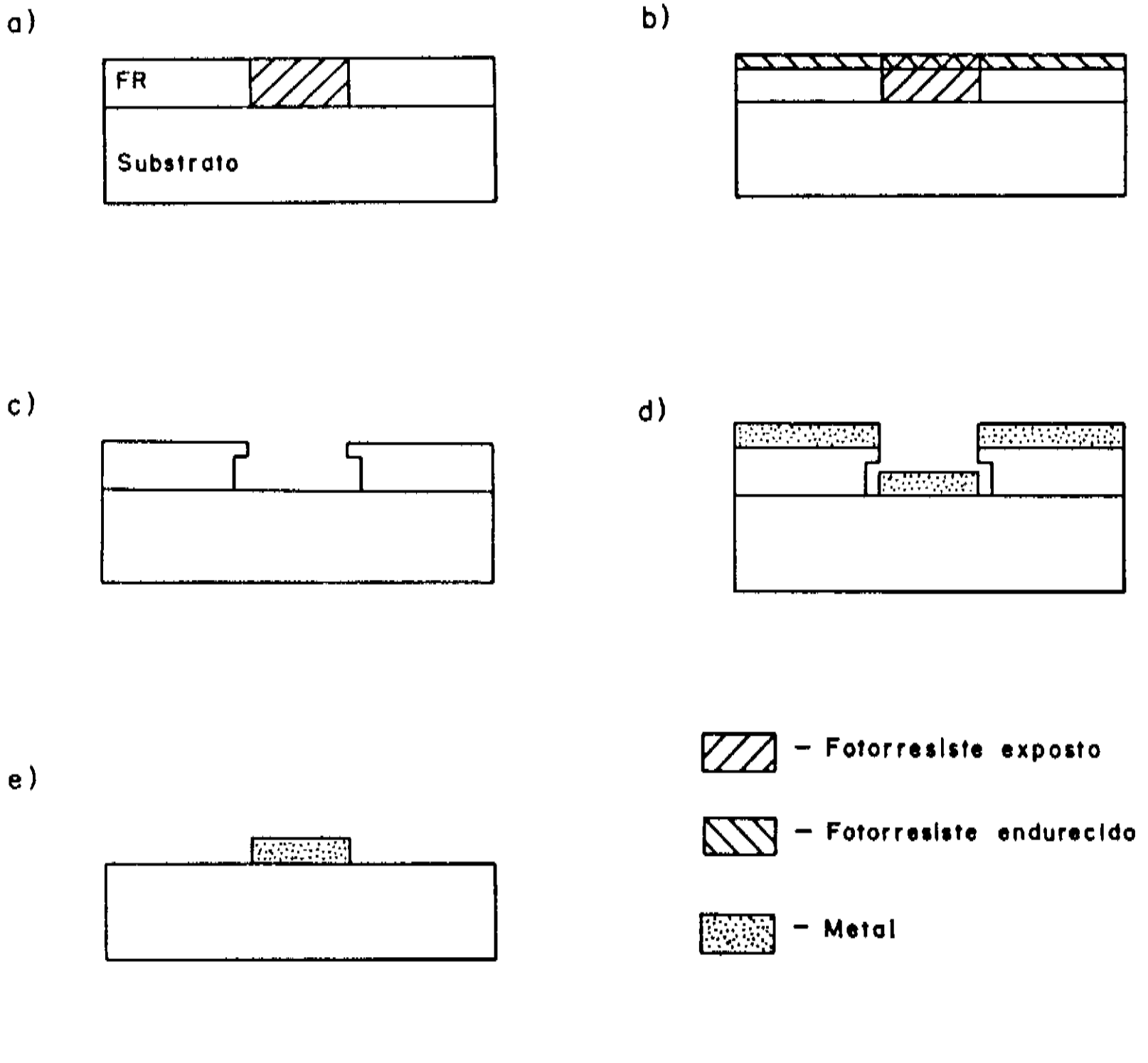

Figura 6.4 - Sequência de fotogravação para "lift-off". a) cristal coberto com "fotoresist" exposto, b) superfície do "fotoresist" endurecida, c) "fotoresist" revelado, d) cristal após evaporação de metal e e) cristal após "lift-off".

A metalização é feita logo após a revelação do cristal, através da evaporação térmica do metal (a ser depositado) próximo ao cristal. Isso provoca condensação sobre o cristal (regiões reveladas) e sobre o "fotoresist" (regiões não reveladas), como mostra a figura 6.4d, resultando em finas camadas metálicas. A evaporação é feita em vácuo (10^{-6} torr), em uma evaporadora

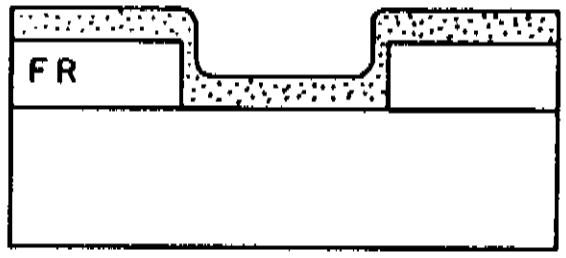

Figura 6.5 - Evaporação de metal em um cristal sem fotogravação para "lift-off".

apropriada (marca Edwards).

O próximo passo é a dissolução do "fotoresist" em acetona (solvente do "fotoresist"). Isso faz com que o metal depositado sobre o "fotoresist" se desprenda do cristal ("lift-off"), deixando apenas o metal depositado sobre a superfície do cristal, como mostra a figura 6.4e.

Podemos observar que o perfil do "fotoresist" da figura 6.4c é necessário, pois como vemos na figura 6.4d, ele favorece a penetração da acetona entre o cristal e o metal depositado sobre o "fotoresist", disolvendo-o e desligando o metal. Na figura 6.5 podemos ver o que aconteceria caso fosse feita uma fotogravação normal e evaporação do metal sobre o cristal. O metal também seria depositado nas paredes do "fotoresist", impedindo a penetração da acetona, o que tornaria difícil, ou mesmo impossível o "lift-off".

6.4 - Formação das "Mesas"

Podemos dizer que são formadas duas estruturas do tipo "mesa" no cristal. As primeiras são as marcas de alinhamento, cuja função é servir de referência para o posicionamento das máscaras seguintes (em número de 4) usadas nos processos. As marcas de alinhamento são os quadrados com quatro pontas em seu interior, como pode ser visto na figura 6.2. Eles estão localizados na parte inferior do "chip" de testes.

A segunda estrutura do tipo "mesa" são as "mesas" propriamente ditas dos transistores, ou região ativa destes.

6.4.1 - Marcas de Alinhamento

As marcas de alinhamento são formadas através de corrosão química localizada do cristal, usado para isso a fotogravação normal, porém com um tempo de exposição (fotogravação) de apenas 6 segundos, em comparação com os 8 segundos normalmente usados. Isso é feito para que as marcas de alinhamento fiquem ligeiramente maiores no cristal fotogravado, facilitando a sua visualização e consequentemente o alinhamento das outras máscaras. Nessa etapa de fotogravação usamos a máscara de marcas de alinhamento (número 1 do conjunto de máscaras).

Nesta etapa o alinhamento da máscara com o cristal é feito posicionando-se a borda clivada do cristal sob um lado do quadrado que envolve os vários circuitos dentro do "chip", já que ainda não

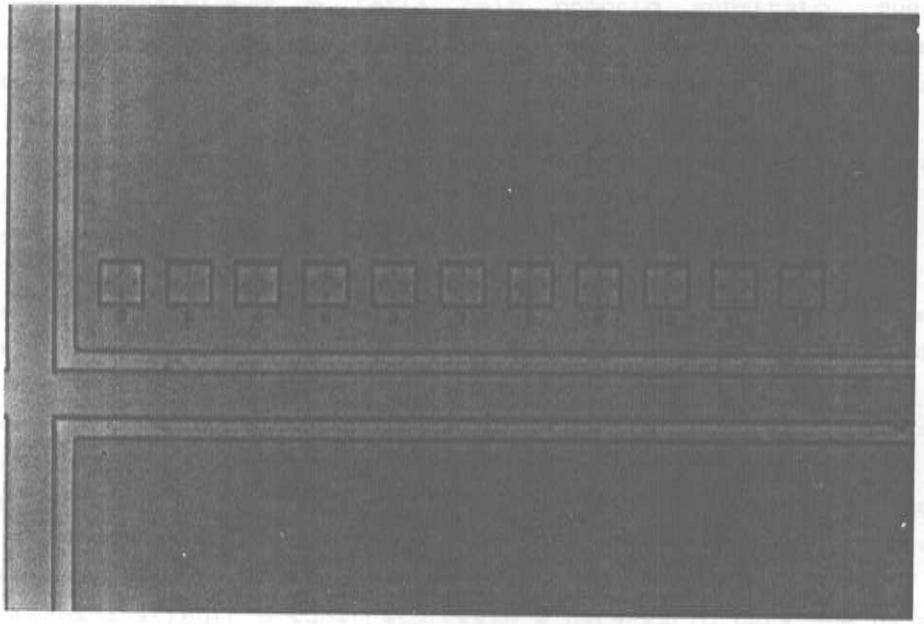

Figura 6.6 - Fotografia das marcas de alinhamento no cristal.

existe nenhuma marca de referência na superfície do cristal.

Após a fotogravação o cristal é atacado por uma solução de ácido sulfúrico, água oxigenada e água DI, numa proporção de 1 parte, 1 parte e 100 partes (por volume), respectivamente (H_2SO_4:H_2O_2:H_2O(DI), 1:1:100). Esse ataque químico é feito por dois minutos e meio, e corroi em média uma camada de 0,6μm de profundidade, criando a estrutura do tipo "mesa" das marcas de alinhamento.

Após o ataque químico o cristal é enxaguado em água DI para que sejam retirados os resíduos da solução ácida, e seco em um jato de nitrogênio. A seguir ele é enxaguado em acetona para que seja retirado o "fotoresist" do cristal, e novamente seco com o nitrogênio. Na figura 6.6 vemos uma fotografia das marcas feitas no cristal.

6.4.2 - "MESAS"

Como já discutimos anteriormente, a isolação elétrica entre os transistores é feita pelo próprio substrato, que é semi-isolante. Isso leva à necessidade de corroer quimicamente o cristal, retirando a camada difundida de onde ela não é necessária, criando, portanto, as "mesas" dos transistores.

Nesta etapa é usada novamente a fotogravação normal para a definição das áreas a serem corroídas. A máscara utilizada é a de "mesas" para os transistores (máscara número 14).

Após a fotogravação o cristal é corroído em uma solução de ácido fluorídrico, água oxigenada e água DI, numa proporção de 1 parte, 1 parte e 5 partes, respectivamente (HF:H_2O_2:H_2O(DI), 1:1:5). Após a corrosão o cristal é enxaguado em água DI e seco em nitrogênio. O tempo de corrosão é de 25 a 30 segundos, o que corresponde a uma camada retirada de aproximadamente 0,6μm de profundidade. Na figura 6.7 podemos ver a sequência de formação das "mesas", e na figura 6.8 vemos uma fotografia das "mesas" sobre o cristal. Observe na figura 6.7 que a corrosão é feita de tal forma a atingir o substrato (isso é necessário para que haja a isolação entre os transistores).

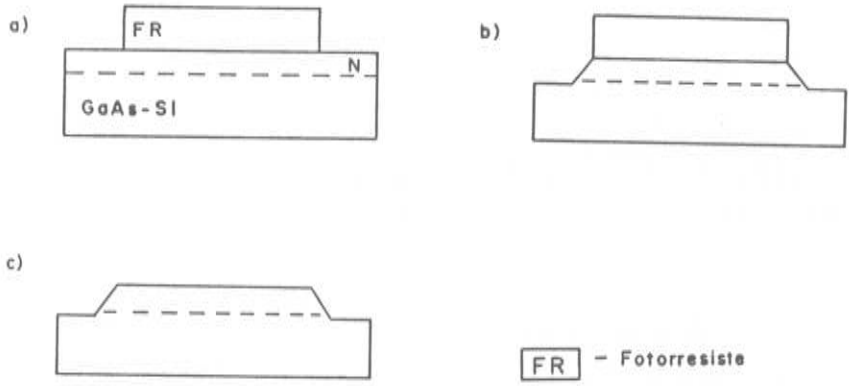

Figura 6.7 - Sequência de formação das "mesas". a) fotogravação, b) corrosão química, e c) "mesa" formada após a dissolução do "fotoresist".

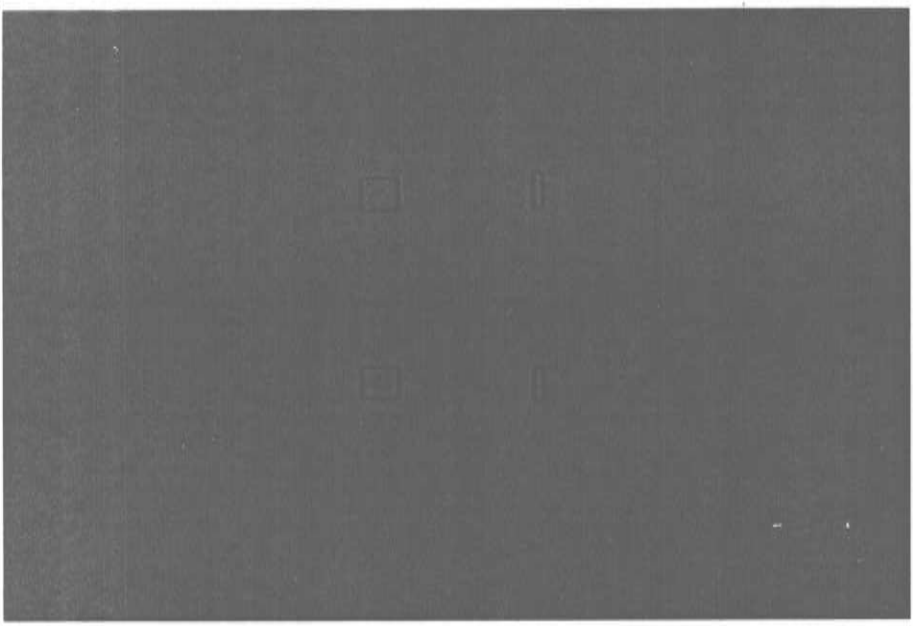

Figura 6.8 - Fotografia mostrando as "mesas" formadas.

Após a corrosão podemos medir a resistência elétrica entre dois pontos do substrato, para determinar se ainda há ou não uma camada condutora. Em caso afirmativo o cristal pode ser atacado pelo ácido por mais alguns segundos. Uma vez estando as "mesas" isoladas eletricamente entre si, o "fotoresist" é retirado com acetona, e passa-se à próxima etapa do processo.

6.5 - Metalização de Contato Ôhmico

O passo seguinte no processamento dos cristais é a formação das metalizações de contato ôhmico de dreno e fonte dos transistores. A função dessa metalização é produzir um contato ôhmico, como o próprio nome já diz, de baixa resistência para os terminais de dreno e fonte, permitindo assim, uma conexão elétrica adequada a esses terminais.

A fotogravação feita é para "lift-off", usando-se a máscara de contato ôhmico (máscara número 6). Após a fotogravação o cristal é submetido a um ataque químico com ácido clorídrico e água DI ($HCl:H_2O$(DI), 1:3) que tem a finalidade de retirar o óxido existente na superfície do cristal, melhorando (diminuindo) a resistência elétrica de contato entre o metal e o cristal.

Após o ataque químico com ácido clorídrico é feita a evaporação do metal de contato e o seu posterior "lift-off". Nas

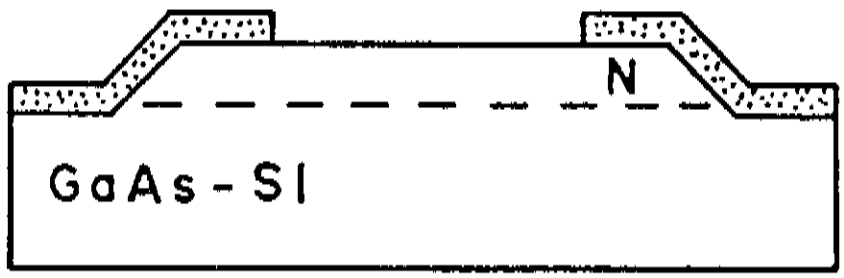

Figura 6.9 - Vista em corte da "mesa" coberta parcialmente pelo metal de contato ôhmico de dreno e fonte.

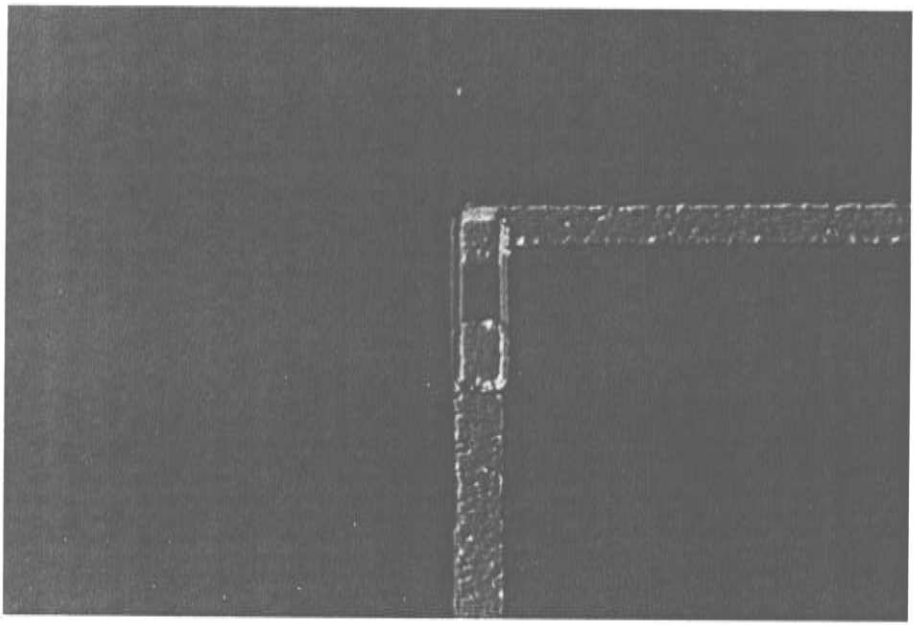

Figura 6.10 - Fotografia mostrando a metalização de contato de dreno e fonte após o "alloy".

figuras 6.9 e 6.10 vemos respectivamente um desenho do transistor em corte, após a metalização, e uma fotografia mostrando a metalização feita. Observe na figura 6.9 que a metalização sobe sobre a "mesa", fazendo contato elétrico com a superfície do cristal, onde a camada difundida possui alta concentração, o que produz um contato de baixa resistência.

6.5.1 - Metalização Utilizada

A metalização utilizada para o contato ôhmico é uma composição de ouro, germânio e níquel. A evaporação desses metais é feita em duas etapas. A primeira camada evaporada é de ouro, germânio e níquel, na proporção de 75,5%, 15,1% e 9,4%, respectivamente. A segunda camada evaporada contém apenas ouro e níquel, na proporção de 92,3% e 7,7%, respectivamente. A espessura total

das duas camadas é de cerca de 5000 Å (as duas camadas tem espessuras aproximadamente iguais).

6.5.2 - "ALLOY" DE CONTATO

O simples fato de depositar o metal de contato ôhmico na superfície do cristal semicondutor não resulta em um contato ôhmico de baixa resistência, ao contrário, produz uma barreira "Schottky". Neste caso, para a formação do contato ôhmico é necessário o aquecimento do cristal. A esse processo de aquecimento damos o nome de "alloy" de contato ou sinterização de contato.

Durante o aquecimento o germânio existente no filme metálico irá difundir no cristal, formando uma camada superficial muito dopada (concentração maior que $1 \times 10^{19} cm^{-3}$), e bastante rasa (da ordem de centenas de angstrons). Essa camada irá permitir o tunelamento dos elétrons entre o metal e o semicondutor, produzindo um contato ôhmico de baixa resistência.

O uso do níquel na metalização de contato tem a finalidade de evitar o aglutinamento do ouro durante o aquecimento. A escolha do ouro para a metalização é devida à sua alta condutividade elétrica.

A temperatura utilizada no "alloy" é de 430°C por 30 segundos. O "alloy" é feito no forno rápido, e é usado como suporte do cristal um "wafer" de silício de duas polegadas de diâmetro por meio milímetro de espessura.

6.6 - METALIZAÇÃO DE "GATE"

A próxima etapa na fabricação dos transistores e a formação do terminal de "gate". Ela começa com a fotogravação para "lift-off" do terminal de "gate" (máscara 12 para transistores de depleção), seguida de um ataque químico que tem a finalidade de ajustar a corrente de saturação do transistor e por fim a evaporação do metal de "gate" e seu posterior "lift-off". A esse

tipo de "gate" é dado o nome "gate recessed".

6.6.1 - Ajuste da Corrente de Saturação

Após a formação das "mesas" e metalização de contato ôhmico já podemos observar a saturação da corrente de dreno dos transistores (não é necessária a presença do "gate" para que ocorra a saturação). Neste caso a curva característica da corrente de dreno em função da tensão de dreno é igual à curva de um FET para um único valor da tensão de "gate".

Como observamos no capitulo 4, a difusão de enxofre cria uma camada difundida, cuja concentração superficial é alta (em torno de $1 \times 10^{18} cm^{-3}$), diminuindo em direção ao centro do cristal. Podemos dessa forma considerar que praticamente temos uma camada superficial do tipo N^+, e abaixo dela uma camada N.

A existência da camada N^+ na região do canal irá fazer com que os transistores apresentem valores altos da corrente de saturação, pois temos uma alta quantidade de cargas no canal (camada N^+). Essa alta corrente de saturação pode provocar a queima do transistor, já que produz uma alta dissipação alta de potência, que não é suportável por ele. Para os cristais processados a corrente de saturação após a metalização de contato ôhmico sempre esteve maior que 10mA. Para diminuir essa corrente é feita a corrosão química da região do canal (região exposta do cristal pela fotogravação), o que elimina a camada N^+, deixando apenas a camada N. Dessa forma o canal fica com um menor número de cargas, o que coresponde a uma menor corrente de saturação.

A corrosão para o ajuste da corrente de saturação é feita com uma solução de ácido sulfúrico, água oxigenada e água DI ($H_2SO_4 : H_2O_2 : H_2O$(DI), 1:1:300). O tempo de corrosão é variável, ficando entre 10 segundos e 45 segundos, tipicamente. A corrosão é feita em passos de 5 segundos, sendo que entre cada passo é feita uma medida da corrente para verificar se ela já atingiu o valor desejado. A medida é feita com o uso de um analisador de parâmetros de dispositivos semicondutores (marca HP, modelo 4145B) e de dois microposicionadores que portam duas agulhas (de ponta

final) usadas para fazer contato elétrico com o transistor, possibilitando a medida da corrente. Após cada ataque químico o cristal é enxaguado em água DI para retirar os resíduos do ácido e seco em um jato de nitrogênio. A corrente de saturação ajustada é de 1mA. Observamos que após a colocação do metal de "gate" a corrente de saturação diminui, devido à formação de uma barreira de potencial entre o metal e o semicondutor, o que depleta parte do canal.

6.6.2 - Metalização Utilizada para o "Gate"

O metal utilizado para o "gate" dos transistores foi o alumínio, que possui uma barreira de potencial intrínseco de 0,7eV com o GaAs. A espessura da camada utilizada para o "gate" foi de aproximadamente 3000Å. Nas figuras 6.11a, b ,c e d vemos a sequência da colocação do "gate". Na figura 6.11a o cristal já

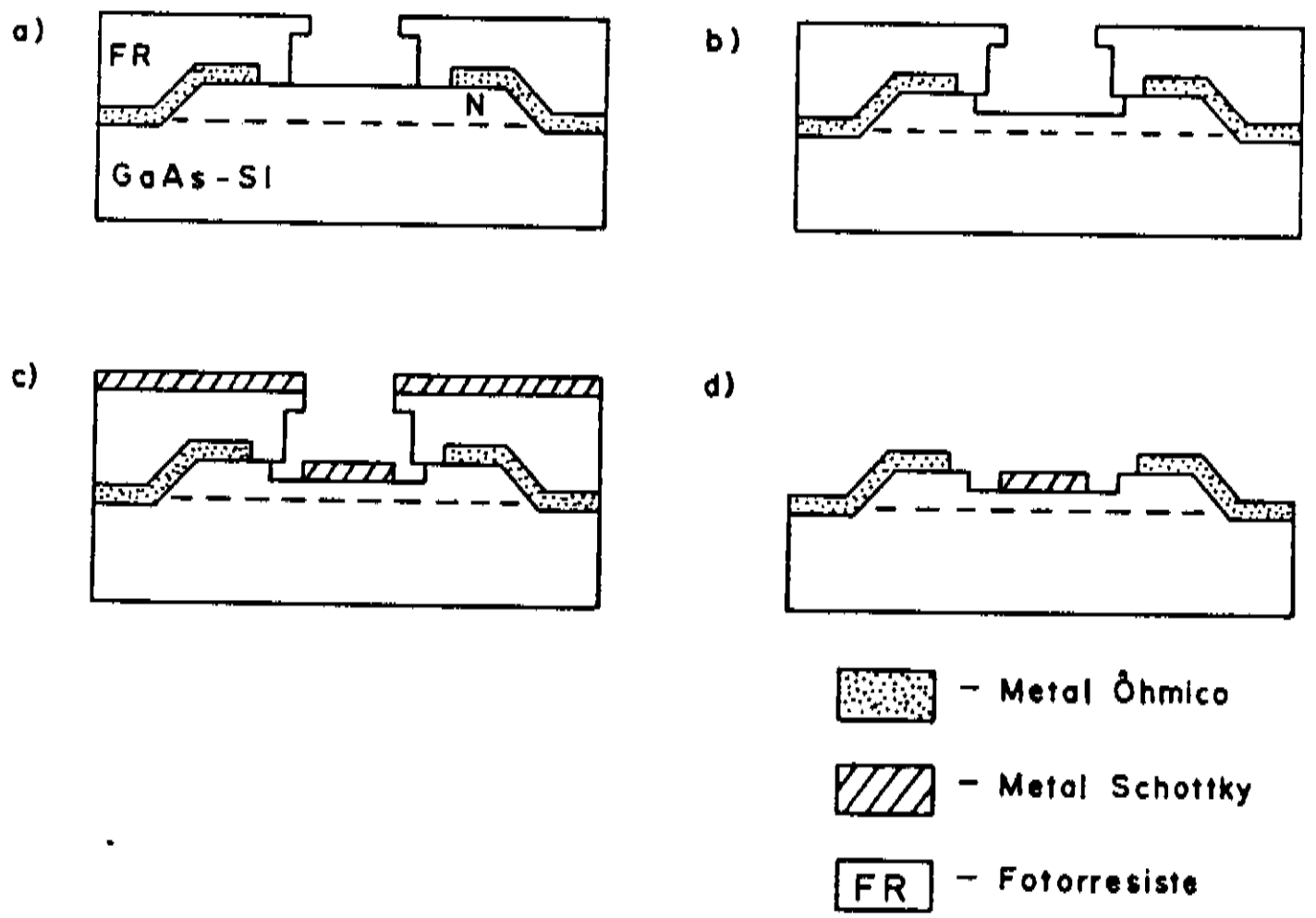

Figura 6.11 - Sequência de formação do terminal de "gate". a) fotogravação para "lift-off", b) cristal após a corrosão química, c) após a evaporação de alumínio e d) após o "lift-off".

está fotogravado, enquanto que em 6.11b, 6.11c e 6.11d vemos respectivamente o cristal já atacado quimicamente, após a evaporação do alumínio, e após o "lift-off". Na figura 6.12 vemos uma fotografia mostrando o "gate" colocado.

Após a formação do "gate" pode ser necessário um pequeno aquecimento do cristal caso as caracteristicas da barreira "Schottky" não estejam boas (alta corrente reversa de fuga e alta tensão de barreira, entre outras). Isso é em geral devido à existência de um óxido entre o alumínio e o cristal, óxido este que é eliminado pelo aquecimento. Nesse caso a temperatura utilizada no aquecimento é de 200°C por 20 minutos. O aquecimento além de eliminar o óxido melhora também a distribuição das tensões de barreira (ela tende a ficar mais centrada em torno do valor ideal, 0,7eV) observadas nos vários transistores de um mesmo cristal.

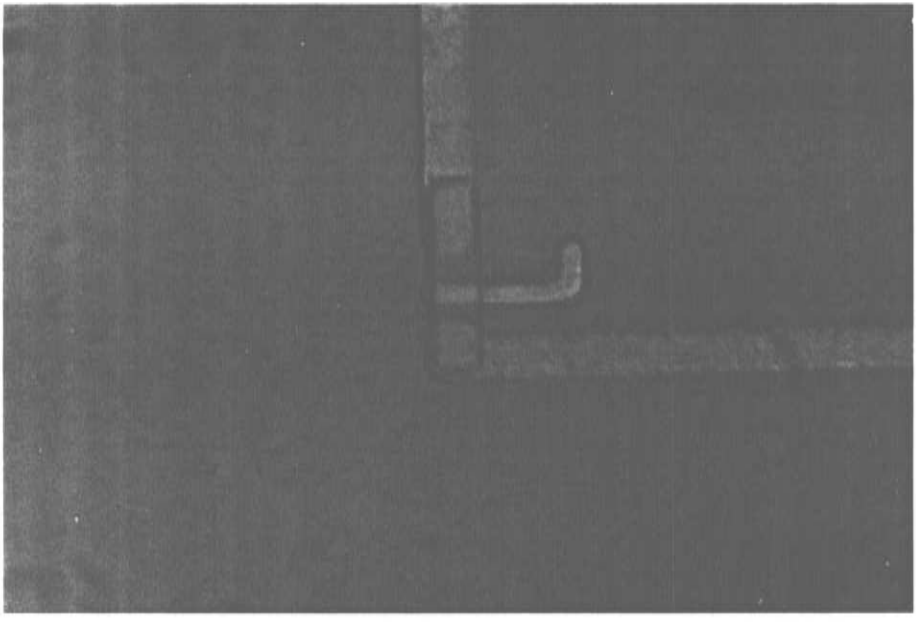

Figura 6.12 - Fotografia dos transistores mostrando o contato de "gate" já formado.

6.7 - METALIZAÇÃO DOS "PADS"

Os "pads" são pequenas áreas metalizadas na superfície do cristal que são ligadas eletricamente (através da sua própria metalização) aos terminais do transistor (dreno, fonte e "gate"). Os "pads" tem a forma de um quadrado de 100μm de lado e sua função é servir como uma área para contato externo do transistor. Os "pads" podem servir como contato elétrico para testes dos transistores (neste caso usamos os microposicionadores e as agulhas para fazer contato com os "pads") ou como áreas para a solda de fios entre os transistores e os terminais de seu encapsulamento.

Na figura 6.10 podemos ver que a metalização de contato ôhmico já forma os "pads" de dreno e fonte dos transistores (os quatro quadrados existentes na fotografia). São esses "pads" os usados para a medida da corrente de saturação dos transistores.

O atual processo de metalização irá formar os "pads" de contato para os "gates" (na figura 6.12 podemos observar que não existem "pads" para os "gates") e irá remetalizar os "pads" de dreno e fonte. Essa remetalização é feita para produzir um contato elétrico de baixa resistência, já que a metalização de contato ôhmico possui uma resistência de placa relativamente alta após a realização do "alloy" (entre 1 e 2Ω/□), o que gera um contato elétrico de resistência alta.

A metalização dos "pads" começa com a fotogravação para "lift-off" usando-se a máscara apropriada (máscara número 7, de metalização dos "pads"), seguida da evaporação do metal dos "pads" e o seu posterior "lift-off".

Para a metalização usamos dois metais. Primeiro é evaporado cromo que forma uma camada de aproximadamente 600Å de espessura, e após o cromo é evaporado ouro, que forma uma camada de 4400Å, aproximadamente. A função do cromo na metalização é melhorar a aderência do ouro ao cristal, já que o cromo possui uma maior aderência ao GaAs do que o ouro. Na figura 6.13 vemos uma fotografia mostrando os transistores completos, com a metalização dos "pads".

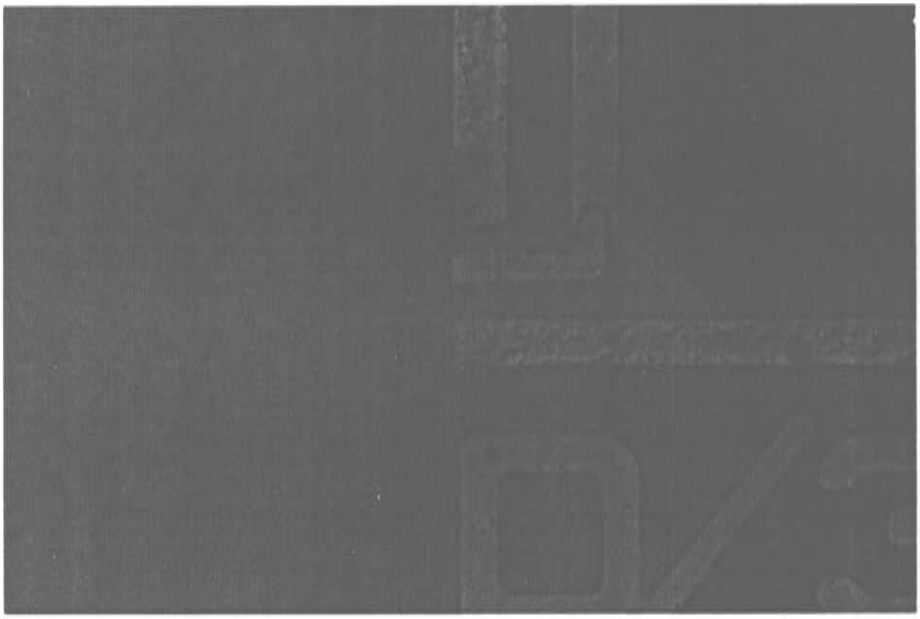

Figura 6.13 - Fotografia dos transistores completos com a metalização para "pads" já colocada.

6.8 - Testes e Caracterização dos Transistores

Os testes e caracterizações foram feitos com um analisador de parâmetros de dispositivos semicondutores, marca HP, modelo HP-4145B. Esse equipamento é constituido basicamente por um conjunto de fontes de tensão e corrente programáveis, um tubo de raios catódicos para visualização das curvas obtidas, e um teclado para entrada de dados e programação.

Os primeiros testes foram realizados com o objetivo de verificar quais transistores dentro do "chip" estavam funcionando. Isso foi feito simplesmente traçando as curvas $I_D \times V_D$ para vários valores de V_G. Uma vez verificado que o transistor está funcionando, passamos a carcterizá-lo. Na caracterização obtemos a transcondutância do transistor em função da tensão de "gate" para

vários valores da tensão de dreno, de onde podemos tirar a máxima transcondutância normalizada, que é o valor deste parâmetro que o transistor teria se a largura de seu "gate" fosse 1mm (é importante usarmos a transcondutância normalizada para podermos comparar MESFET's de larguras de "gate" diferentes). Outra curva obtida na caracterização é a da raiz da corrente de dreno em função da tensão de "gate", que deve ser uma reta, como podemos concluir da equação (5.37). Essa reta deve cortar o eixo V_G no valor V_{th}, de onde podemos extrair V_{po}, que é um importante parâmetro. Por fim traçamos curvas de V_D em função de I_G ("gate" polarizado diretamente) para vários valores de I_D, partindo de zero. Dessas curvas obtemos as resistências série de dreno e fonte [1].

Na figura 6.14 podemos ver as curvas V_DxI_D para vários valores de V_G para um transistor com 3μm de comprimento de "gate". Observe que conseguimos correntes de 900μA com um transistor relativamente estreito, de apenas 5μm de largura. Já na figura 6.15 temos a transcondutância em função da tensão de "gate" para vários valores de V_D. Nesse gráfico as três curvas superiores correspondem à região de saturação, enquanto que as demais à região linear. Observe que a transcondutância máxima é de cerca de 800μS, que normalizada para 1mm equivale a 160mS. Na figura 6.16 temos a raiz da corrente de dreno em função de V_G. Uma reta que passe pela curva do gráfico irá interceptar o eixo V_G em V_{th}. Desse gráfico obtemos um valor de 1,63V para V_{th}. Na figura 6.17 temos um gráfico que mostra a resistência elétrica entre dreno e fonte em função da resistividade do canal (em unidades arbitrárias). Esse gráfico é obtido a partir de cálculos de resitência do canal em função de V_G, usando as curvas de V_DxI_D, para valores de V_D pequenos. A resistividade do canal é obtida por cálculo numérico conhecendo-se o perfil de cargas sob o "gate". Esse perfil é obtido pelo Polaron, para o cristal difundido. Devemos eliminar do perfil a região de superfície ("gate recessed"). Nos apêndices temos o programa FINDVPO que calcula qual é a profundidade da camada retirada, a partir de uma informação do valor de V_{po} e do perfil do Polaron. Conhecendo-se essa profundidade podemos calcular a resistividade do canal. Isso

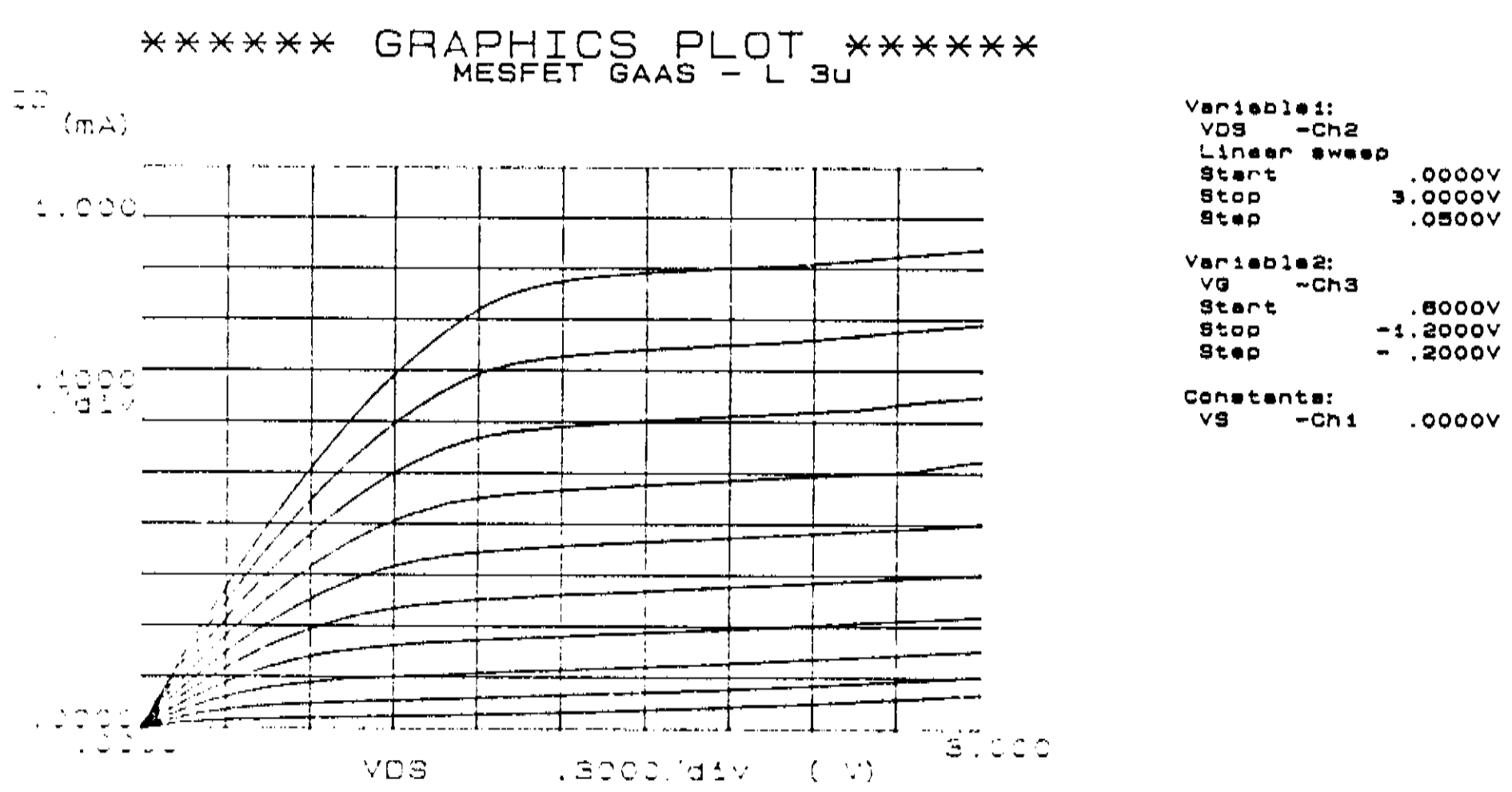

Figura 6.14 - Curvas V_DxI_D para vários valores de V_G obtidas de um transistor de 3μm de comprimento de "gate".

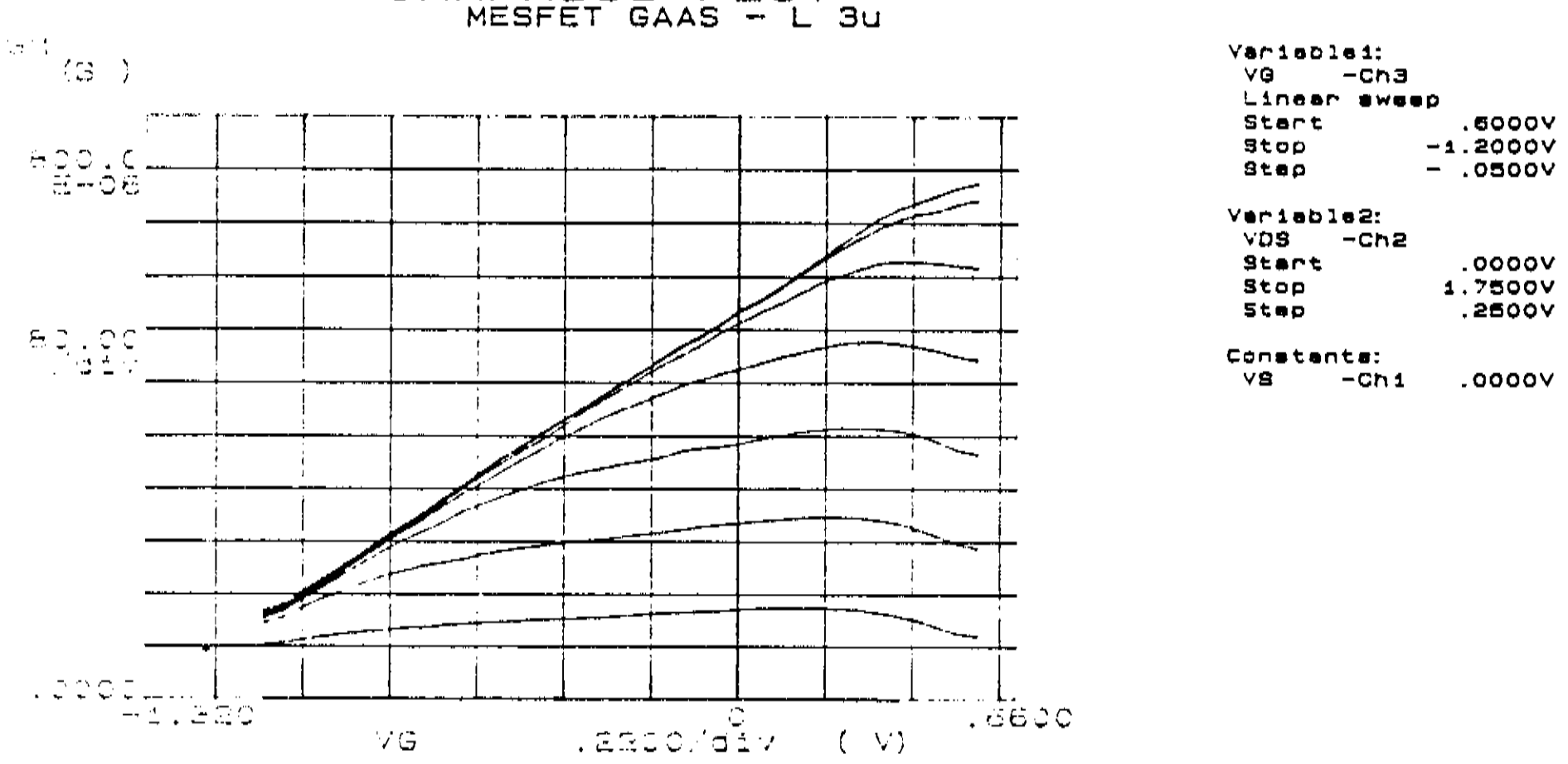

Figura 6.15 - Curvas da transcondutância em função de V_G para vários valores de V_D, de 0V até 1,75V. As três curvas superiores correspondem à região de saturação.

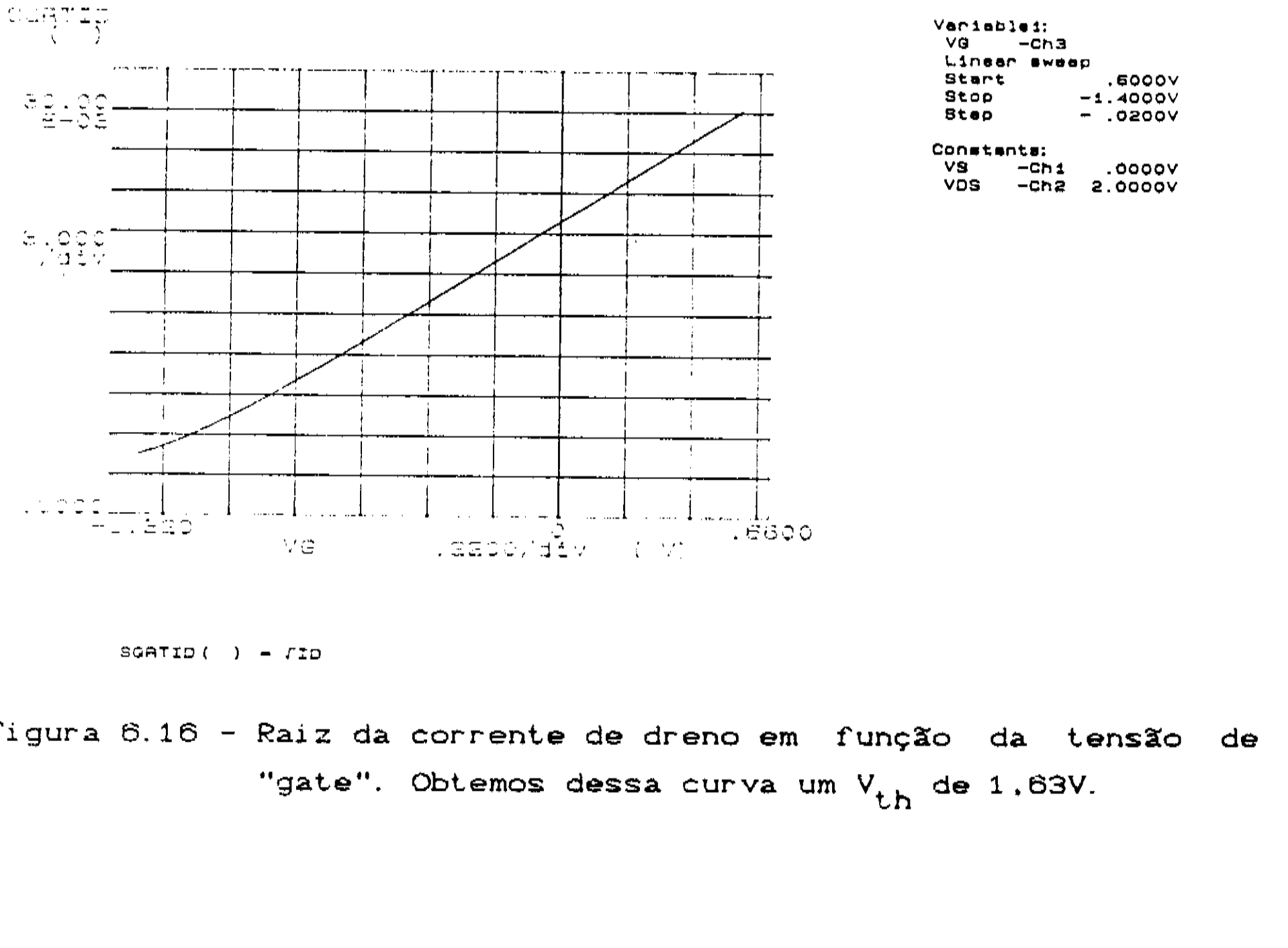

Figura 6.16 - Raiz da corrente de dreno em função da tensão de "gate". Obtemos dessa curva um V_{th} de 1,63V.

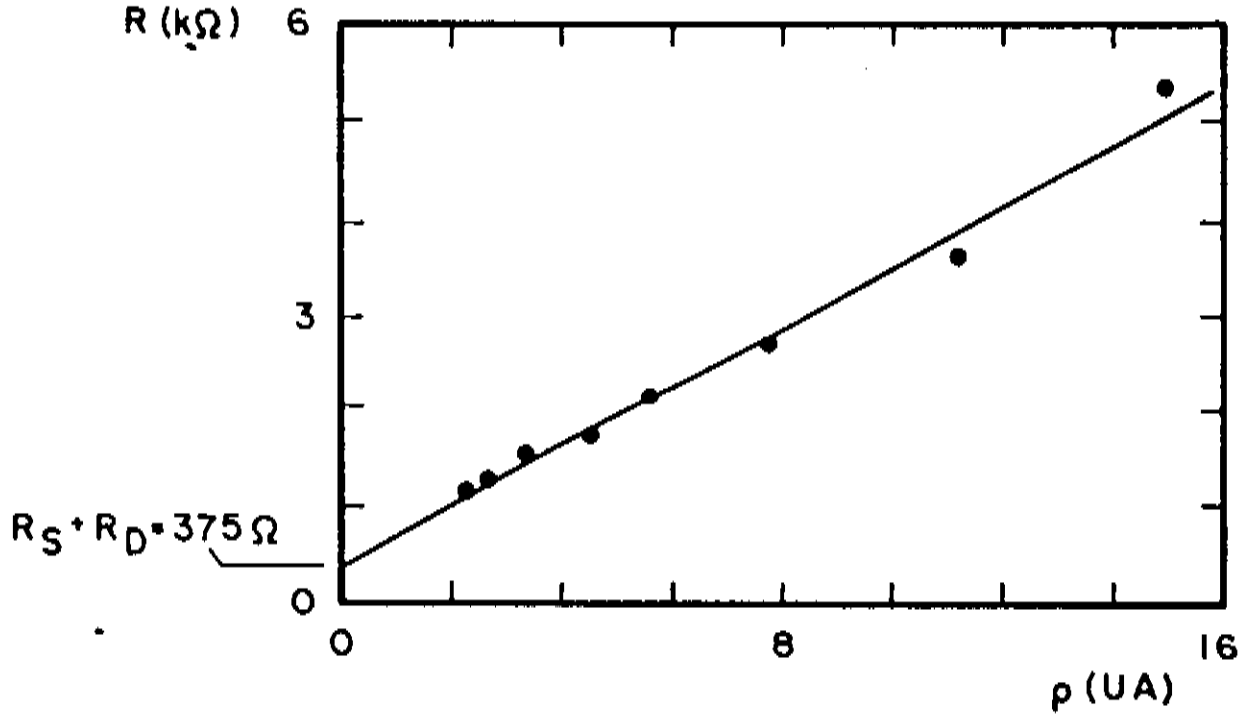

Figura 6.17 - Resistência elétrica entre dreno e fonte em função da resistividade do canal. Obtemos desse gráfico uma soma das resistências série de 375Ω.

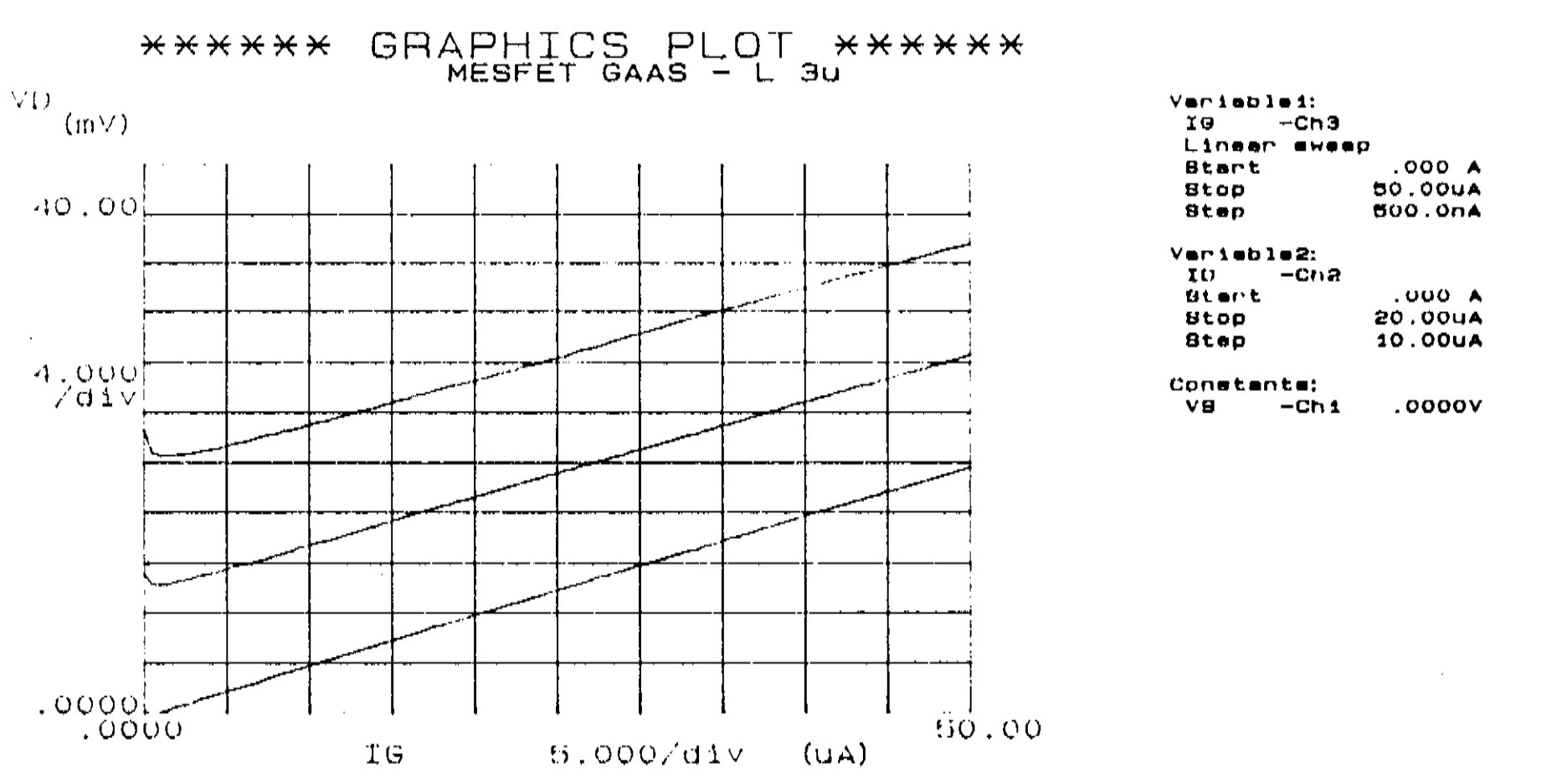

Figura 6.18 - Tensão de dreno em função da corrente de "gate" para vários valores de I_D. Nesse gráfico a fonte está aterrada.

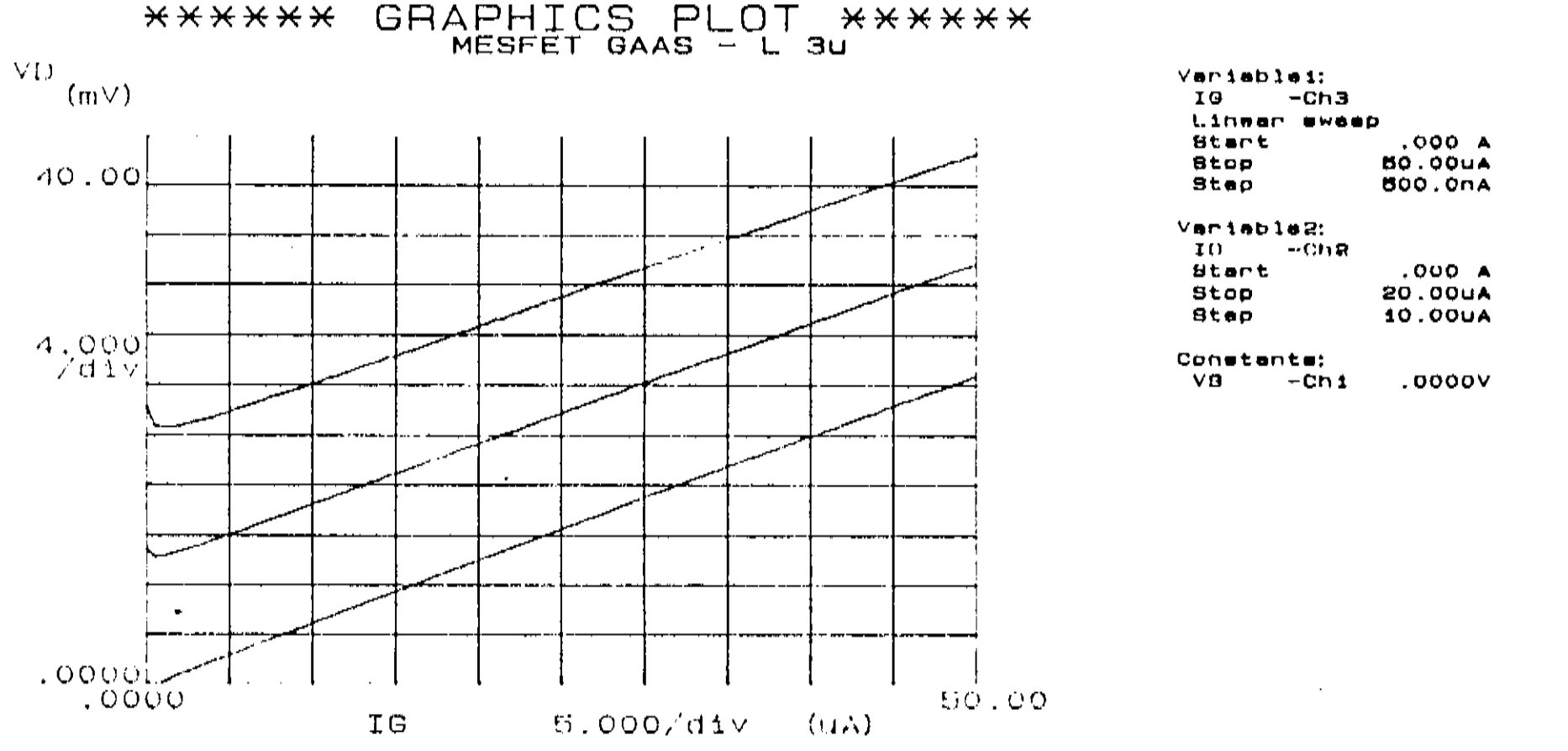

Figura 6.19 - Tensão de dreno em função da corrente de "gate" para vários valores de I_D. Nesse gráfico a fonte foi trocada com o dreno.

é feito pelo programa SIGMA, que a calcula em função da tensão de "gate". Usando os dois programas citados calculamos a resistividade para os mesmos valores de V_G usados na medida de resistência. Obviamente o gráfico obtido deve ser uma reta, que cruza o eixo de resitência na soma das resitências série de dreno e fonte. Se necessário podemos variar V_{po} para produzir uma reta no gráfico. Na figura 6.17 obtemos uma soma de 375Ω. Na figura 6.18 temos o gráfico da tensão de dreno em função da corrente de "gate" para vários valores de I_D. Como mostrado na referência 1 a inclinação das curvas vale $R_S+R_{ch}/2$, onde R_S é a resistência série de fonte e R_{ch} é a resitência do canal. Das curvas para I_D diferente de zero podemos obter a soma $R_S+R_D+R_{ch}$, que sera dada pelo valor de V_D (para I_G igual a zero) dividido por I_D. Na figura 6.19 obtemos as mesmas curvas só que com o dreno e fonte trocados. Da figura 6.18 temos,

$$R_S + R_{ch}/2 = 404\Omega \qquad (6.1a)$$

$$R_S + R_D + R_{ch} = 958\Omega \qquad (6.1b)$$

Da figura 6.19 temos,

$$R_D + R_{ch}/2 = 514\Omega \qquad (6.2a)$$

$$R_S + R_D + R_{ch} = 965\Omega \qquad (6.2b)$$

Subtraindo (6.2a) de (6.1a) obtemos,

$$R_D - R_S = 110\Omega \qquad (6.3)$$

De onde tiramos R_D e R_S já que sabemos a sua soma,

$$R_D = 242\Omega$$

$$R_S = 132\Omega$$

Com a soma de R_D e R_S obtemos de (6.1b) e (6.2b) um valor de médio

de R_{ch} de,

$$R_{ch} = 586\Omega$$

que substituido em (6.1a) e (6.2a) dá,

$$R_D = 221\Omega$$

$$R_S = 111\Omega$$

o que concorda muito bem com os valores anteriores.

Devido a problemas no processamento nos transistores de comprimento de "gate" de $2\mu m$ não obtivemos bons resultados da sua caracterização. Esses transistores apresentaram alta corrente de fuga no "gate", devida a uma alta concentração de elétrons no contato Schottky. Essa alta concentração resultou de problemas na decapagem antes da colocação do "gate". A decapagem não foi completa, pois provavelmente a tensão superficial da solução a impediu de penetrar na abertura do fotorresiste.

6.9 - Conclusões

Mostramos neste capitulo que as camadas tipo N obtidas da difusão de enxofre por processamento térmico rápido podem ser aplicadas na fabricação de transistores MESFET de GaAs. Obtivemos da caracterização uma transcondutância normalizada de 160mS/mm, e a soma das resistências série de dreno e fonte de 375Ω. Comparando os nossos resultados com os de Prince et al. [2], que obtiveram uma transcondutância de 200mS/mm para um transistor de comprimento de "gate" de 2μm, com canal obtido por difusão de enxofre, vemos que a nossa transcondutância é menor, porém temos um "gate" mais longo, e caso escalássemos o seu comprimento para 2μm, obteríamos uma transcondutância de aproximadamente 240mS/mm (ela varia com o inverso do comprimento de "gate"), o que é um valor muito bom. Dos resultados obtidos da caracterização concluimos que os

transistores fabricados são de boa qualidade, já que possuem transcondutância relativamente alta para o seu comprimento de "gate", e baixa resistência série de dreno e fonte.

Podemos concluir que a difusão de enxofre resulta em transistores de maior transcondutância ao compararmos os nossos resultados aos obtidos da fabricação de transistores de canal implantado, referência 3, onde o transistor possui comprimento de "gate" de 1.3μm, e transcondutância de 80mS/mm. Isso pode ser explicado observando-se a barreira Schottky de canal. Para movimentar a interface entre o canal condutor e a região de depleção de uma distância Δy_d, precisamos de uma tensão ΔV (que pode ser de "gate"), que vale (equação (5.8)),

$$\Delta V = \Delta y_d \frac{y_d \rho(y_d)}{\varepsilon} \qquad (6.4)$$

O transistor terá uma transcondutância tanto maior quanto mais fácil conseguirmos movimentar a interface, pois estaremos conseguindo depletar mais o canal, o que corresponde a uma maior variação de corrente. Dessa forma, para um perfil de difusão, que possui maior concentração na superfície, a interface estará mais próxima do "gate", do que no caso de um perfil de implante, que possui maior concentração muito abaixo do "gate". Precisaremos portanto de menor variação de tensão de "gate" no canal difundido para mover a interface de uma mesma distância, o que irá implicar em maior variação de corrente de canal e consequentemente maior transcondutância.

6.10 - Referências

[1] - K. Lee, M. Shur, K. W. Lee, T. Vu, P. Roberts, e M. Helix, "A new Interpretation of "End" Resistence Measurements", *IEEE Electron Device Letters*, vol. EDL-5, n. 1, p. 5 (1984).

[2] - M. Oren e F. C. Prince, "GaAs IC Fabrication Using a Novel Open-Tube Sulphur Diffusion Technique", *Elect. Letters*, vol. 22, n. 4, p. 221 (1986).

[3] - T. H. Chen e M. S. Shur, "Current-Voltage Characteristics of Ion-Implanted GaAs MESFET's", *IEEE Trans. Electron Devices*, vol. ED-32, n. 1, p. 18 (1985).

CAPITULO VII

CONCLUSÕES E TRABALHOS PROPOSTOS

7.1 - INTRODUÇÃO

Passamos a descrever neste capitulo os resultados e conclusões gerais obtidas da montagem do forno, da realização das difusões e da sua aplicação na fabricação de transistores MESFET de GaAs. Baseado nos resultados obtidos da difusão propomos novos trabalhos que podem ser realizados com o uso do forno rápido.

7.2 - CONCLUSÕES

São vários os resultados e conclusões de cada parte deste trabalho. Do projeto e montagem do forno mostramos que é possivel a fabricação de um sistema para RTP com peças e componentes nacionais. O nosso projeto resultou em um forno que comporta amostras de até duas polegadas de diâmetro, e atinge uma máxima temperatura de 950°C com taxas de subida de temperatura de 80°C/s e 125°C/s para um disco de grafite e uma lâmina de silício, respectivamente. Devido ao fato da câmara do forno ser radial, com as lâmpadas relativamente afastadas da amostra, isso resulta em taxas de aquecimento e temperaturas finais mais baixas do que nos sistemas comerciais de RTP. Podemos aumentar a taxas e a máxima temperatura alterando o projeto da câmara, de radial para retangular, com as lâmpadas bem mais próximas da amostra. O controle de temperatura do forno foi implementado com o uso de um microcomputador de 8 bits e uma interface, no forno, para a medida de temperatura e controle de potência das lâmpadas. O "software" de controle usa um algoritmo do tipo PID digital, com a possibilidade de programarmos dois patamares de temperatura. Fora o seu uso na difusão o forno possui aplicações para o recozimento

de wafers implantados e sinterização de contatos.

Das experiências de difusão podemos concluir que foi alcançado o principal objetivo deste trabalho, desenvolver um processo de difusão de enxofre por RTP. Para esse desenvolvimento foi necessário o projeto de um porta-amostra de grafite que recebesse a fonte de difusão e a amostra. Esse porta-amostra possui uma baixa "inércia térmica", caracterizada por taxas de aquecimento em torno de 80°C/s. A fonte para difusão foi preparada misturando-se GaAs e Ga_2S_3, ambos na forma de pó. A exata composição foi escolhida a partir do critério de Matino [1], o que garante difusões reprodutíveis. Mostramos que conforme a composição da fonte existirá uma faixa de temperatura mais adequada para o seu uso, onde conseguimos realizar as difusões sem a criação de depósitos de sulfetos na superfície da amostra. Todas as fontes experimentadas resultaram em amostras dopadas com mobilidade média de $1800 cm^2/Vs$, o que é coerente com a concentração de superfície de cargas, da ordem de 1 a $3x10^{18} cm^{-3}$. Essa concentração de carga corresponde a uma concentração de enxofre da ordem de $10^{19} cm^{-3}$ [2]. A profundidade das camadas obtidas variou entre 0,1μm e 0,5μm, dependendo do tempo e temperatura de difusão. Achamos uma energia de ativação para a fonte de 50% de Ga_2S_3 de 4,46eV, em boa concordância com o trabalho de Matino. Fizemos uma análise de SIMS, que mostrou uma grande diferença entre os perfis de enxofre e de cargas, como já era esperado. O perfil de carga segue uma função erro complementar, enquanto que o de enxofre mostra a existência de duas frentes de difusão, como observado nas experiências de alta temperatura (de difusão de enxofre) por Tuck e Powell [3]. Os nossos resultados mostram que a difusão de enxofre em arseneto de gálio em tempos de 30 a 90 segundos, em temperaturas ao redor de 850°C, que são relativamente baixas, apresentam o mesmo comportamento que difusões em tempos longos e altas temperaturas, o que é um resultado surpreendente. Concluindo podemos dizer que desenvolvemos um processo de difusão de enxofre em GaAs que resulta em camadas rasas do tipo N^+ de alta qualidade. Essas camadas são interessantes para aplicação na fomação de contatos em regiões do tipo N, formação de junções PN rasas, e formação de

regiões de canal para MESFET'S de GaAs.

Concluimos da fabricação dos transistores MESFET que as camadas obtidas por difusão produzem dispositivos de alta transcondutância, 160mS/mm para um comprimento de "gate" de 3μm, e baixa resistência série de dreno e fonte, de 375Ω para uma largura de canal de 5μm. Esse transistor é bastante adequado para a fabricação de circuitos integrados de pequena e média escala (SSI e MSI). A alta transcondutância é devida ao perfil raso e abrupto de difusão, e também à própria forma do perfil, que apresenta uma maior concentração na superfície.

7.3 - Trabalhos Propostos

O desenvolvimento das difusões rápidas de enxofre em arseneto de gálio e a sua aplicação na fabricação de transistores MESFET abre um novo campo de possíveis trabalhos que podem ser realizados. A maioria desses trabalhos estão relacionados ao processo de difusão desenvolvido.

Neste capítulo propomos ao todo sete trabalhos. Eles são a difusão de zinco em GaAs, as difusões de enxofre e zinco em InP, a difusão de zinco em InGaAs, a difusão localizada de enxofre em GaAs, o desenvolvimento de uma fonte sólida de enxofre para uso na difusão, o estudo de processos rápidos de "alloy" de AuGeNi, e a definição de regras de projeto para circuitos integrados de GaAs feitos com camadas difundidas pelo processo rápido.

7.3.1 - Difusão de Zinco em GaAs

Da mesma forma que a difusão de enxofre em GaAs foi feita podemos difundir zinco. O processo de difusão de zinco seria igual ao de enxofre, com apenas algumas diferenças que seriam a temperatura de difusão, entre 550°C e 800°C, a fonte de difusão (Zn_3As_2) e o tipo da camada obtida, que é tipo P para zinco [4,5]. Algumas experiências já foram feitas, e obtivemos como resultados

camadas rasas de 500 Å de profundidade com uma alta concentração de $2x10^{19}$ cm^{-3}.

As aplicações dessas camadas seriam no contato ôhmico para lasers semicondutores, onde é necessária uma camada tipo P bastante rasa e de alta concentração, e a formação do terminal de "gate" para transistores JFET de canal N de GaAs. O uso de uma camada P no "gate" não degrada as características do JFET (as lacunas tem baixa mobilidade, o que faz os transistores de canal P terem baixa frequência de operação), pois ela é muito rasa. A vantagem do uso de transistores JFET é a maior excursão da tensão de "gate" para valores diretos de polarização. Para um transistor MESFET a máxima tensão direta é de 0,6 V (acima deste valor o "gate" começa a conduzir). Já para transistores JFET podemos ter até 1,2 V de polarização direta do terminal de "gate". Isto é particularmente interessante para transistores de enriquecimento, onde o gate recebe polarização direta.

7.3.2- Difusão de Enxofre e Zinco em InP

Da mesma maneira que é feita a difusão de enxofre em GaAs e que é proposta a difusão de zinco, propomos a difusão destes dois materiais no fosfeto de índio (InP). Nesse caso usaríamos um substrato semi-isolante (dopado) InP:Fe para receber a difusão. Temos evidências que as temperaturas de difusão deverão ser menores que as usadas no GaAs [6].

As aplicações dessas difusões seriam no contato ôhmico para lasers semicondutores de InP, fabricação de células solares, e fabricação de JFET's de InP. A vantagem do uso de InP no JFET é a maior mobilidade dos elétrons no InP quando comparado ao arseneto de gálio, o que produz transistores de maior frequência de operação. Não propomos transistores MESFET devido à dificuldade de se obter uma barreira Schottky no InP [7].

7.3.3 - Difusão de Zinco em InGaAs

A maior mobilidade dos elétrons nos compostos ternários, como por exemplo o InGaAs, torna-os de alto interesse na fabricação de transistores do tipo FET. Propomos a realização da difusão de zinco (dopante tipo P) no InGaAs para a aplicação na formação do terminal de "gate" de um transistor JFET de canal N feito com InGaAs. O canal tipo N seria formado pelo crescimento por LPE (epitaxia por fase líquida) de uma camada de InGaAs com dopagem da ordem de $1x10^{17}$ cm^{-3} e profundidade da ordem de 0,2 μm sobre um substrato de fosfeto de índio semi-isolante (InP:Fe). A seguir seria formado o terminal de "gate" através de uma difusão de zinco com 500 Å de profundidade. Completando o dispositivo seriam feitas as metalizações de "gate", dreno e fonte.

7.3.4 - Difusão Localizada de Enxofre em GaAs

Como vimos no capitulo 6, a isolação entre os vários transistores no mesmo substrato é feita através do próprio substrato que é semi-isolante. Para isso é necessária a corrosão química do substrato para retirar a camada difundida onde ela não é desejada. Isso cria estruturas do tipo "mesa" no substrato.

Um tipo de estrutura alternativa e mais vantajosa para a fabricação de circuitos integrados é a chamada "planar". Nesse caso as difusões para a formação dos canais dos transistores são feitas localizadas no substrato, evitando a necessidade da corrosão e consequente formação de "mesas", resultando em um dispositivo bastante plano (daí o nome de "planar").

As difusões localizadas são feitas com a ajuda de uma máscara de difusão. Essa máscara é um filme de nitreto de silício (Si_3N_4) de centenas de angstrons de espessura que é depositado sobre o substrato, e onde são abertas janelas que deixam o substrato exposto [8]. A difusão irá ocorrer apenas onde estiverem as janelas, pois o filme de nitreto de silício irá barrar os vapores de enxofre, impedindo-os de atingir o substrato.

Normalmente os filmes de nitreto de silício são deposita-

dos por CVD (deposição química de vapores) no interior de uma câmara apropriada. Deverão ser tomadas precauções no que diz respeito à aderência do filme ao substrato, pois durante a difusão eles são aquecidos e devido às diferentes dilatações dos dois materiais (filme e substrato) podemos ter o desprendimento do filme, prejudicando a difusão.

7.3.5 - Desenvolvimento de uma Fonte Sólida de Enxofre para Difusão

Vimos no capitulo 4 que a fonte de difusão de enxofre em GaAs é uma mistura de sulfato de gálio e arseneto de gálio, ambos em pó, que é colocada sobre um fino disco de grafite que contém um grande número de furos. A amostra que recebe a difusão é colocada abaixo do disco e o vapor de enxofre a atinge através dos furos no grafite. Embora esse arranjo seja bastante versátil, permitindo que se faça difusão em amostras de tamanhos variados (de diâmetro máximo de duas polegadas), ele possui uma desvantagem que é o uso de pó na fonte de difusão. O manuseio do porta amostra deve ser feito com muito cuidado para que o pó não caia sobre a amostra, e durante a difusão podemos ter a ejeção de partículas do pó sobre a amostra, devido à alta temperatura, o que danifica em parte a sua superfície.

Para evitar os problemas que surgem do uso da fonte em pó sugerimos o desenvolvimento de uma fonte sólida de enxofre para uso na difusão. Essa fonte seria um fino disco de duas polegadas de diâmetro por um ou meio milímetro de espessura, disco este que seria impregnado de enxofre e arsênico (não é necessário o uso de gálio na fonte de difusão). Essa nova fonte de difusão seria a análoga à de boro para difusão em silício [9].

O disco pode ser preparado de duas maneiras diferentes. A primeira seria a sinterização de alumina (Al_2O_3) na forma do disco, com um alto teor de enxofre e arsênico que pode ser adicionado durante a sinterização. A segunda forma de obter o disco seria a fusão de gálio, arsênico e enxofre (em quantidades aproximadamente iguais) em alta temperatura (maior que a de

difusão) para a formação do disco sólido.

7.3.6 - "ALLOY" DE AUGENI

Para transistores MESFET de canal N usa-se normalmente uma metalização de ouro, germânio e níquel para a realização dos contatos ôhmicos no dispositivo. A fabricação dos contatos é feita através da deposição dos metais e posterior aquecimento ("alloy"). É esperado que o "alloy" executado em um forno rápido produza contatos de menor resistência.

Outro trabalho que propomos é o estudo e desenvolvimento do processo de "alloy" rápido. Para isso seria necessário o projeto de um porta amostra para "alloy" (parecido com o usado na difusão). As experiências nesse caso devem ser voltadas para a determinação do tempo e temperatura ideal para o "alloy". Deve ser observada a morfologia do filme de AuGeNi após o "alloy" e devem ser feitas medidas da resistência de contato, com o uso do "chip" de testes projetado no LPD.

7.4 - REFERÊNCIAS

[1] - H. Matino, "Reproducible Sulfur Diffusion into GaAs", *Solid-State Electronics*, vol. 17, p. 35 (1974).

[2] - R. Venkatasubramanian, e S. K. Ghandhi, "Compensation Mechanisms in N^+-GaAs Doped with Sulphur", *Journal of Crystal Growth*, n. 97, p. 827 (1989).

[3] - B. Tuck e R. G. Powell, "High-Temperature Diffusion of Sulphur in GaAs", *J. Phys. D : Appl. Phys.*, vol. 14, p. 1317 (1981).

[4] - M. Oren e A. N. M. Masum Choudhury, "Zn Diffusion in GaAs Obtained by a Simple Open-Tube Technique", *J. Apll. Phys.*, vol. 60, n. 9, p. 3379 (1986).

[5] - Sandip Tiwari et al., "Rapid Thermal Diffusion and Ohmic Contacts Using Zinc in GaAs and GaAlAs", submetido ao *Appl.*

Phys. Lett.

[6] - A. Yamamoto, M. Yamaguchi e C. Uemura, "Fabrication of High Efficiency n^+-p Junction InP Solar Cells by Using Group VIb Element Diffusion into p-Type InP", *IEEE - Transactions on Electron Devices*, vol. ED-32, n. 12, p. 2780 (1985).

[7] - Hadis Morkoç et al., "Schottky Barriers and Ohmic Contacts on N-Type InP Based Compound Semiconductors for Microwave FET's", *IEEE - Transactions on Electron Devices*, vol. ED-28, n. 1, p. 1 (1981).

[8] - F. C. Prince, M. Oren e M. Lam, "High Mobility GaAs Layers Obtained by Open Tube Sulfur Diffusion", *Appl. Phys. Lett.*, vol. 48, n. 8, p. 546 (1986).

[9] - K. T. Kim e C. K. Kim, "Formation of Shallow P^+-N Junctions Using Boron-Nitride Solid Diffusion Source", *IEEE - Electron Device Letters*, vol. EDL-8, n. 12, p. 569 (1987).

APENDICE

PROGRAMAS DESENVOLVIDOS

Apresentamos neste apêndice os programas desenvolvidos para analise do perfil de difusão e dos transistores. Todos os programas foram escritos em Turbo Pascal, e rodam em microcomputadores de 8 bits. Como não usamos nenhuma função ou procedimento caracteristico de um micro de 8 bits, os programas irão funcionar com nenhuma ou pequenas modificações em micros de 16 bits.

Como comentado no capitulo 4, os programas COEF_DIF e ANALISBM são usados na analise dos perfis de difusão. O programa COEF_DIF ajusta por dois pontos dados uma curva do tipo função erro complementar, e nos dá a partir dessa curva a concentração de superficie e o coeficiente de difusão. Já o programa ANALISBM realiza a análise de Boltzmann-Matano, nos dando o coeficiente de difusão em função da concentração. Esses dados são obtidos de um arquivo de entrada, que contem a concentração (de cargas ou atomos) em função da profundidade. Os dados são colocados no arquivo de entrada nessa ordem, com a primeira linha contendo o numero de pontos do arquivo. O programa calcula o coeficiente de difusão usando dois cálculos de derivada. O primeiro com dois pontos do arquivo de entrada (Dponto na listagem de saida), e o segundo com uma exponencial interpolada por quatro pontos.

Temos por fim os programas FINDVPO e SIGMA, que são usados na analise dos transistores. O programa FINDVPO calcula a profundidade de decapagem para a colocação do "gate", a partir do perfil de cargas obtido pelo Polaron, e da medida da tensão de "pinch-off", V_{po}. O ultimo programa, SIGMA, calcula a resistência do canal, em unidades arbitrárias, em função da tensão de "gate" aplicada, e do perfil de cargas sob o "gate", que é obtido do programa FINDVPO. Este calculo é usado para determinarmos as somas das resistências série de dreno e fonte.

```
PROGRAM COEF_DIF(INPUT,OUTPUT);

VAR   C1,C2,X1,X2 :  REAL;
      C1_CALC,PRECIS,CON1,CON2 :  REAL;
      CONM,C1_1_CALC,C1_2_CALC :  REAL;
      KK,D,T :  REAL;



FUNCTION ERFC(X:REAL):REAL;

VAR H,INTE,SOM :  REAL;
    I :  INTEGER;

BEGIN
  H:=X/30.0;
  INTE:=0.0;
  FOR I:=0 TO 30 DO
     BEGIN
     IF (I=0) OR (I=30) THEN
        SOM:=EXP(-SQR(I*H))
       ELSE
        IF (2*INT(I/2)-I)=0 THEN
           SOM:=2.0*EXP(-SQR(I*H))
          ELSE
           SOM:=4.0*EXP(-SQR(I*H));
     INTE:=INTE+SOM;
     END;
  INTE:=INTE*H/3.0;
  ERFC:=1.0-(INTE/0.98622692550);
END;



FUNCTION INV_ERFC(A:REAL):REAL;

VAR VS,VI,VINT,ER :  REAL;

BEGIN
  VS:=0.0;
  VI:=-0.1;
  WHILE NOT((ERFC(VI)>A) AND (ERFC(VS)<=A)) DO
    BEGIN
    VS:=VS+0.1;
    VI:=VI+0.1;
    END;
  WHILE ((VS-VI)>1E-7) DO
    BEGIN
    VINT:= (VS+VI)/2.0;
    ER:=ERFC(VINT);
    IF (ER<A) THEN VS:=VINT
              ELSE VI:=VINT;
    END;
  INV_ERFC:=VINT;
END;
```

```
{Programa Principal}

BEGIN
   WRITELN;
   WRITELN('AJUSTE DO PERFIL DE DIFUSAO POR UM FUNCAO DO TIPO:');
   WRITELN;
   WRITELN('C = CO * erfc(X/(2*SQRT(D*T)))');
   WRITELN;
   WRITELN;
   WRITE(' Entre com C1 (cm-3) e X1 (um) > ');
   READLN(C1,X1);
   WRITE(' Entre com C2 e X2             > ');
   READLN(C2,X2);
   PRECIS:=C2*0.0001; {A precisao e' 0.1 % de C2}
   CON1:=C1;     {CON1 e' maior que CON2}
   CON2:=C1-C2;
   C1_1_CALC:=CON1*ERFC(X1*INV_ERFC(C2/CON1)/X2);
   REPEAT
      CON1:=CON1+C2;
      CON2:=CON2+C2;
      C1_2_CALC:=C1_1_CALC;
      C1_1_CALC:=CON1*ERFC(X1*INV_ERFC(C2/CON1)/X2);
   UNTIL ((C1_1_CALC>C1) AND (C1_2_CALC<C1));
   WRITELN;
   WRITELN(' CO ESTA ENTRE ',CON1:10,' E ',CON2:10);
   REPEAT
      CONM:=(CON1+CON2)/2.0;
      KK:=INV_ERFC(C2/CONM)/X2;
      C1_CALC:=CONM*ERFC(X1*KK);
      IF (C1_CALC>C1) THEN CON1:=CONM
                      ELSE CON2:=CONM;
   UNTIL ((CON1-CON2)<PRECIS);
   WRITELN;
   WRITELN(' CO = ',CONM,' cm-3');
   WRITELN;
   WRITE(' ENTRE COM T (s) > ');
   READLN(T);
   KK:=KK*10000.0;
   D:=0.25/(SQR(KK)*T);
   WRITELN;
   WRITELN(' D  = ',D,' cm2/s');
   WRITELN;
END.
```

```
PROGRAM ANALISBMC(INPUT,OUTPUT);

VAR   ARQ1,ARQ2 : TEXT;
      ARQNOME : STRING[30];
      NPONTOS,I : INTEGER;
      CONC,PROF : ARRAY[1..200] OF REAL;
      TEMPO,B1,B2,AREA,AREAPEQ : REAL;
      C1,C2,X1,X2,D : REAL;
      CB1,CB0,CA,SX,SY,SXX,SYY,SXY : REAL;
      aux,DER,Y1,Y2,Y3,Y4 : REAL;
      OPCAO : CHAR;


BEGIN  (Inicio do Programa Principal)
  WRITELN('Analise de Boltzmann-Matano para Calculo do Coeficiente
de Difusao');
  WRITELN;WRITELN;
  WRITE('Deseja copia dos resultados na impressora (s/n) ? ');
  READLN(OPCAO);
  IF (OPCAO='S') OR (OPCAO='s') THEN
    BEGIN
      OPCAO:='S';
      ASSIGN(ARQ2,'PRINTER.DEL');
      REWRITE(ARQ2);
      WRITELN(ARQ2,'Analise de Boltzmann-Matano para Calculo do
Coeficiente de Difusao');
      WRITELN(ARQ2,'em Funcao da Concentracao.');
    END;
  WRITE('Entre com o nome do arquivo de pontos do perfil,
sem extensao e com o maximo de 5 caracteres > ');
  READLN(ARQNOME);
  ASSIGN(ARQ1,ARQNOME+'.DAT');
  RESET(ARQ1);
  IF OPCAO='S' THEN
    BEGIN
      WRITELN(ARQ2);WRITELN(ARQ2);
      WRITELN(ARQ2,'ARQUIVO -> '.ARQNOME+'.DAT');
    END;
  READLN(ARQ1,NPONTOS);
  IF OPCAO='S' THEN
    BEGIN
      WRITELN(ARQ2);WRITELN(ARQ2,'Numero de pontos = ',NPONTOS);
    END;
  FOR I:=1 TO NPONTOS DO
    READLN(ARQ1,PROF[I],CONC[I]);
  WRITELN;
  WRITE('Entre com o tempo de difusao (s) > ');
  READLN(TEMPO);
  WRITELN;WRITELN;
  IF OPCAO='S' THEN
    BEGIN
      WRITELN(ARQ2);WRITELN(ARQ2,'Tempo de difusao '.TEMPO:6:1,
        ' s.');
    END;
  AREA:=0.0;
  CLOSE(ARQ1);
```

```
ASSIGN(ARQ1,ARQNOME+'DIF.DAT');
REWRITE(ARQ1);
WRITELN(ARQ1,NPONTOS);
FOR I:=NPONTOS-1 DOWNTO 1 DO
  BEGIN
    C1:=CONC[I+1];C2:=CONC[I];X1:=PROF[I+1];X2:=PROF[I];
    B2:=(LN(C2)-LN(C1))/(X1-X2);
    B1:=(X2*B2)+LN(C2);
    AREAPEQ:=((B1*(C2-C1))-(C2*(LN(C2)-1))+(C1*(LN(C1)-1)))/B2;
    AREA:=AREA+AREAPEQ;
    IF I>2 THEN
      BEGIN
        SX:=PROF[I-2]+PROF[I-1]+PROF[I]+PROF[I+1];
        Y1:=LN(CONC[I-2]);Y2:=LN(CONC[I-1]);Y3:=LN(CONC[I]);
        Y4:=LN(CONC[I+1]);
        SY:=Y1+Y2+Y3+Y4;
        SXY:=(Y1*PROF[I-2])+(Y2*PROF[I-1])+(Y3*PROF[I])+
         (Y4*PROF[I+1]);
        SYY:=(Y1*Y1)+(Y2*Y2)+(Y3*Y3)+(Y4*Y4);
        SXX:=SQR(PROF[I-2])+SQR(PROF[I-1])+SQR(PROF[I])+
         SQR(PROF[I+1]);
        CB1:=((4.0*SXY)-(SX*SY))/((4.0*SXX)-(SX*SX));
        CBO:=(SY/4.0)-(CB1*SX/4.0);
        CA:=EXP(CBO);
      END;
    DER:=1/(CA*CB1*EXP(CB1*PROF[I]));
    aux:=-1/(B2*C2);
    IF OPCAO='S' THEN
      BEGIN
        WRITELN(ARQ2);
        WRITE(ARQ2,'dx/dC = ',DER:10,'  dx/dC no ponto = ',
         aux:10);
        WRITELN(ARQ2,' area = ',AREA:10);
      END;
    WRITELN('dx/dC = ',DER,'  dx/dC no ponto = ',aux);
    WRITELN('AREA = ',AREA);
    D:=-1E4*AREA*DER/(2*TEMPO);
    WRITELN(ARQ1,PROF[I],'  ',D*1E-4);
    WRITELN('D(cm2/s) = ',D:10,'  C(cm-3) = ',1E-6*C2:8);
    IF OPCAO='S' THEN
      WRITE(ARQ2,'D(cm2/s) = ',D:10,'  C(cm-3) = ',1E-6*C2:8);
    D:=-1E4*AREA*aux/(2*TEMPO);
    IF OPCAO='S' THEN
      WRITELN(ARQ2,'  Dponto(cm2/s) = ',D:10);
    WRITELN('Dponto(cm2/s) = ',D:10);WRITELN;
  END;
CLOSE(ARQ1);
CLOSE(ARQ2);
END.
```

```
PROGRAM FINDVPO(INPUT,OUTPUT);

CONST

   E = 1.16E-10;
   Q = 1.6E-19;


VAR

   ARQNOME : STRING[20];
   ARQ1 : TEXT;
   I,J,NPONTOS : INTEGER;
   VPO,VPO1,VPO2,PROFREC : REAL;
   PROF,CONC : ARRAY [1..100] OF REAL;
   KR,AR,XEA,XEB,X1,X2,C1,C2,K,A,AUX : REAL;
   C,X,VPO3 : REAL;


BEGIN {Inicio do Programa Principal}
   WRITE('Entre com arquivo que contem o perfil > ');
   READLN(ARQNOME);
   ASSIGN(ARQ1,ARQNOME);
   RESET(ARQ1);
   WRITE('Entre com Vpo (V) > ');READLN(VPO);
   READLN(ARQ1,NPONTOS);
   FOR I:=1 TO NPONTOS DO
      READLN(ARQ1,PROF[I],CONC[I]);
   VPO1:=1000.0;J:=0;
   REPEAT
      J:=J+1;
      PROFREC:=PROF[J];
      VPO2:=VPO1;VPO1:=0.0;
      FOR I:=J TO NPONTOS-1 DO
         BEGIN
         X1:=PROF[I]-PROFREC;
         X2:=PROF[I+1]-PROFREC;
         C1:=CONC[I];
         C2:=CONC[I+1];
         K:=LN(C1/C2)/(X1-X2);
         A:=C1/EXP(K*X1);
         AUX:=(EXP(K*X2)*((K*X2)-1))-(EXP(K*X1)*((K*X1)-1));
         VPO1:=VPO1+(AUX*Q*A/(K*K*E));
         END;
      WRITE('P = ',PROFREC:8);
      WRITELN('   VPO = ',VPO1:7:3);
   UNTIL (J=NPONTOS-1)OR((VPO<VPO2)AND(VPO>VPO1));
   WRITELN;
   WRITELN('Profundidade de gate entre ',PROF[J-1]:8,' e ',PROF[J]:
8);
   X1:=PROF[J-1];X2:=PROF[J];C1:=CONC[J-1];C2:=CONC[J];
   KR:=LN(C1/C2)/(X1-X2);
   AR:=C1/EXP(K*X1);XEA:=X2;XEB:=X1;
   REPEAT
      VPO3:=0.0;
      X:=XEB+((XEA-XEB)/2.0);
```

```
    C:=AR*EXP(KR*X);
    PROF[J-1]:=X;CONC[J-1]:=C;
    FOR I:=J-1 TO NPONTOS-1 DO
      BEGIN
      X1:=PROF[I]-X;
      X2:=PROF[I+1]-X;
      C1:=CONC[I];C2:=CONC[I+1];
      K:=LN(C1/C2)/(X1-X2);
      A:=C1/EXP(K*X1);
      AUX:=(EXP(K*X2)*((K*X2)-1))-(EXP(K*X1)*((K*X1)-1));
      VPO3:=VPO3+(AUX*Q*A/(K*K*E));
      END;
    IF VPO<VPO3 THEN
        XEB:=X
      ELSE
        XEA:=X;
  UNTIL ((XEA-XEB)/X)<0.0002;
  WRITELN('Profundidade de gate ',X:9,'m  com Vpo ',VPO3:8:4,'V');
END.
```

```
PROGRAM SIGMA(INPUT,OUTPUT);

CONST

   E = 1.16E-10;
   Q = 1.6E-19;


VAR

   ARQNOME : STRING[20];
   ARQ1 : TEXT;
   I,J,NPONTOS : INTEGER;
   PROF,CONC : ARRAY [1..100] OF REAL;
   X,XEA,XEB,X1,X2,C1,C2,K,A,AUX : REAL;
   DELTAV,PROFGATE,SIGMA,RESIST : REAL;
   V1,V2,V3,V,VG,VBI :REAL;


BEGIN (Inicio do Programa Principal)
   WRITE('Entre com arquivo que contem o perfil > ');
   READLN(ARQNOME);
   ASSIGN(ARQ1,ARQNOME);
   RESET(ARQ1);
   READLN(ARQ1,NPONTOS);
   FOR I:=1 TO NPONTOS DO
      READLN(ARQ1,PROF[I],CONC[I]);
   WRITE('Entre com a profundidade do gate > ');
   READLN(PROFGATE);
   WRITE('Entre com o valor de VBI (V) > ');
   READLN(VBI);
   I:=0;
   REPEAT
      I:=I+1;
   UNTIL ((PROFGATE<PROF[I+1]) AND (PROFGATE>PROF[I])) OR
     (I=(NPONTOS-1));
   X1:=PROF[I];
   X2:=PROF[I+1];
   C1:=CONC[I];C2:=CONC[I+1];
   K:=LN(C1/C2)/(X1-X2);
   A:=C1/EXP(K*X1);
   PROF[1]:=0.0;
   CONC[1]:=A*EXP(K*PROFGATE);
   FOR J:=I+1 TO NPONTOS DO
      BEGIN
      PROF[J-I+1]:=PROF[J]-PROFGATE;
      CONC[J-I+1]:=CONC[J];
      END;
   NPONTOS:=NPONTOS-I+1;
   WRITELN('Para fim do calculo entre VG > 1V');
   REPEAT
      WRITELN;
      WRITE('Entre com VG > ');
      READLN(VG);V:=-VG+VBI;
      IF VG < 1.0 THEN
         BEGIN
```

```
V1:=0.0;V2:=0.0;
I:=0;
REPEAT
   I:=I+1;
   V2:=V1;
   X1:=PROF[I];X2:=PROF[I+1];
   C1:=CONC[I];C2:=CONC[I+1];
   K:=LN(C1/C2)/(X1-X2);
   A:=C1/EXP(K*X1);
   AUX:=(EXP(K*X2)*((K*X2)-1))-(EXP(K*X1)*((K*X1)-1));
   V1:=V1+(AUX*Q*A/(K*K*E));
UNTIL (I=NPONTOS-1) OR ((V1>V) AND (V2<V));
WRITELN('Profundidade de interface entre ',PROF[I]:8,' e ',
 PROF[I+1]:8);
DELTAV:=V-V2;
XEA:=X2;XEB:=X1;
REPEAT
   X:=XEB+((XEA-XEB)/2.0);
   AUX:=(EXP(K*X)*((K*X)-1))-(EXP(K*X1)*((K*X1)-1));
   V3:=AUX*Q*A/(K*K*E);
   IF DELTAV>V3 THEN
      XEB:=X
   ELSE
      XEA:=X;
UNTIL((XEA-XEB)/XEA)<0.0001;
WRITELN('Profundidade de interface ',X:8);
SIGMA:=(EXP(K*X2)-EXP(K*X))*A/K;
IF I<NPONTOS-1 THEN
  FOR J:=I+1 TO NPONTOS-1 DO
    BEGIN
      X1:=PROF[I];X2:=PROF[I+1];
      C1:=CONC[I];C2:=CONC[I+1];
      K:=LN(C1/C2)/(X1-X2);
      A:=C1/EXP(K*X1);
      SIGMA:=SIGMA+((EXP(K*X2)-EXP(K*X1))*A/K);
      END;
    SIGMA:=SIGMA*Q;
    RESIST:=1/SIGMA;
    WRITELN('Condutancia do canal = ',SIGMA:9,
     ' (UA)    Resistividade = ',RESIST:9,' (UA)');
    END;
  UNTIL (VG>1.0);
END.
```