

wave mechanics applied to semiconductor heterostructures

GERALD BASTARD



Avenue du Hoggar,
Zone Industrielle de Courtabœuf,
B.P. 112,
91944 Les Ulis Cedex, France

Preface

Following the pioneering work of Esaki and Tsu in 1970, the study of two dimensional semiconductor heterostructures, namely modulation-doped heterojunctions, quantum wells and superlattices, has developed rapidly, both from the point of view of basic physics and of applications. For example, modulation-doped heterojunctions are nowadays currently used to investigate the integer and fractional quantum Hall effects as well as to make very fast transistors.

This book, which began as a set of lecture notes for a graduate course given at the Ecole Normale Supérieure, is specifically concerned with the basic electronic and optical properties of two dimensional semiconductor heterostructures based on III-V and (to a lesser extent) II-VI compounds, but applications are not considered. It is intended as an introductory textbook for undergraduate and graduate students and engineers working in this field. It is essentially an attempt of exploring various consequences of one-dimensional size-quantization, a genuine quantum-mechanical effect, on the most basic physical properties of heterolayers. Because one of their degrees of freedom is frozen by this size-quantization, the carriers in heterolayers effectively behave as if their motion were bi-dimensional, a feature which alters a number of physical properties.

The book starts with a chapter recalling a few basic quantum mechanical properties of idealized quantum wells and superlattices (those found in quantum mechanics textbook). Chapter II is a summary of the $k.p$ analysis of the electronic dispersion relations in direct gap bulk III-V and II-VI compounds. In chapter III we show how it is possible, in a simplified approach, to intertwine the results of chapters I and II to obtain the electronic dispersion relations in flat band superlattices. In chapter IV we discuss the occurrence of bound states when the heterostructure is imperfect (hydrogenic impurities or interface defects) or when it is shone with near bandgap light (excitons). Then we switch to charge transfer mechanism and discuss in chapter V the carrier bound states in the quasi-triangular well formed nearby the interface of a modulation-doped heterostructure. Chapter VI is devoted to the in-plane ohmic electrical properties of these modulation-doped heterostructures, focussing the attention on elastic scattering processes. Chapter VII deals with some of the basic optical properties (absorption, photoluminescence) of quantum wells and superlattices. Then, in chapter VIII, we discuss the alterations of heterostructure energy levels by static electric or magnetic fields. I have attempted to make each chapter reasonably self contained, which has led to the occurrence of a few repetitions.

Several difficulties are encountered in the writing of a book on such a rapidly evolving subject. One problem is the large number of papers which are continuously published, providing additional results and leading sometimes to the modification of current ideas. I can only hope that the principles presented in this book will be

helpful and fruitful to understand new results and to generate new ideas. Another difficulty is the large number and the variety of papers already published on semiconductor heterostructures. Instead of trying to include all these papers, a representative bibliography is given at the end of each chapter, and I apologize in advance to all my colleagues whose important contribution has not been quoted. I wish to thank those who granted me permission to reproduce their figures. Finally, the author would like to acknowledge all the physicists with whom he had the privilege to work : Drs. J. A. Brum, C. Delalande, Y. Guldner, M. H. Meynadier, J. Orgonasi, J. P. Vieren, P. Voisin and U. O. Ziemelis at the Ecole Normale Supérieure (Paris) and Drs. C. A. Chang, L. L. Chang, L. Esaki, E. E. Mendez and F. Stern at IBM (Yorktown Heights, U.S.A.). Above all, Michel Voos deserves hearty thanks for having taken the time and displayed enough patience to discuss the physics and to carefully check the manuscript.

Contents

PREFACE	page	III
CHAPTER I : IDEALIZED QUANTUM WELLS AND SUPER-LATTICES. A SUMMARY	page	1
I. <i>Single quantum wells</i>	1	
II. <i>Density of states</i>	11	
III. <i>Tunnel coupling between wells : the symmetric double square well</i>	14	
IV. <i>Superlattices</i>	18	
IV.1. <i>Superlattice dispersion relations</i>	18	
IV.2. <i>Symmetry properties of the eigenfunctions</i>	23	
IV.3. <i>Superlattice density of states</i>	24	
APPENDIX A : Symmetric double square well. Time-dependent aspect	27	
APPENDIX B : Dispersion relations of superlattices with arbitrary potential profiles	28	
CHAPTER II : BAND STRUCTURE OF BULK III-V COMPOUNDS	page	31
I. <i>Crystalline properties</i>	31	
II. <i>Electronic properties</i>	32	
III. <i>Electronic dispersion relations in the vicinity of the zone centre : $k.p$ analysis and Kane model</i>	35	
III.1. <i>$k.p$ analysis and effective masses</i>	36	
III.2. <i>Effective mass tensor. Electrons and holes</i>	38	
III.3. <i>Beyond the quadratic dispersion relations : the Kane model</i>	41	
a) <i>Dispersions relations of the Kane model</i>	41	
b) <i>Band edge effective masses</i>	45	
c) <i>Band non parabolicity</i>	46	
d) <i>Inclusion of higher bands</i>	46	
e) <i>Accuracy of the Kane model</i>	47	
IV. <i>Conclusions</i>	49	
APPENDIX A : Inclusion of remote bands effects	50	
APPENDIX B : Motion of Kane electrons in slowly varying perturbing fields	54	

CHAPTER III : ENVELOPE FUNCTION DESCRIPTION OF HETEROSTRUCTURE ELECTRONIC STATES	page 63
I. <i>Introduction</i>	63
II. <i>The envelope function model</i>	64
II.1. Preliminaries	64
II.2. The envelope function framework	66
II.3. The Ben Daniel-Duke model	73
II.3.1. The Ben Daniel-Duke quantum wells ($m_A m_B > 0$) ...	74
II.3.2. Interface states of Ben Daniel-Duke quantum wells ($m_A m_B < 0 ; k_\perp = 0$)	79
II.4. Quantum wells and superlattices with hosts which display Kane-like bands	83
II.5. Simplified calculations of superlattices and quantum wells states $k_\perp = 0$	88
II.6. Miscellaneous limiting cases	89
II.6.1. Evanescent propagation in one kind of layer. Quantum well bound states	89
II.6.2. Tight-binding expansion of the superlattice states	90
II.6.3. Propagating states in both kinds of layers	91
II.6.4. Heavy hole superlattice states	92
II.7. Specific examples	93
II.8. Labelling and counting superlattice states	100
III. <i>In-plane dispersions in semiconductor heterostructures</i>	101
APPENDIX A : <i>Boundary conditions and stationary states</i>	112
APPENDIX B : <i>Coupling between light and heavy particle states due to inversion asymmetry splitting in bulk materials. Qualitative aspects</i>	113

CHAPTER IV : COULOMBIC BOUND STATES AND INTERFACE DEFECTS IN HETEROSTRUCTURES	page 119
--	-----------------

Coulombic bound states

I. <i>Qualitative aspects</i>	119
I.1. Bulk hydrogenic impurities	119
I.2. Impurities in heterostructures	120
II. <i>Approximate solutions of the hydrogenic impurity problem</i>	121
II.1. Formulation of the problem	121
II.2. Results for the ground impurity state attached to the ground subband	123
II.3. Excited subbands. Continuum	125
II.4. Excited impurity levels attached to the ground subband	127
II.5. Acceptor levels in a quantum well	128

III. <i>Excitons</i>	128
III.1. Excitons in idealized bulk materials	130
III.2. Excitons in idealized quantum well structures	132
III.3. Excitons in actual quantum well structures	139
Interface defects	
I. <i>Graded interfaces</i>	142
II. <i>Quantum well interface defects</i>	143
III. <i>Superlattice defect states</i>	149
APPENDIX : <i>Energy levels in a quantum well with narrow graded interfaces</i>	151
CHAPTER V : ENERGY LEVELS IN MODULATION-DOPED HETEROSTRUCTURES	page 155
I. <i>The modulation doping of heterojunctions. Qualitative aspects</i>	155
I.1. The unperturbed heterojunction (flat band case)	156
I.2. Single heterojunction containing diluted Coulombic impurities ..	157
I.3. Actual heterostructures	160
II. <i>Self consistent calculations of energy levels and charge transfer in single heterojunctions</i>	161
II.1. Energy level calculations	165
II.2. Charge transfer in single heterojunctions	172
III. <i>Energy levels in modulation-doped quantum wells</i>	180
APPENDIX A : <i>The algebra of the modified Fang Howard wavefunction</i>	186
APPENDIX B : <i>Heterojunction energy levels : qualitative aspects</i>	188
CHAPTER VI : ELECTRICAL CONDUCTIVITY OF QUASI BI-DIMENSIONAL ELECTRON GASES	page 193
I. <i>Static conductivity of a quasi bi-dimensional electron gas</i>	194
I.1. Electrical conductivity in the electric quantum limit	199
I.2. Intersubband scattering	200
I.3. Screening in a quasi bi-dimensional electron gas. Application to Coulombic impurities	204
I.4. Discussion of the dielectric function	207
II. <i>Evaluation of some scattering mechanisms</i>	212
II.1. Mobility limited by Coulombic scattering	212
II.2. Alloy scattering	219
II.3. Interface roughness scattering	223
III. <i>Vertical transport</i>	225
IV. <i>Resonant tunnelling</i>	228
APPENDIX : <i>Dimensional dependence of the screening effects</i>	233

CHAPTER VII : OPTICAL PROPERTIES OF QUASI BI-DIMENSIONAL SYSTEMS	page 237
I. <i>Absorption (one-electron approximation)</i>	237
I.1. Intraband transitions	243
I.2. Interband transitions	246
I.2.1. Polarization selection rules	247
I.2.2. Selection rules on the envelope function quantum numbers : evaluation of $\langle f_i/f_f \rangle$	247
I.2.3. Order of magnitude of the absorption coefficient. Comparison between type I and type II systems	251
I.2.4. Interband optical transitions in superlattices	255
II. <i>Absorption : a simplified description of excitonic effects</i>	258
II.1. Absorption in the absence of Coulombic interactions. Equivalence with the one-electron model	260
II.2. Absorption in the presence of electron-electron interactions	262
III. <i>Photoluminescence of quasi bi-dimensional systems</i>	272
III.1. Introduction	272
III.2. Quantum well luminescence (steady state)	281
III.2.1. Band to band emission	282
III.2.2. Excitonic recombination	289
III.2.3. Extrinsic photoluminescence	295
CHAPTER VIII : EFFECT OF STATIC EXTERNAL ELECTRIC AND MAGNETIC FIELDS ON THE ENERGY LEVELS OF QUASI BI-DIMENSIONAL ELECTRON GASES	page 303
I. <i>The Stark effects</i>	303
I.1. Transverse Stark effect in a quantum well ($F//x$)	304
I.2. Longitudinal Stark effect in a quantum well ($F//z$)	308
II. <i>Landau quantization of a quasi bi-dimensional electron gas</i>	317
II.1. Energy levels	317
II.2. Magnetic field dependent density of states	325
II.3. Magnetoconductivity of a quasi bi-dimensional electron gas	329
II.3.1. Macroscopic derivation	329
II.3.2. Microscopic discussion	332
II.4. Cyclotron resonance	339
APPENDIX A : <i>Motion of three-dimensional electrons in crossed electric and magnetic fields</i>	345
AI. The Drude model	346
AII. Quantum motion of bulk electrons in a magnetic field	348
APPENDIX B : <i>Microscopic evaluation of the magnetoconductivity tensor</i>	351

CHAPTER I

Idealized quantum wells and superlattices. A summary

In this introductory chapter we shall briefly recall some of the properties of the solutions of the Schrödinger equation for one dimensional square quantum wells, double square quantum wells and superlattices. By idealized we mean that the carrier mass in such structures is both position and energy independent. In principle, this mass should be taken as equal to the bare electron mass m_0 . However, for the sake of numerical examples, we shall often use an effective mass m^* .

This differs from m_0 and will be taken as equal to either the conduction or to the heavy hole effective mass of GaAs. ($0.067 m_0$ and $\sim 0.4 m_0$ respectively)

I. Single quantum wells.

Let us consider the one-dimensional motion of a particle of mass m^* which is subjected to a potential energy $V_b(z)$ (see Fig. 1) which is such that

$$V_b(z) = \begin{cases} 0 & |z| > \frac{L}{2} \\ -V_b & |z| < \frac{L}{2} \end{cases} \quad (1)$$

where L is the thickness of the potential well (hereafter termed quantum well).

I.1 BOUND MOTIONS ($-V_b \leq \varepsilon < 0$).

1) *Classical analysis* : For energies ε such that $-V_b \leq \varepsilon < 0$, the classical solutions consist of bound particle motions which take place between $\pm \frac{L}{2}$. The particle never

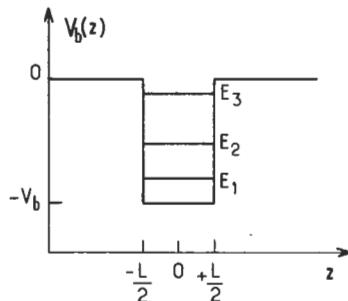


Fig. 1. — Potential energy profile of a square quantum well.

penetrates the barrier since $\epsilon \leq 0$ and $|z| > \frac{L}{2}$ would correspond to an imaginary velocity, which is meaningless in classical mechanics. Inside the well the carrier velocity is constant and equal to

$$v_z = \pm \sqrt{\frac{2}{m^*} (\epsilon + V_b)}. \quad (2)$$

At the well boundaries the carrier is perfectly reflected and its velocity changes instantaneously from $\pm |v_z|$ to $\mp |v_z|$. The classical motion is thus periodic in time with a period of

$$T = \frac{2L}{|v_z|} \quad (3)$$

Any energy $\epsilon \geq -V_b$ is allowed, i.e. it can be associated with a possible motion.

2) *Quantum-mechanical analysis* : The quantum-mechanical motion is described by a wavefunction $\psi(z, t)$ solution of the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = \mathcal{H} \left(z, p_z = \frac{\hbar}{i} \frac{\partial}{\partial z} \right) \psi, \quad (4)$$

where \mathcal{H} is the Hamiltonian of the classical problem :

$$\mathcal{H} = \frac{p_z^2}{2m^*} + V_b(z), \quad (5)$$

since \mathcal{H} does not depend explicitly on time t , $\psi(z, t)$ factorizes into

$$\psi(z, t) = \chi(z) \exp \left(-i \frac{\epsilon t}{\hbar} \right), \quad (6)$$

provided that $\chi(z)$ satisfies the eigenvalue problem

$$\mathcal{H} (z, p_z) \chi(z) = \epsilon \chi(z), \quad (7)$$

or more explicitly

$$\left[-\frac{\hbar^2}{2m^*} \frac{d^2}{dz^2} + V_b(z) \right] \chi(z) = \epsilon \chi(z). \quad (8)$$

$\chi(z)$ must fulfill the following boundary conditions :

- i) $\chi(z)$ is continuous everywhere ;
- ii) by integrating equation (8) around any z_0 , and for the specific potential of the square quantum well, $\frac{d\chi}{dz}$ is continuous everywhere ;
- iii) $\lim_{z \rightarrow \pm \infty} |\chi(z)|$ is finite.

Since $V_b(z)$ is piecewise constant, exact solutions of equation (8) can be obtained. For both $|z| < \frac{L}{2}$ and $|z| > \frac{L}{2}$, $\chi(z)$ is the sum of two plane waves of opposite wavevectors. Inside the quantum well these waves propagate and the characteristic wavevectors are $\pm k_w$ where

$$k_w = \sqrt{\frac{2m^*}{\hbar^2} (\varepsilon + V_b)}. \quad (9)$$

The waves are evanescent outside the quantum well and thus the modulus of the imaginary wavevectors associated with the evanescent waves are

$$\kappa_b = \sqrt{-\frac{2m^* \varepsilon}{\hbar^2}}. \quad (10)$$

It should be noted that $V_b(z)$ is even in z . Thus the wavefunction $\chi(z)$ can be chosen either even or odd in z . To account for this symmetry property, $\chi(z)$ can be written

$$\chi(z) = A \cos k_w z; \quad \varepsilon = -V_b + \frac{\hbar^2 k_w^2}{2m^*} \quad \text{for even states} \quad (11)$$

$$\chi(z) = A \sin k_w z; \quad \varepsilon = -V_b + \frac{\hbar^2 k_w^2}{2m^*} \quad \text{for odd states,} \quad (12)$$

inside the well. Outside the well the form of the wavefunction is

$$\chi(z) = B \exp \left[-\kappa_b \left(z - \frac{L}{2} \right) \right] + C \exp \left[+\kappa_b \left(z - \frac{L}{2} \right) \right]; \quad z \geq \frac{L}{2} \quad (13)$$

$$\chi(z) = E \exp \left[\kappa_b \left(z + \frac{L}{2} \right) \right] + D \exp \left[-\kappa_b \left(z + \frac{L}{2} \right) \right]; \quad z \leq -\frac{L}{2}. \quad (14)$$

However, since $|\chi(z)|$ should not diverge when $z \rightarrow \pm \infty$, $C = D = 0$.

For even states $B = E$ and for odd states $B = -E$. If we apply the boundary conditions at $z = \frac{L}{2}$, we easily find that the energy ε satisfies the transcendental equations

$$k_w \tan \left(k_w \frac{L}{2} \right) = \kappa_b \quad \text{for even states} \quad (15)$$

$$k_w \cotan \left(k_w \frac{L}{2} \right) = -\kappa_b \quad \text{for odd states} \quad (16)$$

if a motion is going to be described by equation (8). Equations (15, 16) are satisfied only for discrete values of the energy ε . Thus, in marked contrast with the continuous classical spectrum, the quantum mechanical spectrum ($\varepsilon < 0$) is discrete. This discreteness is the key reason for having quasi bi-dimensional behaviour in the III-V heterostructures. If the energy spacing between the discrete energy levels is much

larger than the thermal broadening (~ 25 meV at room temperature) or the collision broadening Γ (~ 0.17 meV if $m^* = 0.067 m_0$, if the electronic mobility is $\mu = 10^5$ cm²/Vs and if one identifies the velocity relaxation time with the level lifetime), the carriers will effectively behave as if their motion was bi-dimensional in the (x, y) plane, the z motion being locked in one of the bound states characterized by $\chi(z)$. Another feature of the solutions of equations (11-14) is the finite probability of finding the carrier in the barrier: the well-known tunnel effect. It is of purely quantum-mechanical origin, has no classical counterpart, and arises from the wave-like nature of any material particle.

Let us define P_b as the integrated probability of finding the carrier in the barrier ($|z| > \frac{L}{2}$) while it is in the ground state ($\epsilon = E_1$):

$$P_b = \int_{|z| > \frac{L}{2}} \chi_1^2(z) dz. \quad (17)$$

In figure 2 P_b is plotted *versus* L for a quantum well problem corresponding to $V_b = 224$ meV, $m^* = 0.067 m_0$ and $V_b = 150$ meV, $m^* = 0.4 m_0$ respectively. It can be seen that the penetration of the barrier by the carrier is very small when $L > 100$ Å. For these L 's one may consider to a good approximation that the carrier is completely confined in the well.

Equations (15, 16) can only be solved numerically. However, one may easily show algebraically that the solutions E_{n+1} of these equations ($n = 0, 1, \dots$) are such that $dE_{n+1}/dL < 0$. We also notice that the solution $E_{n+1} = 0$, i.e. $\kappa_b(E_{n+1}) = 0$ which happens just as the $(n+1)^{th}$ bound state enters the well, corresponds to

$$k_w(E_{n+1} = 0) L = n\pi, \quad n = 0, 1, \dots \quad (18)$$

These solutions are thus evenly spaced in L with a period $\pi/k_w(E = 0)$. In other words a quantum well of thickness L admits $n(L)$ bound states where :

$$n(L) = 1 + \text{Int} \left[\sqrt{\frac{2m^* V_b L^2}{\pi^2 \hbar^2}} \right] \quad (19)$$

$\text{Int}(x)$ denotes the integer part of x (see Fig. 3). A one dimensional quantum well supports at least one bound state, irrespective of the height of the confining barrier. It supports an infinite number of bound levels if V_b is infinite. Under this condition the solutions of equations (15, 16) are

$$k_w L = p\pi, \quad p = 1, 2, \dots \quad (20)$$

and thus, with the energy zero coinciding with the bottom of the well,

$$E_p = \frac{\hbar^2 \pi^2}{2m^* L^2} p^2; \quad p \geq 1 \quad (21)$$

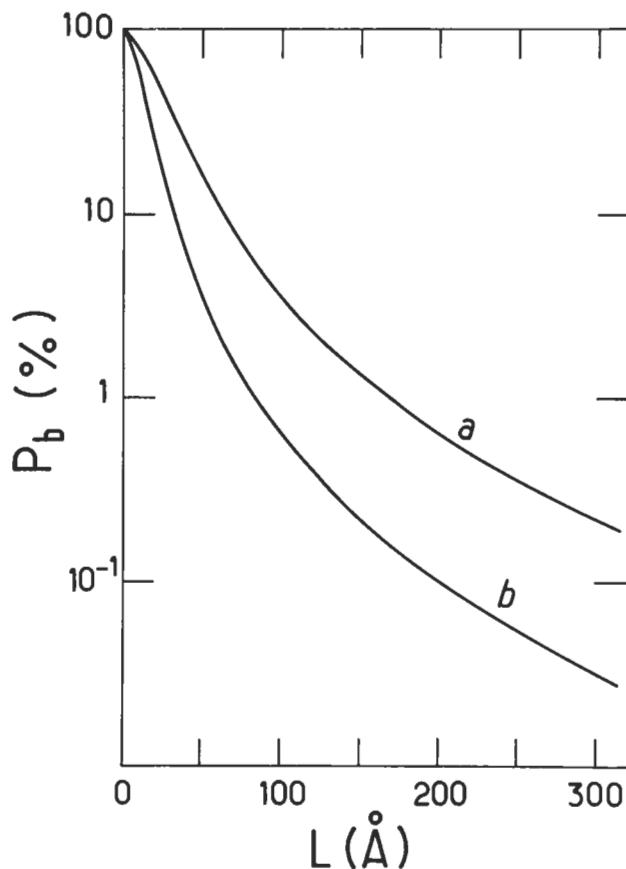


Fig. 2. — Quantum well thickness (L) dependence of the integrated probability P_b of finding the particle in the barrier of a square quantum well. The carrier state is assumed to be the lowest ground bound state of the well. Curve (a) : $V_b = 224$ meV, $m^* = 0.067 m_0$. Curve (b) $V_b = 150$ meV, $m^* = 0.4 m_0$

The quantity $\frac{\hbar^2 \pi^2}{2m^* L^2}$ is equal to ~ 56 meV if $L = 100$ Å and $m^* = 0.067 m_0$ (GaAs conduction band mass).

The normalized even and odd wavefunctions of the infinitely deep well are :

$$\chi_{2p+1}(z) = \sqrt{\frac{2}{L}} \cos \left[(2p+1) \frac{\pi z}{L} \right]; \quad |z| \leq \frac{L}{2}; \quad p \geq 0 \quad (22)$$

$$\chi_{2p+2}(z) = \sqrt{\frac{2}{L}} \sin \left[(2p+2) \frac{\pi z}{L} \right]; \quad |z| \leq \frac{L}{2}; \quad p \geq 0 \quad (23)$$

* Classical limit ($p \rightarrow \infty$) of the infinitely deep well.

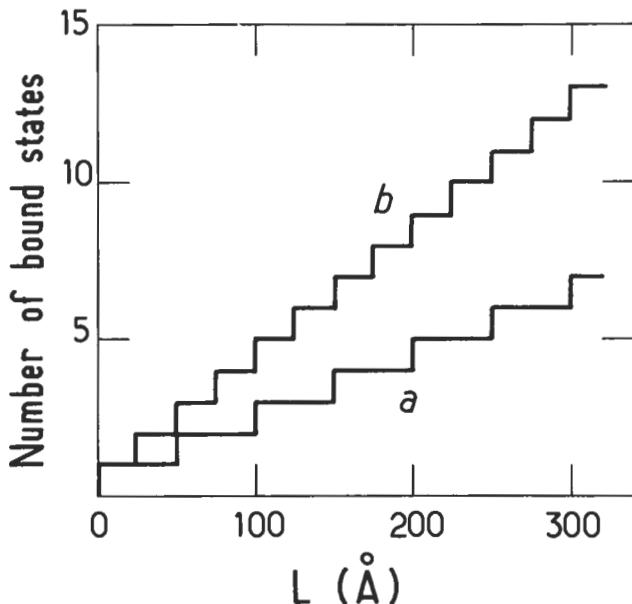


Fig. 3.— The number of quantum well bound states of a square well are plotted *versus* the quantum well thickness L . Curve (a) : $V_b = 224$ meV, $m^* = 0.067 m_0$. Curve (b) : $V_b = 150$ meV, $m^* = 0.4 m_0$

In the limit of large p we study the probability $P_p(z_0)$ of finding the carrier between $z_0 - \frac{dz}{2}$ and $z_0 + \frac{dz}{2}$ while having the energy E_p . It is equal to

$$P_p(z_0) dz = \frac{1}{L} \left\{ 1 + (-1)^{p+1} \cos \left[2 \frac{p \pi z_0}{L} \right] \right\} dz ; \quad p \geq 1 \quad (24)$$

i.e. the carrier becomes uniformly distributed over the well (the cosine term averages to zero when $p \rightarrow \infty$ for any $z_0 \neq 0$). This is just the classical result : during a period T of the classical motion, one may define the probability of finding the particle between $z_0 - \frac{dz}{2}$ and $z_0 + \frac{dz}{2}$ as the ratio between the duration Δt spent by the particle on the segment dz which is centred at z_0 , to the period of the motion. This yields :

$$P_{\text{classical}}(z_0) dz = \frac{\Delta t}{T} = \frac{1}{T} \frac{2dz}{|v_z|} = \frac{1}{L} dz \quad (25)$$

The coincidence between the quantum description at high quantum numbers and the classical description is a general property which holds, irrespective of the exact shape of the Hamiltonian \mathcal{H} .

* Confinement energy. Links with the localization.

We have seen that the lowest lying allowed energy in the quantum description is not $-V_b$ as in the classical description but differs from $-V_b$ by some positive energy ($+\frac{\hbar^2 \pi^2}{2m^* L^2}$ in the specific case of a quantum well of infinite barrier height). This phenomenon is quite general and may be qualitatively derived using the Heisenberg relations.

Consider a quantum system with eigenstates $|n\rangle$. The root-mean square deviation of a quantity a when the system is in the state $|n\rangle$ can be defined as :

$$\Delta_n a = \sqrt{\langle a^2 \rangle_n - \langle a \rangle_n^2} \quad (26)$$

where

$$\langle a^p \rangle_n = \langle n | a^p | n \rangle = \int \chi_n^* a^p \chi_n dz. \quad (27)$$

Then, the Heisenberg relations, when applied to the position of a particle with exhibits a one-dimensional motion, state that :

$$\Delta_n z \Delta_n p_z \geq \hbar/2, \quad (28)$$

where $\hbar = 2\pi\hbar$ is the Planck constant. The inequality in equation (28) means that it is impossible to know simultaneously the position and the momentum of the particle with an arbitrary accuracy. Rather, if the position of the particle is known exactly, its momentum is undefined and *vice versa*. Let us apply the inequality equation (28) to the square quantum well problem by taking $n = 1$ (ground state). First we notice that $\langle z \rangle_1 = \langle p_z \rangle_1 = 0$ since $\chi_1(z)$ is even in z . As $\chi_1(z)$ is essentially localized over the quantum well width L , the quantity $\langle z^2 \rangle_1$ is of the order of L^2 . By using equation (28), we get

$$\langle p_z^2 \rangle_1 \sim \frac{\hbar^2}{4L^2}. \quad (29)$$

The average energy of the particle is $\langle \mathcal{H} \rangle_1$:

$$E_1 = \langle \mathcal{H} \rangle_1 = -V_b(1 - P_b) + \left\langle \frac{p_z^2}{2m^*} \right\rangle_1 \sim -V_b(1 - P_b) + \frac{\hbar^2}{8m^* L^2}, \quad (30)$$

where P_b has been defined in equation (17). Thus $E_1 > -V_b$. The difference between the quantum and classical results does not originate from the term $V_b P_b$ which can be made arbitrarily small if V_b (or L) is large enough. Rather, the difference lies in the confinement energy term $\hbar^2/8m^* L^2$, which is of kinetic origin and is a genuine quantum feature in as much as it results from the Heisenberg relationships which involve \hbar . Whenever the carrier position is known with an accuracy of L , (which is the meaning of the statement that the particle is localized in a region of size L), the carrier momentum is larger than or of the order of

\hbar/L , and the kinetic energy is of the order of $\hbar^2/2m^* L^2$. The reader will check by direct computations that for a quantum well with infinite barrier height one finds

$$\Delta_1 z \Delta_1 p_z = \frac{\hbar}{2} \left[\frac{\pi^2}{3} - 2 \right]^{1/2} \sim 1.14 \frac{\hbar}{2}, \quad (31)$$

whereas for the ground state of the one-dimensional harmonic oscillator ($V_b(z) = \frac{1}{2} m^* \omega^2 z^2$) whose wavefunction has a Gaussian shape, one finds $\Delta_1 z \Delta_1 p_z = \frac{\hbar}{2}$.

I.2 UNBOUND MOTIONS ($\epsilon \geq 0$). — For positive energies, the classical motion is unbound. A carrier moving from the left to the right of figure 4 has a constant velocity $\hbar k_b/m^*$ in the left-hand side barrier. At $z = -L/2$ it accelerates instantaneously ; it moves across the well with the constant velocity $\hbar k_w/m^*$ until it reaches the interface $z = L/2$ where it decelerates instantaneously and finally escapes to $z = +\infty$ at the constant velocity $\hbar k_b/m^*$. A second classical motion occurs at the same energy : the carrier coming from the right hand side barrier impinges on the quantum well and finally escapes to $z = -\infty$. In this motion the velocities are the opposite of those found in the previous description.

The quantum-mechanical motion is characterized by a continuous spectrum of allowed energy states. Each eigenvalue ϵ is twice degenerate because the carrier motion occurs either from the left to the right or from the right to the left of the quantum well. The carrier wavevectors in the well and in the barrier are real, corresponding in both cases to propagating states :

$$k_b = \sqrt{\frac{2m^*}{\hbar^2} \epsilon}, \quad k_w = \sqrt{\frac{2m^*}{\hbar^2} (\epsilon + V_b)}. \quad (32)$$

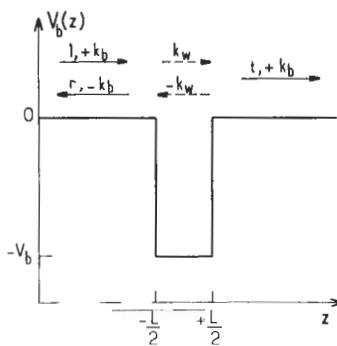


Fig. 4. — Schematic representation of the reflection-transmission phenomena experienced by an electron whose energy falls in the continuous part of the energy spectrum of a square quantum well.

An electron which impinges from $z = -\infty$ on the quantum well is partially reflected, partially transmitted at the interface $z = -L/2$. Inside the well, the eigenstate is a combination of plane waves characterized by wavevectors $\pm k_w$. The occurrence of a wave with wavevector $-k_w$ accounts for the partial reflection at the second interface $z = +L/2$. For $z \geq L/2$ the carrier is also partly transmitted and escapes towards $z = \infty$ with a wavevector $+k_b$. Thus

$$\chi(z) = \begin{cases} \exp\left[i k_b \left(z + \frac{L}{2}\right)\right] + r \exp\left[-i k_b \left(z + \frac{L}{2}\right)\right] & z \leq -\frac{L}{2} \\ \alpha \exp(i k_w z) + \beta \exp(-i k_w z) & |z| \leq \frac{L}{2} \\ t \exp\left[i k_b \left(z - \frac{L}{2}\right)\right] & z \geq \frac{L}{2} \end{cases} \quad (33)$$

Writing the continuities of $\chi(z)$ and $\frac{d\chi}{dz}(z)$ at both interfaces we obtain, after some manipulations :

$$t(\varepsilon) = \left\{ \cos k_w L - \frac{i}{2} \left(\xi + \frac{1}{\xi} \right) \sin k_w L \right\}^{-1} \quad (34)$$

$$r(\varepsilon) = \frac{i}{2} \left(\xi - \frac{1}{\xi} \right) \sin k_w L \left\{ \cos k_w L - \frac{i}{2} \left(\xi + \frac{1}{\xi} \right) \sin k_w L \right\}^{-1} \quad (35)$$

where

$$\xi = k_w/k_b \quad (36)$$

Let $T(\varepsilon)$ and $R(\varepsilon)$ denote the transmission and reflection coefficients of the well :

$$T(\varepsilon) = |t(\varepsilon)|^2; \quad R(\varepsilon) = |r(\varepsilon)|^2 \quad (37)$$

Then

$$R(\varepsilon) + T(\varepsilon) = 1 \quad (38)$$

$$T(\varepsilon) = \left[1 + \frac{1}{4} \left(\xi - \frac{1}{\xi} \right)^2 \sin^2 k_w L \right]^{-1} \quad (39)$$

Equations (38, 39) look familiar. Actually they are nothing but the reflection and transmission coefficients of a Fabry-Pérot dielectric slab. This is not accidental but arises from the close analogy between the one-dimensional Schrödinger equation and the equation governing the propagation of an electromagnetic wave in a medium characterized by a position-dependent refractive index $n(z)$.

The transmission coefficient has two remarkable features :

i) If $Lk_w(\varepsilon = 0) \neq p\pi$, the transmission coefficient vanishes at the onset of the continuum. We might have expected that a carrier with a vanishingly small velocity in the barrier would easily fall into the potential well. However, just the reverse is true in that the electron does not penetrate the well ($\alpha = \beta = 0$). This behaviour reflects the wave-like nature of the electron, which is most strikingly revealed when the potential has sharp discontinuities.

ii) The transmission coefficient is not a smooth function of the energy. Instead

$T(\varepsilon)$ reaches unity whenever

$$k_w L = p\pi. \quad (40)$$

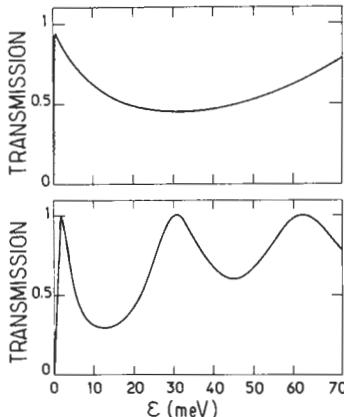


Fig. 5. — Energy dependence of the transmission coefficient $T(\varepsilon)$ in a square quantum well of thickness $L = 250 \text{ \AA}$. Curve (a) : $V_b = 224 \text{ meV}$, $m^* = 0.067 m_0$. Curve (b) : $V_b = 150 \text{ meV}$, $m^* = 0.48 m_0$

This corresponds to constructive interferences inside the quantum well slab, whose effective thickness $2L$ should fit an integer number of carrier wavelengths. The discrete energies which fulfill equation (40) are called transmission resonances. They correspond to an enhanced probability of finding the carrier *inside* the quantum well (for the same reason that lasing action, as seen from outside the laser, corresponds to a very large amplitude of the electromagnetic wave inside the cavity). The transmission resonances may equally be viewed as virtual bound states of the quantum well. Like regular bound states, the virtual ones occur for discrete energies. True bound states however, are stationary solutions of the Schrödinger equation, which decay far from the quantum well, whereas virtual bound states are associated with extended wavefunctions. Another way to describe these states is to consider them as bound but non-stationary (decaying) states, i.e. to add to their energy a finite, negative, imaginary part. This view is supported by the time-dependent analysis of the transmission of a wave-packet by the quantum well (see e.g. [2]). When equation (40) is not fulfilled, the delay experienced by the wave-packet when it travels across the well is negative, as expected classically (the particle has a larger velocity in the well). On the other hand, when equation (40) is fulfilled, one finds that the carrier spends a long time inside the quantum well slab, swinging back and forth between $z = \pm L/2$, as in a regular bound level. However the particle finally escapes from the well. The delay time is longer for the narrower transmission resonances. This means that long-lived virtual bound states can be seen only if $k_b/k_w \ll 1$, i.e. if $\varepsilon \ll V_b$.

Finally at $\varepsilon = 0$ the resonance peaks ($k_w L = p\pi$) becomes infinitely narrow. This in turn implies that the virtual bound state has become a true bound state. Indeed, the conditions $\varepsilon = 0$ and $Lk_w(0) = p\pi$ are those which express the binding of a new bound state inside the quantum well (see Eq. (18)). Thus, true and virtual bound states match at the onset of the quantum well continuum. The latter are the continuation of the former when their confinement energies exceed the height of the confining barrier.

II. Density of states.

It is important to know how many quantum states $|\nu\rangle$ per unit energy are available around a given energy ε . This quantity, the density of states, is equal to :

$$\rho(\varepsilon) = \sum_{\nu} \delta(\varepsilon - \varepsilon_{\nu}), \quad (41)$$

where ε_{ν} is the energy associated with the state $|\nu\rangle$. It often proves convenient to rewrite equation (41) in slightly different forms :

$$\rho(\varepsilon) = \sum_{\nu} \delta(\varepsilon - \varepsilon_{\nu}) = \sum_{\nu} \langle \nu | \delta(\varepsilon - \mathcal{H}) | \nu \rangle = \text{Trace } \delta(\varepsilon - \mathcal{H}), \quad (42)$$

where \mathcal{H} is the Hamiltonian operator whose eigenvalues are ε_{ν} . The advantage of the Trace is that it can be calculated on any basis $|l\rangle$, even if $|l\rangle$ are not eigenstates of \mathcal{H} .

In some instances one can also rewrite $\delta(\varepsilon - \mathcal{H})$ as :

$$\delta(\varepsilon - \mathcal{H}) = -\frac{1}{\pi} \lim_{\eta \rightarrow 0} \text{Im} (\varepsilon - \mathcal{H} + i\eta)^{-1}, \quad (43)$$

where $\text{Im } z$ stands for "imaginary part of z ". Using equations (42, 43), $\rho(\varepsilon)$ is found to be equal to :

$$\rho(\varepsilon) = -\frac{1}{\pi} \lim_{\eta \rightarrow 0} \text{Im} \sum_l \langle l | (\varepsilon - \mathcal{H} + i\eta)^{-1} | l \rangle. \quad (44)$$

For the square quantum well problem, the discrete spectrum is labelled by the index n of the bound state and, since electrons have a spin, by a spin quantum number $\sigma_z = \pm \frac{1}{2}$. Thus $|\nu\rangle = |n, \sigma_z\rangle$ and

$$\rho(\varepsilon) = 2 \sum_n \delta(\varepsilon - \varepsilon_n), \quad \varepsilon < 0 \quad (45)$$

where the factor 2 has accounted for the spin degeneracy of each level n .

We know in fact that a carrier motion is three dimensional. Thus a more realistic description of one-dimensional quantum well structures is provided by the solutions

of :

$$-\frac{\hbar^2}{2m^*} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] \psi(\mathbf{r}) + V_b(z) \psi(\mathbf{r}) = \epsilon \psi(\mathbf{r}). \quad (46)$$

As the Hamiltonian is the sum of x , y and z dependent contributions, we know that we can look for eigenfunctions which are separable in x , y , z . Moreover, the carrier motion is free along the x and y directions. Thus

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{S}} \exp[i k_x x + i k_y y] \chi(z) \quad (47)$$

where $S = L_x L_y$ is the sample area. With equation (47) we find

$$\epsilon = E_n + \frac{\hbar^2 k_\perp^2}{2m^*}, \quad (48)$$

where E_n is one of the eigenvalues of the one-dimensional Schrödinger equation of the quantum well problem (Eq. (8)) and $k_\perp^2 = k_x^2 + k_y^2$. Thus, one may associate a two-dimensional subband, which represents the kinetic energy arising from the in-plane motion of the carrier with each of the quantum well bound states E_n . The density of states associated with the motion described by equation (46) is :

$$\rho(\epsilon) = 2 \sum_{k_x k_y n} \delta \left[\epsilon - E_n - \frac{\hbar^2 k_\perp^2}{2m^*} \right], \quad E_n < 0. \quad (49)$$

The sample area S is assumed to be of macroscopic size. Applying cyclic boundary conditions to the x and y motions, we find that

$$k_x = m \frac{2\pi}{L_x}, \quad k_y = n \frac{2\pi}{L_y} \quad (50)$$

Since L_x , L_y are very large, any summation $\sum_{k_x k_y} a(k_x, k_y)$ can be converted into an integration :

$$\sum_{k_x k_y} a(k_x, k_y) \xrightarrow[L_x, L_y \rightarrow \infty]{} \frac{L_x L_y}{(2\pi)^2} \int dk_x dk_y a(k_x, k_y), \quad (51)$$

and finally

$$\rho(\epsilon) = \frac{m^* S}{\pi \hbar^2} \sum_n Y[\epsilon - E_n]; \quad E_n < 0, \quad (52)$$

where $Y(x)$ is the step function :

$$Y(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0. \end{cases} \quad (53)$$

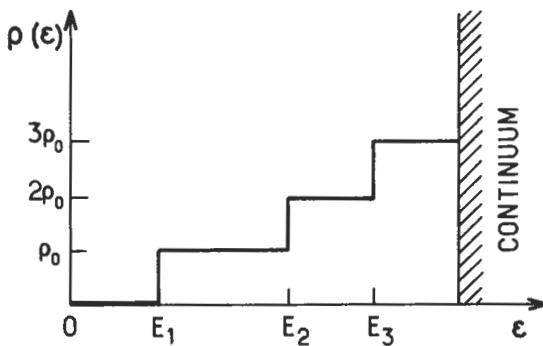


Fig. 6. — Density of states associated with the quasi bi-dimensional motion of a carrier whose z motion corresponds to bound states of a square quantum well ; $\rho_0 = m^*S/\pi\hbar^2$. The energy zero is taken at the bottom of the well.

The density of states is thus staircase-shaped. (see Fig. 6) Two facts are worth noting :

i) $\rho(\epsilon) = 0$ for $\epsilon < E_1$. A classical description would have led to allowed motions for any ϵ . Here again we find the effect of the confinement energy associated with the localization of the particle along the z direction

ii) $\rho(\epsilon)$ exhibits jumps of finite amplitude whenever the energy passes through the edge of a two-dimensional subband. Such behaviour contrasts with the smoothly varying $\rho(\epsilon)$ which is found for the free particles ($\rho(\epsilon) = A\sqrt{\epsilon}$). The reduction of the dimensionality (i.e. two *versus* three) is witnessed by the increasingly singular behaviour of the density of states.

In practice each step is smoothed out by defects. An empirical description of the broadening effects is obtained by replacing the step function $Y(\epsilon - E_n)$ by :

$$Y(\epsilon - E_n) \xrightarrow{\text{broadening}} \frac{1}{\pi} \left\{ \frac{\pi}{2} + \text{Arctan} \left[\frac{(\epsilon - E_n)}{\Gamma_n} \right] \right\} \quad (54)$$

where Γ_n is the broadening parameter.

As will be shown in other chapters, points i) and ii) have important practical consequences :

i) If a bulk material has a bandgap ϵ_g it absorbs light with an energy $\hbar\omega \geq \epsilon_g$. A slab of the same material, clad between confining barriers to make a quantum well structure, has an absorption edge which is blue-shifted (with respect to the bulk value) by the sum of the confinement energies of the electrons and the holes. By adjusting the quantum well thickness L , this blue shift can be varied from zero to a few hundredths of a milli-electron volt.

ii) If at low temperatures, the carriers only populate the lowest lying subband E_1 of a quantum well, their scattering by static defects only occurs through intra-subband (E_1) mechanisms. If, by some means, some carriers become energetic

enough to have energies which exceed E_2 , elastic inter-subband mechanisms become allowed. The onset of these new scattering channels is rather sharp in high quality materials and is witnessed by a sudden drop in the carrier mobility.

* Classical limit of the density of states

When the quantum well thickness L becomes very large, many states are bound in the well $\left[\frac{L}{\pi} k_w(0) \rightarrow \infty \right]$ and their energy separation becomes small (decreasing like L^{-2}). Let us suppose that we are interested in evaluating $\rho(\varepsilon)$ for energy ε such that

$$-V_b < \varepsilon \ll 0. \quad (55)$$

Since $L \rightarrow \infty$ the energy levels such that $E_n \ll 0$ nearly coincide with those of an infinitely deep well and $\rho(\varepsilon)$ can be approximated by the expression :

$$\rho(\varepsilon) \approx \frac{m^* S}{\pi \hbar^2} \sum_{n=1}^{\infty} Y[\varepsilon + V_b - n^2 E_1], \quad (56)$$

where E_1 is the confinement energy of the ground state (see Eq. (21)). Since L is assumed to be large E_1 is small. Thus, by converting the summation over n into an integration we obtain for $\varepsilon + V_b \gg E_1$

$$\rho(\varepsilon) \approx \frac{m^* S}{\pi \hbar^2} \left\{ \sqrt{\frac{\varepsilon + V_b}{E_1}} - 1 \right\} \approx \frac{m^* S}{\pi \hbar^2} \sqrt{\frac{\varepsilon + V_b}{E_1}} \quad (57)$$

or

$$\rho(\varepsilon) \underset{L \rightarrow \infty}{\approx} \frac{m^* S L}{\pi^2 \hbar^2} \sqrt{\frac{2m^*}{\hbar^2} (\varepsilon + V_b)}. \quad (58)$$

Equation (58) coincides with the result obtained by neglecting the effect of the quantum well potential on the energy levels of the particle, i.e. equation (58) is the density of states of a particle of mass m^* moving freely in a bulk sample having a volume LS . In practice, the classical limit equation (58) is reached if $L \geq 10^3$ Å in GaAs-Ga_{1-x}Al_xAs quantum wells.

In table I a comparison between the expressions of some physical quantities calculated for bulk materials (volume Ω) and quasi bidimensional structures is summarized.

III. Tunnel coupling between wells : the symmetric double square well.

Let us consider two equivalent one-dimensional quantum wells of height V_b , width L and separated by a distance h (Fig. 7). Each of the wells possesses n_{\max} bound states when they are isolated ($n_{\max} \geq 1$). A localised wavefunction, which may be associated with each of these bound states, exponentially decays far away from the well. In the limit of infinite h , the bound states of the discrete spectrum

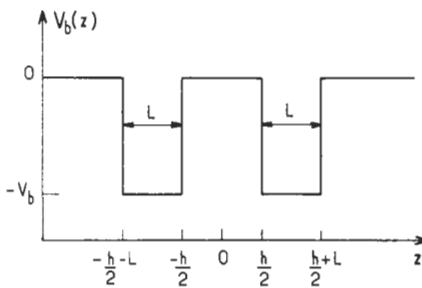


Fig. 7.— Potential energy profile of a double, symmetric, quantum well structure.

$(0 \geq \varepsilon \geq -V_b)$ are twofold degenerate : the particle can be found either in one well, or in the other. At finite h , the previous eigenstates are no longer eigenstates of the coupled wells Hamiltonian :

$$\mathcal{H} = \frac{p_z^2}{2m^*} + V_b(z - z_1) + V_b(z - z_2) \quad (59)$$

where

$$V_b(z - z_i) = \begin{cases} 0 & |z - z_i| \geq \frac{L}{2} \\ -V_b & |z - z_i| \leq \frac{L}{2} \end{cases} \quad i = 1, 2 \quad (60)$$

Let $\chi_1(z - z_2)$ be the ground state eigenfunction of the isolated well centred at $z = z_2$. Despite the exponential decay displayed by $\chi_1(z - z_2)$, when $|z - z_2| \geq \frac{L}{2}$, $\chi(z_1 - z_2)$ is different from zero. Thus $\mathcal{H}\chi_1(z - z_2)$ is not proportional to $\chi_1(z - z_2)$. However if h is large enough, we may expect the coupling between the wells (due to the tunnel effect in the middle barrier), to be small enough to admix only those states which are twofold degenerate at $h = \infty$ and not the ground states with excited states. From the exact solution of equation (59) :

$$\psi(z) = \sum_v a_v \chi_v(z - z_1) + b_v \chi_v(z - z_2), \quad (61)$$

where v runs over the discrete and continuous spectra, we only retain for the lowest lying states of the double quantum well a linear combination of the ground states of the isolated wells :

$$\psi(z) = \alpha \chi_1(z - z_1) + \beta \chi_1(z - z_2), \quad (62)$$

we obtain :

$$\begin{bmatrix} E_1 + s - \varepsilon & (E_1 - \varepsilon) r + t \\ (E_1 - \varepsilon) r + t & E_1 + s - \varepsilon \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = 0 \quad (63)$$

Table I

	quasi 2D	bulk
eigenstates	$ n, \mathbf{k}_\perp, \sigma_z\rangle, \mathbf{k}_\perp = (k_x, k_y)$	$ \mathbf{k}, \sigma_z\rangle; \mathbf{k} = (k_x, k_y, k_z)$
density of states	$\sum_n \rho_n(\varepsilon); \rho_n(\varepsilon) = \frac{m^* S}{\pi \hbar^2} Y(\varepsilon - E_n)$	$\frac{m^* \Omega}{\pi^2 \hbar^2} \sqrt{\frac{2m^* \varepsilon}{\hbar^2}}$
Fermi energy ε_F versus number of carriers(n_c). $T = 0$ K	$n_c = \frac{m^* S}{\pi \hbar^2} \sum_n [\varepsilon_F - E_n] Y[\varepsilon_F - E_n]$	$n_c = \frac{\Omega}{3\pi^2} \left[\frac{2m^*}{\hbar^2} \varepsilon_F \right]^{3/2}$
average kinetic energy per particle : $\frac{1}{n_c} \left\langle \frac{\hbar^2 k^2}{2m^*} \right\rangle, T = 0$ K	$\frac{1}{2} \frac{\sum_n (\varepsilon_F - E_n)^2 Y[\varepsilon_F - E_n]}{\sum_n (\varepsilon_F - E_n) Y[\varepsilon_F - E_n]}$	$\frac{3}{5} \varepsilon_F$
$n_c(\mu, T)$ μ : chemical potential k_B :Boltzmann constant	Finite temperature, no broadening	
	$n_c = \frac{m^* S}{\pi \hbar^2} k_B T \sum_n \log \left[1 + \exp \left[\frac{\mu - E_n}{k_B T} \right] \right]$	$n_c = \frac{m^* \Omega k_B T}{\pi^2 \hbar^2} \sqrt{\frac{2m^* k_B T}{\hbar^2}} \int_0^\infty \frac{x^{1/2} dx}{1 + \exp[x - \mu/k_B T]}$
high temperature limit	$n_c \approx \frac{m^* S}{\pi \hbar^2} k_B T \sum_n \exp \left[\frac{\mu - E_n}{k_B T} \right]$	$n_c \approx \frac{m^* \Omega k_B T}{2\pi^{3/2} \hbar^2} \sqrt{\frac{2m^* k_B T}{\hbar^2}} \exp(\mu/k_B T)$
density of states	Phenomenological broadening, $T = 0$ K	$\frac{m^* \Omega}{\pi^2 \hbar^2} \sqrt{\frac{2m^* \Gamma}{\hbar^2}} \frac{\pi}{\pi^2 \hbar^2} \int_0^\infty \frac{\sqrt{x} dx}{\Gamma^2 + (x - \varepsilon)^2}$

or :

$$\varepsilon = E_1 \mp \frac{t}{1 \mp r} + \frac{s}{1 \mp r} \quad (64)$$

where E_1 is the energy of the ground bound state of the wells when they are isolated and

$$\begin{aligned} r &= \langle \chi_1(z - z_1) | \chi_1(z - z_2) \rangle \\ &= \langle \chi_1(z - z_2) | \chi_1(z - z_1) \rangle \end{aligned} \quad (65)$$

$$\begin{aligned} s &= \langle \chi_1(z - z_1) | V_b(z - z_2) | \chi_1(z - z_1) \rangle \\ &= \langle \chi_1(z - z_2) | V_b(z - z_1) | \chi_1(z - z_2) \rangle \end{aligned} \quad (66)$$

$$\begin{aligned} t &= \langle \chi_1(z - z_1) | V_b(z - z_1) | \chi_1(z - z_2) \rangle \\ &= \langle \chi_1(z - z_2) | V_b(z - z_2) | \chi_1(z - z_1) \rangle \end{aligned} \quad (67)$$

The quantities r , s , t are called overlap (r), shift (s) and transfer (t) integrals respectively. The overlap integral is positive whereas both the transfer and the shift integral are negative. It is usual to neglect r , although it is not quantitatively a very good approximation. In this case :

$$\varepsilon = E_1 \mp t + s, \quad (68)$$

s is thus interpreted as the shift of the ground state of each well due to the presence of the other well. This shift does not lift the twofold degeneracy prevailing at infinite h (see Fig. 8). This degeneracy is lifted by the transfer term t . The ground state corresponds to the symmetric combination ($\alpha = \beta$) of isolated quantum well wavefunctions, whereas the excited state corresponds to $\alpha = -\beta$.

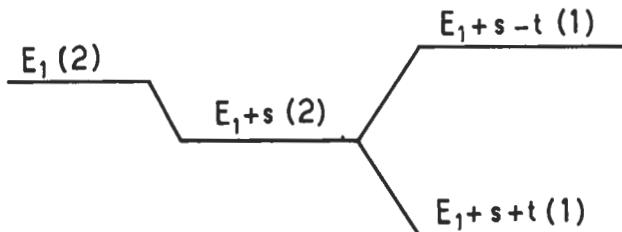


Fig. 8. — Shift and lifting of the degeneracy of the twofold degenerate ground state of isolated quantum wells due to the non-zero coupling between the wells.

Finally let us give the exact solution of the double, symmetric, quantum well problem. The eigenstates are the solutions of the equation :

$$2 \cos k_w L + \left(\xi - \frac{1}{\xi} \right) \sin k_w L \pm \left(\xi + \frac{1}{\xi} \right) e^{-\kappa_b h} \sin k_w L = 0. \quad (69)$$

where :

$$\xi = \frac{\kappa_b}{k_w}; \quad \kappa_b = \sqrt{\frac{-2m^* \epsilon}{\hbar^2}}; \quad k_w = \sqrt{\frac{2m^*}{\hbar^2} (\epsilon + V_b)}. \quad (70)$$

The minus (respectively plus) sign in equation (69) refers to the symmetric (respectively antisymmetric) states with respect to the centre of the structure.

The previous equation admits for any h at least one solution (like any one-dimensional problem with attractive and symmetric potential). When h decreases from infinity to zero this bound state, which is a symmetric state, continuously evolves from the ground state of a well with thickness L to the ground state of a well with thickness $2L$. An antisymmetric state remains bound for any h if the quantum well of thickness $2L$ binds at least two states (i.e. if $Lk_w(\epsilon = 0) \geq \pi$). If the quantum well of thickness $2L$ only binds one state, the ground antisymmetric state of the double well, bound at very large h , becomes unbound when $h < h_c$, where :

$$2 \cotan [Lk_w(\epsilon = 0)] = h_c k_w(\epsilon = 0) \quad (71)$$

As in the single well problem, the bound states of the double well transform into transmission resonances when their confinement energy exceeds the top of the confining barrier. The resonances can be narrow when the ratio of the wavevectors k_b/k_w is $\ll 1$.

IV. Superlattices.

IV.1 SUPERLATTICE DISPERSION RELATIONS. — Let us consider an infinite sequence of quantum wells of thickness L , separated by barriers of thickness h (Fig. 9). The potential energy $V_b(z)$ is a periodic function of z with a period of d ($d = L + h$) :

$$V_b(z) = \sum_{n=-\infty}^{+\infty} V_b(z - nd), \quad (72)$$

$$V_b(z - nd) = \begin{cases} -V_b & \text{if } |z - nd| \leq \frac{L}{2} \\ 0 & \text{if } |z - nd| > \frac{L}{2} \end{cases} \quad (73)$$

The solutions of the one-dimensional Schrödinger equation is known exactly for each layer : well-acting (A) layer or barrier-acting (B) layer. For positive energies we have :

$$\chi(z) = \alpha \exp[i k_w(z - nd)] + \beta \exp[-i k_w(z - nd)]; \quad |z - nd| \leq \frac{L}{2} \quad (74)$$

$$\chi(z) = \gamma \exp\left[i k_b\left(z - nd - \frac{d}{2}\right)\right] + \delta \exp\left[-i k_b\left(z - nd - \frac{d}{2}\right)\right] \\ \left|z - nd - \frac{d}{2}\right| \leq \frac{h}{2} \quad (75)$$

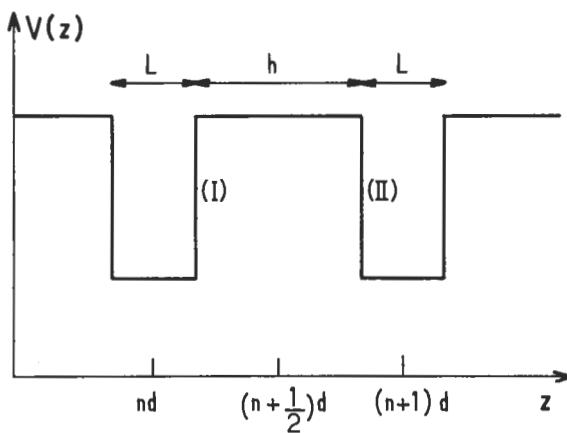


Fig. 9.— Potential energy profile of a segment of a superlattice which consists of an infinite sequence of square quantum wells.

and :
$$\varepsilon = \frac{\hbar^2 k_b^2}{2m^*} = -V_b + \frac{\hbar^2 k_w^2}{2m^*}. \quad (76)$$

We shall now exploit the periodicity of $V_b(z)$. Consider the translation operator \mathcal{T}_d which is such that for any function $f(z)$, $\mathcal{T}_d f(z) = f(z + d)$. The operator \mathcal{T}_d commutes with the Hamiltonian \mathcal{H} of the particle. We can thus find the eigenfunctions of \mathcal{H} which are also the eigenfunctions of \mathcal{T}_d . The \mathcal{T}_d eigenvalues can be written $\exp(iqd)$ where q is an arbitrary complex number. Moreover $\mathcal{T}_{nd} = [\mathcal{T}_d]^n$. Thus the functions $\chi_q(z)$, which are eigenfunctions of both \mathcal{H} and \mathcal{T}_d , must fulfill not only the Schrödinger equation but also the equation :

$$\mathcal{T}_{nd} \chi_q(z) = \chi_q(z + nd) = \exp[iqnd] \chi_q(z). \quad (77)$$

Equivalently to equation (77), the χ_q 's can be written in the form :

$$\chi_q(z) = u_q(z) \exp[iqz]; u_q(z + d) = u_q(z), \quad (78)$$

i.e. the χ_q 's are the products of a plane wave term by a periodic function in z . Equations (77, 78) constitute the Bloch-Floquet theorem ; χ_q is the Bloch function and u_q is its associated periodic part.

Let us impose cyclic boundary conditions on the eigenstates of \mathcal{H} . We imagine that the length of the crystal Nd is very large and we assume that the crystal maps onto itself end to end. Thus $\chi_q(z)$ must be such that :

$$\chi_q(z + Nd) = \chi_q(z), \quad (79)$$

which means that :

$$qNd = 2p\pi, \quad p \text{ integer.} \quad (80)$$

Consequently q must be real. N independent values of q are allowed, the spacing between any two of these consecutive values being $2\pi/Nd$. Without loss of generality we may restrict q to the segment $[-\frac{\pi}{d}, +\frac{\pi}{d}]$. The q space is called the reciprocal space and the segment $[-\frac{\pi}{d}, +\frac{\pi}{d}]$, which is a particular unit cell of the reciprocal space, is the first Brillouin zone. In the limit of large N , the discreteness of the allowed q values becomes unimportant and we may convert any summation of the form $\sum_{q \in \text{first B.Z.}} g(q)$ into an integral :

$$\sum_{q \in \text{first B.Z.}} g(q) \xrightarrow{N \rightarrow \infty} \frac{Nd}{2\pi} \int_{-\frac{\pi}{d}}^{+\frac{\pi}{d}} g(q) dq, \quad (81)$$

where B.Z. stands for Brillouin zone. For a more complete treatment of reciprocal space, Brillouin zone, etc... the reader is for example referred to Ashcroft and Mermin's textbook [11].

Let us express the continuity of χ_q and $\frac{d\chi_q}{dz}$ at the interface I of figure 9 :

$$\begin{cases} \alpha \exp\left(ik_w \frac{L}{2}\right) + \beta \exp\left(-ik_w \frac{L}{2}\right) = \\ \quad = \gamma \exp\left(-ik_b \frac{h}{2}\right) + \delta \exp\left(ik_b \frac{h}{2}\right) \\ ik_w \left[\alpha \exp\left(ik_w \frac{L}{2}\right) - \beta \exp\left(-ik_w \frac{L}{2}\right) \right] = \\ \quad = ik_b \left[\gamma \exp\left(-ik_b \frac{h}{2}\right) - \delta \exp\left(ik_b \frac{h}{2}\right) \right]. \end{cases} \quad (82)$$

At the interface II of figure 9 we have :

$$\chi_q \left[z = (n+1)d - \frac{L}{2} + 0 \right] = \gamma \exp\left(ik_b \frac{h}{2}\right) + \delta \exp\left(-ik_b \frac{h}{2}\right). \quad (83)$$

To express $\chi_q \left[z = (n+1)d - \frac{L}{2} + 0 \right]$ we take advantage of the Bloch theorem and write

$$\chi_q \left[z = (n+1)d - \frac{L}{2} + 0 \right] = e^{iqd} \chi_q \left[z = nd - \frac{L}{2} + 0 \right]. \quad (84)$$

Using the same method for the derivative, we finally end up with a 4×4 homogeneous system in $\alpha, \beta, \gamma, \delta$. Non trivial solutions exist only if the following equation is satisfied :

$$\cos(qd) = \cos(k_w L) \cos(k_b h) - \frac{1}{2} \left(\xi + \frac{1}{\xi} \right) \sin(k_w L) \sin(k_b h) \quad (85)$$

where :

$$\xi = k_b/k_w. \quad (86)$$

Equations (85, 86) are implicit equations linking the allowed energy ε and the wave vector q , i.e. of the dispersion relations $\varepsilon(q)$ of the Bloch waves. For a given value of q , these equations admit an infinite number of solutions. Thus, in order to label the various solutions, a subband index n is affixed to ε and χ . A one-dimensional Bloch state depends on two orbital quantum numbers n, q and will thus be written as $|nq\rangle$. The allowed subbands ε_{nq} are separated by forbidden gaps. For energies corresponding to these energy gaps, there are no allowed Bloch states, i.e. there are no allowed eigenstates of \mathcal{H} which fulfill the cyclic boundary conditions of equation (79). Evanescent states, corresponding to imaginary q 's, do however exist and are only allowed if the superlattice has a finite length along the z axis.

The magnitude of the forbidden gaps which separate two consecutive superlattice subbands $|n, q\rangle, |n+1, q\rangle$ decreases when n increases. In fact at large $\varepsilon (\varepsilon \gg V_b)$, $k_b \sim k_w$ and equation (85) reduces to $k_w = q + 2j \frac{\pi}{d}$. The superlattice potential is hardly felt by highly energetic electrons, or equivalently the electron wavelength $(2\pi/k_w)$ becomes much smaller than the characteristic length of the obstacles (d) and thus the effects of diffraction become negligible.

It can be noticed that the energies ε , which are such that $k_w L = p\pi$, always correspond to an allowed superlattice state (see Eq. (85)). The superlattice states which correspond to the continuum states of the isolated wells can thus be viewed as the result of the hybridization of the virtual bound states of isolated wells.

We have up to now only considered the case of Bloch states associated with propagating states in both kinds of layers (wells and barriers) : $\varepsilon_{nq} \geq 0$. Let us now look at the case corresponding to $-V_b \leq \varepsilon_{nq} \leq 0$. The wave is evanescent in the barrier and thus :

$$k_b = i\kappa_b; \quad \kappa_b = \sqrt{\frac{-2m^* \varepsilon}{\hbar^2}} \quad (87)$$

$$\xi \rightarrow i\tilde{\xi}; \quad \tilde{\xi} = \frac{\kappa_b}{k_w}. \quad (88)$$

Allowed Bloch states must then fulfill :

$$\cos(qd) = \cos(k_w L) \cosh(\kappa_b h) - \frac{1}{2} \left[-\tilde{\xi} + \frac{1}{\tilde{\xi}} \right] \sin(k_w L) \sinh(\kappa_b h). \quad (89)$$

In the limit of infinitely thick barriers the right-hand side of equation (89) will diverge like $\exp(\kappa_b h)$, unless the multiplicative coefficient in front of it vanishes. This occurs if :

$$\cos(k_w L) - \frac{1}{2} \left[-\tilde{\xi} + \frac{1}{\tilde{\xi}} \right] \sin(k_w L) = 0, \quad (90)$$

which coincides with the transcendental equation whose solutions are the bound states of isolated quantum wells of thickness L . Actually equation (90) is the product of equation (15) by equation (16), i.e. we obtain all the levels, odd or even, at the same time. This is because we directly exploited the symmetry property (evenness in z) of $V_b(z)$ in the single well analysis.

Thus, the superlattice subbands of negative energies appear as the hybridization of isolated quantum well bound states due to the tunnel coupling between the wells across the finite barriers. If these barriers are thick (and/or high) but finite, the product $\kappa_b h$ is large and the subbands are narrow (the subband width which is proportional to the transfer integral, decreases roughly exponentially with h at fixed κ_b). The function $F(\epsilon)$ which appears on the right-hand side of equation (89) shows large variations with ϵ and the energy segments where $|F(\epsilon)| < 1$ are narrow (see Fig. 10). In order to obtain approximate, but convenient, subband dispersion relations, let us denote by E_1, E_2, \dots, E_n , the isolated quantum well bound states solutions of equation (90). In addition, let us expand $F(\epsilon)$ in the vicinity of the E 's, say E_j . To the first order in $\epsilon - E_j$ we obtain the dispersion relation of the j^{th} subband in the form :

$$\epsilon_j(q) = E_j + s_j + 2t_j \cos(qd), \quad (91)$$

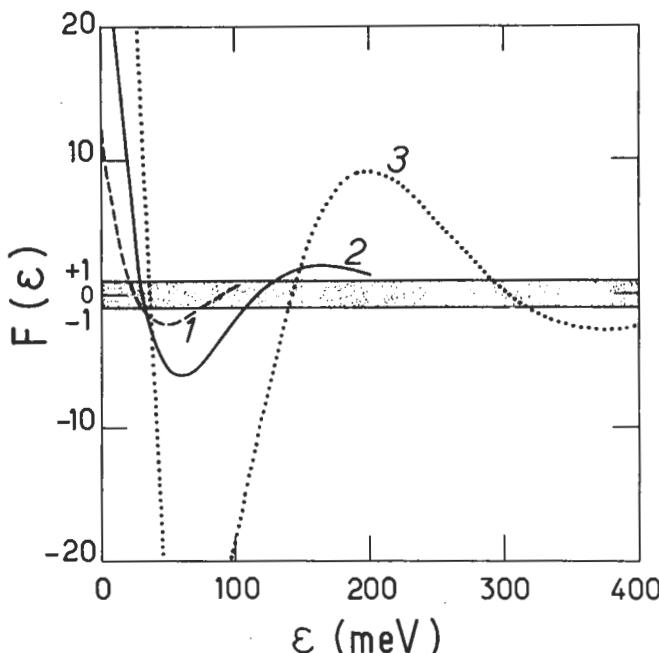


Fig. 10. — The right-hand side of equation (89) is plotted *versus* the energy ϵ . Three barrier heights are considered: curve (1) : $V_b = 0.1$ eV ; curve (2) : $V_b = 0.2$ eV ; curve (3) : $V_b = 0.4$ eV. $m^* = 0.067 m_0$, $L = 100$ Å, $h = 50$ Å. In all three curves the energy zero is taken at the bottom of the well. Each curve is interrupted at the top of the well. The shaded area corresponds to the allowed superlattice states : $|F(\epsilon)| < 1$.

where :

$$s_j = \frac{-F(E_j)}{[F'(\varepsilon)]_{\varepsilon=E_j}} ; \quad 2t_j = \frac{1}{[F'(\varepsilon)]_{\varepsilon=E_j}} \quad (92)$$

and where the prime denotes the derivative with respect to ε . Equation (91) is simply a tight-binding result (see e.g. [11]). Explicit expressions for t_j and s_j can be obtained

i) by expanding $\chi_{jq}(z)$ in terms of the "atomic" functions $\varphi_{\text{loc}}^{(j)}(z - nd)$, where $\varphi_{\text{loc}}^{(j)}(z - nd)$ is the j^{th} bound state wavefunction of the quantum well centred at $z_n = nd$ when it is considered as isolated :

$$\chi_{jq}(z) = \frac{1}{\sqrt{N}} \sum_n \exp(iqnd) \varphi_{\text{loc}}^{(j)}(z - nd) \quad (93)$$

and

ii) by retaining only nearest-neighbour interactions :

$$t_j = \int \varphi_{\text{loc}}^{(j)}(z) V_b(z) \varphi_{\text{loc}}^{(j)}(z - d) \quad (94)$$

$$s_j = \sum_{n \neq 0} \int \varphi_{\text{loc}}^{(j)}(z) V_b(z - nd) \varphi_{\text{loc}}^{(j)}(z). \quad (95)$$

IV.2 SYMMETRY PROPERTIES OF THE EIGENFUNCTIONS. — For $q = 0$ or $q = \frac{\pi}{d}$, the Bloch wave is stationary [$\chi_{nq=0}(z + d) = \chi_{nq=0}(z)$; $\chi_{n,q=\frac{\pi}{d}}(z + d) = -\chi_{n,q=\frac{\pi}{d}}(z)$]. The wavefunction repeats itself identically ($q = 0$) or changes sign ($q = \frac{\pi}{d}$) when going from one elementary cell to the next. Let us consider the ratio $r = \alpha/\beta$ (see Eq. (74)). We find that :

$$r = -\frac{(1 - \xi)}{(1 + \xi)} \times \frac{[\exp(i k_b h) - \exp(i (k_w L + qd))]}{[\exp(i (k_w L + k_b h)) - \exp(i qd)]} \quad (96)$$

One can show that for $q = 0$, $q = \frac{\pi}{d}$ the ratio r is real and such that $|r|^2 = 1$. Thus $r = \pm 1$. This shows that for $q = 0$ and $q = \frac{\pi}{d}$ the Bloch function is either odd or even with respect to the centres of the well acting layers. The same conclusion holds for the parity with respect to the centre of the barrier-acting layers. These analytical results arise from the symmetry properties of the superlattice Hamiltonian : let Δ_A, Δ_B be two parallel axes separated by $\frac{d}{2}$. The product of two symmetries with respect to these two axes is equal to a translation \mathcal{C}_d , the vector $\mathbf{d}(|\mathbf{d}| = d)$ being perpendicular to Δ_A, Δ_B . Let us apply this geometrical property to the axis Δ_A, Δ_B which bisects two consecutive A and B segments of a superlattice (Fig. 11). We have

$$R_B R_A = \mathcal{C}_d ; \quad R_A R_B = \mathcal{C}_{-d} \quad (97)$$

where \mathcal{C}_d is the translation characterized by $\mathbf{d} = d\hat{\mathbf{z}}$ with $d = L + h$. We know that the Bloch states are eigenstates of \mathcal{C}_d with an eigenvalue e^{iqd} . In addition however

$$[V_b(z), R_A] = [V_b(z), R_B] = 0 \quad (98)$$

and also

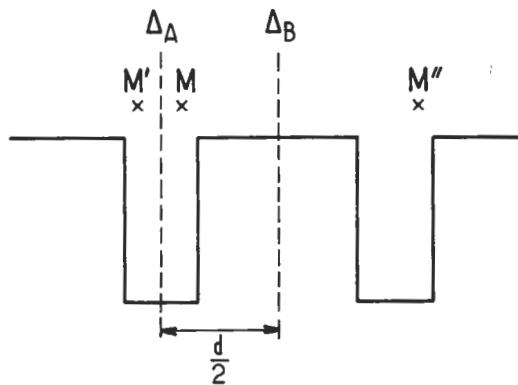
$$\left[\frac{P_z^2}{2m^*}, R_A \right] = \left[\frac{P_z^2}{2m^*}, R_B \right] = 0 \quad (99)$$

Thus R_A, R_B commute with the superlattice Hamiltonian \mathcal{K} . However R_A, R_B do not commute with each other since

$$[R_A, R_B] = -\mathcal{C}_d + \mathcal{C}_{-d} \quad (100)$$

It is thus impossible to diagonalize simultaneously \mathcal{K} , \mathcal{C}_d , R_A and R_B in the general case. The points $q = 0$ and $q = \frac{\pi}{d}$ of the Brillouin zone are noticeable exceptions since the commutator $[R_A, R_B]$, evaluated over the Bloch states, vanishes for these two values of q . For $q = 0$, $R_B R_A \chi_{n, q=0}(z) = +\chi_{n, q=0}(z)$. If we know that $\chi_{n, q=0}(z)$ is even (odd) with respect to the centre of the A layers, then it should also be even (odd) with respect to the centres of the B layers. For $q = \frac{\pi}{d}$, $R_B R_A [\chi_{n, q = \frac{\pi}{d}}] = -\chi_{n, q = \frac{\pi}{d}}(z)$. Thus if $\chi_{n, q = \frac{\pi}{d}}(z)$ is even (odd) with respect to the centre of the A layers, it should be odd (even) with respect to the centres of the B layers.

IV.3 SUPERLATTICE DENSITY OF STATES. — Let us reintroduce the free motion in the layer plane (x, y directions) and label a superlattice state with $n, q, \mathbf{k}_\perp, \sigma_z$ where n is the subband index, q the superlattice wavevector ($|q| < \frac{\pi}{d}$), $\mathbf{k}_\perp = (k_x, k_y)$ the



$$M' = R_A(M), \quad M'' = R_B(M') = \mathcal{C}_d(M)$$

Fig. 11.

carrier wavevector in the layer plane and σ_z the electron spin orientation ($\sigma_z = \pm \frac{1}{2}$). The eigenenergies

$$\varepsilon(n, q, \mathbf{k}_\perp, \sigma_z) = \frac{\hbar^2 k_\perp^2}{2m^*} + \varepsilon_n(q) \quad (101)$$

are σ_z -independent and are the sum of the energies corresponding to the motions parallel ($\varepsilon_n(q)$) and perpendicular ($\frac{\hbar^2 k_\perp^2}{2m^*}$) to the z axis. Nd denotes the length of the superlattice along the z direction and S the area of the layers. Both Nd and $S = L_x L_y$ are macroscopic quantities. The density of states $\rho(\varepsilon)$ is then equal to

$$\rho(\varepsilon) = \sum_{n, q, \mathbf{k}_\perp, \sigma_z} \delta \left[\varepsilon - \varepsilon_{nq} - \frac{\hbar^2 k_\perp^2}{2m^*} \right] \quad (102)$$

or

$$\rho(\varepsilon) = \frac{SNd}{\pi^2} \frac{m^*}{\hbar^2} \sum_n \int_0^{\frac{\pi}{d}} dq Y[\varepsilon - \varepsilon_{nq}] = \sum_n \rho_n(\varepsilon) \quad (103)$$

where $\rho_n(\varepsilon)$ is the density of states associated with the n^{th} subband and $Y(x)$ is the step function.

In general the subband ε_{nq} has a finite energy width when q describes the first Brillouin zone: $\varepsilon_{\min} \leq \varepsilon_{nq} \leq \varepsilon_{\max}$. For $\varepsilon < \varepsilon_{\min}$, $\rho_n(\varepsilon) = 0$ whereas for $\varepsilon > \varepsilon_{\max}$, $\rho_n(\varepsilon)$ is a constant :

$$\rho_n(\varepsilon) = N \frac{m^* S}{\pi \hbar^2} \quad \varepsilon > \varepsilon_{\max}. \quad (104)$$

We observe that the plateau value $N \frac{m^* S}{\pi \hbar^2}$ is equal to N times the density of states associated with a single bound state of a given quantum well. From equation (104) we deduce that when ε falls into the bandgaps of the superlattice dispersion relations, the density of states is quantized in units of $N m^* S / \pi \hbar^2$. To study what happens when ε corresponds to the allowed superlattice states, let us take the simple case :

$$\varepsilon_{nq} = E_n + s_n - 2 |t_n| \cos(qd), \quad (105)$$

which, as shown previously, corresponds to a tight-binding description of the superlattice subbands. In the vicinity of $q = 0$ we find :

$$\varepsilon_{nq} \approx E_n + s_n - 2 |t_n| + |t_n| q^2 d^2, \quad (106)$$

whereas in the vicinity of $q = \frac{\pi}{d}$, we find :

$$\varepsilon_{nq} = E_n + s_n + 2 |t_n| - |t_n| \left[q - \frac{\pi}{d} \right]^2 d^2. \quad (107)$$

Notice that in the vicinity of both $q = 0$ and $q = \frac{\pi}{d}$ the carrier dispersion relation is quadratic upon q . In both cases the effective mass along the \hat{z} axis is proportional to the modulus of the transfer integral $|t_n|$. The smaller $|t_n|$ (thick and/or high barrier), the heavier the effective mass. In the limit of vanishing $|t_n|$ (isolated quantum wells) the subband dispersion is flat (discrete level). For uncoupled discrete levels the effective mass for the propagation along \hat{z} is infinite, although the carrier is characterized in each layer by a finite mass m^* . The reason for this is that the carrier oscillates back and forth in the well and on average does not move along the \hat{z} axis. Thus its *effective* mass is infinite. Using equation (105) the density of states is readily evaluated :

$$\rho_n(\varepsilon) = \begin{cases} 0 & \text{if } \varepsilon < E_n + s_n - 2|t_n| \\ \frac{\rho_0}{\pi} \arccos \left\{ \frac{-\varepsilon + E_n + s_n}{2|t_n|} \right\} & \text{if } |\varepsilon - E_n - s_n| < 2|t_n| \\ \rho_0 & \text{if } \varepsilon > E_n + s_n + 2|t_n| \end{cases} \quad (108)$$

where :

$$\rho_0 = \frac{m^* NS}{\pi \hbar^2}. \quad (109)$$

In the vicinity of $E_n + s_n \pm 2|t_n|$ the derivatives of $\rho_n(\varepsilon)$ are infinite. These two van Hove singularities respectively correspond to a three-dimensional minimum ($q = 0$) where $\frac{\partial^2 \varepsilon}{\partial q^2} > 0$, $\frac{\partial^2 \varepsilon}{\partial k_x^2} > 0$, $\frac{\partial^2 \varepsilon}{\partial k_y^2} > 0$ and to a saddle point $\left(q = \frac{\pi}{d}\right)$ where $\frac{\partial^2 \varepsilon}{\partial q^2} < 0$, $\frac{\partial^2 \varepsilon}{\partial k_x^2} > 0$, $\frac{\partial^2 \varepsilon}{\partial k_y^2} > 0$. The energy dependence of $\rho_n(\varepsilon)$ is sketched in figure 12.

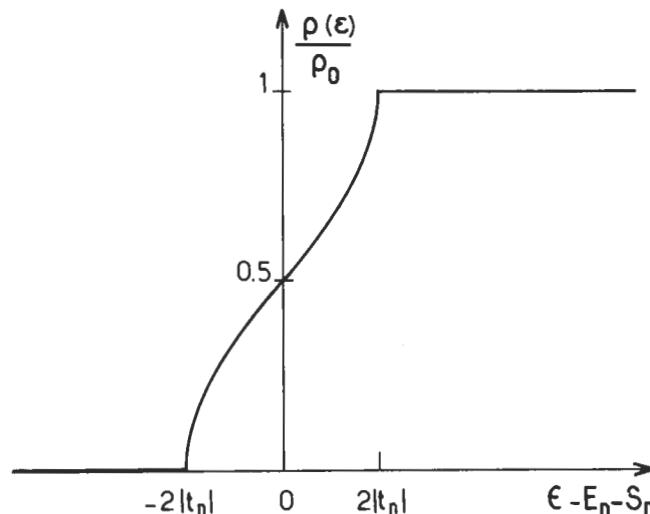


Fig. 12. — Energy dependence of the density of states of the superlattice subband corresponding to equation (105). $\rho_0 = m^* NS / \pi \hbar^2$

Appendix A

Symmetric double square well. Time-dependent aspect.

Let us consider a symmetric double square well and assume that at time $\tau = 0$ the particle is in the state $|\chi_1\rangle$: the ground state of the well "1". We want to evaluate the probability $P(T)$ of finding the particle at time T in the state $|\chi_2\rangle$: the ground state of the well "2". If $|\Psi(\tau)\rangle$ denotes the state of the particle at time τ , the rules of the quantum mechanics assert that :

$$\mathcal{P}(T) = |\langle\chi_2|\Psi(T)\rangle|^2. \quad (\text{A1})$$

To evaluate $|\Psi(\tau)\rangle$ we have to solve the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial\tau} |\Psi(\tau)\rangle = (T + V_1 + V_2) |\Psi(\tau)\rangle, \quad (\text{A2})$$

with the boundary condition :

$$|\Psi(\tau = 0)\rangle = |\chi_1\rangle. \quad (\text{A3})$$

In equation (A2) we have :

$$V_1 = V_b(z - z_1); V_2 = V_b(z - z_2), \quad (\text{A4})$$

and similarly :

$$\langle z|\chi_1\rangle = \chi_1(z - z_1); \quad \langle z|\chi_2\rangle = \chi_1(z - z_2). \quad (\text{A5})$$

E_1 denotes the eigenenergy of the ground state of an isolated well. To solve equation (A2) we make the same assumptions as when we derived equations (61-68), i.e. we write :

$$|\Psi(\tau)\rangle = \{a_1(\tau)|\chi_1\rangle + a_2(\tau)|\chi_2\rangle\} e^{-iE_1\tau/\hbar}. \quad (\text{A6})$$

If we insert equation (A6) into equation (A2) and project them onto the $|\chi_1\rangle$ and $\langle\chi_2|$ states, we obtain :

$$i\hbar \frac{da_1}{d\tau} = sa_1 + ta_2 \quad (\text{A7})$$

$$i\hbar \frac{da_2}{d\tau} = sa_2 + ta_1, \quad (\text{A8})$$

where s is the shift integral :

$$s = \langle\chi_1|V_2|\chi_1\rangle = \langle\chi_2|V_1|\chi_2\rangle, \quad (\text{A9})$$

and t is the transfer integral :

$$t = \langle\chi_1|V_1|\chi_2\rangle = \langle\chi_2|V_2|\chi_1\rangle. \quad (\text{A10})$$

and where we have neglected the overlap $\langle\chi_1|\chi_2\rangle$ for simplicity.

When we account for the boundary condition we obtain :

$$|\Psi(T)\rangle = \left[\cos\left(\frac{tT}{\hbar}\right) |\chi_1\rangle - i \sin\left(\frac{tT}{\hbar}\right) |\chi_2\rangle \right] \exp\left[-i(E_1 + s)\frac{T}{\hbar}\right], \quad (\text{A11})$$

and

$$P(T) = \sin^2\left[\frac{tT}{\hbar}\right]. \quad (\text{A12})$$

The physical interpretation of the transfer integral t is thus such that $\frac{2\pi\hbar}{4|t|}$ is the jump time from site "1" to site "2". The probability $P(T)$ oscillates with time with a period of $2\pi\hbar/2|t|$ which can be interpreted as the time needed for a round trip excursion from site "1" to site "2".

Appendix B

Dispersion relations of superlattices with arbitrary potential profiles.

In this appendix we show that the dispersion relations $\epsilon(q)$ of a superlattice with an arbitrary potential shape always takes the form

$$\cos(qd) = F(\epsilon). \quad (\text{B1})$$

The Schrödinger equation is :

$$\left[\frac{-\hbar^2}{2m^*} \frac{d^2}{dz^2} + V(z) \right] \chi(z) = \epsilon \chi(z), \quad (\text{B2})$$

where the potential energy $V(z)$ is periodic in z with a periodicity d :

$$V(z + d) = V(z). \quad (\text{B3})$$

Generally speaking we can consider the segment $[0, d]$ as a unit cell of the superlattice. Inside this segment the Schrödinger equation, (which is a second order differential equation), admits two independent solutions $f(z)$ and $g(z)$ which are implicitly ϵ -dependent. Thus, the superlattice wave function associated with the energy ϵ is a linear combination of these two solutions :

$$\chi(z) = \alpha f(z) + \beta g(z); \quad 0 \leq z \leq d. \quad (\text{B4})$$

We apply the Bloch theorem at $z = d$ and write on the one hand :

$$\chi(d) = \alpha f(d) + \beta g(d) \quad (\text{B5})$$

$$\chi'(d) = \alpha f'(d) + \beta g'(d), \quad (\text{B6})$$

and on the other :

$$\chi(d) = e^{iqd} \chi(0) = e^{iqd} [\alpha f(0) + \beta g(0)] \quad (\text{B7})$$

$$\chi'(d) = e^{iqd} \chi'(0) = e^{iqd} [\alpha f'(0) + \beta g'(0)], \quad (\text{B8})$$

by virtue of the Bloch theorem. In equations (B5-B8) the prime denotes the differentiation with respect to z . Equating expressions (B5) to (B7) and (B6) to (B8) we finally obtain :

$$\cos(qd) = \frac{1}{2w} [f(d)g'(0) - f'(0)g(d) + f(0)g'(d) - f'(d)g(0)], \quad (B9)$$

where w is the Wronskian of f and g

$$w(z) = f(z)g'(z) - f'(z)g(z), \quad (B10)$$

which is a constant ($w'(z) = 0$). Thus, irrespective of the potential shape, the superlattice dispersion relations have the same form as stated in equation (B1). A similar conclusion was established in a slightly different way in reference [11]. Equation (B9) is of little use unless we know $f(z)$ and $g(z)$ explicitly. These favourable circumstances are met in some practical cases ; for instance when $V(z)$ linearly or quadratically varies with z inside the unit cell.

References

Many excellent textbooks on Quantum Mechanics and Solid State Physics are available. Among others the following may be consulted

1) Quantum Mechanics

- [1] A. MESSIAH, *Mécanique Quantique*, Vols 1, 2 (Dunod, Paris) 1959.
- [2] D. BOHM, *Quantum Theory*, Prentice-Hall (New-York) 1951.
- [3] L.I. SCHIFF, *Quantum Mechanics* (Mc Graw-Hill, New-York) 1955.
- [4] L.LANDAU and E. LIFSHITZ, *Quantum Mechanics* (Addison-Wesley, New-York) 1958.
- [5] C. COHEN-TANNOUDJI, B. DIU, F. LALOE, *Mécanique Quantique*, Vols 1, 2 (Hermann, Paris) 1977.

2) Solid State Physics

- [6] C. KITTEL, *Introduction to Solid State Physics* (Wiley, New-York) 1956.
- [7] C. KITTEL, *Quantum Theory of Solids* (Wiley, New-York) 1963.
- [8] G.H. WANNIER, *Elements of Solid State Theory* (Cambridge University Press) 1960.
- [9] W.A. HARRISON, *Solid State Theory* (Mc.Graw Hill, New-York) 1970.
- [10] S.M. SZE, *Physics of Semiconductor Devices* (Wiley, New-York) 1969.
- [11] N.W. ASHCROFT and N.D. MERMIN, *Solid State Physics* (Holt, Rinehart and Winston, New-York) 1976.
- [12] J.C. PHILLIPS, *Bonds and Bands in Semiconductors* (Academic Press, London) 1973.

An up to date and invaluable reference for electronic properties of two dimensional systems is T. ANDO, A.B. FOWLER, F. STERN, *Rev. Mod. Phys.* **54** (1982) 437.

CHAPTER II

Band structure of bulk III-V compounds

In this second chapter we will recall some of the electronic properties of the bulk materials involved in the III-V or II-VI heterostructures.

I. Crystalline properties.

The III-V compounds crystallize in the zinc-blende (spharelite) structure. This lattice consists of two interpenetrating, face-centred cubic (f.c.c.) lattices, displaced from one another by a fourth of one of the cube main diagonals (see Fig. 1). The zinc blende lattice is not a Bravais lattice because the elementary cell contains two atoms, e.g. one located at the origin and the other at $\left(\frac{a}{4}, \frac{a}{4}, \frac{a}{4}\right)$ where a is the cube side.

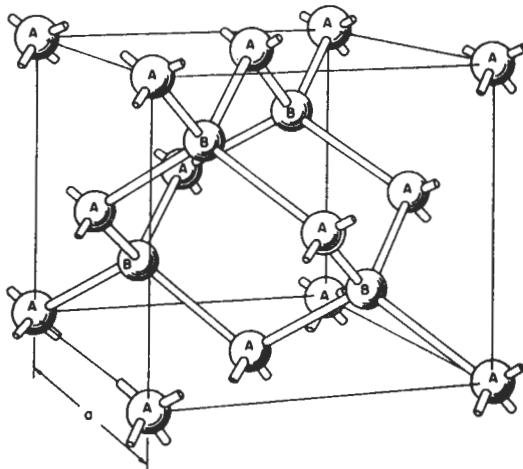


Fig. 1. — Spatial arrangement of atoms in a zinc-blende lattice.

The reciprocal lattice of the Bravais lattice underlying the zinc blende lattice (i.e. a f.c.c. lattice) is a body centred cubic lattice. The first Brillouin zone of the reciprocal lattice is a truncated octahedron (Fig. 2). Several high symmetry points or lines of this first Brillouin zone have received specific notations, e.g the Γ , X or L points.

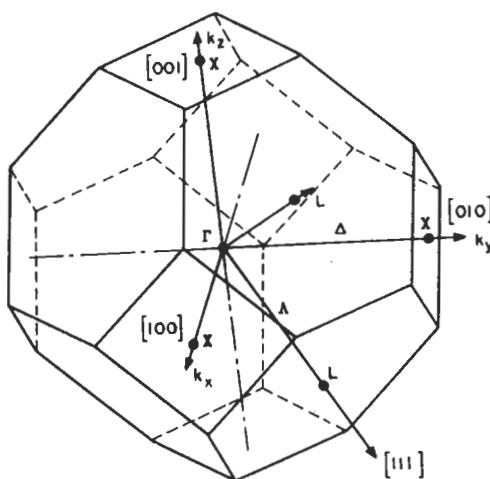


Fig. 2. — First Brillouin zone of face-centred cubic lattice.

II. Electronic properties.

In III-V binary compounds like GaAs, there are 8 outer electrons per unit cell which contribute to the chemical bonds. The other electrons of each kind of atom are "frozen" in closed shell configurations and their wavefunctions are highly bound around the Ga or As nuclei. They do not contribute to the electronic properties investigated here: electrical transport or near bandgap optical properties. The 8 outermost electrons (3 from Ga and 5 from As) hybridize to form tetrahedral bonds between one kind of atom (say Ga) and its four nearest neighbours (As). In basic terms one could say that the orbitals of every atom (s-like or p-like) hybridize with an orbital of the neighbouring atom, thus producing two levels; one bonding and one antibonding. Because there is a large number of unit cells, bonding and antibonding levels broaden into bands. The bonding s levels are deeply bound and always occupied by two electrons per unit cell. The remaining six electrons per unit cell completely fill the three bonding p orbitals. The bands originating from the antibonding orbitals are all empty, the lowest lying (often an s band) forming the conduction band of the material.

In all the III-V materials the top of the valence band occurs at the centre of the Brillouin zone (Γ point). In the absence of spin-orbit coupling, the three valence bands (which originate from bonding p orbitals) are degenerate at Γ . The spin-orbit coupling lifts this sixfold degeneracy and gives rise to a quadruplet (symmetry Γ_8) which corresponds to $J = \frac{3}{2}$, and to a doublet (symmetry Γ_7) which corresponds to

$J = \frac{1}{2}$, where J is the total angular momentum. Since the spin-orbit coupling is essentially an atomic property associated with the anion (As for GaAs) of the binary compound, one should expect to find that the energy separation between the spin-

orbit split-off bands is almost independent of the cation. This is indeed the case (see Table I). In addition, the fact that the valence bands mostly originate from the p orbitals of the anion leads one to expect that the top of the valence band of binary materials which have the same lattice parameter and the same anion should almost line up. This common anion rule, introduced by Harrison [1], is a useful guideline for predicting rough band lineups between two materials which form a given heterostructure. Note however that the absolute accuracy of the common anion rule is only several tenths of an electron-volt.

The conduction band edge of the III-V materials is found either at the Γ point or near the L or the X point. In general, the heavier the cation of the binary, the more likely it is to find the lowest lying conduction band edge at the Γ point. For instance, in the Sb family the conduction band edge is near the X point for AlSb and $\varepsilon_X - \varepsilon_\Gamma \sim 650$ meV. However, it is barely at the Γ point for GaSb ($\varepsilon_L - \varepsilon_\Gamma \sim 95$ meV) whereas the InSb conduction band edge is at the Γ point and $\varepsilon_L - \varepsilon_\Gamma \sim 0.63$ eV. The same trend is found for the phosphide and arsenide series.

At the Γ point, the bandgap which separates the topmost valence band from the lowest lying conduction band increases as the cation becomes lighter. In the Sb family [2] $\varepsilon_0^{(\Gamma)} = 0.235$ eV for InSb ; 0.811 eV for GaSb and 2.32 eV for AlSb. All these trends are not fortuitous but can be explained by considering the ionicity of the chemical bonds which exist between the anion and the cation. For a detailed account of this approach, the reader is referred to Phillip's textbook [3].

It is possible to form ternary or quaternary solid solutions between III-V or II-VI binary compounds. From the point of view of the electronic structure, these alloys are not crystalline : the potential energy felt by the electrons has no translational symmetry due to the random distribution of the atoms at the sites of the zinc-blende lattice. However, in many cases one may use the virtual crystal approximation to describe the alloy electronic states. In such a model the actual aperiodic potential is replaced by an average one. For example, consider a solid solution $A B_{1-x} C_x$. The A atoms sit on the sites of one of the f.c.c. lattices of the zinc-blende structure, whereas the B and C atoms randomly occupy the sites of the second f.c.c. lattice. In the virtual crystal approximation, the random potential, created by the B(V_B) and C(V_C) atoms is replaced by a periodic one, whose strength is the weighted average of V_B and V_C : $\langle V \rangle = V_A + xV_C + (1-x)V_B$. Using this procedure, the translational invariance is restored. Thus, Bloch waves can be defined as well as bandgaps, effective masses etc... Of course disorder effects play some part, e.g. blurring fine structures, scattering Bloch waves, creating bandtails etc... However, in most instances, the very notion of composition dependence of, say, the fundamental bandgap, $\varepsilon_0(x)$ in a random alloy, is relevant. If two binaries have similar band structures and are closely lattice-matched, the fundamental bandgap of their alloys varies almost linearly with x . A well-known example pertains to the II-VI family : the $Hg_{1-x} Cd_x Te$ alloys bandgap interpolates almost linearly between those of HgTe and CdTe [4]. On the other hand, if the binaries are significantly lattice-mismatched (even though they have similar band structures), the alloy bandgap $\varepsilon_0(x)$ is strongly non linear upon x ; see the example of the solid solutions $InAs_{1-x} Sb_x$. If the binaries, although reasonably lattice-matched, have dissimilar

Table I. — Parameters of several III-V compounds compiled from reference [2]. The brackets give the low temperature (upper figure) and the room temperature (lower figure) values. The lattice parameter a is a room temperature value. The symbol \leftrightarrow indicates that several figures of a given quantity appear in the Landolt-Börnstein tables [2].

band structures, e.g. the solid solutions $\text{Ga}_{1-x}\text{Al}_x\text{As}$ where the conduction band edge occurs at the Γ point for GaAs and near the X point for AlAs , the same non linear variation of the fundamental bandgap is observed.

III. Electronic dispersion relations in the vicinity of the zone centre : $k.p$ analysis and Kane model.

In a bulk crystal the one-electron Schrödinger equation which has to be solved is :

$$\left[\frac{p^2}{2m_0} + V(\mathbf{r}) + \frac{\hbar}{4m_0^2 c^2} (\boldsymbol{\sigma} \times \nabla V) \cdot \mathbf{p} + \delta \mathcal{H}_r \right] \psi(\mathbf{r}) = \varepsilon \psi(\mathbf{r}), \quad (1)$$

where m_0 is the free electron mass and $V(\mathbf{r})$ is the crystalline potential. The latter includes some average of the electron-electron interaction and is periodic with the periodicity of the underlying Bravais lattice. The third and fourth terms in equation (1) are relativistic corrections. The third one is the spin-orbit coupling ($\boldsymbol{\sigma}$ is the electron spin) and the fourth one, which is only important for very heavy atoms, includes the so-called mass-velocity and Darwin terms. If \mathbf{a}_i , $i = 1, 2, 3$ are three basis vectors of the Bravais lattice, we have :

$$V \left(\mathbf{r} + \sum_i n_i \mathbf{a}_i \right) = V(\mathbf{r}), \quad (2)$$

for any \mathbf{r} and any relative integers n_i . Thus the total potential energy which appears in equation (1) is periodic in \mathbf{r} , and we may look for eigenfunctions of the Hamiltonian which are also eigenfunctions of the translation operators \mathcal{T}_d where $\mathbf{d} = \sum_i n_i \mathbf{a}_i$:

$$\mathcal{T}_d \psi(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{d}) = \exp(i \mathbf{k} \cdot \mathbf{d}) \psi(\mathbf{r}). \quad (3)$$

Equivalently, the solutions of equations (1, 3) can be written in the Bloch form :

$$\psi_{n\mathbf{k}}(\mathbf{r}) = N u_{n\mathbf{k}}(\mathbf{r}) \exp(i \mathbf{k} \cdot \mathbf{r}), \quad (4)$$

where N is a normalization coefficient and $u_{n\mathbf{k}}(\mathbf{r})$ a periodic function of \mathbf{r} with the periodicity of the lattice :

$$u_{n\mathbf{k}} \left(\mathbf{r} + \sum_i n_i \mathbf{a}_i \right) = u_{n\mathbf{k}}(\mathbf{r}). \quad (5)$$

If we normalize $\psi_{n\mathbf{k}}$ over the whole crystal (volume Ω), which consists of N elementary cells (volume Ω_0), and assume \mathbf{k} to be real we get $N = \Omega^{-1/2}$, provided the $u_{n\mathbf{k}}$'s are normalized to Ω_0 in a unit cell.

A Bloch state $|n\mathbf{k}\rangle$ is thus labelled by a discrete band index n and a crystal wave vector \mathbf{k} , which can be restricted to the first Brillouin zone of the reciprocal lattice (the spin index is omitted for brevity). Applying periodic boundary conditions to a macroscopic crystal

$$\psi_{n\mathbf{k}}(\mathbf{r} + N_i \mathbf{a}_i) = \psi_{n\mathbf{k}}(\mathbf{r}), \quad N_i \rightarrow \infty \quad (6)$$

we find that \mathbf{k} must be real and such that

$$\mathbf{k} \cdot \mathbf{N}_i \mathbf{a}_i = 2\pi p_i, \quad i = 1, 2, 3 \quad (7)$$

where p_i are integers. The allowed values of \mathbf{k} thus form a quasicontinuum. Any summation of the form

$$S = \sum_{\mathbf{k} \in \text{1}^{\text{st}} \text{ B.Z.}} f(\mathbf{k})$$

can be converted into an integration :

$$S \xrightarrow[\Omega \rightarrow \infty]{\Omega}{(2\pi)^3} \int_{\mathbf{k} \in \text{1}^{\text{st}} \text{ B.Z.}} d^3k f(\mathbf{k}). \quad (8)$$

The Bloch functions are seldom known explicitly. However, at the high symmetry points of the first Brillouin zone (noticeably the Γ point), the way in which the Bloch functions change under the action of the symmetry operations belonging to the crystal structure point group can be analysed using group theory arguments. Four Γ point Bloch functions and their linear combinations will repeatedly appear in the following discussions. They describe the crystal states for energies which correspond either to the top of the occupied valence bands or to the bottom of the lowest-lying, empty, conduction band. The states are labelled $|S\rangle$, $|X\rangle$, $|Y\rangle$, $|Z\rangle$ and their associated wavefunctions transform in the same way as the atomic s, x, y, z functions under the symmetry operations which map the local tetrahedron onto itself.

III.1 $\mathbf{k.p}$ ANALYSIS AND EFFECTIVE MASSES. — Global descriptions of the dispersion relations of bulk materials are available (pseudo-potentials, tight-binding) and quite accurate. However, they rarely provide explicit expressions for the quantities of common use in semiconductor physics : effective masses, wavefunctions etc... In fact, for many aspects of semiconductor electronic properties, such global descriptions of the dispersion relations over the whole Brillouin zone are unnecessary. What is needed is a knowledge of the ϵ_{nk} relationship over a small k range around the band extrema. Thus, in the following, we describe the results of a local description of the band structure : the $\mathbf{k.p}$ method.

The periodic parts of the Bloch functions u_{nk} are the solutions of

$$\left\{ \frac{p^2}{2m_0} + V(\mathbf{r}) + \frac{\hbar}{4m_0^2 c^2} (\boldsymbol{\sigma} \times \nabla V) \cdot \mathbf{p} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar \mathbf{k}}{m_0} \left(\mathbf{p} + \frac{\hbar}{4m_0 c^2} \boldsymbol{\sigma} \times \nabla V \right) \right\} u_{nk} = \epsilon_{nk} u_{nk} \quad (9)$$

where the relativistic corrections have been dropped for simplicity. Equation (9) can be formally rewritten as :

$$[H(\mathbf{k} = \mathbf{0}) + W(\mathbf{k})] u_{nk} = \epsilon_{nk} u_{nk}, \quad (10)$$

where $H(\mathbf{k} = \mathbf{0})$ is the crystal Hamiltonian whose eigenfunctions are ψ_{n0} (or u_{n0} equivalently) :

$$H(\mathbf{k} = \mathbf{0}) u_{n0} = \varepsilon_{n0} u_{n0}. \quad (11)$$

The \mathbf{k} -dependent operator $W(\mathbf{k})$ vanishes at $\mathbf{k} = \mathbf{0}$ and commutes with the translation operator \mathcal{T}_d . Thus we can write :

$$u_{nk} = \sum_m c_m(\mathbf{k}) u_{m0}. \quad (12)$$

By inserting equation (12) into equation (10), multiplying by u_{m0}^* and integrating over a unit cell we obtain :

$$\sum_m \left\{ \left(\varepsilon_{n0} - \varepsilon_{nk} + \frac{\hbar^2 k^2}{2m_0} \right) \delta_{nm} + \frac{\hbar \mathbf{k}}{m_0} \cdot \langle n0 | \mathbf{p} + \frac{\hbar}{4m_0 c^2} (\boldsymbol{\sigma} \times \nabla V) | m0 \rangle \right\} c_m(\mathbf{k}) = 0 \quad (13)$$

where :

$$\langle n0 | A | m0 \rangle = A_{nm} = \int_{\text{unit cell}} u_{n0}^* A u_{m0} d^3r. \quad (14)$$

Note in equation (13) the absence of diagonal (in n) terms proportional to \mathbf{k} . One can easily show that such terms vanish identically. Equation (13), which is equivalent to equation (10), is well suited for a perturbative approach. Let us suppose that the n^{th} band edge (energy ε_{n0}) is non degenerate (apart from spin). We can thus assume that for small \mathbf{k} :

$$c_n(\mathbf{k}) \sim 1; \quad c_m(\mathbf{k}) = \alpha \cdot \mathbf{k} \quad (15)$$

since $c_m(\mathbf{0}) = \delta_{nm}$. In fact we have :

$$c_m(\mathbf{k}) = \frac{\hbar \mathbf{k}}{m_0} \cdot \boldsymbol{\pi}_{nm} \frac{1}{\varepsilon_{n0} - \varepsilon_{m0}} \quad (16)$$

which, when inserted into equation (13), gives the second order correction to ε_{n0} :

$$\varepsilon_{nk} = \varepsilon_{n0} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{m \neq n} \frac{|\boldsymbol{\pi}_{nm} \cdot \mathbf{k}|^2}{\varepsilon_{n0} - \varepsilon_{m0}}. \quad (17)$$

In equations (16, 17) the vector $\boldsymbol{\pi}$ is defined as :

$$\boldsymbol{\pi} = \mathbf{p} + \frac{\hbar}{4m_0 c^2} (\boldsymbol{\sigma} \times \nabla V). \quad (18)$$

Thus, as long as k is small (i.e. that $\varepsilon_{nk} - \varepsilon_{n0}$ remains much smaller than all the

band edge gaps $\varepsilon_{n0} - \varepsilon_{m0}$) the dispersion relations of the non degenerate bands are parabolic in \mathbf{k} in the vicinity of the Γ point.

$$\varepsilon_{n\mathbf{k}} = \varepsilon_{n0} + \frac{\hbar^2}{2} \sum_{\alpha, \beta} k_\alpha \frac{1}{\mu_n^{\alpha\beta}} k_\beta ; \quad \alpha, \beta = x, y, z \quad (19)$$

where

$$\frac{1}{\mu_n^{\alpha\beta}} = \frac{1}{m_0} \delta_{\alpha\beta} + \frac{2}{m_0^2} \sum_{m \neq n} \frac{\pi_{mn}^\alpha \pi_{nm}^\beta}{\varepsilon_{n0} - \varepsilon_{m0}} . \quad (20)$$

III.2 EFFECTIVE MASS TENSOR. ELECTRONS AND HOLES. — $\mu_n^{\alpha\beta}$ is the effective mass tensor which describes the carrier kinematics in the vicinity of the zone centre and for energy close to the n^{th} band edge. If the carrier energy is such that equation (20) is valid, the overall effects of the band structure (i.e. the fact that the carrier experiences the periodic potential $V(\mathbf{r})$ instead of moving in the vacuum) are embodied in the use of an effective mass instead of the free electron mass. The notion of the effective mass is at the heart of the semiclassical description of the carrier motion in semiconductor. It can be shown [5] that, if an external force \mathbf{F} is applied to the carrier, the crystal wavevector changes according to the law

$$\hbar \frac{d\mathbf{k}}{dt} = \mathbf{F} . \quad (21)$$

The carrier velocity in a Bloch state $|n\mathbf{k}\rangle$ is exactly equal to

$$\mathbf{v} = \frac{1}{\hbar} \frac{\partial \varepsilon_n}{\partial \mathbf{k}} . \quad (22)$$

This relation together with the quadratic dispersion law equation (19), leads to :

$$v_\alpha = \frac{\hbar}{2} \sum_\beta \left(\frac{1}{\mu_n^{\alpha\beta}} + \frac{1}{\mu_n^{\beta\alpha}} \right) k_\beta \quad (23)$$

which is reminiscent of the familiar relation $\mathbf{v} = \frac{\mathbf{p}}{m_0}$ for free particles in vacuum. The semiclassical assumption is that the external force is weak enough to preclude any interband transition. In this way, despite the increment of its \mathbf{k} vector through equation (21), the carrier velocity remains equal to equation (22). This is because the band index is constant. A more complete treatment of the semiclassical motion in solids can be found in [5]. For the conduction band edge associated with the antibonding s orbitals (Γ_6 symmetry) $\mu_{\Gamma_6}^{\alpha\beta}$ is the simplest since it is a scalar :

$$\frac{1}{\mu_{\Gamma_6}^{\alpha\beta}} = \frac{1}{m_{\Gamma_6}} \delta_{\alpha\beta} ; \quad \alpha, \beta = x, y, z \quad (24)$$

$$\frac{1}{m_{\Gamma_6}} = \frac{1}{m_0} + \frac{2}{m_0^2} \sum_{m \neq \Gamma_6} \frac{|\pi_{\Gamma_6 m}^z|^2}{\varepsilon_{\Gamma_6} - \varepsilon_{m0}} . \quad (25)$$

Equations (20, 25) explicitly show that the sign and magnitude of an effective mass are governed by two factors :

- i) the squared strength of the π matrix element between the Bloch functions of the edge of interest and those of all the other $k = 0$ edges.
- ii) the sign and magnitude of the Γ bandgaps separating the edge of interest and all the other edges.

In GaAs and InP for instance, the Γ edges nearest to the Γ_6 conduction band edge are associated with the topmost valence bands v_t (bonding p orbitals), and the dominant term in equation (25) is $|\pi_{\Gamma_6, v_t}^z|^2/\varepsilon_{\Gamma_6} - \varepsilon_{v_t}$, which is positive. The Γ_6 effective mass is therefore positive. It should also be noted that if the v_t edges alone contributed to $\frac{1}{m_{\Gamma_6}}$ and if π_{Γ_6, v_t} was material-independent, $m_{\Gamma_6}^{-1} - m_0^{-1}$ would be proportional to the inverse of the bandgap which separates the bonding states and the antibonding states, i.e. to the fundamental bandgap ε_0 of direct gap III-V or II-VI compounds. In fact, it is true that the smaller ε_0 , the smaller m_{Γ_6} (see Tab. I). However the strict proportionality between m_{Γ_6} and ε_0 is only observed in very narrow gap materials ($\varepsilon_0 \leq 0.1$ eV as in $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ alloys with $0.16 \leq x \leq 0.21$ at low temperature). In wider gap materials, the other Γ edges give a significant contribution to m_{Γ_6} . Finally it is important to notice that the m_0^{-1} term in the expression of $\mu_{\alpha\beta}^{-1}$ ensures that the effective mass coincides with m_0 as it should for an empty lattice characterized by $V(\mathbf{r}) = 0$.

If the dominant contributions to the effective mass of a given edge arise from $\mathbf{k} \cdot \mathbf{p}$ coupling between this edge and *higher* band edges, this effective mass is negative. This is what happens to the effective mass of the topmost occupied valence band v_t . The analysis of the semiclassical motion of a particle with negative mass is not easy, as we are more accustomed to working with particles of positive mass. Let us assume for simplicity that the effective mass tensor is diagonal : $\mu_n^{\alpha\beta} = -m_v \delta_{\alpha\beta}$, $m_v > 0$. Equation (23) then shows that the carrier velocity \mathbf{v} is opposite to the carrier wavevector \mathbf{k} . Newton's law, equation (21), reads :

$$-m_v \frac{d\mathbf{v}}{dt} = \mathbf{F}. \quad (26)$$

Under many circumstances \mathbf{F} has an electromagnetic origin :

$$\mathbf{F} = -e \left[\mathbf{E} + \frac{\mathbf{v}}{c} \wedge \mathbf{B} \right]. \quad (27)$$

Equations (26, 27) can be rewritten

$$+m_v \frac{d\mathbf{v}}{dt} = +e \left[\mathbf{E} + \frac{\mathbf{v}}{c} \wedge \mathbf{B} \right], \quad (28)$$

which shows that the motions of valence electrons (negative charge, negative effective mass) when subjected to electromagnetic forces can be analysed just as well

as the motions of fictitious particles which are characterized by a positive effective mass and a positive charge.

Having considered the carrier motion, let us turn to the electrical current. If the carrier velocity in the Bloch state $|n\mathbf{k}\rangle$ is \mathbf{v}_k , the electrical current density is

$$\mathbf{J}_k = -\frac{e}{\Omega} \mathbf{v}_k. \quad (29)$$

The macroscopic current is the sum of \mathbf{J}_k over all the occupied states. Let $f_e(\mathbf{k})$ denote the distribution function of these occupied electron states (at thermal equilibrium, $f_e(\mathbf{k})$ would be the Fermi-Dirac distribution function). Then :

$$\mathbf{J} = -\frac{e}{\Omega} \sum_{\mathbf{k}} \mathbf{v}_k f_e(\mathbf{k}), \quad (30)$$

where we have assumed that a single band was populated. If this band is a partially filled valence band ν_t , the electrical current carried by electrons in such a band can be rewritten :

$$\mathbf{J} = \mathbf{J}_f + \mathbf{J}_h \quad (31)$$

where :

$$\mathbf{J}_f = -\frac{e}{\Omega} \sum_{\mathbf{k}} \mathbf{v}_k \quad (32)$$

$$\mathbf{J}_h = +\frac{e}{\Omega} \sum_{\mathbf{k}} \mathbf{v}_k [1 - f_e(\mathbf{k})] = +\frac{e}{\Omega} \sum_{\mathbf{k}} f_h(\mathbf{k}) \mathbf{v}_k \quad (33)$$

The first term, \mathbf{J}_f , is the contribution due to a filled band ($f_e(\mathbf{k}) = 1$ for all \mathbf{k}). This contribution vanishes identically owing to equation (22). Thus the current carried by electrons which occupy a partially filled valence band is equivalent to that caused by fictitious particles of positive charge whose distribution function is that of unoccupied states. The fictitious particles are called holes. They are similar to free electrons in that they have a positive effective mass but, unlike the free electrons, their electrical charge is positive. In addition, the distribution function of the holes $f_h(\mathbf{k})$ is the same as that of the unoccupied electron states : $1 - f_e(\mathbf{k})$. Finally let us compare the equilibrium distribution function (Fermi-Dirac) of electrons and holes in isotropic parabolic bands. We have :

$$\varepsilon_c(\mathbf{k}) = \varepsilon_c + \frac{\hbar^2 k^2}{2m_c}; \quad \varepsilon_v(\mathbf{k}) = -\varepsilon_v - \frac{\hbar^2 k^2}{2m_v} \quad (34)$$

for the electron energy spectra. The Fermi Dirac distribution function of an electron is

$$f_e(\mathbf{k})_{\text{therm. equil.}} = \left\{ 1 + \exp \left[\frac{1}{k_B T} \left(\frac{\hbar^2 k^2}{2m_c} + \varepsilon_c - \mu \right) \right] \right\}^{-1}, \quad (35)$$

where k_B is the Boltzmann constant, T the temperature and μ the chemical potential.

The hole distribution function of a partially filled valence band is therefore

$$f_h(\mathbf{k})|_{\text{therm. equil.}} = [1 - f_e(k)]_{\text{therm. equil.}} = \left\{ 1 + \exp \left[-\frac{1}{k_B T} \left(\epsilon_v - \frac{\hbar^2 k^2}{2m_v} - \mu \right) \right] \right\}^{-1} \quad (36)$$

which can be rewritten

$$f_h(k)|_{\text{therm. equil.}} = \left\{ 1 + \exp \left[\frac{1}{k_B T} \left(\frac{\hbar^2 k^2}{2m_v} - \xi \right) \right] \right\}^{-1} \quad (37)$$

with

$$\xi = \epsilon_v - \mu. \quad (38)$$

Thus the hole distribution function of a partially filled valence band coincides with the electron distribution function of a partially filled conduction band. Both f 's decay exponentially at large wave vectors. Note however that the hole chemical potential measured from the valence edge is minus the electron chemical potential measured from the same edge.

III.3 BEYOND THE QUADRATIC DISPERSION RELATIONS : THE KANE MODEL. — For a given \mathbf{k} , the lighter the effective mass, the larger the kinetic energy term $\epsilon_{n\mathbf{k}} - \epsilon_{n0}$. Sometimes this term ceases to be negligible with respect to the various Γ bandgaps and consequently, a more detailed analysis of the $\epsilon_{n\mathbf{k}}$ relations is required. The first way to improve our calculations would be to extend the perturbative treatment of $W(\mathbf{k})$ beyond the second order. This procedure, seldom used, is very cumbersome. In a celebrated paper [6], Kane took a different approach from that of the perturbative expansion of $\epsilon_{n\mathbf{k}}$ in ascending powers of k . He noticed that for InSb ($\epsilon_0 \sim 0.23$ eV), the topmost valence states v_t and the lowest-lying conduction band Γ_6 were very close and well separated from all the other bands. Thus Kane diagonalized *exactly* $W(\mathbf{k})$ within a limited set of band edges (Γ_6, v_t), and afterwards introduced the $W(\mathbf{k})$ coupling between (Γ_6, v_t) and the other Γ edges within the framework of a second order perturbative treatment. For a detailed account of the Kane model the reader is referred to his review papers [7, 8]. Here we restrict our considerations to the first step of Kane's analysis.

a) *Dispersion relations of the Kane model* : Since the spin-orbit coupling is non-zero in III-V compounds it is desirable to have a $\mathbf{k} = 0$ basis in which this term is already diagonal. Thus, instead of using the 8 band edge Bloch functions $|S\uparrow\rangle, |X\uparrow\rangle, |Y\uparrow\rangle|Z\uparrow\rangle, |S\downarrow\rangle, |X\downarrow\rangle, |Y\downarrow\rangle, |Z\downarrow\rangle$, we form linear combinations of these functions. These are such that the total angular momentum $J = \mathbf{L} + \boldsymbol{\sigma}$ and its projection J_z along the z axis are diagonal in the new basis. For the S edge, the addition of $L = 0$ and $\sigma = 1/2$ only gives $J = 1/2$ (Γ_6 symmetry). For the P edges, adding $L = 1$ to $\sigma = 1/2$ gives either $J = 3/2$ or $J = 1/2$. In III-V compounds the quadruplet $J = 3/2$ is always higher in energy than the doublet $J = 1/2$. The

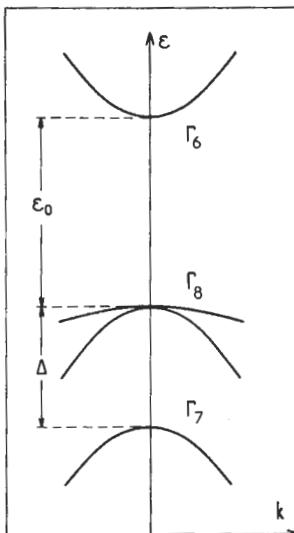


Fig. 3. — Band structure of a direct gap III-V compound in the vicinity of the zone centre. Only the states (Γ_6 , Γ_7 , Γ_8) near the Fermi energy at $T = 0$ K for a perfect crystal are presented.

quadruplet corresponds to Γ_8 symmetry and the doublet to Γ_7 symmetry. Their energy separation is noted Δ (see Fig. 3). The basis wavefunctions of the Γ_6 , Γ_7 , Γ_8 edges which we shall use are labelled in table II. Note that each edge is twice degenerate and corresponds to the two opposite values of m_j which are the eigenvalues of J_z . Taking the matrix elements of $W(\mathbf{k})$ between these 8 Bloch functions and dropping the relativistic corrections together with the \mathbf{k} -dependent spin-orbit term $(\mathbf{\pi} - \mathbf{p})$, we find that the energies $\varepsilon(\mathbf{k})$ are the eigenvalues of the following 8×8 matrix :

$ iS^\dagger\rangle$	$ \frac{3}{2}, \frac{1}{2}\rangle$	$ \frac{3}{2}, \frac{3}{2}\rangle$	$ \frac{1}{2}, \frac{1}{2}\rangle$	$ iS^\dagger\rangle$	$ \frac{3}{2}, -\frac{1}{2}\rangle$	$ \frac{3}{2}, -\frac{3}{2}\rangle$	$ \frac{1}{2}, -\frac{1}{2}\rangle$
$\langle iS^\dagger $	$\frac{\hbar^2 k^2}{2m_0}$	$-\sqrt{\frac{2}{3}} P\hbar k_z$	$P\hbar k_+$	$\frac{1}{\sqrt{3}} P\hbar k_z$	0	$-\frac{1}{\sqrt{3}} P\hbar k_-$	$-\sqrt{\frac{2}{3}} P\hbar k_-$
$\left\langle \frac{3}{2}, \frac{1}{2} \right $	$-\sqrt{\frac{2}{3}} P\hbar k_z - \varepsilon_0 + \frac{\hbar^2 k^2}{2m_0}$	0	0	$\frac{P}{\sqrt{3}} \hbar k_-$	0	0	0
$\left\langle \frac{3}{2}, \frac{3}{2} \right $	$P\hbar k_-$	0	$-\varepsilon_0 + \frac{\hbar^2 k^2}{2m_0}$	0	0	0	0
$\left\langle \frac{1}{2}, \frac{1}{2} \right $	$\frac{1}{\sqrt{3}} P\hbar k_z$	0	0	$-\varepsilon_0 - \Delta + \frac{\hbar^2 k^2}{2m_0}$	$\sqrt{\frac{2}{3}} P\hbar k_-$	0	0
$\langle iS^\dagger $	0	$\frac{P}{\sqrt{3}} \hbar k_+$	0	$\sqrt{\frac{2}{3}} P\hbar k_+$	$\frac{\hbar^2 k^2}{2m_0}$	$-\sqrt{\frac{2}{3}} P\hbar k_z$	$\frac{1}{\sqrt{3}} P\hbar k_z$
$\left\langle \frac{3}{2}, -\frac{1}{2} \right $	$-\frac{1}{\sqrt{3}} P\hbar k_+$	0	0	0	$-\sqrt{\frac{2}{3}} P\hbar k_z - \varepsilon_0 + \frac{\hbar^2 k^2}{2m_0}$	0	0
$\left\langle \frac{3}{2}, -\frac{3}{2} \right $	0	0	0	$P\hbar k_+$	0	$-\varepsilon_0 + \frac{\hbar^2 k^2}{2m_0}$	0
$\left\langle \frac{1}{2}, -\frac{1}{2} \right $	$-\sqrt{\frac{2}{3}} P\hbar k_+$	0	0	0	$\frac{P}{\sqrt{3}} \hbar k_z$	0	$-\varepsilon_0 - \Delta + \frac{\hbar^2 k^2}{2m_0}$

(39)

Table II. — $\Gamma_6, \Gamma_7, \Gamma_8$ periodic parts of the Bloch functions. The notations J, m_J label the eigenvalues of the operators J and J_z , respectively.

u_i	$ J, m_J\rangle$	ψ_{J, m_J}	$\epsilon_i(k=0)$
u_1	$\left \frac{1}{2}, \frac{1}{2} \right\rangle$	$i S^\uparrow\rangle$	0
u_3	$\left \frac{3}{2}, \frac{1}{2} \right\rangle$	$-\sqrt{\frac{2}{3}} Z^\uparrow\rangle + \frac{1}{\sqrt{6}} (X+iY)^\downarrow\rangle$	$-\epsilon_0$
u_5	$\left \frac{3}{2}, \frac{3}{2} \right\rangle$	$\frac{1}{\sqrt{2}} (X+iY)^\uparrow\rangle$	$-\epsilon_0$
u_7	$\left \frac{1}{2}, \frac{1}{2} \right\rangle$	$\frac{1}{\sqrt{3}} (X+iY)^\downarrow\rangle + \frac{1}{\sqrt{3}} Z^\uparrow\rangle$	$-\epsilon_0 - \Delta$
u_2	$\left \frac{1}{2}, -\frac{1}{2} \right\rangle$	$i S^\downarrow\rangle$	0
u_4	$\left \frac{3}{2}, -\frac{1}{2} \right\rangle$	$-\frac{1}{\sqrt{6}} (X-iY)^\uparrow\rangle - \sqrt{\frac{2}{3}} Z^\downarrow\rangle$	$-\epsilon_0$
u_6	$\left \frac{3}{2}, -\frac{3}{2} \right\rangle$	$\frac{1}{\sqrt{2}} (X-iY)^\downarrow\rangle$	$-\epsilon_0$
u_8	$\left \frac{1}{2}, -\frac{1}{2} \right\rangle$	$-\frac{1}{\sqrt{3}} (X-iY)^\uparrow\rangle + \frac{1}{\sqrt{3}} Z^\downarrow\rangle$	$-\epsilon_0 - \Delta$

In equation (39) $k_\pm = \frac{1}{\sqrt{2}} (k_x \pm ik_y)$ and the energy zero is taken at the Γ_6 edge. The eigenenergies depend on three parameters : the fundamental bandgap ϵ_0 , the Γ spin-orbit coupling of the topmost valence band Δ

$$\epsilon_0 = \epsilon_{\Gamma_6} - \epsilon_{\Gamma_8} \quad ; \quad \Delta = \epsilon_{\Gamma_8} - \epsilon_{\Gamma_7} \quad (40)$$

and the interband matrix element of the velocity operator :

$$P = \frac{-i}{m_0} \langle S | p_x | X \rangle = \frac{-i}{m_0} \langle S | p_y | Y \rangle = \frac{-i}{m_0} \langle S | p_z | Z \rangle. \quad (41)$$

All the other p matrix elements vanish by symmetry.

If $\lambda(\mathbf{k})$ denotes the energy

$$\lambda(\mathbf{k}) = \varepsilon(\mathbf{k}) - \frac{\hbar^2 k^2}{2m_0}. \quad (42)$$

We find from equation (39) that $\lambda(\mathbf{k})$ is the solution of

$$\lambda(\mathbf{k}) = -\varepsilon_0 \quad (43)$$

or :

$$\lambda(\mathbf{k})[\lambda(\mathbf{k}) + \varepsilon_0][\lambda(\mathbf{k}) + \varepsilon_0 + \Delta] = \hbar^2 k^2 P^2 \left[\lambda(\mathbf{k}) + \varepsilon_0 + \frac{2\Delta}{3} \right]. \quad (44)$$

Each of the solutions of equations (43, 44) is twice degenerate. We notice that the dispersion relations are isotropic in \mathbf{k} , i.e. $\varepsilon(\mathbf{k})$ depends only on $k = |\mathbf{k}|$. This occurs in spite of the fact that a preferential axis seems to be present : the \mathbf{J} quantization axis. Actually, band warping [i.e. anisotropic ε_{nk}] does exist in III-V materials, noticeably in the valence bands. However it results from $\mathbf{k} \cdot \mathbf{p}$ interactions between the $\Gamma_6, \Gamma_7, \Gamma_8$ edges and the remote bands. In the present analysis, restricted to $\Gamma_6, \Gamma_7, \Gamma_8$ bands, we could have chosen to quantize \mathbf{J} along the \mathbf{k} vector instead of along a fixed reference axis. Such a procedure renders the matrix (Eq. (39)) block-diagonal ($k_{\pm} = 0, k = k_z$) and, in addition, the solutions of equations (43, 44) correspond for any k_z to a given value of m_J : $m_J = \pm \frac{3}{2}$ for equation (43) and $m_J = \pm \frac{1}{2}$ for equation (44). When \mathbf{J} is quantized along a fixed direction, the eigenstates are in general a mixture of $m_J = \pm \frac{1}{2}$ and $m_J = \pm \frac{3}{2}$ components. Clearly in a bulk material the quantization of \mathbf{J} along \mathbf{k} is the simplest solution, due to the isotropicity of the dispersion relations. Later on, we shall see that in the case of heterostructures it is convenient to quantize \mathbf{J} along the growth (z) axis of the heterostructure. The dispersion relations will become anisotropic and will be simply obtained only if $\mathbf{k}/\mathbf{J}/\hat{z}$. This is because the $m_J = \pm \frac{1}{2}, m_J = \pm \frac{3}{2}$ decoupling will remain valid. On the other hand if k_{\pm} is different from zero, the states of different m_J 's will become admixed, leading to complicated algebra.

b) *Band edge effective masses* : If we take $\mathbf{k}/\mathbf{J}/\hat{z}$ we notice that the $m_J = \pm \frac{3}{2}$ states (Eq. (43)) correspond to the heavy particles. In fact the lack of $k_z p_z$ coupling between Γ_8 ($m_J = \pm \frac{3}{2}$) and Γ_6 [$\langle S|p_z|X + iY \rangle = 0$] leads to an effective mass which coincides with the bare electron mass.

The $m_J = \pm \frac{1}{2}$ states, however, are associated with light particles (Γ_6 electrons, Γ_8 light holes, Γ_7 holes) : their effective mass is already lighter than m_0 even if $\mathbf{k} \cdot \mathbf{p}$ coupling is allowed only within the $\Gamma_6, \Gamma_7, \Gamma_8$ subspace. In the vicinity of the band

edges ($\varepsilon = 0$, $\varepsilon = -\varepsilon_0$, $\varepsilon = -\varepsilon_0 - \Delta$) we may expand the solutions of equation (44) to the second order in k , thus obtaining explicit expressions for the band edge ($k \rightarrow 0$) effective masses of the light particles. We find :

$$\frac{1}{m_{\Gamma_6}} = \frac{1}{m_0} + \frac{4P^2}{3\varepsilon_0} + \frac{2P^2}{3(\varepsilon_0 + \Delta)} \quad (45)$$

$$\frac{1}{m_{\Gamma_8^l}} = \frac{1}{m_0} - \frac{4P^2}{3\varepsilon_0} \quad (46)$$

$$\frac{1}{m_{\Gamma_7}} = \frac{1}{m_0} - \frac{2P^2}{3(\varepsilon_0 + \Delta)} \quad (47)$$

i.e. $m_{\Gamma_6} > 0$ and $m_{\Gamma_8^l}, m_{\Gamma_7} < 0$. Equations (45, 47) illustrate the previous general considerations on band edge effective masses. For instance $m_{\Gamma_8^l}$ (l stands for light) does not depend on Δ , which is quite natural since there is no $\mathbf{k.p}$ interaction between the Γ_8 and Γ_7 edges. For the same reason $m_{\Gamma_7}^l$ depends only on $\varepsilon_0 + \Delta$, which is the magnitude of the bandgap separating Γ_7 from the Γ_6 band with which it is $\mathbf{k.p}$ coupled.

The mirror effects for $\mathbf{k.p}$ couplings in equations (45, 47) should also be noted. The Γ_7 and Γ_8 $\mathbf{k.p}$ contributions to m_{Γ_6} are equal in magnitude and opposite in sign to the Γ_6 $\mathbf{k.p}$ contributions to m_{Γ_7} and $m_{\Gamma_8^l}$ respectively.

c) *Band non-parabolicity* : Equation (44) is an implicit equation for $\lambda(\mathbf{k})$ versus \mathbf{k} , but is explicit for \mathbf{k} versus $\lambda(\mathbf{k})$. In figure 4 we show a plot of $\lambda(\mathbf{k})$ versus $\frac{\hbar^2 k^2}{2m_{\Gamma_6} \varepsilon_0}$ for both GaAs and InSb. We notice that the parabolic dispersion laws $\lambda(k) \sim k^2$ only hold when close to the band edges. By increasing k , the conduction band effective mass increases. This phenomenon, called the band non-parabolicity, is for a given k larger with a smaller ε_0 . Indeed if $\varepsilon_0 \rightarrow 0$, $\lambda(k)$ tends towards a linear instead of a quadratic dependence upon the wavevector. Another way to depict band non-parabolicity is to rewrite the Γ_6 dispersion relation in the form

$$\lambda = \frac{\hbar^2 k^2}{2m_{\Gamma_6}(\lambda)} \quad (48)$$

where from equations (44, 45) :

$$\frac{m_{\Gamma_6}(\lambda)}{m_{\Gamma_6}(0)} = \frac{(\lambda + \varepsilon_0)(\lambda + \varepsilon_0 + \Delta)(\varepsilon_0 + 2\Delta/3)}{\varepsilon_0(\varepsilon_0 + \Delta)(\lambda + \varepsilon_0 + 2\Delta/3)} \quad (49)$$

In this way it is clearly revealed that the mass increases as energy increases.

d) *Inclusion of higher bands* : One of the main drawbacks of restricting the $\mathbf{k.p}$ interaction to the $\Gamma_6, \Gamma_7, \Gamma_8$ edges is that the heavy particle states are dispersionless, apart from the free electron term. For \mathbf{k}/\mathbf{J} it is easy to cure this

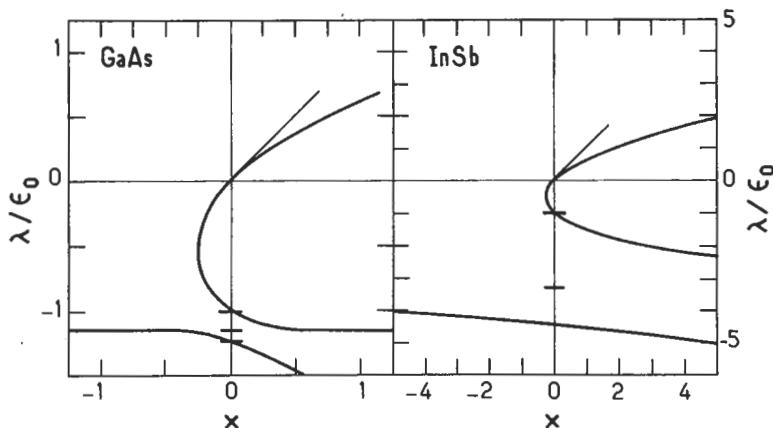


Fig. 4.— The light particle energy $\lambda(k)$ scaled to the bandgap ϵ_0 is plotted against $x = \frac{\hbar^2 k^2}{2m_{\Gamma_6}\epsilon_0}$ for GaAs (left panel) and InSb (right panel). In both figures the matrix element P is fitted (Eq. (45)) to reproduce the actual effective mass m_{Γ_6} . Notice that negative x 's correspond to evanescent states.

weakness. Since both the light and heavy states decouple, we can take care of the $\mathbf{k.p}$ coupling between Γ_6 , Γ_7 , Γ_8 and the remote bands only for the heavy particle states. In doing so, the $m_j = \pm 3/2$ states acquire a finite effective mass. Since this arises from an interaction with the remote bands, a parabolic treatment suffices. Thus :

$$\epsilon(k) = -\epsilon_0 - \frac{\hbar^2 k^2}{2m_{hh}}, \quad m_j = \pm \frac{3}{2} \quad (50)$$

with

$$\frac{1}{m_{hh}} = -\frac{1}{m_0} + \frac{2}{m_0^2} \sum_{m \neq \Gamma_6, \Gamma_7, \Gamma_8} \frac{\left| \left\langle \frac{3}{2}, \pm \frac{3}{2} | p_z | u_m \right\rangle \right|^2}{\epsilon_m + \epsilon_0}. \quad (51)$$

As the energy denominators in equation (51) are large, the heavy hole effective mass is indeed heavy : usually $m_{hh}/m_0 \sim 0.4 - 0.6$ in all the III-V compounds. Another way exists for including the remote bands effects inside the Γ_6 , Γ_7 , Γ_8 basis. It is based on the construction of an effective $\mathbf{k.p}$ operator inside the Γ_6 , Γ_7 , Γ_8 basis. For details see Appendix A.

e) *Accuracy of the Kane model* : Because the Kane model goes beyond the simple parabolic dispersion laws, it represents a significant improvement on the latter ones, especially for narrow gap materials like InSb or InAs. However the Kane model has

obvious limitations. For instance, the Γ_6 band dispersion relation asymptotically approaches the expression

$$\varepsilon_{\Gamma_6}(\mathbf{k}) \rightarrow \frac{\hbar^2 k^2}{2m_0} + \hbar k P, \quad (52)$$

when $\lambda(\mathbf{k}) \gg \varepsilon_0, \Delta$. This is incorrect when $\varepsilon_{\Gamma_6}(\mathbf{k})$ is no longer negligible compared with the bandgap which separates Γ_6 and the nearest remote band. To ascertain the accuracy of the Kane model it is useful to compare its predictions with those derived from a global description of the band structure, for example, the empirical tight binding method. The latter model builds the crystal band structure from a knowledge of the atomic properties of the constituting atoms. It is, in its construction, a global band structure calculation, valid over the whole Brillouin zone. The adjective "empirical" comes from the fact that some of the atomic properties, which are the inputs of the calculation, are fitted to the experiments to reproduce exactly some of the important features of the Brillouin zone, e.g. Γ bandgaps, effective masses etc.. Figure 5 presents a comparison between the dispersion relations obtained from the Kane model and the results of empirical tight binding calculations for GaAs and AlAs and $k/[100]$ [9]. The right hand part of each panel corresponds to propagating states (real wave vector) whereas the left hand part corresponds to evanescent states (imaginary wave vectors: $k = i\kappa$ in equations (44, 50)). The evanescent states do not play any part in the bulk materials (their wavefunctions cannot be normalized).

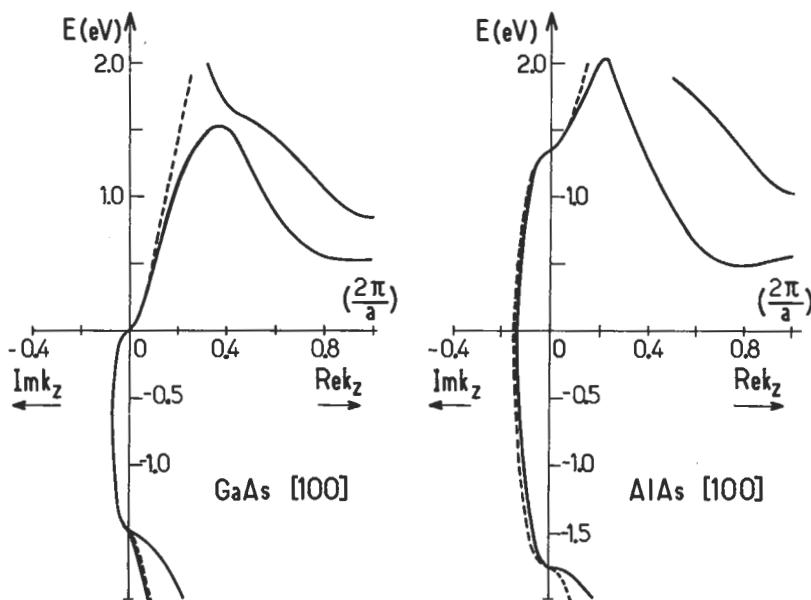


Fig. 5. — Comparison between the electronic dispersion relations of GaAs and AlAs calculated either by the empirical tight-binding method (solid line), or by the Kane model (dashed line). After reference [9].

They are of paramount importance in heterostructures made of bulk layers of finite thicknesses. It is remarkable that the Kane model and global band structure descriptions coincide so well for energies which correspond to the entire GaAs and AlAs bandgaps and up to ~ 0.3 eV in the GaAs conduction band. It is quite reassuring, with a future use of the Kane model in mind, that the heterostructure energy levels, which are the most important for practical realizations (e.g. quantum well lasers or two-dimensional field effect transistors), are precisely those which are close enough from the Γ point where the Kane description is at its best.

IV. Conclusions.

Since we are going to rely on the Kane model, it is instructive to know the value of P or $E_p = 2m_0 P^2$ in III-V materials. This value can be extracted from reliable experiments and the use of equation (45). The bandgap values ϵ_0 and Δ are accurately known from optical experiments. The conduction band effective mass is also well known from cyclotron resonance experiments. Thus with these values and equation (45) we can extract an "empirical" Kane matrix element E_p . It is empirical in the sense that higher band effects are present in the experimental m_{Γ_6} but not in the theoretical expression. With this reservation in mind, table III, which shows the E_p values for several III-V compounds, demonstrates that E_p is mostly constant over the whole III-V family : $E_p \sim 22.3$ eV if InP is excluded, $E_p \sim 21.2$ eV if InP is included. This means that an atomic like property

$$\int_{\Omega_0} u_s(\mathbf{r}) p_x u_X(\mathbf{r}) d^3r$$

is merely material independent. Thus the Bloch functions of two different III-V materials, which can be written (see Eq. (4)) as the products of a slowly varying envelope function $\exp i\mathbf{k} \cdot \mathbf{r}$. ($\mathbf{k} \ll$ Brillouin zone boundary) by a rapidly varying (atomic-like) function $u_{\mathbf{k}}(\mathbf{r})$, are such that the atomic-like functions of the host Γ_6 , Γ_7 , Γ_8 band edges are quite similar. This result will lead us in the following chapters to analyse the heterostructure eigenstates in terms of the products of the Γ

Table III. — The values of ϵ_0 , Δ , E_p , m_{Γ_6} are listed for several III-V compounds. The quantity E_p is equal to $2m_0 P^2$ and is fitted to reproduce the experimental m_{Γ_6} via equation (45).

	InP	InAs	InSb	GaAs	GaSb
ϵ_0 (eV)	1.4236	0.418	0.2352	1.5192	0.811
Δ (eV)	0.108	0.38	0.81	0.341	0.752
m_{Γ_6}/m_0	0.079	0.023	0.0139	0.0665	0.0405
E_p (eV)	17	21.11	22.49	22.71	22.88

points atomic-like functions which are taken as *identical* in both host materials by functions which slowly vary at the scale of the hosts elementary cells and which will alone experience the differences between the band structures of these host materials.

Appendix A.

Inclusion of remote bands effects.

In this appendix we attempt to present a more complete discussion of the remote bands contributions to the $8 \times 8 \Gamma_6 \Gamma_7 \Gamma_8$ effective Hamiltonian. This topic has been the subject of many papers and complete reviews. Thus we will only describe the method of handling the remote band effects and then give the result of the simplest approach. We are interested in constructing an effective $\mathbf{k} \cdot \mathbf{p}$ operator which acts inside the $\Gamma_6 \Gamma_7 \Gamma_8$ basis by allowing at the lowest order for the presence of $k = 0$ edges other than $\Gamma_6 \Gamma_7 \Gamma_8$. Our problem for the u_{nk} functions can be formally written as :

$$(H_0 + W)|\Psi\rangle = \varepsilon |\Psi\rangle \quad (A1)$$

where H_0 is the crystal Hamiltonian and W the operator

$$W = \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p} + \frac{\hbar^2 k^2}{2m_0} \quad (A2)$$

The energy spectrum of H_0 consists of a set of closely spaced levels $|l\rangle$ ($1 \leq l \leq 8$ to span the $\Gamma_6 \Gamma_7 \Gamma_8$ basis) which are well energy-separated from the other eigenstates $|\nu\rangle$ of H_0 . We are interested in the eigenvalues ε which are all of the order of $\bar{\varepsilon}$, where $\bar{\varepsilon}$ is the average energy of the states $|l\rangle$, and thus very different from ε_ν for all $|\nu\rangle$. We expand $|\Psi\rangle$ in the form

$$|\Psi\rangle = \sum_l c_l |l\rangle + \sum_\nu c_\nu |\nu\rangle \quad (A3)$$

Inserting equation (A3) into (A1) and multiplying by $\langle m |$ or $\langle \mu |$ we obtain

$$\sum_l c_l [(\varepsilon_l - \varepsilon) \delta_{lm} + \langle m | W | l \rangle] + \sum_\nu c_\nu \langle m | W | \nu \rangle = 0 \quad (A4)$$

$$\sum_\nu c_\nu [(\varepsilon_\nu - \varepsilon) \delta_{\mu\nu} + \langle \mu | W | \nu \rangle] + \sum_l c_l \langle \mu | W | l \rangle = 0 \quad (A5)$$

In equations (A4-A5) we know that all the $c_l \sim 1$ and all the $c_\nu \sim \frac{1}{\varepsilon_\nu}$. From equation (A5) we can thus write

$$c_\nu \sim \frac{1}{\varepsilon - \varepsilon_\nu} \sum_l c_l \langle \nu | W | l \rangle \quad (A6)$$

Inserting this value into equation (A4) we obtain the required diagonalization of W inside the $|l\rangle$ basis :

$$\sum_l c_l \left\{ (\varepsilon_l - \varepsilon) \delta_{lm} + \langle m | W | l \rangle + \langle m | W \sum_{\nu} \frac{| \nu \rangle \langle \nu |}{\varepsilon - \varepsilon_{\nu}} W | l \rangle = 0 \right\} \quad (A7)$$

The virtual transitions from $|l\rangle$ to the remote bands $|\nu\rangle$ modify the value of the operator W inside the $|l\rangle$ basis :

$$W \rightarrow \tilde{W} = W + W \sum_{\nu} \frac{| \nu \rangle \langle \nu |}{\varepsilon - \varepsilon_{\nu}} W \quad (A8)$$

In principle \tilde{W} is energy-dependent. However, this dependence is very weak since, by assumption, $\varepsilon - \varepsilon_{\nu}$ is always large. In fact one replaces ε by $\bar{\varepsilon}$ in equation (A8). In equation (34) we saw that W has no non-vanishing matrix elements in the $\Gamma_7 \Gamma_8$ subspace (the corresponding u_l 's, $l > 2$, all display a p-like symmetry). The adjunction of the correcting term $\tilde{W} - W$ to the effective $\mathbf{k} \cdot \mathbf{p}$ matrix equation (A7) cures this shortcoming. Thus, the Γ_8 heavy particle states, which were dispersionless (apart from the free electron term), now acquire a dispersion. The price we pay for this improvement is that the diagonalization of equation (A7) can only be achieved numerically. The group theory analysis of $\tilde{W} - W$ helps to reduce the number of independent constants. [10-12]. However since the zinc-blende lattice lacks for inversion symmetry (the two atoms of the basis being different), this number is still too large. Inversion asymmetry-induced terms are usually very small. Thus, in a simplified approach we may discard them. Four independent parameters survive to this "quasi-Ge" assumption [10]. The 8×8 matrix $\tilde{W} - W$ has been derived by several authors. Here we follow Bir and Pikus [13] to obtain :

$ iS^\dagger\rangle$	$ \frac{3}{2}, \frac{1}{2}\rangle$	$ \frac{3}{2}, \frac{3}{2}\rangle$	$ \frac{1}{2}, \frac{1}{2}\rangle$	$\langle iS^\dagger $	$ \frac{3}{2}, -\frac{1}{2}\rangle$	$ \frac{3}{2}, -\frac{3}{2}\rangle$	$ \frac{1}{2}, -\frac{1}{2}\rangle$
$\langle iS^\dagger $	$\frac{\hbar^2 k^2}{2} \left(\frac{1}{m_c} - \frac{1}{m_0} \right)$	0	0	0	0	0	0
$\left\langle \frac{3}{2}, \frac{1}{2} \right $	0	$G - \frac{\hbar^2 k^2}{2m_0}$	iH^*	$-\frac{1}{\sqrt{2}}(G - F)$	0	0	$i\sqrt{\frac{3}{2}}H$
$\left\langle \frac{3}{2}, \frac{3}{2} \right $	0	$-iH$	$F - \frac{\hbar^2 k^2}{2m_0}$	$\frac{iH}{\sqrt{2}}$	0	$-i$	$-i\sqrt{2}$
$\left\langle \frac{1}{2}, \frac{1}{2} \right $	0	$-\frac{1}{\sqrt{2}}(G - F)$	$-\frac{iH^*}{\sqrt{2}}$	$\frac{1}{2}(F + G) - \frac{\hbar^2 k^2}{2m_0}$	0	$-i\sqrt{\frac{3}{2}}H$	0
$\langle iS^\dagger $	0	0	0	$\frac{\hbar^2 k^2}{2} \left(\frac{1}{m_c} - \frac{1}{m_0} \right)$	0	0	0
$\left\langle \frac{3}{2}, -\frac{1}{2} \right $	0	0	$-I^*$	$iH^* \sqrt{\frac{3}{2}}$	0	$G - \frac{\hbar^2 k^2}{2m_0}$	$-iH$
$\left\langle \frac{3}{2}, -\frac{3}{2} \right $	0	I^*	0	$\sqrt{2}I^*$	0	iH^*	$F - \frac{\hbar^2 k^2}{2m_0}$
$\left\langle \frac{1}{2}, -\frac{1}{2} \right $	0	$-i\sqrt{\frac{3}{2}}H^*$	$-I^* \sqrt{2}$	0	0	$\frac{-1}{\sqrt{2}}(G - F)$	$\frac{iH}{\sqrt{2}}$
						$\frac{1}{2}(F + G) - \frac{\hbar^2 k^2}{2m_0}$	

$$\tilde{W} - W =$$

(A9)

where the polynomials F, G, H, I are given by :

$$F(\mathbf{k}) = Ak^2 + \frac{B}{2} (k^2 - 3k_z^2) \quad (A10)$$

$$G(\mathbf{k}) = Ak^2 - \frac{B}{2} (k^2 - 3k_z^2) \quad (A11)$$

$$H(\mathbf{k}) = -Dk_z (k_x + ik_y) \quad (A12)$$

$$I(\mathbf{k}) = \frac{\sqrt{3}}{2} B (k_x^2 - k_y^2) - iD k_x k_y \quad (A13)$$

with :

$$A = \frac{L + 2M}{3}, \quad B = \frac{L - M}{3}, \quad C^2 = D^2 - 3B^2, \quad D = \frac{N}{\sqrt{3}} \quad (A14)$$

$$L = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{\nu} \frac{|\langle X|p_x|\nu \rangle|^2}{\bar{\varepsilon} - \varepsilon_{\nu}} \quad (A15)$$

$$M = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{\nu} \frac{|\langle X|p_y|\nu \rangle|^2}{\bar{\varepsilon} - \varepsilon_{\nu}} \quad (A16)$$

$$N = \frac{\hbar^2}{m_0^2} \sum_{\nu} \frac{\langle X|p_x|\nu \rangle \langle \nu|p_y|Y \rangle + \langle X|p_y|\nu \rangle \langle \nu|p_x|Y \rangle}{\bar{\varepsilon} - \varepsilon_{\nu}} \quad (A17)$$

$$\frac{1}{m_c} = \frac{2}{m_0^2} \sum_{\nu} \frac{|\langle S|p_x|\nu \rangle|^2}{\bar{\varepsilon} - \varepsilon_{\nu}}. \quad (A18)$$

In contrast to equation (43), the eigenvalues of the full $H_0 + W$ matrix are anisotropic in \mathbf{k} (warped energy surfaces). This effect is most significant for the Γ_8 bands.

When ε_0 and Δ are large enough a parabolic description of the Γ_8 dispersion relations is sufficient. To obtain it one must include the Γ_6 levels back in the $|\nu\rangle$ states in equations (A15-A17). Then, one ends up with a 4×4 matrix Hamiltonian \mathcal{H}_{Γ_8} . Rearranging the lines and columns we get finally :

$$\mathcal{H}_{\Gamma_8}(\mathbf{k}) = \begin{array}{c|ccccc} & \left| \frac{3}{2}, \frac{3}{2} \right\rangle & \left| \frac{3}{2}, \frac{1}{2} \right\rangle & \left| \frac{3}{2}, -\frac{1}{2} \right\rangle & \left| \frac{3}{2}, -\frac{3}{2} \right\rangle \\ \left\langle \frac{3}{2}, \frac{3}{2} \right| & F - \varepsilon_0 & -iH & -I & 0 \\ \hline & \left\langle \frac{3}{2}, \frac{1}{2} \right| & iH^* & G - \varepsilon_0 & 0 & I \\ \hline & \left\langle \frac{3}{2}, -\frac{1}{2} \right| & -I^* & 0 & G - \varepsilon_0 & -iH \\ \hline & \left\langle \frac{3}{2}, -\frac{3}{2} \right| & 0 & I^* & iH^* & F - \varepsilon_0 \end{array} \quad (A19)$$

The dispersion relations of the Γ_8 bands are exactly obtained in the parabolic limit :

$$\varepsilon_{\Gamma_8}(\mathbf{k}) = -\varepsilon_0 + \frac{1}{2}(F + G) \pm \sqrt{\left(\frac{F - G}{2}\right)^2 + |I|^2 + |H|^2} \quad (\text{A20})$$

or :

$$\varepsilon_{\Gamma_8}(\mathbf{k}) = -\varepsilon_0 + Ak^2 \pm \sqrt{B^2 k^4 + C^2 [k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2]} \quad (\text{A21})$$

The relation (A21) explicitly reveals the cubic (but not spherical) symmetry of the iso-energy surfaces. The band warping disappears when the parameter C is set equal to zero. This spherical approximation is often done for computational conveniences, but, as seen from table I, is not very well justified.

Note that if $\mathbf{k} \parallel \hat{z}$ equation (A21) simplifies into

$$\varepsilon_{\Gamma_8}(k_z) = -\varepsilon_0 + (A \pm B) k_z^2 \quad (\text{A22})$$

which for the $|3/2, \pm 3/2\rangle$ states coincides with equation (50).

Appendix B.

Motion of Kane electrons in slowly varying perturbing fields.

Up to now we have focused our attention on the energy spectra of perfect III-V compounds. In this appendix we attempt to describe the quantum motion of an electron moving in an imperfect III-V crystal. We look for the solutions of

$$[\mathcal{H}_0 + \phi(\mathbf{r})] \Psi(\mathbf{r}) = \varepsilon \Psi(\mathbf{r}) \quad (\text{B1})$$

where \mathcal{H}_0 is the Hamiltonian of the perfect crystal and $\phi(\mathbf{r})$ is a non periodic potential which we assume to be slowly varying at the scale of the primitive cell [14]. Let us first establish a property linked to the slowly varying nature of a function $f(\mathbf{r})$. We consider the integral

$$I = \int_{\Omega} f(\mathbf{r}) \alpha(\mathbf{r}) d^3r \quad (\text{B2})$$

where $\alpha(\mathbf{r})$ is a periodic function with the periodicity of the direct lattice :

$$\alpha\left(\mathbf{r} + \sum_i n_i \mathbf{a}_i\right) = \alpha(\mathbf{r}) \quad (\text{B3})$$

where \mathbf{a}_i , $i = 1, 2, 3$ are the basis vectors of the direct lattice, n_i are relative integers,

and where $f(\mathbf{r})$ decays fast enough at large \mathbf{r} to ensure that I converges. The periodicity of $\alpha(\mathbf{r})$ allows the Fourier expansion.

$$\alpha(\mathbf{r}) = \sum_{\mathbf{K}} \alpha(\mathbf{K}) \exp(i\mathbf{K}\cdot\mathbf{r}) \quad (B4)$$

where the \mathbf{K} 's are vectors belonging to the reciprocal lattice :

$$\mathbf{K} = \sum_i m_i \mathbf{b}_i \quad (B5)$$

the \mathbf{b} 's being the basis vector of the reciprocal lattice. The Fourier coefficients $\alpha(\mathbf{k})$ are given by :

$$\alpha(\mathbf{K}) = \frac{1}{\Omega} \int_{\Omega} \alpha(\mathbf{r}) \exp[-i\mathbf{K}\cdot\mathbf{r}] d^3r = \frac{1}{\Omega_0} \int_{\Omega_0} \alpha(\mathbf{r}) \exp[-i\mathbf{K}\cdot\mathbf{r}] d^3r \quad (B6)$$

where Ω_0 is the volume of the unit cell of the direct lattice. Thus, the integral I is equal to

$$I = \sum_{\mathbf{K}} \alpha(\mathbf{K}) \int_{\Omega} d^3r f(\mathbf{r}) \exp[i\mathbf{K}\cdot\mathbf{r}]. \quad (B7)$$

If we decompose the integral over Ω into a sum of integrals over the unit cells Ω_i centred at the lattice points \mathbf{R}_i , we get :

$$\int_{\Omega} d^3r f(\mathbf{r}) \exp[i\mathbf{K}\cdot\mathbf{r}] = \sum_{\mathbf{R}_i} \int_{\Omega_i} d^3r f(\mathbf{R}_i + \mathbf{r}) \exp[i\mathbf{K}\cdot\mathbf{r}] \quad (B8)$$

Here, we have made use of the property $\mathbf{K}\cdot\mathbf{R}_i = 2n\pi$ where n is a relative integer. We now exploit the slowly varying nature of $f(\mathbf{r})$ in order to write $f(\mathbf{R}_i + \mathbf{r}) \approx f(\mathbf{R}_i)$ (\mathbf{r} is restricted to a unit cell). Thus :

$$\int_{\Omega} d^3r f(\mathbf{r}) \exp[i\mathbf{K}\cdot\mathbf{r}] \sim \sum_{\mathbf{R}_i} f(\mathbf{R}_i) \int_{\Omega_i} \exp[i\mathbf{K}\cdot\mathbf{r}] d^3r. \quad (B9)$$

The last integral is zero (unless $\mathbf{K} = \mathbf{0}$). Thus :

$$I \approx \alpha(0) \int_{\Omega} d^3r f(\mathbf{r}) \quad (B10)$$

or :

$$\int_{\Omega} f(\mathbf{r}) \alpha(\mathbf{r}) d^3r \approx \frac{1}{\Omega} \int_{\Omega} \alpha(\mathbf{r}) d^3r \int_{\Omega} d^3r f(\mathbf{r}). \quad (B11)$$

The Bloch states $|n\mathbf{k}\rangle$ are normalized

$$\langle n\mathbf{k} | n' \mathbf{k}' \rangle = \delta_{nn'} \delta_{\mathbf{k}\mathbf{k}'}$$

This means that

$$\delta_{nn'} = \frac{1}{\Omega} \int u_{nk}^*(\mathbf{r}) u_{n'k}(\mathbf{r}) d^3r = \frac{1}{\Omega_0} \int_{\Omega_0} u_{nk}^*(\mathbf{r}) u_{n'k}(\mathbf{r}). \quad (B12)$$

To solve equation (B1) we assume that $\phi(\mathbf{r})$ only couples the $\Gamma_6 \Gamma_7 \Gamma_8$ states between themselves. Thus we look for a solution of (B1) in the form of :

$$\psi(\mathbf{r}) = \sum_{l=1}^8 c_l f_l(\mathbf{r}) u_{l0}(\mathbf{r}), \quad (B13)$$

where the $f_l(\mathbf{r})$'s are *envelope* functions which slowly vary on the scale of Ω_0 and where the u_{l0} 's are the Γ point Bloch functions for the 8 band edges (see Table II). For instance, if $\phi(\mathbf{r})$ vanishes, the f_l are the plane waves $\Omega^{-1/2} \exp \mathbf{k} \cdot \mathbf{r}$. As long as the Kane model is justified ($k \sim \frac{1}{10}$ of the Brillouin zone, see Fig. 5) they vary slowly. The wave function $\psi(\mathbf{r})$ is normalized :

$$1 = \int_{\Omega} |\psi(\mathbf{r})|^2 d^3r = \sum_{l,m} c_l c_m^* \int_{\Omega} f_m^* f_l u_{m0}^* u_{l0} d^3r. \quad (B14)$$

Making use of equations (B11, B13) we obtain

$$1 = \sum_l |c_l|^2 \int_{\Omega} |f_l(\mathbf{r})|^2 d^3r. \quad (B15)$$

Thus, the envelope functions have to be normalized to 1 in the volume Ω , the c_l being such that $\sum_l |c_l|^2 = 1$.

By inserting equation (B13) into equation (B1), multiplying by $u_{m0}^* f_m^*$ and integrating over Ω we obtain :

$$0 = \sum_{l=1}^8 c_l \left\{ (\varepsilon_l - \varepsilon) \int_{\Omega} f_m^* f_l u_{l0} u_{m0}^* d^3r + \int_{\Omega} u_{m0}^* u_{l0} f_m \frac{p^2}{2m_0} f_l d^3r + \frac{1}{m_0} \int_{\Omega} u_{m0}^* \mathbf{p} u_{l0} \cdot f_m^* \mathbf{p} f_l d^3r + \int_{\Omega} f_m^* \phi(\mathbf{r}) f_l u_{m0}^* u_{l0} d^3r \right\} \quad (B16)$$

where the \mathbf{k} -dependent spin-orbit coupling has been dropped. Making use of equation (B11) we finally obtain

$$0 = \int_{\Omega} f_m^*(\mathbf{r}) d^3r \sum_{l=1}^8 c_l \left\{ \delta_{lm} \left[\varepsilon_l - \varepsilon + \frac{p^2}{2m_0} + \phi(\mathbf{r}) \right] + \mathbf{p}_{ml} \cdot \mathbf{p} \right\} f_l(\mathbf{r}) \quad (B17)$$

where :

$$\mathbf{p}_{ml} = \frac{1}{m_0} \frac{1}{\Omega} \int_{\Omega} u_{m0}^* \mathbf{p} u_{l0} d^3r = \frac{1}{m_0 \Omega_0} \int_{\Omega_0} u_{m0}^* \mathbf{p} u_{l0} d^3r. \quad (B18)$$

It is worth noting that the u_{l0} 's no longer explicitly appear in the coupled system (Eq. (B17)). Their existence is embodied into the matrix elements p_{ml} . This is a direct consequence of the slow spatial variation of the perturbing potential $\phi(\mathbf{r})$. It implies that $\phi(\mathbf{r})$ has only intra-edge matrix elements, whereas all the mixed terms $\int f_l f_m^* \phi u_m^* u_l, l \neq m$, vanish. The system (B17) can be formally rewritten as

$$\langle \mathbf{f} | D + (\phi(\mathbf{r}) - \varepsilon) \mathbf{1} | \mathbf{f} \rangle = 0, \quad (\text{B19})$$

where $|\mathbf{f}\rangle$ is a (8×1) column vector, D an 8×8 matrix and $\mathbf{1}$ the 8×8 identity matrix. Non-vanishing solutions of the system exist if \mathbf{f} is the eigensolution of the problem

$$(D + \phi(\mathbf{r})) \mathbf{f} = \varepsilon \mathbf{1} \mathbf{f}. \quad (\text{B20})$$

The differential operator D is simply the (8×8) matrix of equation (39) where $\hbar \mathbf{k}$ has been replaced by the differential operator $\mathbf{p} = \frac{\hbar}{i} \nabla$. Indeed equation (39) is recovered if $\phi(\mathbf{r})$ is zero. In this case, $\mathbf{f} = \Omega^{-1/2} \exp i \mathbf{k} \cdot \mathbf{r} \mathbf{C}$, where \mathbf{C} is a 8×1 c -column vector.

Instead of handling the $8 \times 8 D$, it may prove useful to eliminate six of the f 's to obtain a 2×2 effective D which acts on the two remaining envelopes [15]. For instance, if $\phi(\mathbf{r})$ is the potential created by a Coulombic donor, which we expect to give rise to bound states below the Γ_6 edge, we will get rid of $f_3, f_4 \dots f_8$ to the benefit of f_1, f_2 which are the envelopes associated with the Γ_6 edge.

If we introduce

$$p_{\pm} = \frac{1}{\sqrt{2}} (p_x \pm i p_y) \quad (\text{B21})$$

and drop the free electron term, we obtain :

$$\begin{aligned} f_3(\mathbf{r}) &= [\varepsilon_0 + \varepsilon - \phi(\mathbf{r})]^{-1} \left[-\sqrt{\frac{2}{3}} P p_z f_1 + \frac{P}{\sqrt{3}} p_- f_2 \right]. \\ f_4(\mathbf{r}) &= [\varepsilon_0 + \varepsilon - \phi(\mathbf{r})]^{-1} \left[-\frac{P}{\sqrt{3}} p_+ f_1 - \sqrt{\frac{2}{3}} P p_z f_2 \right] \\ f_5(\mathbf{r}) &= [\varepsilon_0 + \varepsilon - \phi(\mathbf{r})]^{-1} [P p_- f_2] \\ f_6(\mathbf{r}) &= [\varepsilon_0 + \varepsilon - \phi(\mathbf{r})]^{-1} [P p_+ f_2] \\ f_7(\mathbf{r}) &= [\varepsilon_0 + \Delta + \varepsilon - \phi(\mathbf{r})]^{-1} \left[\frac{P}{\sqrt{3}} p_z f_1 + P \sqrt{\frac{2}{3}} p_- f_2 \right] \\ f_8(\mathbf{r}) &= [\varepsilon_0 + \Delta + \varepsilon - \phi(\mathbf{r})]^{-1} \left[-\sqrt{\frac{2}{3}} P p_+ f_1 + \frac{P}{\sqrt{3}} p_z f_2 \right] \end{aligned} \quad (\text{B22})$$

Notice that in equations (B22) the two brackets do *not* commute. The effective

2×2 system, which is a projection of the 8×8 system onto the Γ_6 edges, is finally rewritten as :

$$\begin{bmatrix} \mathcal{H}_d & \mathcal{H}_{nd} \\ \mathcal{H}_{nd}^+ & \mathcal{H}_d \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \varepsilon \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \quad (\text{B23})$$

where :

$$\begin{aligned} \mathcal{H}_d = & \frac{2P^2}{3} p_z [\varepsilon + \varepsilon_0 - \phi(\mathbf{r})]^{-1} p_z + P^2 p_+ [\varepsilon + \varepsilon_0 - \phi(\mathbf{r})]^{-1} p_- + \frac{P^2}{3} \times \\ & p_- [\varepsilon + \varepsilon_0 - \phi(\mathbf{r})]^{-1} p_+ + \frac{2P^2}{3} p_- [\varepsilon + \varepsilon_0 + \Delta - \phi(\mathbf{r})]^{-1} p_+ + \frac{P^2}{3} p_z \times \\ & [\varepsilon + \varepsilon_0 + \Delta - \phi(\mathbf{r})]^{-1} p_z + \phi(\mathbf{r}) \end{aligned} \quad (\text{B24})$$

$$\begin{aligned} \mathcal{H}_{nd} = & \frac{\sqrt{2}}{3} P^2 [p_- \{ [\varepsilon + \varepsilon_0 - \phi(\mathbf{r})]^{-1} - [\varepsilon + \varepsilon_0 + \Delta - \phi(\mathbf{r})]^{-1} \} p_z - \\ & - p_z \{ [\varepsilon + \varepsilon_0 - \phi(\mathbf{r})]^{-1} - [\varepsilon + \varepsilon_0 + \Delta - \phi(\mathbf{r})]^{-1} \} p_-] \end{aligned} \quad (\text{B25})$$

Let us recall that $f_1(f_2)$ is associated with the $S^\uparrow(S^\downarrow)$ band edge Bloch states. Equations (B23-B25) show that a scalar (i.e. spin independent) potential $\phi(\mathbf{r})$ mixes the $| \uparrow \rangle$ and $| \downarrow \rangle$ states. This is not surprising since the host crystalline matrix displays, even in the absence of a perturbing potential, a non zero spin-orbit coupling. For perfect crystals we have shown that it is always possible to obtain eigenstates of D which are also, at finite \mathbf{k} , eigenstates of J_z and thus do not display S^\uparrow and S^\downarrow mixing. To obtain such states we only had to rotate \mathbf{J} in order to line it up along the \mathbf{k} vector. The envelope functions of the perturbed crystal are linear combinations of plane waves. Thus in the presence of a perturbing potential $\phi(\mathbf{r})$ it is in general impossible to achieve a simultaneous diagonalization of $D + \phi(\mathbf{r})$ and of J_z . The eigenstates of a perturbed semiconductor are, in general, a mixture of different m_j 's.

It should be noted that \mathcal{H}_{nd} vanishes identically if $\Delta = 0$, i.e. if the host matrix has no spin-orbit effect. We also note that \mathcal{H}_{nd} involves the energy denominators $[\varepsilon + \varepsilon_0 - \phi(\mathbf{r})]^{-1}$ and $[\varepsilon + \varepsilon_0 + \Delta - \phi(\mathbf{r})]^{-1}$. Thus the smaller ε_0 favours the larger spin mixing effects. In fact, striking evidences of the influence of \mathcal{H}_{nd} have been found in narrow gap materials with large spin orbit effects, e.g. InSb or $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ alloys ($x \sim 0.2$). For instance, strong electric dipole spin-flip magneto-optical transitions have been observed [4], as well as the anomalous Hall effect, i.e. a Hall conductivity which is induced by the spin magnetization of the carriers instead of by the external magnetic field [17, 18].

Finally it is interesting to notice that the eigenvalue equation (B23) is non-linear in ε as far as \mathcal{H}_d and \mathcal{H}_{nd} explicitly depend on ε . This situation is not unique as it is also found in the quantum treatment of relativistic particles [16, 19].

The hydrogenic donor problem.

It may be illustrative to discuss more thoroughly a specific case. We take

$$\phi(r) = \frac{-e^2}{\kappa r} \quad (B26)$$

where κ is the static dielectric constant of the semiconductor. We are interested in deriving the first order corrections to the usual hydrogenic Hamiltonian which are induced by the band non-parabolicity.

We notice that for almost all r $|\phi(r)| \ll \varepsilon_0$. When this inequality is not fulfilled, r is so small that $\phi(r)$ can no longer be considered as slowly varying. Under most circumstances this range of the r values plays a negligible part in the binding energy. Since $|\phi(r)| \ll \varepsilon_0$, the energy ε which is of interest to us is also much smaller than ε_0 . Thus we can expand \mathcal{H}_d and \mathcal{H}_{nd} to the lowest order in $(-\phi(r) + \varepsilon)/\varepsilon_0$, $(-\phi(r) + \varepsilon)/(\varepsilon_0 + \Delta)$.

We obtain :

$$\begin{aligned} \mathcal{H}_d = & \frac{p^2}{2m_{\Gamma_6}} + \phi(r) - \frac{2P^2}{3\varepsilon_0^2} \varepsilon p^2 - \frac{P^2 \varepsilon}{3(\varepsilon_0 + \Delta)^2} p^2 + \frac{2P^2}{3\varepsilon_0^2} p_z \phi(r) p_z + \frac{P^2}{\varepsilon_0^2} p_+ \phi(r) p_- \\ & + \frac{P^2}{3\varepsilon_0^2} p_- \phi(r) p_+ + \frac{P^2}{3(\varepsilon_0 + \Delta)^2} p_z \phi(r) p_z + \frac{2P^2}{3(\varepsilon_0 + \Delta)^2} p_- \phi(r) p_+ \end{aligned} \quad (B27)$$

$$\mathcal{H}_{nd} = \frac{\sqrt{2}}{3} P^2 \left[\frac{1}{\varepsilon_0^2} - \frac{1}{(\varepsilon_0 + \Delta)^2} \right] [p_- \phi(r) p_z - p_z \phi(r) p_-]. \quad (B28)$$

The zero order term

$$\mathcal{H}_d^{(0)} = \frac{p^2}{2m_{\Gamma_6}} + \phi(r) \quad (B29)$$

is the "parabolic" hydrogenic donor Hamiltonian whose solutions are the same as those of the hydrogenic atom problem, the only exception being that the Bohr radius and Rydberg energy are scaled according to

$$\begin{aligned} a_0^* &= 0.53 \text{ \AA} \times \kappa \times \frac{m_0}{m_{\Gamma_6}} \\ R_0^* &= 13.6 \text{ eV} \times \frac{m_{\Gamma_6}}{\kappa^2 m_0} \end{aligned} \quad (B30)$$

The next two terms to $\mathcal{H}_d^{(0)}$ in \mathcal{H}_d are kinetic energy corrections induced by the band non-parabolicity : they are of the order of $\frac{\varepsilon}{\varepsilon_0} \frac{P^2}{m_{\Gamma_6}}$. To construct the hydrogenic

wavefunctions we have to build a localized wave packet by picking up states of finite \mathbf{k} in the conduction band. As soon as \mathbf{k} is non zero, the effective mass attached to this

\mathbf{k} state becomes heavier than $m_{F_6}^*$ (see Eq. (49)). The parabolic term $p^2/2m_{F_6}$ thus overestimates the kinetic energy of the electron. Consequently, the non-parabolicity-induced kinetic energy corrections can only *lower* the energy. This explains the presence of a minus sign in front of the two terms. The other correcting terms of $\mathcal{H}_d^{(0)}$ in \mathcal{H}_d involve $\phi(\mathbf{r})$. Nevertheless, they are also of kinetic origin. Locally, at the \mathbf{r} point, the kinetic energy of the carrier is not ϵ (as would be the case if $\phi(\mathbf{r})$ were zero) but rather $\epsilon - \phi(\mathbf{r})$. Thus in establishing the previous kinetic energy corrections we have gone too far and must allow for the spatial non-uniformity of the non-parabolicity effects. This explains why the ϕ -dependent terms are all affected by a sign which is opposite to that of the first set of corrections.

As for \mathcal{H}_{nd} , it has no non zero-order terms but results from the combined actions of the external potential $\phi(\mathbf{r})$, the band non-parabolicity $\left(\mathcal{H}_{nd} \sim \frac{\phi}{\epsilon_0} \right)$ and the host spin-orbit coupling ($\mathcal{H}_{nd} = 0$ if $\Delta = 0$).

The eigenstates of $\mathcal{H}_d^{(0)} = \mathcal{H}_d^{(0)} \mathbf{1}$ are the two sets of column vectors

$$\begin{bmatrix} \varphi_{nlm}(\mathbf{r}) \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ \varphi_{nlm}(\mathbf{r}) \end{bmatrix}, \quad (B31)$$

where n, l, m are the usual hydrogenic quantum numbers [19]. To the first order in $\mathcal{H} - \mathcal{H}_d^{(0)}$ only the diagonal corrections, i.e. $\mathcal{H}_d - \mathcal{H}_d^{(0)}$ appear.

Higher order corrections involve \mathcal{H}_{nd} which admixes, say, the $|1s\uparrow\rangle$ level with $|ndm\downarrow\rangle$ levels. For the p states, which are degenerate at the parabolic approximation, one should diagonalize $\mathcal{H} - \mathcal{H}_d^{(0)}$ exactly. Clearly \mathcal{H}_{nd} acts as a spin-orbit effect and partially removes the sixfold degeneracy of the $|n, p, m, \pm \frac{1}{2}\rangle$ levels.

References

- [1] W.A. HARRISON, *J. Vac. Sci. Technol.* **14** (1977) 1016 and **B 3** (1985) 1231.
- [2] LANDOLT-BORNSTEIN, *Numerical Data and Functional Relationships in Science and Technology*, group III, Volume 17 (Springer Verlag, Berlin) 1982.
- [3] J.C. PHILLIPS, *Bonds and Bands in Semiconductors* (Academic Press, London) 1973.
- [4] see, e.g., M.H. WEILER in *Semiconductors and Semimetals* Volume 16, edited by R.K. Willardson and A.C. Beer (Academic Press New, York) 1981.
- [5] N.W. ASHCROFT and N.D. MERMIN, *Solid State Physics* (HoltRinehart and Winston, New York) 1976.
- [6] E.O. KANE, *J. Phys. Chem. Sol.* **1** (1957) 249.
- [7] see, e.g., E.O. KANE in *Physics of III-V Compounds*, Volume 1, edited by R.K. Willardson and A.C. Beer (Academic Press, New York) 1966.
- [8] E.O. KANE in *Narrow Gap Semiconductors. Physics and Applications* edited by W. Zawadzki, *Lect. Notes Phys.*, Volume 133 (Springer Verlag, Berlin) 1980.
- [9] J.N. SCHULMAN and T.C. Mc GILL, in *Synthetic Modulated Structure Materials*, edited by L.L. Chang and B.C. Giessen (Academic Press, New York) 1984.
- [10] C.R. PIDGEON and R.N. BROWN, *Phys. Rev.* **146** (1966) 575.
- [11] J.M. LUTTINGER, *Phys. Rev.* **102** (1956) 1030.

- [12] G. DRESSELHAUS, F. KIP and C. KITTEL, *Phys. Rev.* **98** (1955) 368.
 - [13] G.L. BIR and G.E. PIKUS, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York) 1974.
 - [14] J.M. LUTTINGER and W. KOHN, *Phys. Rev.* **97** (1955) 869.
 - [15] D.M. LARSEN, *J. Phys. Chem. Solids* **29** (1968) 271.
 - [16] W. ZAWADZKI, *Adv. Phys.* **23** (1974) 45.
 - [17] Y. YAFET, *Solid State Phys.* **14** (1965) 1.
 - [18] P. NOZIERES and C. LEWINER, *J. Phys. France* **34** (1973) 901.
 - [19] A. MESSIAH, *Mécanique Quantique* (Dunod, Paris) 1959.
-

CHAPTER III

Envelope function description of heterostructure electronic states

I. Introduction.

From Chapter II we know which are the eigenstates of bulk III-V compounds in the vicinity of the zone centre. In this chapter we shall deal with the determination of the eigenstates in heterostructures. The emphasis will be placed on a simple description of these eigenstates, i.e. a description based on the Kane analysis of the dispersion relations of the host materials. In this envelope function description [1-5], our problem will be to find the boundary conditions which the slowly varying parts of the heterostructure wavefunctions must fulfil at the hetero-interfaces.

It should be stressed that other approaches to the heterostructure energy levels have been proposed which are more microscopic in essence than the envelope function scheme. In the empirical tight-binding calculations for instance (see e.g. [6]), one begins with a series of energies which are characteristic of the sp^3 bonds linking one atom to its neighbours. The heterostructure wavefunction is then built atom after atom. In other words a heterostructure is nothing but a bulk material with a very large unit cell (e.g. the superlattice unit cell in the case of a periodic stacking of the A and B layers). It was previously thought that the size of the computer calculations necessary to handle such large cells would become rapidly prohibitive, restricting the tight-binding analysis to short period superlattices where the unit cell is not too large. Chang and Schulman have however shown how to cure this drawback [7] and the tight-binding model is nowadays successfully used for heterostructures of any size, although it may have difficulties in handling self-consistent calculations which arise when charges are present in the heterostructure.

Another microscopic approach is the pseudo-potential formalism, which is very successful in bulk materials. Recently, it has been applied to a variety of heterostructures by Jaros *et al.* [8, 9]. Here, the core of the model is to consider, say, a periodic stacking of GaAs and AlAs slabs as a perturbation over the zero-order situation, which in this case would be the bulk GaAs. In other words a calculation is made analogous to the deep level ones in bulk semiconductors. The advantage of these microscopic approaches is their capacity to handle any heterostructure energy levels, i.e. those close to or far from the Γ edge. This occurs because these models reproduce the whole bulk dispersion relations. The model that we shall develop has no such generality. Basically, it is restricted to the vicinity of the high-symmetry points in the host's Brillouin zone (Γ , X, L). We feel however that it is invaluable due to its simplicity and versatility. It often leads to analytical results and leaves the user with the feeling that he can trace back, in a relatively transparent way, the physical origin of the numerical results. Besides, most of the heterostructures' energy levels

relevant to actual devices are relatively close to a high symmetry point in the hosts' Brillouin zone.

We shall first present the assumptions used to derive the envelope function model. As the reader may have perceived in chapter II, the algebra is relatively easy when the in-plane wavevector \mathbf{k}_\perp of the carrier is zero. The heterostructure states can then be classified according to m_J , the z projection of the total angular momentum J , and two kinds of energy levels result. The first kind corresponds to light particle states which are hybrids of the Γ_6 , Γ_8' and Γ_7 host states, whereas the second kind corresponds to the heavy hole levels. These $\mathbf{k}_\perp = \mathbf{0}$ states in quantum wells and superlattices will be extensively discussed and our considerations will be illustrated by specific examples. The third part of this chapter will be devoted to a discussion of the in-plane dispersion relations in the heterostructures. This is a topic which is currently being actively researched. Because of the mixing between the $\mathbf{k}_\perp = \mathbf{0}$ light and heavy particle states it has proved, up to now, impossible to obtain analytical solutions of the $\mathbf{k}_\perp \neq \mathbf{0}$ heterostructure problem. Therefore, we shall only present the assumptions used for the calculations and some examples of the in-plane dispersion relations in several heterostructures. Unless otherwise specified, the heterostructures are assumed to be under flat band conditions, i.e. contain no charges. In the presence of charges, self consistent calculations of energy levels are required. They are the subject of chapter V.

II. The envelope function model.

II.1 PRELIMINARIES. — Advanced epitaxial techniques, such as molecular beam epitaxy or metal-organic chemical vapour deposition, have made it possible to grow interfaces between two semiconductors which are flat up to one atomic monolayer (2.83 Å in GaAs), which is the ultimate resolution which can be achieved. It is common to represent such an ideal interface in terms of a continuously varying position-dependent band edge (Fig. 1). Basically speaking the sketch shown in figure 1 means that for $z > 0$ the electron experiences a one-electron potential which is identical to that of a perfect bulk A material, whereas for $z < 0$ it experiences a one-electron potential which is the same as is found in a perfect bulk B layer. Of course, this scheme that we shall follow later on is only approximately true. Its accuracy is limited by several factors, both fundamental and technological in origin. On the one hand the sketch shown in figure 1 by-passes any ambiguity concerning the description of the electron states which originates from the border atoms, i.e. that hybrid interface bonds exist which are present in neither the bulk A layer nor in the bulk B layer. On the other hand, figure 1 tacitly assumes a perfectly bi-dimensional growth, i.e. that all over the area S the structure grows monolayer after monolayer. This bi-dimensional growth is, in practice never completely achieved and for that reason the interface location along the z axis varies with the in-plane coordinates. In practice, we feel it is better to envisage the interface as having a finite thickness to account for the effects previously mentioned. Fortunately, most of the electronic states which we shall discuss in the rest of the book hardly experience the interfaces, as they have small probability amplitudes of being found in these regions. Thus, in a first

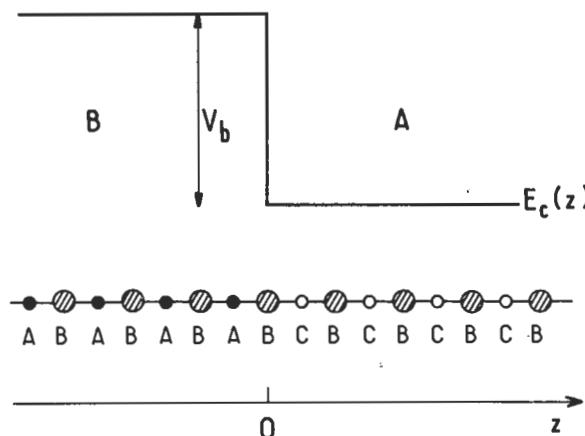


Fig. 1. — One-dimensional sketch of a heterojunction formed between two perfectly lattice-matched A and B semiconductors with chemical formulae CB and AB respectively. Upper part : representation in terms of a position-dependent conduction band edge. Lower part : actual bonds. Notice the formation of the hybrid bonds A-B-C at the interface.

approximation and for technologically abrupt interfaces, we shall retain the notion of mathematically abrupt interfaces. Consequences of the deviations in the actual interfaces with respect to this idealized model will be examined in chapter IV.

For certain heterostructures, the notion of abrupt interfaces is irrelevant from the growth point of view due, for example, to an excessive and uncontrolled interdiffusion between the A and B materials. For those heterostructures with severely graded interfaces, one should resort to modelization of the grading effect. To our knowledge, little effort has been devoted to this technologically important problem.

Since the notion of interfaces is somewhat fuzzy, the definition of layer thickness is also rather imprecise. However in the case of periodically arranged heterostructures (superlattices) the X-ray determination of the superlattice period is precise (see e.g. [10]). On the other hand, the thickness of an individual layer is rarely known to an accuracy of more than one atomic monolayer. In-situ measurements such as Reflection High Energy Electron Diffraction oscillations [11] have proved to be very valuable in ascertaining the layer thickness. Simpler techniques such as measuring the growth duration and using calibration obtained on thick ($\sim 1\mu\text{m}$) layers are less precise for the narrow layers ($\sim 100\text{ \AA}$) which interest us.

The abrupt interface model sketched in figure 1 makes a clear distinction between barrier-acting (B) and well-acting (A) materials. One may however raise questions on the significance of barrier and well denominations since we have seen that the interface notion is not very well defined. The two notions call for different properties. The interface is ill-defined because it is difficult to know at the atomic scale what kind of environment an electron sees when it is in the transition region between the A and B layers. The notion of wells and barriers calls, on the contrary, for the asymptotic behaviour of the carrier wavefunction occurring far from the

interface. The exact heterostructure wavefunction corresponding to ε_ν , where $0 \leq \varepsilon_\nu \leq V_b$ oscillates deep in the A layer but is exponentially damped in the B layer. On the other hand if $\varepsilon_\nu \geq V_b$, the wavefunction would oscillate in both kinds of layers.

With respect to the valence and conduction edges, two possible band edge profiles exist for an AB heterostructure. In figure 2a B is a barrier for both the valence and conduction electrons (type I configuration). Another situation is also found (type II, also called staggered configuration) where one material acts as a well for conduction electrons but as a barrier for valence electrons (Fig. 2b). Examples of the type I configuration are GaAs-Ga(Al)As, $\text{Ga}_{0.47}\text{Al}_{0.53}\text{As-InP}$, GaSb-AlSb etc..., whereas InAs-GaSb is an example of a heterostructure which displays a type II configuration.

II.2 THE ENVELOPE FUNCTION FRAMEWORK. — In the following we shall assume that the A and B materials constituting the heterostructures are perfectly lattice-matched and that they crystallize with the same crystallographic structure (in most cases the zinc-blende structure). The approximation of perfect lattice matching is relatively well justified for GaAs-Ga(Al)As and $\text{Ga}_{0.47}\text{In}_{0.53}\text{As-InP}$ (if the In mole

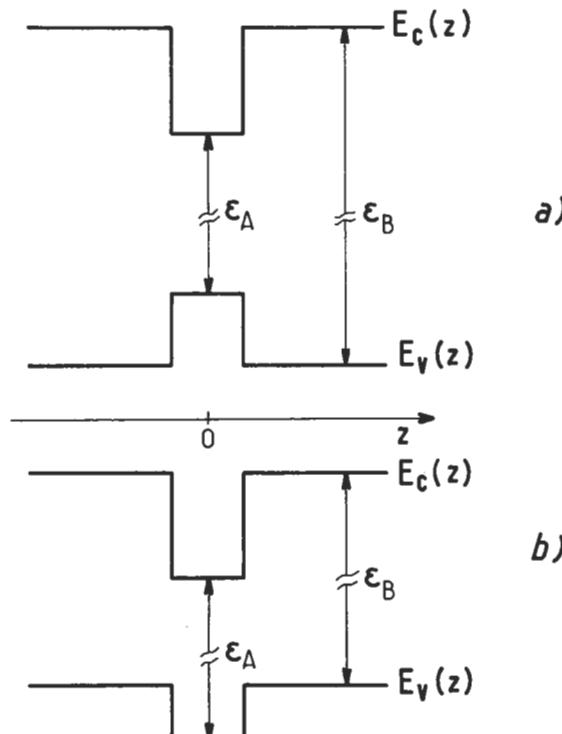


Fig. 2. — Conduction and valence band profiles in a type I heterostructure (Fig 2a) and in a type II heterostructure (Fig. 2b).

fraction is really 0.53) : the relative lattice mismatch $\delta a/a = (a_B - a_A)/a_A$ between both host materials is between 0 and 0.1 %. For other materials (e.g. GaSb-AlSb) $\delta a/a$ is too large ($\geq 0.6 \%$) to be neglected. If the layers constituting the heterostructures are thin enough, the lattice mismatch is accommodated by strain and stress effects in each kind of layer in order to achieve a common in-plane lattice parameter [12-14]. The envelope function framework can be suitably generalized to include the stress effects. We shall not discuss these effects here, as [12-14] can be referred to for such a generalization.

In the envelope function model the similarity found between the Kane matrix elements of the various III-V or II-VI compounds is exploited. Two key assumptions are made :

i) Inside each layer the wavefunction is expanded on the periodic parts of the Bloch functions of the edges under consideration

$$\psi(\mathbf{r}) = \sum_l f_l^{(A)}(\mathbf{r}) u_{l, k_0}^{(A)}(\mathbf{r}) \quad (1)$$

if \mathbf{r} corresponds to an A layer and,

$$\psi(\mathbf{r}) = \sum_l f_l^{(B)}(\mathbf{r}) u_{l, k_0}^{(B)}(\mathbf{r}) \quad (2)$$

if \mathbf{r} corresponds to a B layer. In equations (1, 2), k_0 is the point in the Brillouin zone around which the heterostructure states are built. The summation over l runs over as many edges as are included in the analysis.

ii) The periodic parts of the Bloch functions are assumed to be the same in each kind of layer which constitutes the heterostructure :

$$u_{l, k_0}^{(A)}(\mathbf{r}) \equiv u_{l, k_0}^{(B)}(\mathbf{r}). \quad (3)$$

Thus our heterostructure wavefunction will be written as

$$\psi(\mathbf{r}) = \sum_l f_l^{(A, B)}(\mathbf{r}) u_{l, k_0}(\mathbf{r}) \quad (4)$$

and our objective will be to determine $f_l^{(A, B)}(\mathbf{r})$. But let us first examine some of the implications of assumptions (i) and (ii).

The fact that we have truncated the summation over l to a finite number of band edges means that we have tacitly assumed that the heterostructure states are built with the host wavevectors $\mathbf{k}_A, \mathbf{k}_B$ which, when measured from k_0 , are small. More precisely this means that the ranges $\delta k_A = |\mathbf{k}_A - \mathbf{k}_0|$, $\delta k_B = |\mathbf{k}_B - \mathbf{k}_0|$ which are required to build $\psi(\mathbf{r})$ are such that the actual dispersion relations of the hosts are well described by the approximate $\epsilon_l^{(A)}(\mathbf{k}_A)$, $\epsilon_l^{(B)}(\mathbf{k}_B)$. The latter result from the diagonalization of the $\delta \mathbf{k} \cdot \mathbf{p}$ bulk Hamiltonians inside the restricted set of edges retained in equations (1, 2). To discover such a range, one should (in principle) compare the approximate $\epsilon_l^{(A)}(\mathbf{k}_A)$, $\epsilon_l^{(B)}(\mathbf{k}_B)$ with more exact computations. Such a

comparison has been made for the GaAs-AlAs pair in the previous chapter where the results of the Kane model, designed to be valid in the vicinity of the zone centre, were compared to a tight binding calculation [15] valid over the whole Brillouin zone. This comparison demonstrates that the conduction band states are fairly well reproduced by the Kane model in *both* GaAs and AlAs materials up to ~ 0.3 eV in the GaAs conduction band. Valence states are reproduced with comparable accuracy.

As the summations over l in equations (1, 2) are the same for both the A and B layers, it is assumed that the heterostructure state is built from \mathbf{k}_A , \mathbf{k}_B wavevectors close to the same \mathbf{k}_0 edge in each layer. To illustrate this point we shall suppose that we want to apply the envelope function model to a GaAs-AlAs quantum well grown along the [001] direction. Let us set $k_x = k_y = 0$. The model will give results for the energy levels related to either the Γ extremum ($\mathbf{k}_0 = \mathbf{0}$) or to the X extremum ($\mathbf{k}_0 = \left(0, 0, \frac{2\pi}{a_0}\right)$) *separately*. The upper lying valence states will always be Γ -like and essentially confined in GaAs. The lowest lying conduction states will be Γ -like if the GaAs well is wide enough. Some of the excited conduction states will be Γ -like whereas others will be X -like. If the GaAs well is narrow enough the lowest lying Γ states will coincide in energy with the lowest lying X state (this is because the X band has a much heavier mass along the [001] axis than the Γ mass). The envelope function scheme will predict that these two states are degenerate in energy and consequently will fail. The degeneracy is actually lifted and the two levels anticross which leads to a strong mixing between the Γ -related and X -related wavefunctions. More generally, for heterostructures which are such that the widely separated extrema of the host Brillouin zone are involved in the heterostructure wavefunctions, the envelope function model will, in essence, be unable to account for the couplings between the various valleys. To fully describe the latter, models involving the entire hosts' Brillouin zones will be required, e.g. tight-binding [6-7] or pseudo-potential methods [8-9]. On the other hand, for heterostructure states which are built from the same extrema in both types of layers, the envelope function model should work well. It has been successfully implemented for the Γ -related states of a variety of III-V heterostructures : GaAs-Ga(Al)As, Ga_{0.47}In_{0.53}As-InP, GaSb-InAs... as well as for the L -related states of PbTe-Pb(Sn)Te heterostructures (see [16] for more details). From now on we shall restrict our considerations to the Γ -related extrema.

The assumption of identical $u_{l,0}^{(A)}(\mathbf{r})$, $u_{l,0}^{(B)}(\mathbf{r})$ implies that the interband p_x matrix element $\langle S | p_x | X \rangle$ is the same in the A and B layers. Let the plane $z = z_0$ be the interface separating the A and B layers (the growth axis is assumed to be the \hat{z} axis). Since the $u_{l,0}$ are linearly independent and since $\psi(\mathbf{r})$ has to be continuous at $z = z_0$ it results that

$$f_l^{(A)}(\mathbf{r}_\perp, z_0) = f_l^{(B)}(\mathbf{r}_\perp, z_0) \quad (5)$$

where \mathbf{r}_\perp is a two-dimensional position vector. Since the lattice constants of the host layers are assumed to be the same (at least in the layer plane in order to include the strained layer materials in our analysis), the heterostructure becomes translationally

invariant in the layer plane. Thus, the f_l 's can be factorized into :

$$f_l^{(A)}(\mathbf{r}_\perp, z) = \frac{1}{\sqrt{S}} \exp(i \mathbf{k}_\perp \cdot \mathbf{r}_\perp) \chi_l^{(A)}(z) \quad (6)$$

$$f_l^{(B)}(\mathbf{r}_\perp, z) = \frac{1}{\sqrt{S}} \exp(i \mathbf{k}_\perp \cdot \mathbf{r}_\perp) \chi_l^{(B)}(z) \quad (7)$$

or in short :

$$f_l^{(A, B)}(\mathbf{r}_\perp, z) = \frac{1}{\sqrt{S}} \exp(i \mathbf{k}_\perp \cdot \mathbf{r}_\perp) \chi_l^{(A, B)}(z) \quad (8)$$

where S is the sample area and $\mathbf{k}_\perp = (k_x, k_y)$ is a bi-dimensional wavevector which is the same in the A and B layers in order to comply with the in-plane translational invariance.

Although \mathbf{k}_\perp could theoretically span the whole in-plane section of the hosts' Brillouin zone, it is in practice seldom larger than $\sim 1/10$ of its size.

We shall denote by $\varepsilon_{l0}^{(A)}$, $\varepsilon_{l0}^{(B)}$ the energies of the l^{th} band edge at the zone centre of the A and B materials respectively. We shall also assume that for all l $\chi_l^{(A, B)}(z)$ varies slowly at the scale of the hosts' unit cell.

Thus, the heterostructure wavefunction $\psi(\mathbf{r})$ is a sum of the products of rapidly varying functions : the $u_{l,0}$'s, which are periodic with the hosts' periodicity by slowly varying *envelope functions* : the f_l 's.

To lighten the notations, we write the heterostructure Hamiltonian in the form

$$H = \frac{p^2}{2m_0} + V_A(\mathbf{r}) Y_A + V_B(\mathbf{r}) Y_B, \quad (9)$$

where Y_A (Y_B) are step functions which are unity if \mathbf{r} corresponds to an A layer (to a B layer). We have

$$H u_{l0}(\mathbf{r}) = (\varepsilon_{l0}^{(A)} Y_A + \varepsilon_{l0}^{(B)} Y_B) u_{l0}(\mathbf{r}). \quad (10)$$

We now let H act upon $\psi(\mathbf{r})$, multiply by $u_{m0}^*(\mathbf{r}) \exp(-i \mathbf{k}_\perp \cdot \mathbf{r}_\perp) \chi_m^{*(A, B)}(z)$ and integrate over space. Following the same procedure as in Appendix B of chapter II we find that $\chi_l^{(A, B)}(z)$ must fulfil the set of eigenvalue equations :

$$Q^{(0)} \left(z, -i\hbar \frac{\partial}{\partial z} \right) \chi = \varepsilon \chi. \quad (11)$$

In equation (11) χ is a N dimensional column vector and Q a $N \times N$ matrix, where N is the number of bands edges which are retained in equations (1, 2). The $Q^{(0)}$ matrix elements $D_{lm}^{(0)}$ are functions of z and $\partial/\partial z$:

$$D_{lm}^{(0)} \left(z, \frac{\partial}{\partial z} \right) = \left[\varepsilon_{l,0}^{(A)} Y_A + \varepsilon_{l,0}^{(B)} Y_B + \frac{\hbar^2 k_\perp^2}{2m_0} - \frac{\hbar^2}{2m_0} \frac{\partial^2}{\partial z^2} \right] \delta_{l,m} + \frac{\hbar \mathbf{k}_\perp}{m_0} \cdot \langle l | \mathbf{p}_\perp | m \rangle - \frac{i\hbar}{m_0} \langle l | p_z | m \rangle \frac{\partial}{\partial z} \quad (12)$$

where :

$$\langle l | \mathbf{p} | m \rangle = \int_{\Omega_0} u_{l0}^* \mathbf{p} u_{m0} d^3r \quad (13)$$

and Ω_0 is the unit cell of the host layers.

By comparing equations (12, 13) with the results obtained in chapter II, we see that $\mathcal{Q}^{(0)}$ is nothing but the $\mathbf{k} \cdot \mathbf{p}$ matrix of bulk materials, except that

i) k_z is replaced by $-i \partial/\partial z$

ii) the band edges ϵ_{l0} are now position-dependent and vary in a steplike manner.

Instead of using $\epsilon_{l0}^{(A)}$, $\epsilon_{l0}^{(B)}$ we may denote by V_l the algebraic energy shift of the l^{th} band edge when going from the A to the B material and define the step functions $V_l(z)$ by

$$V_l(z) = 0 \quad \text{if } z \text{ corresponds to an A layer} \quad (14)$$

$$V_l(z) = \epsilon_l^{(B)} - \epsilon_l^{(A)} \quad \text{if } z \text{ corresponds to a B layer} \quad (15)$$

In the case where an external potential $\varphi(z)$ slowly varying at the scale of the host unit cell is superimposed on the heterostructure potential (e.g. a band bending potential arising from charges), equation (10) is modified by the adjunction of a term $\varphi(z) \mathbb{1}$, where $\mathbb{1}$ is the $N \times N$ identity matrix. Similarly, if the external potential is a slowly varying function of \mathbf{r} , equation (10) should be rewritten as

$$\left[\mathcal{Q}^{(0)} \left(\mathbf{r}, \frac{\partial}{\partial \mathbf{r}} \right) + \varphi(\mathbf{r}) \mathbb{1} \right] \mathbf{f} = \epsilon \mathbf{f}, \quad (16)$$

where :

$$D_{lm}^{(0)} \left(\mathbf{r}, \frac{\partial}{\partial \mathbf{r}} \right) = \left[\epsilon_{l0}^{(A)} + V_l(z) - \frac{\hbar^2}{2m_0} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] \delta_{l,m} - \frac{i\hbar}{m_0} \langle l | \mathbf{p} | m \rangle \cdot \frac{\partial}{\partial \mathbf{r}} \quad (17)$$

In equation (11) the larger N , the more accurate the results will be (granted that the underlying assumptions of this equation are well justified). In practice, we shall restrict N to 8, i.e. we shall study the heterostructure states which are attached to the Γ_6 , Γ_7 , Γ_8 bands of the host materials. To the extent that the other host's bands are far from the Γ_6 , Γ_7 , Γ_8 edges, their effect on the related envelope functions can either be neglected or taken into account only up to the second order in \mathbf{p} . In the latter case this amounts to replacing the matrix $\mathcal{Q}^{(0)}$ in equation (11) by the effective \mathcal{Q} given by :

$$\mathcal{Q} = \mathcal{Q}^{(0)} - \frac{\hbar^2}{2} \sum_{\alpha, \beta} \frac{\partial}{\partial r_\alpha} \underbrace{\frac{1}{M^{\alpha\beta}} \frac{\partial}{\partial r_\beta}}_{\mathcal{M}^{\alpha\beta}} \quad (18)$$

where $\alpha, \beta = x, y, z$ and $\mathcal{M}^{\alpha\beta}$ is a 8×8 matrix whose coefficients express the indirect Γ_6 , Γ_7 , Γ_8 couplings via a single excursion outside the Γ_6 , Γ_7 , Γ_8 multiplet (see

Appendix A of the previous chapter) :

$$\frac{m_0}{M_{lm}^{\alpha\beta}} = \frac{2}{m_0} \sum_{\nu} \langle l | p_{\alpha} | \nu \rangle \frac{1}{\bar{\varepsilon} - \varepsilon_{\nu 0}^{(A)} - V_{\nu}(z)} \langle l | p_{\beta} | \nu \rangle \quad (19)$$

where ν labels the remote Γ edges of the host layers, whose energies are position-dependent in the heterostructure (as expressed by the piecewise constant functions $V_{\nu}(z)$) and where $\bar{\varepsilon}$ is an average energy of the $\Gamma_6, \Gamma_7, \Gamma_8$ set in the heterostructure.

To summarize, the $\chi_l(z)$ envelope functions are the solutions of a 8×8 second order differential system :

$$\sum_{m=1}^8 \left\{ \left[\varepsilon_{m0}^{(A)} + V_m(z) + \frac{\hbar^2 k_{\perp}^2}{2m_0} - \frac{\hbar^2}{2m_0} \frac{\partial^2}{\partial z^2} \right] \delta_{lm} - \frac{i\hbar}{m_0} \langle l | p_z | m \rangle \frac{\partial}{\partial z} + \right. \\ + \frac{\hbar k_{\perp}}{m_0} \cdot \langle l | p_{\perp} | m \rangle - \frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{M_{lm}^{zz}} \frac{\partial}{\partial z} - \frac{i\hbar^2}{2} \sum_{\alpha=x, y} \left[k_{\alpha} \frac{1}{M_{lm}^{\alpha z}} \frac{\partial}{\partial z} + \frac{\partial}{\partial z} \frac{1}{M_{lm}^{z\alpha}} k_{\alpha} \right] + \\ \left. + \frac{\hbar^2}{2} \sum_{\alpha, \beta=x, y} k_{\alpha} \frac{1}{M_{lm}^{\alpha\beta}} k_{\beta} \right\} \chi_m = \varepsilon \chi_l ; \quad 1 \leq l \leq 8 . \quad (20)$$

As can be seen from equation (20) the problem of the heterostructure energy levels has been reduced to the solution of a set of second order differential equations which govern the spatial behaviour of the slowly varying envelope functions. The microscopic details of the heterostructures, i.e. all the parameters which depend on the rapidly varying phenomena at the scale of the hosts' unit cells, have *explicitly* disappeared from the equations of motion. In equation (20) they only survive through *effective* parameters : the interband matrix elements $\langle l | p | m \rangle$, the effective-mass tensor $M_{lm}^{\alpha\beta}$ and the band offsets V_l . Either these parameters are *a priori* known and equation (20) enables the heterostructure energy levels to be determined, or the band offsets V_l are unknown (as is often the case) and, by comparing the measurements and the energy levels deduced from equation (20), the V_l 's being treated as adjustable parameters, these band offsets can be determined. First principle or semi-empirical calculations of the band offsets are available [see e.g. 17-19] and for a given A-B pair the V_l 's are determined. In particular, the common anion rule derived by Harrison [17], states that two semiconductors which have a common anion (e.g. GaAs, AlAs), the same crystallographic structure and are lattice matched should nearly have a zero valence band (V_p) offset. This rule however is accurate only up to several tenths of an eV, as any other first principle determination of the V_l 's. This is *insufficient* for most practical purposes and one is forced to consider the V_l 's as adjustable parameters.

Knowledge of the V_l 's is of paramount importance in the heterostructures, as the band offsets decisively influence their energy levels and therefore their electronic properties (see Fig. 3 for an illustration). One of the first tasks when dealing with a given heterostructure is to design experiments which would enable the V_l 's to be determined. The reader is referred to [20, 21] for a thorough discussion of the GaAs-AlAs band offsets.

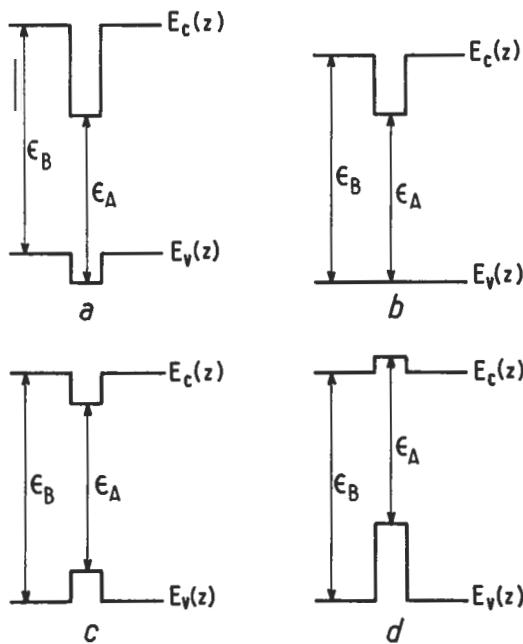


Fig. 3. — Illustration of the part played by different apportionments between the valence and conduction bands of the bandgap energy differences $\epsilon_B - \epsilon_A$ of two semiconductors A and B on the electronic states of a BAB rectangular quantum well a) electrons are confined in the A layer, holes in the B layer (type II quantum well) ; b) same as a) except that the A layer is no longer a barrier for holes ; c) both electrons and holes are essentially confined in the A layer (type I quantum well) ; d) the A layer is a barrier for electrons and a well for holes. This structure is thus a type II quantum well, inverted with respect to a). Out of these four structures only c) will display significant excitonic absorption, photoluminescence and stimulated emission. Structure c) is a good candidate for low threshold lasing action.

At the hetero-interface $z = z_0$ we saw that χ is continuous. If we now integrate equation (20) across the interface, we obtain a second set of boundary conditions. These can be cast in the form :

$$\mathcal{A}^{(A)} \chi^{(A)}(z = z_0) = \mathcal{A}^{(B)} \chi^{(B)}(z = z_0) \quad (21)$$

where \mathcal{A} is a 8×8 matrix whose elements are, at most, of the first order in $\frac{\partial}{\partial z}$:

$$A_{lm} = \frac{-\hbar^2}{2m_0} \left[\left(\delta_{lm} + \frac{m_0}{M_l^{zz}} \right) \frac{\partial}{\partial z} + \frac{2i}{\hbar} \langle l | p_z | m \rangle + i \sum_{\alpha=x,y} \frac{m_0}{M_{lm}^{z\alpha}} k_{\alpha} \right] \quad (22)$$

where we have made use [22-24] of the property

$$M_{lm}^{zz} = M_l^{zz} \delta_{lm}. \quad (23)$$

It is worth pointing out that A_{lm} is *not* diagonal in (l, m) and, *a fortiori*, does not reduce to $-\frac{\hbar^2}{2m_0} \delta_{lm} \frac{\partial}{\partial z}$. Although familiar, the continuity condition of $\frac{d\chi}{dz}$ is only valid in very special circumstances. In fact it is the limit of equation (22) in the empty lattice approximation, where all the u_{l0} 's and $u_{\alpha 0}$'s are constant, which leads to vanishing $(M_{lm}^{\alpha\beta})^{-1}$ and $\langle l | p_z | m \rangle$. We show in Appendix A that equation (22), when applied to a heterostructure whose host layers are both characterized by non degenerate parabolic and isotropic bands but by different effective masses (Ben Daniel-Duke model [25]), is compatible with the conservation of the probability current across the heterostructure and thus with the stationarity of the heterostructure state. On the other hand, the continuity condition of $\frac{d\chi}{dz}$ is incompatible with these conservation laws and should therefore be rejected.

In the following we shall apply the envelope function machinery to various examples. We shall deal either with single quantum wells or with superlattices. The latter heterostructures are obtained by infinitely repeating a sequence of two adjacent A and B layers. The A(B) thickness will be denoted by L_A (L_B) and the superlattice period by d ($d = L_A + L_B$). The growth axis will be taken along the \hat{z} direction and the first Brillouin zone of the superlattice will correspond to the superlattice wavevector q such that

$$-\frac{\pi}{d} < q \leq +\frac{\pi}{d} \quad (24)$$

Although equations [5, 21, 22] tell us how to match χ and $\frac{d\chi}{dz}$ across the interfaces, we need to know the asymptotic behaviour of χ at large z to complete the determination of the eigensolutions. This asymptotic behaviour depends on the heterostructure under consideration. For superlattices, the band edge profile is periodic upon z with a periodicity d . Thus χ is a Bloch wave

$$\chi(z + d) = \exp(iqd) \chi(z) \quad (25)$$

For the bound states of a quantum well, χ should tend to zero at large (z) :

$$\lim_{|z| \rightarrow \infty} \chi(z) = 0. \quad (26)$$

We shall first discuss the Ben Daniel-Duke model, then the $\mathbf{k}_\perp = \mathbf{0}$ – quantum well levels and the superlattice dispersion relations of heterostructures whose host energy levels are well described by the Kane model. The in-plane dispersion relations will finally be considered in section III.

II.3 THE BEN DANIEL-DUKE MODEL. — This model of heterostructure energy levels is the simplest one and works qualitatively for the lowest conduction states of GaAs-Ga(Al)As heterostructures with GaAs layer thickness larger than ~ 100 Å and for the heavy hole levels at $\mathbf{k}_\perp = \mathbf{0}$ in any heterostructure. It amounts to assuming that the heterostructure envelope function is built from host quantum states which

belong to a single parabolic band. For simplicity we shall take an isotropic band and, for definiteness, a conduction band. The effective masses in the A (B) layers will be denoted by m_A (m_B). Each of the levels is twice degenerate (Kramers degeneracy). Equation (20) simplifies considerably as all the host bands are remote for the conduction edge under consideration. Thus, we can write

$$\left[\varepsilon_s + V_s(z) - \frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{\mu(z)} \frac{\partial}{\partial z} + \frac{\hbar^2 k_\perp^2}{2\mu(z)} \right] \chi(z) = \varepsilon \chi(z) \quad (27)$$

$$\mu^{-1}(z) = m_0^{-1} + (M_{ss}^{zz})^{-1} \quad (28)$$

$$\mu(z) = \begin{cases} m_A & \text{if } z \text{ corresponds to an A layer} \\ m_B & \text{if } z \text{ corresponds to a B layer} \end{cases} \quad (29)$$

$$V_s(z) = \begin{cases} 0 & \text{if } z \text{ corresponds to an A layer} \\ V_s & \text{if } z \text{ corresponds to a B layer} \end{cases} \quad (30)$$

and V_s is the algebraic energy shift of the S band edge when going from the A to the B material. The boundary conditions at the A-B interfaces are also very simple. They are such that :

$$\chi(z) \quad \text{and} \quad \frac{1}{\mu(z)} \frac{d\chi}{dz} \quad (31)$$

are both continuous. It should be noticed that the effective mass mismatch contributes to the total confining barrier by a term which is k_\perp -dependent and, like $V_s(z)$, exhibits step-like variations. This extra term is however small in most instances (e.g. conduction states in GaAs-Ga_{1-x}Al_xAs, Ga_{0.47}In_{0.53}As-InP quantum wells) although it leads to a decreasing effective barrier height with increasing k_\perp if $m_B > m_A > 0$.

It is also interesting to notice that the effective mass mismatch leads to a *discontinuity in the derivative* of the envelope function at the interfaces. In the extreme case where m_A and m_B are of opposite signs, this discontinuity causes a cusp at the interfaces of the envelope function. The latter situation occurs in HgTe-CdTe heterostructures (but only at $k_\perp = 0$) (see section II 3.2).

II.3.1 The Ben Daniel-Duke quantum wells ($m_A m_B > 0$). — The k_\perp -dependent potential energy $V_s(z) + \frac{\hbar^2 k_\perp^2}{2\mu(z)}$ is even with respect to the middle of the A layer. Thus, as in chapter I, one can look for bound states solutions in the following forms :

$$\begin{aligned} \chi_{\text{even}}(z) &= A \cos(k_A z) & |z| &\leq \frac{L_A}{2} \\ \chi_{\text{even}}(z) &= B \exp\left[-\kappa_B\left(z - \frac{L_A}{2}\right)\right] & z &\geq \frac{L_A}{2} \\ \chi_{\text{even}}(-z) &= \chi_{\text{even}}(z) \end{aligned} \quad (32)$$

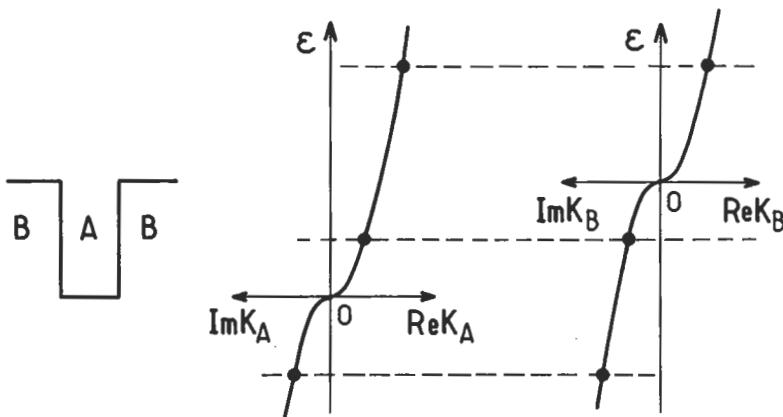


Fig. 4. — Dispersion relations *versus* the real and imaginary wavevectors in the A and B layers stacked to form a BAB quantum wells. The three dashed lines, drawn for three energies in the heterostructure and for $k_{\perp} = 0$, show which wavevectors in each kind of layers participate in the heterostructure state. The upper line corresponds to a delocalized quantum well state and the middle line to a quantum well bound state. No heterostructure state can be associated with the lower line as the carrier effective masses are assumed to have the same sign in this particular Ben Daniel-Duke quantum well.

or :

$$\begin{aligned}
 \chi_{\text{odd}}(z) &= A \sin(k_A z) & |z| &\leq \frac{L_A}{2} \\
 \chi_{\text{odd}}(z) &= B \exp\left[-\kappa_B \left(z - \frac{L_A}{2}\right)\right] & z &\geq \frac{L_A}{2} \\
 \chi_{\text{odd}}(-z) &= -\chi_{\text{odd}}(z)
 \end{aligned} \tag{33}$$

with :

$$\varepsilon - \varepsilon_s = \frac{\hbar^2 k_A^2}{2m_A} + \frac{\hbar^2 k_{\perp}^2}{2m_A} = V_s - \frac{\hbar^2 \kappa_B^2}{2m_B} + \frac{\hbar^2 k_{\perp}^2}{2m_B}. \tag{34}$$

Equations (32-34) hold if $\kappa_B^2 > 0$, i.e. if the heterostructure state is built from the evanescent states of the B layers (see Fig. 4).

By matching $\chi(z)$ and $\mu^{-1}(z) \frac{d\chi}{dz}$ at the interface $z = \frac{L_A}{2}$ (or equivalently $z = -\frac{L_A}{2}$), one obtains the implicit equations whose roots are the bound solutions of the Ben Daniel-Duke quantum well problem. These are :

$$\cos \varphi_A - \frac{m_B}{m_A} \frac{k_A}{\kappa_B} \sin \varphi_A = 0 \quad \text{for even states} \tag{35}$$

$$\cos \varphi_A + \frac{m_A \kappa_B}{m_B k_A} \sin \varphi_A = 0 \quad \text{for odd states} \quad (36)$$

$$\varphi_A = \frac{1}{2} k_A L_A \quad (37)$$

A comparison with the results obtained in chapter I shows that equations (35-37) are the same as those of text book quantum wells except that the wavevectors k_A , κ_B have been replaced by k_A/m_A and κ_B/m_B . This is a direct consequence of the matching conditions of $\mu^{-1}(z) \frac{d\chi}{dz}$ at the interfaces. In all other instances however, many of the results obtained in chapter I can be applied to equations (35-37). In particular, the number of states bound by the well (at $k_\perp = 0$) is equal to

$$\mathcal{N} = 1 + \text{Int} \left[\left(\frac{2m_A}{\hbar^2 \pi^2} V_s L_A^2 \right)^{1/2} \right] \quad (38)$$

In figure 5 are shown the variations of the confinement energies E_1 , E_2 , E_3 at $k_\perp = 0$ of a Ben Daniel-Duke quantum well ($L_A = 100 \text{ \AA}$, $V_s = 0.3 \text{ eV}$,

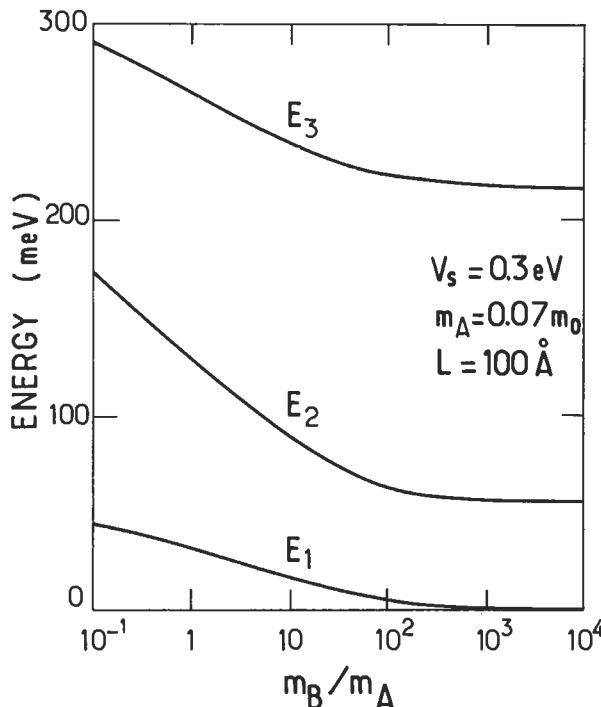


Fig. 5. — Evolution of the confinement energies E_1 , E_2 , E_3 with the mass ratio $\frac{m_B}{m_A}$ in a Ben Daniel-Duke quantum well. $L = 100 \text{ \AA}$; $V_s = 0.3 \text{ eV}$; $m_A = 0.07 m_0$.

$m_A = 0.07 m_0$) with the mass ratio $\frac{m_B}{m_A}$. It can be seen that all the E_i 's decrease with increasing $\frac{m_B}{m_A}$ and tend to values E_{p+1}^∞ which are such that :

$$k_A (E_{p+1}^\infty) L_A = p\pi \quad p = 0, 1, 2, \dots \quad (39)$$

This equation resembles the bound states equation in a quantum well with an infinite barrier height. It expresses however quite a different physical situation. In a quantum well with infinite V_s , the envelope function vanishes at the interface. Besides, according to equation (38) \mathcal{N} is infinite. If V_s diverges, κ_B also diverges. This leads either to $\cos \varphi_A = 0$ (Eq. (35)) or $\sin \varphi_A = 0$ (Eq. (36)). The ground state solution, which is nodeless, fulfills $L_A k_A (V_s = \infty) = \pi$, the associated envelope function having a finite slope at $z = \pm \frac{L_A}{2}$. On the other hand, when V_s is kept fixed but m_B/m_A increases to infinity, \mathcal{N} remains unchanged. Moreover, if the ground states envelope function is to be nodeless, it has to be a cosine in the well. It should barely penetrate the barrier (since κ_B increases) and, in addition, should have a derivative at $z = \pm \frac{L_A}{2}$ whose modulus becomes smaller and smaller to comply with the continuity of $\mu^{-1}(z) \frac{d\chi}{dz}$ at the interface. Thus at infinite m_B/m_A ratio the only possible wave function is constant in the well and zero in the barrier, so that both the envelope function and its derivative are also zero in the barrier. The only $\cos(k_A z)$ function which is constant corresponds to $k_A = 0$ and thus to $E_1 = 0$... This is why when m_B/m_A diverges and V_s is fixed the even states of the well fulfill $\sin \varphi_A = 0$ (and not $\cos \varphi_A = 0$), i.e. admit $E_1 = 0$ as an acceptable solution.

Symmetrically, the odd states fulfill $\cos \varphi_A = 0$ and finally the series of levels (Eq. (39)) is recovered. The ground state envelope function of a quantum well with either infinite V_s or finite V_s but infinite m_B is shown in figure 6 in order to depict the differences between the two physical situations.

When the effective masses m_A and m_B are not widely different, as for instance in GaAs-Ga(Al)As heterostructures, the in-plane dispersions of the subbands attached to the $k_\perp = 0$ bound states of a quantum well are nearly parabolic in k_\perp :

$$E_n (k_\perp) \approx \epsilon_s + E_n(0) + \frac{\hbar^2 k_\perp^2}{2m_n} \quad (40)$$

The in-plane mass m_n should, in principle, be obtained by numerically solving equations (32-37). However when k_\perp is small enough, an approximate scheme can be designed in the following way. The term $\hbar^2 k_\perp^2 / 2\mu(z)$ in equation (27) is formally rewritten :

$$\frac{\hbar^2 k_\perp^2}{2\mu(z)} = \frac{\hbar^2 k_\perp^2}{2m_n} + \frac{\hbar^2 k_\perp^2}{2} \left[\frac{1}{\mu(z)} - \frac{1}{m_n} \right] \quad (41)$$

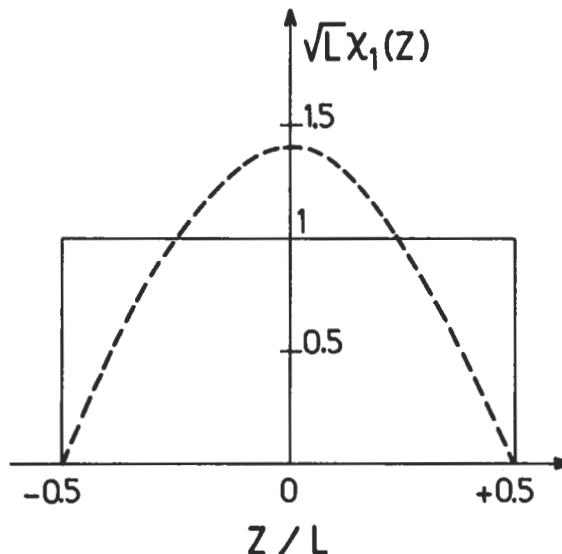


Fig. 6. — Ground state envelope functions for a quantum well with infinite V_s (dashed line) or for a quantum well with finite V_s but infinite $\frac{m_B}{m_A}$ (solid line).

and the second term on the right-hand side of equation (41) is considered as a perturbation to \mathcal{H}_0 where

$$\mathcal{H}_0 = \varepsilon_s + V_s(z) - \frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{\mu(z)} \frac{\partial}{\partial z} + \frac{\hbar^2 k_\perp^2}{2m_n}, \quad (42)$$

whose eigenstates are in the form given by equation (40). The first order corrections to these eigenstates are given by :

$$\Delta E_n = \frac{\hbar^2 k_\perp^2}{2} \left[\frac{1}{m_A} [1 - P_b(E_n)] + \frac{1}{m_B} P_b(E_n) - \frac{1}{m_n} \right] \quad (43)$$

where :

$$P_b(E_n) = 2 \int_{\frac{L_A}{2}}^{\infty} \chi_n^2(z) dz = \frac{B^2}{\kappa_B} \quad (44)$$

is the integrated probability of finding the electron in the barriers while in the E_n state. The first order energy shift will vanish if :

$$\frac{1}{m_n} = \frac{1}{m_A} [1 - P_b(E_n)] + \frac{1}{m_B} P_b(E_n). \quad (45)$$

Equation (45) defines the in-plane effective mass of the n^{th} subband in the vicinity of $k_{\perp} = 0$. It may be remarked that if $m_B > m_A$, as is the case in GaAs-Ga(Al)As or $\text{Ga}_{0.47}\text{In}_{0.53}\text{As-InP}$, this in-plane mass m_n will increase with increasing subband index n .

Using the approximately parabolic in-plane dispersion laws (equation (40)) it is very easy to calculate the density of states $\rho(\varepsilon)$ associated with the bound states E_n . Proceeding exactly as in chapter I we obtain :

$$\rho(\varepsilon) = \sum_n \rho_n(\varepsilon) \quad (46a)$$

$$\rho_n(\varepsilon) = \frac{m_n S}{\pi \hbar^2} Y(\varepsilon - E_n) \quad (46b)$$

where $Y(x)$ is the step function. We recover the familiar staircase density of states. The properties of a Ben Daniel-Duke quantum well are summarized in figure 7.

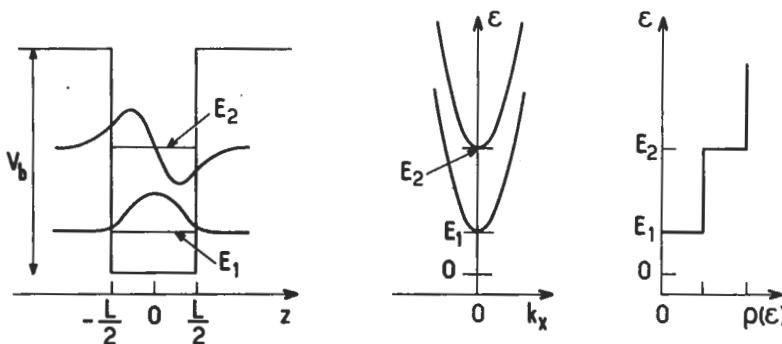


Fig. 7. — A recollection of the main properties of the quantum well bound states, solutions of a Ben Daniel-Duke Hamiltonian. From left to right : conduction band edge profile, energy levels E_1 and E_2 and their associated envelope functions ; in-plane dispersion relations of the E_1 and E_2 subbands ; energy dependence of the heterostructure density of states $\rho(\varepsilon)$.

II.3.2 Interface states of Ben Daniel-Duke quantum wells ($m_A m_B < 0$; $\mathbf{k}_{\perp} = \mathbf{0}$). — The case $m_A m_B < 0$ is practically realized in HgTe-CdTe heterostructures [26] (see Fig. 8). CdTe is a conventional open gap semiconductor whose level ordering is the same as is found in GaAs. HgTe is a symmetry-induced zero gap semiconductor. The Γ_6 band, which is a conduction band in most III-V and II-VI semiconductors, is a light hole band in HgTe. The Γ_6 edge lies ~ 0.3 eV below the Γ_8 edges. As the Γ_8 light band and Γ_6 band are nearly mirror-like, the Γ_8 light band is a conduction band in HgTe, degenerate at the zone centre with the Γ_8 heavy hole band (inversion asymmetry splitting having been neglected).

Ignoring the absence of centro-symmetry of the zinc-blende lattice, we shall see in section (II.4) that the light particle and heavy hole states decouple at $\mathbf{k}_{\perp} = \mathbf{0}$. We can

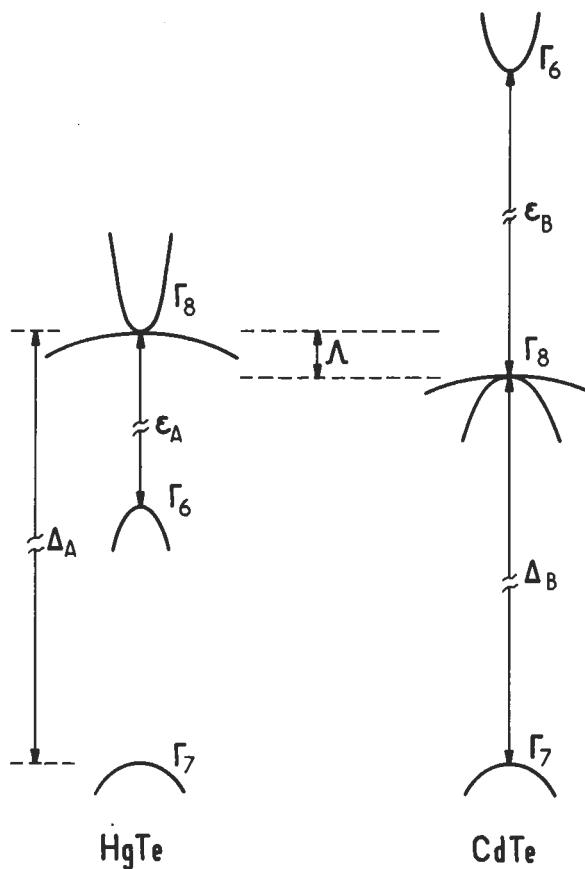


Fig. 8. — Band structures of bulk HgTe (left panel) and CdTe (right panel) in the vicinity of the Γ point (schematic).

thus treat the problem of the light particle states associated with a Γ_8 edge as if we were considering a single band. The interesting feature of the HgTe-CdTe heterostructure is that the light particle changes the sign of its effective mass across the interfaces, being electron-like in the HgTe layer and light hole-like in the CdTe layers. To be specific, let us consider a CdTe-HgTe-CdTe double heterostructure.

According to [27] the bottom of the HgTe Γ_8 conduction band lies at an energy $\Lambda \sim 40$ meV above the top of the CdTe Γ_8 valence band. Thus, bound states of the heterostructure only exist if $\epsilon \geq -\Lambda$ (the energy zero being taken at the Γ_8 edge in HgTe). If $-\Lambda \leq \epsilon \leq 0$, the states are evanescent in both kinds of layers while if $\epsilon \geq 0$, the carrier wavevector is real (imaginary) in the HgTe (CdTe) layers. Clearly, bound states of positive energies will exist (an infinite number in the one-band description of each host layer). Proceeding as in section II 3.1 their energies will

fulfil

$$\cos \varphi_A + \frac{|m_B|}{m_A} \frac{k_A}{\kappa_B} \sin \varphi_A = 0 \quad \text{for even states} \quad (47)$$

$$\cos \varphi_A - \frac{m_A}{|m_B|} \frac{\kappa_B}{k_A} \sin \varphi_A = 0 \quad \text{for odd states} \quad (48)$$

$$\varphi_A = \frac{1}{2} k_A L_A \quad (49)$$

$$k_A = \sqrt{\frac{2m_A}{\hbar^2} \varepsilon}; \quad \kappa_B = \sqrt{\frac{2|m_B|}{\hbar^2} (\varepsilon + \Lambda)} \quad (50)$$

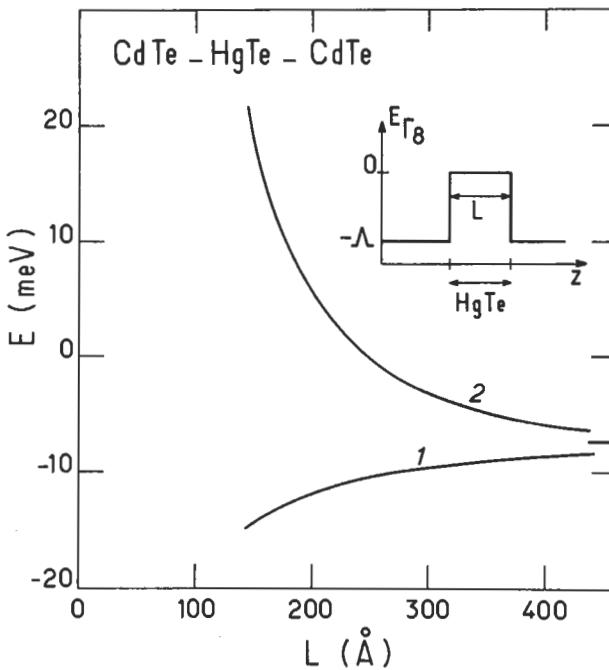


Fig. 9. — Evolution of the ground and first excited bound states (labelled 1 and 2 respectively) versus the HgTe slab thickness in a CdTe-HgTe-CdTe double heterostructure.

The bound state wavefunctions are all characterized by *cusps* at the interfaces due to the change in the carrier effective mass at the hetero-interfaces. This sign reversal also implies that equation (48) can be fulfilled at $\varepsilon = 0$ for a certain L_A while equation (47) can not. This means that at least one state (even in z) should lie below the bottom of the HgTe conduction band edge. This state is an interface level, built from evanescent states in *each* of the host layers, whose wavefunction peaks at the

interface. More precisely, we can write :

$$\chi_1(z) = A \cosh(\kappa_A z) \quad |z| \leq \frac{1}{2} L_A \quad (51)$$

$$\chi_1(z) = B \exp\left[-\kappa_B\left(z - \frac{1}{2} L_A\right)\right] \quad z \geq \frac{1}{2} L_A \quad (52)$$

$$\chi_1(-z) = \chi_1(z) \quad (53)$$

with : $\kappa_A = \sqrt{\frac{2m_A}{\hbar^2}(-\varepsilon)}; \quad \kappa_B = \sqrt{\frac{2|m_B|}{\hbar^2}(\varepsilon + \Lambda)}.$ (54)

By matching $\chi_1(z)$ and $\mu^{-1}(z) \frac{d\chi_1}{dz}$ at $z = \frac{1}{2} L_A$, we find that ε should be the root of the implicit equation

$$\tanh\left(\frac{1}{2}\kappa_A L_A\right) = \frac{m_A}{|m_B|} \frac{\kappa_B}{\kappa_A} \quad (55)$$

It is very easy to check that equation (55) always admits one solution E_1 (and only one) which extrapolates to $-\Lambda$ when $L_A \rightarrow 0$. A second state may actually exist in the energy segment $[-\Lambda, 0]$ if the HgTe layer is thick enough. It corresponds to an odd envelope function :

$$\chi_2(z) = A \sinh(\kappa_A z) \quad ; \quad |z| \leq \frac{1}{2} L_A \quad (56)$$

$$\chi_2(z) = B \exp\left[-\kappa_B\left(z - \frac{1}{2} L_A\right)\right]; \quad z \geq \frac{1}{2} L_A \quad (57)$$

$$\chi_2(-z) = -\chi_2(z) \quad (58)$$

The E_2 energy is the solution of the implicit equation :

$$\coth\left(\frac{1}{2}\kappa_A L_A\right) = \frac{m_A}{|m_B|} \frac{\kappa_B}{\kappa_A} \quad (59)$$

which admits a solution if

$$L_A > \frac{2|m_B|}{m_A} \sqrt{\frac{\hbar^2}{2|m_B| \Lambda}} \quad (60)$$

Again, the solution of equation (59), if it exists, is unique. When L_A becomes very large the energies E_1 and E_2 converge to the value :

$$E_\infty = -\frac{\Lambda}{1 + \frac{|m_B|}{m_A}} \quad (61)$$

which is the energy position of the interface state in a single HgTe-CdTe heterojunction [28, 29]. Clearly, at large L_A (i.e. $\kappa_A L_A \gg 1$) the two states E_1 and E_2 are very well approximated by the symmetric and antisymmetric combinations of the two interface states centred at $\pm \frac{1}{2} L_A$ respectively. The

behaviour of E_1 and E_2 versus L_A is presented in figure 9 to illustrate the previous discussion. In figure 10 we show the calculated $\chi_1(z)$ envelope functions in $\text{Hg}_{1-x}\text{Cd}_x\text{Te}-\text{HgTe}-\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ quantum wells to illustrate the interface nature of the E_1 state. Although the existence of the interface state relies only on the relative position of the Γ_8 edges of HgTe and CdTe, their actual energy position, as well as their behaviour at $\mathbf{k}_\perp \neq \mathbf{0}$ (where they strongly couple to the heavy hole states), remains a subject of active research.

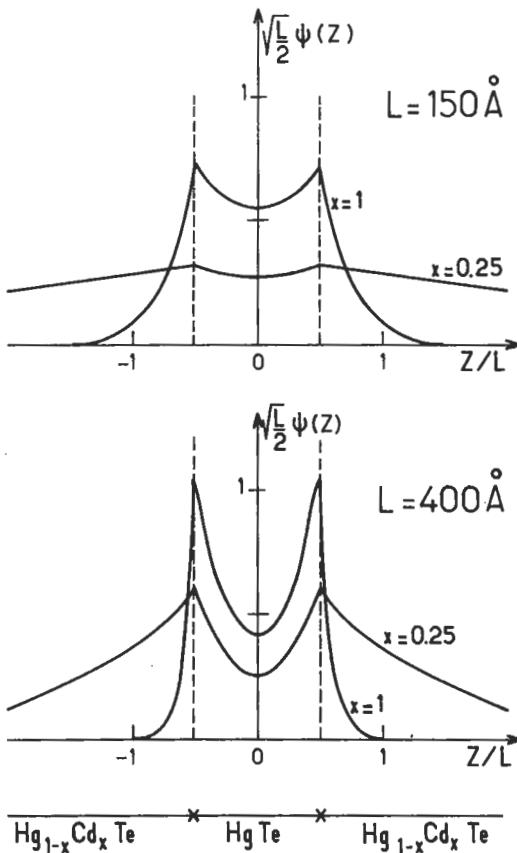


Fig. 10. — Dimensionless envelope functions of the ground states in $\text{Hg}_{1-x}\text{Cd}_x\text{Te}-\text{HgTe}-\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ double heterostructures ($x = 1$ and $x = 0.2$) for two different HgTe slab thicknesses.

II.4 QUANTUM WELLS AND SUPERLATTICES WITH HOSTS WHICH DISPLAY KANE-LIKE BANDS. — This situation is found in most of the III-V and II-VI heterostructures, e.g. GaAs-Ga(Al)As, $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}-\text{InP}$, HgTe-CdTe, $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}-\text{Al}_{0.48}\text{In}_{0.52}\text{As}$ etc... At $\mathbf{k}_\perp = \mathbf{0}$ the 8×8 differential system defined in equation (20) is block-diagonal if the u_{10} basis diagonalizes both the total angular momentum

\mathbf{J} and its z projection J_z . Specifically we shall take

$$\begin{aligned}
 u_{10} &= \left| S, \frac{1}{2}, \frac{1}{2} \right\rangle = i |S \uparrow\rangle \\
 u_{30} &= \left| P, \frac{3}{2}, \frac{3}{2} \right\rangle = \frac{1}{\sqrt{2}} |(X + iY) \uparrow\rangle \\
 u_{50} &= \left| P, \frac{3}{2}, \frac{1}{2} \right\rangle = -\sqrt{\frac{2}{3}} |Z \uparrow\rangle + \frac{1}{\sqrt{6}} |(X + iY) \downarrow\rangle \\
 u_{70} &= \left| P, \frac{1}{2}, \frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} |(X + iY) \downarrow\rangle + \frac{1}{\sqrt{3}} |Z \uparrow\rangle \\
 u_{20} &= \left| S, \frac{1}{2}, -\frac{1}{2} \right\rangle = i |S \downarrow\rangle \\
 u_{40} &= \left| P, \frac{3}{2}, -\frac{3}{2} \right\rangle = \frac{1}{\sqrt{2}} |(X - iY) \downarrow\rangle \\
 u_{60} &= \left| P, \frac{3}{2}, -\frac{1}{2} \right\rangle = -\frac{1}{\sqrt{6}} |(X - iY) \uparrow\rangle - \sqrt{\frac{2}{3}} |Z \downarrow\rangle \\
 u_{80} &= \left| P, \frac{1}{2}, -\frac{1}{2} \right\rangle = -\frac{1}{\sqrt{3}} |(X - iY) \uparrow\rangle + \frac{1}{\sqrt{3}} |Z \downarrow\rangle
 \end{aligned} \tag{62}$$

The differential system $\mathcal{D}\chi = \varepsilon\chi$ factorizes into :

$$\begin{bmatrix} \mathcal{D}_+ & 0 \\ 0 & \mathcal{D}_- \end{bmatrix} \begin{bmatrix} \chi_+ \\ \chi_- \end{bmatrix} = \varepsilon \begin{bmatrix} \chi_+ \\ \chi_- \end{bmatrix} \tag{63}$$

where \mathcal{D}_+ , \mathcal{D}_- , χ_+ , χ_- are 4×4 matrices and 4×1 column vectors respectively. \mathcal{D}_+ and \mathcal{D}_- are identical and therefore each eigenenergy will be twice degenerate. \mathcal{D}_+ (and \mathcal{D}_-) are equal to :

$$D_+ \left(z, p_z = -i\hbar \frac{\partial}{\partial z} \right) =$$

$$\begin{bmatrix}
 V_s(z) + \frac{1}{2m_0} p_z F p_z & 0 & -\sqrt{\frac{2}{3}} P p_z & \frac{P}{\sqrt{3}} p_z \\
 0 & -\varepsilon_A + V_p(z) & 0 & 0 \\
 -\sqrt{\frac{2}{3}} P p_z & 0 & -\varepsilon_A + V_p(z) & \frac{\sqrt{2}}{m_0} p_z \gamma_2 p_z \\
 \frac{P}{\sqrt{3}} p_z & 0 & \frac{\sqrt{2}}{m_0} p_z \gamma_2 p_z & -\varepsilon_A - \Delta_A + V_\delta(z) - \frac{1}{2m_0} p_z \gamma_1 p_z
 \end{bmatrix} \tag{64}$$

where $V_p(z)$ and $V_s(z)$ are the step functions which describe the algebraic shifts of the Γ_8 and Γ_7 edges when going from the A to the B materials and :

$$P = \frac{-i}{m_0} \langle S | p_x | X \rangle = \frac{-i}{m_0} \langle S | p_y | Y \rangle = \frac{-i}{m_0} \langle S | p_z | Z \rangle \quad (64a)$$

In equation (64) we have included the free electron term in the definition of F , γ_1 , γ_2 :

$$F = \frac{2}{m_0} \sum_{\nu} \langle S | p_x | \nu \rangle \frac{1}{\bar{\varepsilon} - \varepsilon_{\nu 0}^{(A)} - V_{\nu}(z)} \langle \nu | p_x | S \rangle + 1 \quad (64b)$$

$$\gamma_1 = \frac{-2}{3m_0} \sum_{\nu} \frac{\langle X | p_x | \nu \rangle \langle \nu | p_x | X \rangle + 2 \langle X | p_y | \nu \rangle \langle \nu | p_y | X \rangle}{\bar{\varepsilon} - \varepsilon_{\nu 0}^{(A)} - V_{\nu}(z)} - 1 \quad (64c)$$

$$\gamma_2 = \frac{-1}{3m_0} \sum_{\nu} \frac{\langle X | p_x | \nu \rangle \langle \nu | p_x | X \rangle - \langle X | p_y | \nu \rangle \langle \nu | p_y | X \rangle}{\bar{\varepsilon} - \varepsilon_{\nu 0}^{(A)} - V_{\nu}(z)} \quad (64d)$$

The explicit z dependences of F , γ_1 , γ_2 (via the $V_{\nu}(z)$ functions) has led us to use combinations like $p_z F p_z$, $p_z \gamma_1 p_z$, $p_z \gamma_2 p_z$ in equation (64) to ensure the hermiticity of D_+ .

It is important to notice that the $\left| P, \frac{3}{2}, \pm \frac{3}{2} \right\rangle$ lines are uncoupled to the others in equation (64). The associated envelope functions correspond to the heavy hole states which can therefore be treated separately from the other light particle states. The latter are hybrids of Γ_6 , Γ_8 light and Γ_7 host states which are admixed by the non-parabolicity of the host materials and the $V_s(z)$, $V_p(z)$, $V_s(z)$ functions (see Appendix B of the previous chapter). Notice however that the decoupling between heavy holes and light particles is only approximate as, even at $\mathbf{k}_\perp = 0$, the effect of the non centro-symmetry of the host zinc blende lattices in equation (64) has been neglected. This effect is usually very small (smaller in III-V than in II-VI compounds) and may eventually be treated in perturbation (see Appendix B). Its main effect is to prevent any crossing between light and heavy particle states. The latter would occur, for instance, in a quantum well when the confinement energy of a heavy hole state HH_m equals that of a light hole state LH_n for certain thicknesses of the well-acting material. Notice that the lack of centro-symmetry appears automatically in tight-binding calculations [6] where the difference between the anion and the cation in the host's unit cell is naturally taken into account.

Across a A-B interface, the boundary conditions are such that $\chi_+(\chi_-)$ and $\mathcal{A}_+(\mathcal{A}_-)$ are continuous where \mathcal{A}_+ equals \mathcal{A}_- and

$$\mathcal{A}_+ \left(z, \frac{\partial}{\partial z} \right) = \begin{bmatrix} F \frac{\partial}{\partial z} & 0 & 0 & 0 \\ 0 & -(\gamma_1 - 2\gamma_2) \frac{\partial}{\partial z} & 0 & 0 \\ 0 & 0 & -(\gamma_1 + 2\gamma_2) \frac{\partial}{\partial z} & 2\sqrt{2} \gamma_2 \frac{\partial}{\partial z} \\ 0 & 0 & 2\sqrt{2} \gamma_2 \frac{\partial}{\partial z} & -\gamma_1 \frac{\partial}{\partial z} \end{bmatrix} \quad (65)$$

where we have made use of the χ_+ continuity to simplify \mathcal{A}_+ .

As expected, \mathcal{A}_+ only involves *small* terms, which arise either from the free electron contribution or from the $\mathbf{k} \cdot \mathbf{p}$ interaction between Γ_6 , Γ_7 , Γ_8 and the remote edges. This interaction is very important for heavy hole states since the heavy hole bands would be dispersionless in the hosts if this interaction was neglected. On the other hand, for light particle states, the F , γ_1 , γ_2 parameters contribute very little to the host effective mass. Their presence allows a much better fit of these effective masses than was obtained by retaining only a single parameter (P). In fact, we have to account for four band edge effective masses. With P , F , γ_1 and γ_2 we obtain four adjustable parameters, which is the required number.

The reasons why we wish to discard the remote band parameters for light particle states are twofold. Firstly, without F , γ_1 and γ_2 , the differential system for the three light particle components of χ_+ becomes of the first order in $\frac{\partial}{\partial z}$ and therefore easier to handle, to the extent that all the results can be obtained in closed forms. Secondly, by retaining F , γ_1 and γ_2 , we face a difficulty termed by White and Sham [1] the "wing bands", which are extensively discussed by Schuurman and t'Hooft [3].

Let us try to pinpoint the origins of this difficulty by using a simplified model : we shall neglect the Γ_7 band, i.e. we shall assume that the spin-orbit energy is very large compared with the light hole confinement energy. In each type of layer, the heterostructure states of a given energy ε are linear combinations of bulk solutions. These solutions are either propagating (k_A , k_B real) or evanescent (k_A real, k_B imaginary) in the A and/or B layers. Thus in order to know our heterostructure states, we simply have to calculate the k_A , k_B corresponding to the energy ε . Therefore to calculate k_A , for example, we have to find the roots of :

$$\text{Det} [D_+ (p_z = \hbar k_A) - \varepsilon \mathbf{1}] = 0 \quad (66)$$

In our three band (Γ_{8h} , Γ_{8l} , Γ_6) model, equation (66) factorizes into

$$\left(\varepsilon + \varepsilon_A + \frac{\hbar^2 k_A^2}{2M_{hh}^{(A)}} \right) \left[\left(\varepsilon - \frac{F\hbar^2 k_A^2}{2m_0} \right) \left(\varepsilon + \varepsilon_A + (\gamma_1 + 2\gamma_2) \frac{\hbar^2 k_A^2}{2m_0} \right) - \frac{2\hbar^2}{3} P^2 k_A^2 \right] = 0 \quad (67)$$

The first root is obvious and describes the heavy hole branch. At the moment however, it is of no concern to us. The two other k_A^2 roots, which are always real, describe the dispersions of either the *coupled* ($\Gamma_6 + \Gamma_8'$) bands or the wing band (see Fig. 11). The former, which has a clear physical significance is i) small compared with the size of the A Brillouin zone along the \hat{z} axis and ii) almost insensitive to any change in the small parameters F and $\gamma_1 + 2\gamma_2$. k_A is real for both $\varepsilon < -\varepsilon_A$ and $\varepsilon > 0$ and imaginary for $-\varepsilon_A \leq \varepsilon \leq 0$: the propagating light hole state ($\varepsilon \leq -\varepsilon_A$) becomes evanescent as ε enters into the bandgap $\left(-\varepsilon_A \leq \varepsilon \leq -\frac{\varepsilon_A}{2} \right)$ and progressively transforms into an evanescent Γ_6 electron $\left(-\frac{\varepsilon_A}{2} \leq \varepsilon \leq 0 \right)$, ending up as a propagating electron state ($\varepsilon \geq 0$). The other k_A^2 root is unphysical as i) it has a

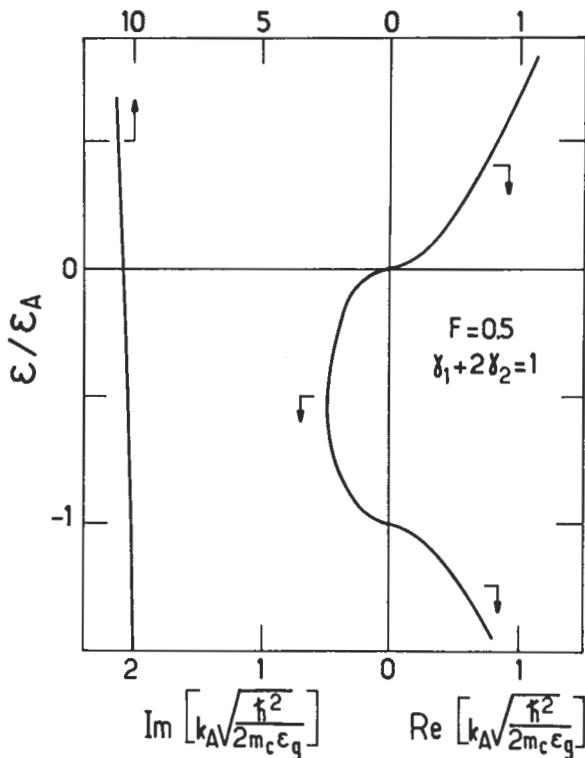


Fig. 11. — Dispersion relations of the light particles states *versus* the real and imaginary wavevectors of a bulk material treated in the Kane model with the remote band effects included up to the second order in $k\Delta = \infty$; $F = 0.5$; $\gamma_1 + 2\gamma_2 = 1$; $m_c = 0.067 m_0$. Notice the different scales for the upper and lower horizontal axis of the left panel.

very large magnitude, eventually larger than the A Brillouin zone size along the \hat{z} axis and ii) is unstable with respect to any change in the small parameters F , $\gamma_1 + 2\gamma_2$. This unphysical solution is a remnant of the remote band and should in fact correspond to their dispersions if they were correctly described (by using a basis much larger than Γ_6 and Γ_8). Clearly, there is no hope of describing the remote bands dispersion relations in the vicinity of Γ_6 and Γ_8 , i.e. far from their own edges, by a quadratic law. The wing band seen in figure 11 is just the mathematical transcription of this gross physical inaccuracy. Note however that, in principle, the wing bands are harmless when they correspond to evanescent states which decay very rapidly in space (a few Angströms) on each side of the interfaces. This happens when $F(\gamma_1 + 2\gamma_2) > 0$. They prove troublesome however in the numerical computations. Therefore as the introduction of higher bands adds little physical information to the Γ_6 , Γ_7 , Γ_8 bands (except of course for the heavy hole band) and produces artefacts like the wing bands, we feel that it is reasonable to get rid of them wherever possible.

II.5 SIMPLIFIED CALCULATIONS OF SUPERLATTICE AND QUANTUM WELL STATES ($\mathbf{k}_\perp = \mathbf{0}$). — At $\mathbf{k}_\perp = \mathbf{0}$, setting F , γ_1 , γ_2 equal to zero and discarding the heavy hole line, we obtain a 3×3 first order differential system which can be transformed into a scalar, non-linear in ε , second order differential equation. The latter is obtained by eliminating χ_5 and χ_7 to the benefit of χ_1 in equation (64) (and similarly in the D_- matrix). Proceeding as in Appendix B of the previous chapter, we readily obtain :

$$\left\{ \frac{P^2}{3} p_z \left[\frac{2}{\varepsilon + \varepsilon_A - V_p(z)} + \frac{1}{\varepsilon + \varepsilon_A + \Delta_A - V_\delta(z)} \right] p_z + V_s(z) \right\} \chi_1(z) = \varepsilon \chi_1(z) \quad (68)$$

together with a similar expression for $\chi_2(z)$. The boundary conditions that χ_1 fulfills at the A-B interfaces are such that

$$\chi_1(z) \text{ and } \frac{1}{\mu(\varepsilon, z)} \frac{d\chi_1}{dz} \text{ are both continuous} \quad (69)$$

where :

$$\frac{1}{\mu(\varepsilon, z)} = \frac{2P^2}{3} \left[\frac{2}{\varepsilon + \varepsilon_A - V_p(z)} + \frac{1}{\varepsilon + \varepsilon_A + \Delta_A - V_\delta(z)} \right] \quad (70)$$

$\mu(\varepsilon, z)$ is nothing other than the energy-dependent effective mass which appears in the Kane three-band model when the dispersion relations in each kind of layer are written in the implicit form

$$\begin{aligned} \varepsilon &= \frac{\hbar^2 k_A^2}{2\mu_A(\varepsilon)} && \text{in the A layer} \\ \varepsilon - V_s &= \frac{\hbar^2 k_B^2}{2\mu_B(\varepsilon)} && \text{in the B layer} \end{aligned} \quad (71)$$

the energy being set at the Γ_6 edge of the A material.

We are now in position to calculate the dispersion relations of an A-B superlattice. Let us consider a superlattice unit cell (thickness $d = L_A + L_B$) containing two interfaces. At each interface we apply the continuity conditions given by equation (69). In addition, the band edges $V_s(z)$, $V_p(z)$, $V_\delta(z)$ and the effective mass are periodic functions of z with periodicity d . Thus $\chi_1(z)$ may be written as a Bloch wave :

$$\chi_1(z + d) = \exp(iqd) \chi_1(z) \quad (72)$$

with

$$-\frac{\pi}{d} < q \leq \frac{\pi}{d} \quad (73)$$

Inside the A and B layers χ_1 is a linear combination of incoming and outgoing plane waves

$$\begin{aligned}\chi_1(z) &= \alpha \exp(ik_A z) + \beta \exp(-ik_A z); & z \in A \\ \chi_1(z) &= \gamma \exp(ik_B z) + \delta \exp(-ik_B z); & z \in B\end{aligned}\quad (74)$$

Four linear equations are obtained (two boundary conditions per interface) for four unknowns ($\alpha, \beta, \gamma, \delta$). These equations can be satisfied only if the determinant of the associated matrix vanishes, which in turn leads to the superlattice dispersion relations.

$$\cos(qd) = \cos(k_A L_A) \cos(k_B L_B) - \frac{1}{2} \left(\xi + \frac{1}{\xi} \right) \sin(k_A L_A) \sin(k_B L_B) \quad (75)$$

with :

$$\xi = \frac{k_A}{\mu_A(\varepsilon)} \frac{\mu_B(\varepsilon)}{k_B} \quad (76)$$

$$\varepsilon(\varepsilon + \varepsilon_A)(\varepsilon + \varepsilon_A + \Delta_A) = \hbar^2 k_A^2 P^2 \left(\varepsilon + \varepsilon_A + \frac{2\Delta_A}{3} \right) \quad (77)$$

$$(\varepsilon - V_s)(\varepsilon - V_s + \varepsilon_B)(\varepsilon - V_s + \varepsilon_B + \Delta_B) = \hbar^2 k_B^2 P^2 \left(\varepsilon - V_s + \varepsilon_B + \frac{2\Delta_B}{3} \right) \quad (78)$$

Compared with the results obtained in chapter I, we see that the same kind of superlattice dispersion relations are obtained in idealized situations (a single band, a single effective mass etc...) and in the envelope function description of semiconductor superlattices. This is, after all, not very surprising. All our efforts have been put into neglecting the atomic-like details specific to solids. Some of the important features associated with them have, however, survived :

- the wavevectors k_A and k_B in equations (75-78) are related to the energy via expressions which are more complicated than those found in vacuum. This accounts for the multiband nature of solids.
- ξ is no longer given by k_A/k_B but should be corrected by the effective mass ratio μ_B/μ_A to account for the effective mass mismatch at the interface.

II.6 MISCELLANEOUS LIMITING CASES.

II.6.1 *Evanescent propagation in one kind of layer. Quantum well bound states.* — Suppose that $V_s > 0, V_p < 0, V_\delta < 0$ as found in type I heterostructures. Thus for energies ε such that $0 \leq \varepsilon \leq V_s, V_p - \varepsilon_A \leq \varepsilon \leq -\varepsilon_A$ or that $V_\delta - \varepsilon_A - \Delta_A \leq \varepsilon \leq -\varepsilon_A - \Delta_A$, the wavevector k_B is imaginary. The corresponding superlattice states are derived from the isolated quantum well levels which, due to the non-zero tunnel coupling across the B layer, have hybridized to form superlattice bands. To obtain their dispersion relations we set $k_B = i\kappa_B$ in equations (75, 76). The quantity ξ

changes into $-i\tilde{\xi}$ and the superlattice dispersion relations becomes :

$$\cos(qd) = \cos(k_A L_A) \cosh(\kappa_B L_B) - \frac{1}{2} \left(\tilde{\xi} - \frac{1}{\tilde{\xi}} \right) \sin(k_A L_A) \sinh(\kappa_B L_B) \quad (79)$$

Moreover, if the barrier thickness L_B becomes infinitely thick, suppressing the tunnel coupling between the wells, the isolated quantum well bound states are recovered

$$\cos(k_A L_A) - \frac{1}{2\tilde{\xi}} \left(\tilde{\xi} - \frac{1}{\tilde{\xi}} \right) \sin(k_A L_A) = 0 \quad (80)$$

In equation (80) we have not made use of the parity property of $V_s(z)$, $V_p(z)$, $V_\delta(z)$ with respect to the centre of the A layer. If we do so, we easily find that equation (80) factorizes into two equations corresponding to even and odd χ_1 's respectively :

$$\begin{aligned} \cos\left(\frac{1}{2}k_A L_A\right) - \tilde{\xi} \sin\left(\frac{1}{2}k_A L_A\right) &= 0 \\ \cos\left(\frac{1}{2}k_A L_A\right) + \frac{1}{\tilde{\xi}} \sin\left(\frac{1}{2}k_A L_A\right) &= 0 \end{aligned} \quad (81)$$

It should be noted that these equations are the same as those obtained in the case of the Ben Daniel-Duke quantum well equations (35, 36), except that the masses μ_B , μ_A which enter into equations (80, 81) are energy-dependent. Thus, the quantum well states in host displaying Kane-like bands admit the Ben Daniel-Duke solutions as a limiting case.

II.6.2 Tight-binding expansion of the superlattice states. — Let us denote the right hand side of equation (79) by $f(\varepsilon)$. If the barriers are thick, $f(\varepsilon)$ will display large variations and the band widths, which are the energy segments where $|f(\varepsilon)| < 1$, will be narrow. If L_B is infinite we know that the superlattice states reduce to the bound states ε_n of the isolated wells. Thus, in the limit of the thick barriers we may expand $f(\varepsilon)$ in the vicinity of one of these bound states ε_j in order to obtain the superlattice band which originates from the hybridized ε_j 's. We find that equation (79) simplifies into :

$$\varepsilon = \varepsilon_j + s_j + 2t_j \cos(qd) \quad (82)$$

with :

$$s_j = -\frac{f(\varepsilon_j)}{f'(\varepsilon_j)} ; \quad 2t_j = \frac{1}{f'(\varepsilon_j)} \quad (83)$$

These are results which are analogous to a tight-binding formulation of a superlattice state. In fact, the shift and transfer integrals s_j and t_j can be analytically obtained in terms of isolated well wavefunctions only. With a tight-binding expansion in mind, let us write the superlattice state as :

$$\chi_1(z) = \frac{1}{\sqrt{2N+1}} \sum_{n=-N}^{+N} \chi_{1,j}(z - nd) \exp(i qnd) ; \quad N \rightarrow \infty \quad (84)$$

where $\chi_{1,j}(z-nd)$ is the j^{th} bound state wavefunction of the well centred at nd . By inserting equation (84) into equation (68), we obtain, after some manipulations,

$$s_j \simeq -2 V_s \int_{d - \frac{1}{2} L_A}^{d + \frac{1}{2} L_A} \chi_{1,j}^2(z) dz \quad (85)$$

$$t_j \simeq -V_s \int_{-\frac{1}{2} L_A}^{\frac{1}{2} L_A} \chi_{1,j}(z) \chi_{1,j}(z-d) dz \quad (86)$$

In equations (85, 86) only nearest neighbour interactions have been retained and the non-parabolicity effects have been taken into account only in the evaluation of ϵ_j . Equations (85, 86) hold for Γ_6 -related superlattice states. For light hole states similar expression could be derived with $\chi_{1,j}$ replaced by $\chi_{5,j}$ and V_s by V_p . The orders of magnitude of s_j and t_j are :

$$s_j \simeq -V_s P_b(\epsilon_j) \exp(-2 \kappa_B L_B) [1 - \exp(-2 \kappa_B L_A)] \quad (87)$$

$$t_j \simeq (-1)^j V_s P_b(\epsilon_j) \exp(-\kappa_B L_B) \frac{\kappa_B^2}{\kappa_B^2 + k_A^2} \times \left\{ 1 + \frac{\mu_A(\epsilon_j)}{\mu_B(\epsilon_j)} + (-1)^j \left[1 - \frac{\mu_A(\epsilon_j)}{\mu_B(\epsilon_j)} \right] \exp(-\kappa_B L_A) \right\} \quad (88)$$

where $j = 1, 2, \dots$; κ_B is the magnitude of the imaginary wavevector in the barrier for the j^{th} bound state of the isolated well and $P_b(\epsilon_j)$ is the integrated probability of finding the electron in the barrier while in this j^{th} bound state.

Although approximate, equations (87, 88) are helpful for anticipating the trend in superlattice bandwidths and shift integrals with varying barrier or well thicknesses. We notice that s_j decays more rapidly with L_B than t_j . In addition t_j increases in magnitude when ϵ_j increases, i.e. when the well thickness is decreased for a given j or when dealing with quantum well excited states for a given L_A . Both of these increases are due to decreasing κ_B and increasing P_b . More importantly, we see that t_j and thus the j^{th} bandwidth decrease exponentially with increasing L_B , the decay being controlled by κ_B , i.e. it is faster when V_s, L_A are larger. To a very good accuracy, one may write for the ground superlattice bandwidth ΔE_1

$$\Delta E_1 = V_{\text{eff}} \exp(-\kappa_B L_B) \quad (89)$$

Some numerical examples will be given below (section II.7)

II.6.3 Propagating states in both kinds of layers. — We have seen that the superlattice states corresponding to the imaginary wavevectors in one kind of layer can be analyzed in terms of hybridized bound levels of isolated wells. It could be asked if a similar link can be established between the superlattice states which propagate in both kinds of layers (i.e. k_A, k_B real, which means $\epsilon \geq V_s$,

$\varepsilon \leq -\varepsilon_A - V_p$, $\varepsilon \leq -\varepsilon_A - \Delta_A - V_\delta$ in type I heterostructures) and certain continuum states of isolated wells. We recall that the continuum of a single quantum well is not structureless (see chapter I). On the contrary, virtual bound states take place in the continuum when

$$k_A L_A = p \pi ; \quad p = 1, 2, \dots \quad (90)$$

The virtual bound states are the continuation of the true quantum well bound states when their confinement energies exceed the barrier height. Both true and virtual bound states are matched at the onset of the continuum : it can immediately be checked that equation (90), with k_A given by equation (77) in terms of the energies corresponding to the onset of a continuum (i.e. $\varepsilon = V_s$, $\varepsilon = -\varepsilon_A + V_p$, $\varepsilon = -\varepsilon_A - \Delta_A + V_\delta$) is indeed a solution of the quantum well bound state equation (at the onset of the continuum $k_B \rightarrow 0$ and ξ diverges in Eq. (80)).

By examining equation (75) we notice that the energy of a virtual bound level of an isolated well always corresponds to an allowed superlattice state with q equal to

$$q = \pm \left(k_B \frac{L_B}{L_A + L_B} + \frac{p \pi}{L_A + L_B} \right) \quad (91)$$

To some extent therefore, the superlattice states which propagate in each kind of layer can be viewed as the hybridization of the virtual bound levels of the isolated wells.

II.6.4 Heavy hole superlattice states. — If we were to use our simplified Kane model, the heavy hole superlattice states would be dispersionless and only the energies $-\varepsilon_A$, $-\varepsilon_A + V_p$ would be allowed. However since the heavy holes decouple from the light particles states, we can re-introduce the coupling between Γ_{8h} and the remote bands to correctly describe the heavy hole curvature. The problem becomes identical to a Ben Daniel-Duke one because the heavy hole envelope functions are the eigenfunctions of :

$$\left[-\varepsilon_A - p_z \frac{1}{2M_{hh}(z)} p_z + V_p(z) \right] \chi_3(z) = \varepsilon \chi_3(z) \quad (92)$$

$$V_p(z + d) = V_p(z) \quad (93)$$

where : $\frac{m_0}{M_{hh}^{(A)}} = \gamma_1^{(A)} - 2\gamma_2^{(A)} ; \frac{m_0}{M_{hh}^{(B)}} = \gamma_1^{(B)} - 2\gamma_2^{(B)}$ (94)

We can proceed along the same lines as before to calculate $\chi_3(z)$. Writing $\chi_3(z)$ as the sum of the incoming and outgoing plane waves in each kind of layer, we obtain the heavy hole superlattice states in the same form as equation (75), except that :

$$\xi = \frac{k_A}{M_{hh}^{(A)}} \frac{M_{hh}^{(B)}}{k_B} \quad (95)$$

for heavy holes. In equation (95), there is

$$k_A = \sqrt{\frac{2M_{hh}^{(A)}}{\hbar^2} [-\varepsilon_A - \varepsilon]} ; \quad k_B = \sqrt{\frac{2M_{hh}^{(B)}}{\hbar^2} [-\varepsilon_A - \varepsilon + V_p]} \quad (96)$$

In the energy segment $-\varepsilon_A + V_p < \varepsilon < -\varepsilon_A$, one of the wavevectors (Eq (96)) becomes imaginary (say k_B). In this case equation (75) should be replaced by equation (79), k_B by $i\kappa_B$ and ξ by $-\tilde{\xi}$. Qualitatively therefore both light particle and heavy hole superlattice states are quite similar. Quantitatively, however, the heavy hole superlattice bands are much narrower than the light particle ones due to the larger heavy hole mass. For numerical examples, see section (II.7).

II.7 SPECIFIC EXAMPLES. — To substantiate the previous considerations figures (12-20) present the calculated energy levels of several quantum wells and superlattices.

Quantum wells

The dominant effect in quantum well energy levels is the finiteness of the barrier height. This is illustrated in figures (12-15) in the case of GaAs-Ga(Al)As where we have plotted :

i) the dependence of the confinement energies E_1 , LH_1 , HH_1 of the ground bound states upon the barrier height V_s and $|V_p|$ for electron, light and heavy holes

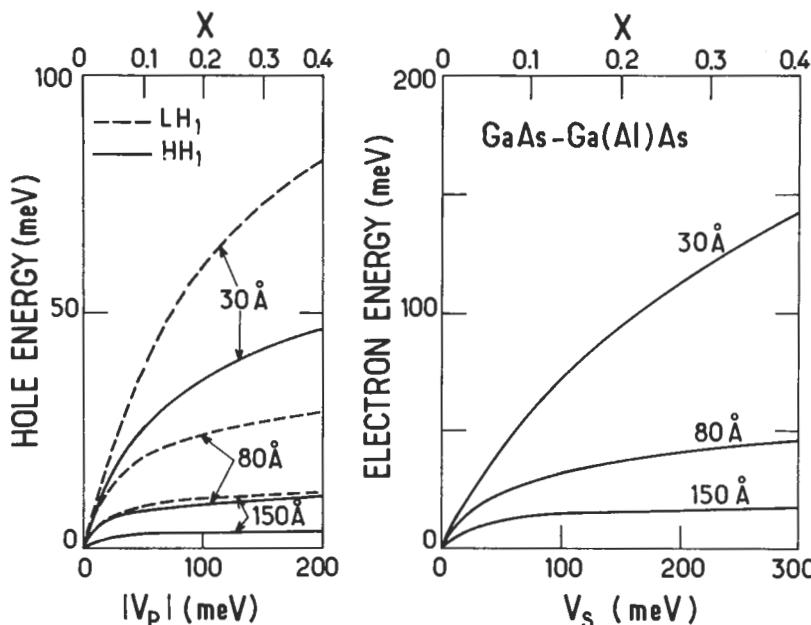


Fig. 12. — Evolution of the ground states for electron (E_1), heavy hole (HH_1) and light hole (LH_1) versus the barrier heights V_s and $|V_p|$ in GaAs-Ga(Al)As quantum wells. Three well thicknesses have been considered $L = 30 \text{ \AA}$, 80 \AA and 150 \AA respectively.

respectively, i.e. upon the Al mole fraction x , for three GaAs slab thicknesses : $L = 30 \text{ \AA}$, 80 \AA and 150 \AA respectively. The calculations have been performed using a 60 % – 40 % split between the conduction and valence bands of the bandgap energy difference ΔE_g between $\text{Ga}_{1-x}\text{Al}_x\text{As}$ and GaAs. Taking

$$\Delta E_g(x) = 1247 x \text{ meV} \quad (97)$$

we are left with

$$V_s(x) = 748.2 x \text{ meV} \quad (98)$$

$$V_p(x) = -498.8 x \text{ meV} \quad (99)$$

Figure 12 shows that the asymptotic values of E_1 , LH_1 , HH_1 , which correspond to a perfect confinement (i.e. $k_A L = \pi$, where k_A is the appropriate wavevector in GaAs), are barely obtained, apart from heavy holes in thick ($L > 100 \text{ \AA}$) wells. On

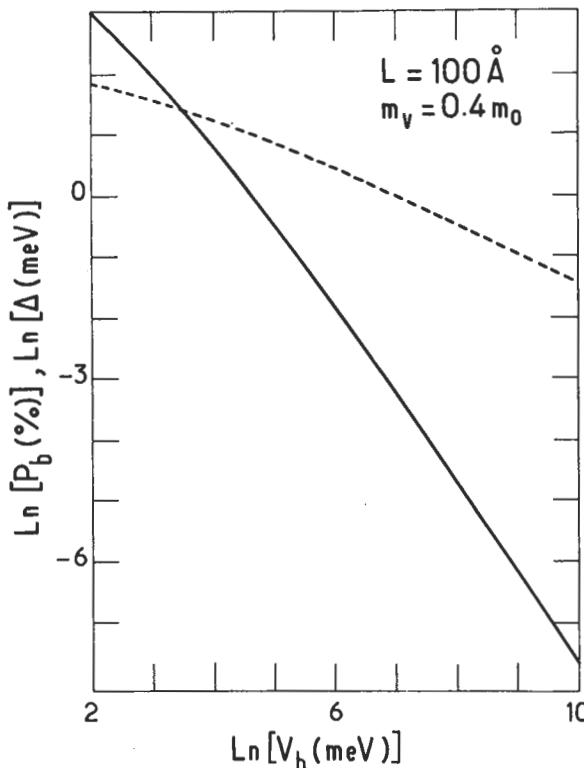


Fig. 13. — The natural logarithms of the energy difference $\Delta = \frac{\hbar^2 \pi^2}{2M_{hh}L^2} - HH_1$ (dashed line) and of the integrated probability P_b of finding the hole in the barriers (solid line) are plotted versus the hole barrier height ($-V_p$) for a quantum well thickness $L = 100 \text{ \AA}$. $M_{hh} = 0.4 m_0$.

the one hand, it should be noticed that a better justification can be obtained for the same approximation (i.e. perfect confinement) in the case of the wavefunction. This is illustrated in figure 13 for heavy holes and can be analytically traced back. The convergence of $E_1(HH_1, LH_1)$ towards $\hbar^2\pi^2/2m L^2$, where m is the effective mass appropriate for the electrons (light hole, heavy hole), is slow since it is proportional to $V_s^{-1/2}(|V_p|^{-1/2})$. On the other hand, the spatial confinement of the carrier in the well is completed more rapidly : the integrated probability of finding the particle outside the well drops to zero like $V_s^{-3/2}(|V_p|^{-3/2})$.

ii) the energy level diagrams *versus* the GaAs slab thickness L for an Al mole fraction of 0.3 (Figs. 14, 15). The solid lines represent the true quantum well bound states whereas the dashed lines represent the locii of the virtual bound states (or transmission resonances) which fulfil

$$k_A(\epsilon_p) L = p\pi; \quad p = 1, 2, \dots \quad (100)$$

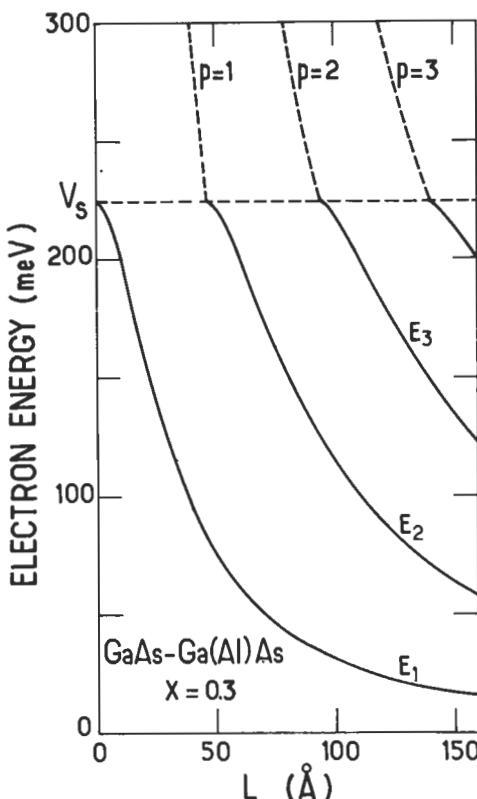


Fig. 14. — Energy level diagram for electrons in $\text{GaAs}-\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ quantum wells *versus* the GaAs slab thickness. The solid lines correspond to bound states and the dashed lines to virtual bound states. $V_s = 224.5$ meV.

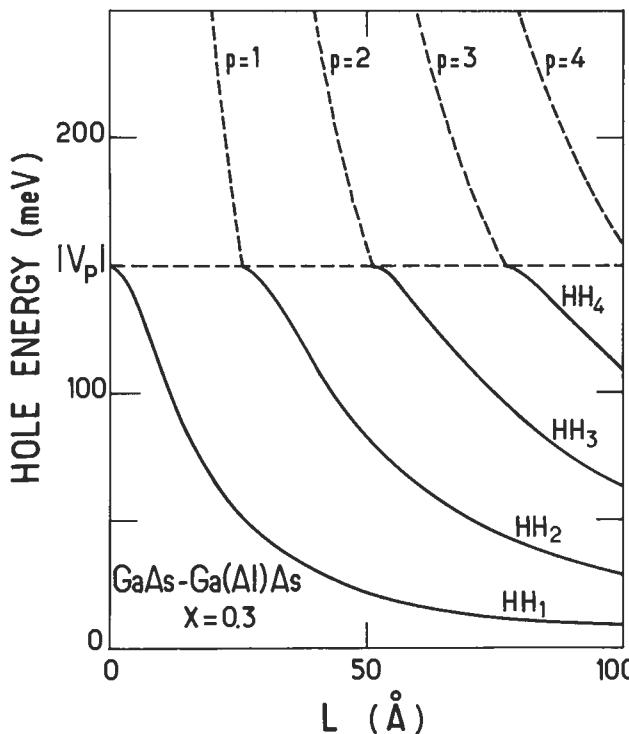


Fig. 15. — Energy level diagram for heavy holes in $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ quantum wells *versus* the GaAs slab thickness. The solid lines correspond to bound states and the dashed lines to virtual bound states. $|V_p| = 149.6$ meV. $M_{hh}^{\text{GaAs}} = 0.38 m_0$. $M_{hh}^{\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}} = 0.45 m_0$.

for energies exceeding the top of the confining barrier. In equation (100) k_A denotes the carrier wavevector inside the well. It is related to ϵ_p either by equation (77) (light particle) or by equation (96) (heavy hole). It should be pointed out that semiconductor quantum wells always admit one bound state, as the ground states $E_1(HH_1, LH_1)$ only reach the top of their respective barriers at $L = 0$.

The band non-parabolicity, which only affects the light particle states, plays a relatively minor part in GaAs-Ga(Al)As quantum wells. This is due to the large GaAs bandgap. Basically speaking, band non-parabolicity amounts to replacing the GaAs band edge mass by an energy-dependent mass whose relative increment is equal to the ratio between the kinetic energy and the bandgap. As the confinement energy varies roughly in the same way as the inverse of the carrier effective mass, one expects a non-parabolicity-induced correction of the order of the confinement energy (calculated with the band edge mass) divided by the GaAs bandgap. This qualitative prediction is more or less supported by the calculations.

Obviously, the smaller the well-acting material bandgaps, the larger the non-parabolicity effects. Again, qualitatively, the non-parabolicity effects scale the same

way as ε_A^{-2} because the band edge mass itself is proportional to ε_A . Therefore, the non-parabolicity is more important in the evaluation of the electron confinement energies of $\text{Ga}_{0.47}\text{In}_{0.53}\text{As-InP}$ or GaSb-AlSb quantum wells than in GaAs-Ga(Al)As ones. Also it is more important for excited levels than for the ground states.

Superlattices

The superlattice states corresponding to evanescent propagation in one kind of layer present few surprises (Figs. 16-18). One notices that the isolated quantum well levels broaden to form the superlattice bands. Their widths increase i) when the well-acting layer narrows at a fixed barrier thickness, ii) when at fixed L_A, L_B one looks at the bands originating from the excited levels of the isolated quantum wells, and iii) when, at fixed L_A , the barrier thickness decreases.

As anticipated in our tight-binding description of the superlattice states, the E_1 bandwidth decays exponentially with the barrier thickness (Fig. 19). Furthermore, it is clear from figures 16 and 18 that the superlattice bandwidth decreases with the increasing carrier effective mass, to the point where the heavy hole bandwidths are almost negligible for $L_A \geq 100 \text{ \AA}$.

The pattern of the superlattice bands which originate from the propagating states in both kinds of layers is perhaps more intricate. This is particularly clear for the

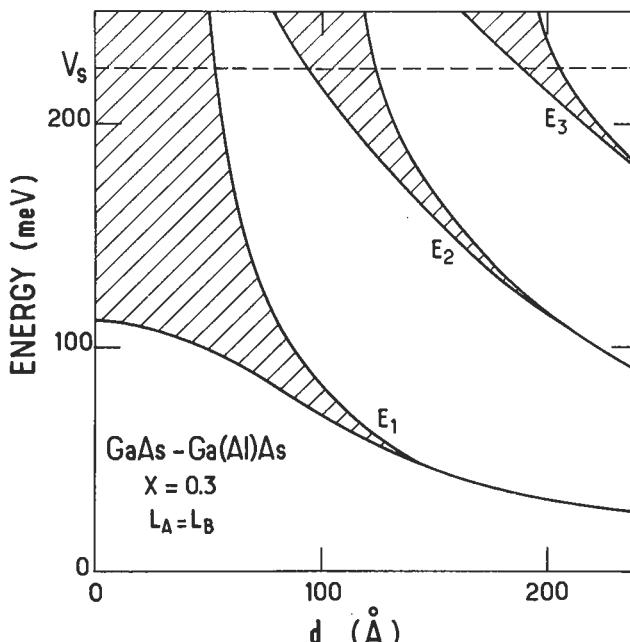


Fig. 16. — Superlattice band structure for electrons in $\text{GaAs-Ga}_{0.7}\text{Al}_{0.3}\text{As}$ superlattices *versus* period d . Equal layer thicknesses are assumed in the calculations. The allowed energy states are hatched.

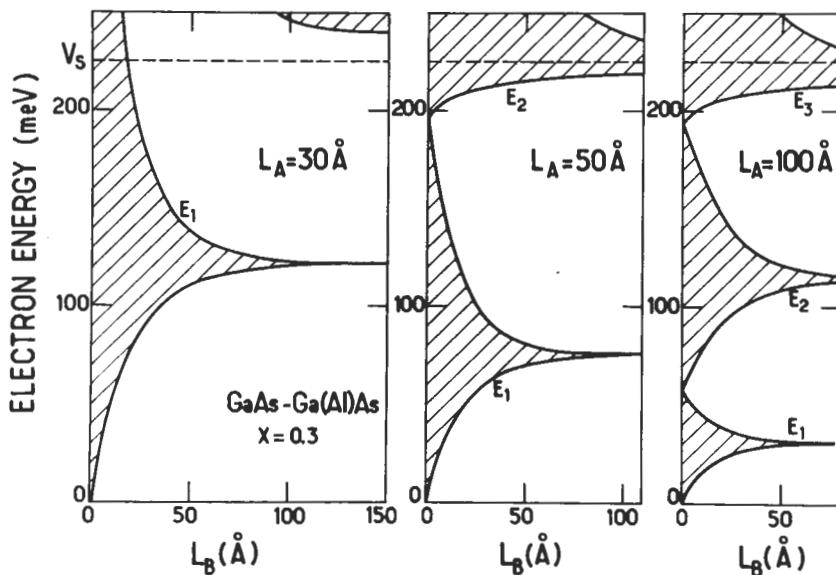


Fig. 17. — Evolution of the superlattice band structure for electrons in $\text{GaAs-Ga}_{0.7}\text{Al}_{0.3}\text{As}$ superlattices when the $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ layer thickness L_B is increased and the GaAs layer thickness L_A is kept fixed. The allowed energy states are hatched.

heavy hole bands (Fig. 18) where a fairly complex band pattern above the barrier height contrasts with the simple features found below $|V_p|$. In particular, the bands HH_2 and HH_3 cross at one point, as do HH_3 and HH_4 etc... It is possible to check that these degeneracy points take place when both $k_A L_A$ and $k_B L_B$ are integer multiples of π . For these particular energies and thicknesses there is a Fabry-Pérot effect for the hole waves in *both* kinds of layers. It can be inferred from this particular example that the superlattice states which correspond to the propagating states in both kinds of layers should strongly depend upon the barrier thickness L_B , while those originating from the hybridization of quantum well bound states are less dependent upon this quantity.

The InAs-GaSb superlattices are the extreme case of type II heterostructures [30]. The latter are characterized by the fact that V_s and V_p have the same sign, so that one kind of layer attracts electrons whereas the other kind is a potential well for the holes. In InAs-GaSb the top of the GaSb valence band is located in energy above the bottom of the InAs conduction band. When the superlattice period increases, as L_A is equal to L_B , the E_1 band moves towards the InAs conduction band edge while HH_1 approaches the top of the GaSb valence band. There should thus exist a critical period d_c beyond which HH_1 is located at a higher energy than E_1 (Fig. 20). This situation is, to our knowledge, unique in semiconductor superlattices. For $d \geq d_c$ one should expect some kind of a semiconductor \rightarrow semimetal transition whereby electrons flow from the GaSb to the InAs layers leaving holes behind them.

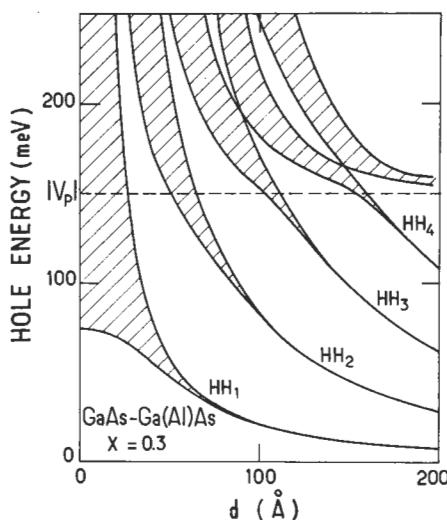


Fig. 18. — Superlattice band structure for heavy holes in $\text{GaAs}-\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ superlattices versus period d . Equal layer thicknesses are assumed in the calculations. The allowed energy states are hatched.

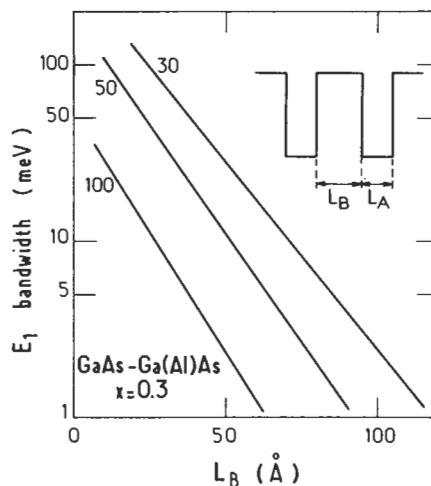


Fig. 19. — The E_1 bandwidth is plotted versus the $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ layer thickness L_B in $\text{GaAs}-\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ superlattices with three different GaAs layer thicknesses (30 \AA , 50 \AA , 100 \AA). Notice the logarithmic scale for the E_1 bandwidth.

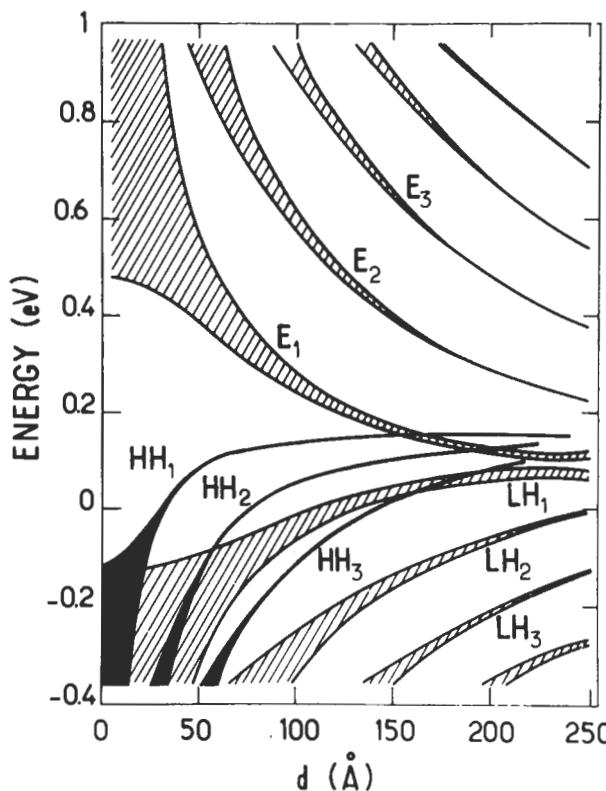


Fig. 20. — Evolution of the InAs-GaSb superlattice band structure with increasing period d ($L_A = L_B$). The allowed energy levels are hatched (light particles) or black (heavy holes). The energy zero is taken at the bottom of the bulk InAs conduction band.

Actually, the situation is more complicated since the band crossings which take place at $\mathbf{k}_\perp = \mathbf{0}$ are replaced by anticrossings if the in plane wavevector is non zero [2, 32]. This leads to the formation of small gaps between E_1 and HH_1 . Our present knowledge of this fascinating system is still incomplete as the combination of technological problems (residual impurities in InAs and GaSb layers) and theoretical complexity (non-negligible stress effects due to the imperfect lattice matching, charge redistribution between InAs and GaSb layers, many-body effects ?) conspire to elude clear cut answers.

II.8. LABELLING AND COUNTING SUPERLATTICE STATES. — A superlattice state is labelled by two wavevectors : $q\hat{z}$ and \mathbf{k}_\perp (which express the delocalized nature of the wavefunctions along and perpendicular to the growth axis), and by a subband index n ($n = 1, 2\dots$). A label E , HH , $LH\dots$ is also affixed to a superlattice state to indicate from which host band the superlattice state is principally derived. Notice that this

label has little physical meaning if significant band mixing occurs in the host layer. This happens at $\mathbf{k}_\perp = \mathbf{0}$ if the host bandgaps are very narrow or in the Γ_8 bands at $\mathbf{k}_\perp \neq \mathbf{0}$ (see below). In addition, there is a "spin" quantum number which can take two values (say $\pm \frac{1}{2}$) and which reduces to the z component of the spin quantum number σ_z if there are no spin-orbit effects. Due to the finite spin-orbit splitting, σ_z is no longer a good quantum number. However, each state of the Hamiltonian (Eq. (16)) remains twice degenerate at $\mathbf{k}_\perp = \mathbf{0}$ (since the inversion asymmetry terms have been omitted). This Kramers degeneracy is a direct consequence of the invariance of $\mathcal{D}^{(0)} + \varphi(\mathbf{r}) \mathbb{1}$ under a time-reversal operation [23]. We shall see in section III that this twofold degeneracy is lifted in heterostructures which lack inversion symmetry. This happens for instance when $\varphi(z) \neq \varphi(-z)$ (asymmetric band bending) or $V_s(z) \neq V_s(-z)$ etc... (asymmetrically designed heterostructures).

We have seen that the superlattice states are Bloch waves (Eq. (72)). Assuming that the superlattice is made of an infinite sequence of blocks containing N superlattice periods where N is very large, we may apply the Born-von Karman boundary conditions, i.e. we require the superlattice wavefunction to be the same at both ends of a given block :

$$\chi_1(z + Nd) = \chi_1(z) \quad (101)$$

$$\text{or :} \quad qNd = 2\pi p \quad p = 0, 1, 2, \dots, N-1 \quad (102)$$

Thus, there are N independent q values, i.e. as many q values as superlattice unit cells in a block. These q values are equidistant and separated from each other by $\frac{2\pi}{Nd}$. Any summations of the form $\sum_q \alpha(q)$ will be converted into an integration :

$$\sum_q \alpha(q) \rightarrow \frac{Nd}{2\pi} \int_{q_1}^{q_2} \alpha(q) dq \quad (103)$$

The upper and lower limits of the integral can be any number provided they differ by $\frac{2\pi}{d}$, i.e. the size of the superlattice Brillouin zone along the growth axis. Often the segment $\left[-\frac{\pi}{d} + \frac{\pi}{d} \right]$ is chosen to emphasize the parity property

$$\varepsilon(q) = \varepsilon(-q) \quad (104)$$

fulfilled by any superlattice band.

III. In-plane dispersions in semiconductor heterostructures.

When the in-plane wavevector \mathbf{k}_\perp is not equal to zero, the heavy hole and light particle states become coupled. This is simply the consequence of the degenerate nature of the Γ_8 hosts' band edges, which originates from the P symmetry of the orbital parts of the band edge Bloch functions. In fact, even the simplest (i.e.

isotropic) quadratic Hamiltonian for Γ_8 bands in a bulk material is not a scalar, which is the case for Γ_6 bands (S symmetry). As shown by Luttinger [22] the isotropic Γ_8 Hamiltonian should be written :

$$\mathcal{H} = \alpha k^2 + \beta (\mathbf{k} \cdot \mathbf{J})^2 \quad (105)$$

where \mathbf{J} is the total angular momentum $(J = \frac{3}{2})$. From equation (105) it can be seen that, if the \mathbf{J} quantization axis is given (say the \hat{z} axis), the decoupling between the $m_J = \pm \frac{3}{2}$ and $m_J = \pm \frac{1}{2}$ states is only possible if $\mathbf{k} \parallel \mathbf{J} \parallel \hat{z}$ while, if \mathbf{k} is arbitrary, the eigenstates of equation (105) are no longer eigenfunctions of J_z . However, it is always possible to rotate the reference frame in a bulk material in such a way that the new direction z along which \mathbf{J} will be quantized coincides with \mathbf{k} , thus obtaining the simultaneous diagonalizations of J_z and \mathcal{H} .

In heterostructures the potential term $V(z)$, which should be added to the right hand side of equation (105), gives rise to anisotropy. The interface planes have the simple expressions $z = \text{constant}$ and the wavefunctions are separable in z and (x, y) . If we quantize \mathbf{J} along the growth axis, which is what we did in section II, the $m_J = \pm \frac{3}{2}$ and $m_J = \pm \frac{1}{2}$ states become decoupled if $\mathbf{k} \parallel \mathbf{J}$, i.e. if $\mathbf{k}_\perp = \mathbf{0}$. If

$\mathbf{k}_\perp \neq \mathbf{0}$ light particle and heavy hole states will hybridize. However, contrary to the bulk case, \mathbf{J} can no longer be rotated to line up along \mathbf{k} , thereby recovering a simple situation. This is because \mathbf{k} is no longer a good quantum number due to the potential term $V(z)$. If we attempt to rotate the (xyz) frame to obtain a (XYZ) frame where Z and J_z are parallel to \mathbf{k}_A (the carrier wavevector in the A material), it will not lead to the lining up between \mathbf{J} and \mathbf{k}_B (the carrier wavevector in the B material). The decoupling will thus never be complete. The best one can do is to choose a \mathbf{J} quantization axis which renders the valence Hamiltonian less complicated. To quantize \mathbf{J} along the growth axis seems the most natural choice. However, if band warping is neglected, significant simplifications are obtained by quantizing \mathbf{J} in a direction perpendicular to the growth axis [2]. Obviously, if we add the Γ_8 band warping (i.e. the non-sphericity) to equation (105) and include the Γ_6 edge in the analysis (to account for non-parabolicity) the complexity is increased, making it impossible to obtain analytical results. To our knowledge, the only circumstances where analytical results for which Γ_6 or Γ_8 in-plane dispersions have been obtained are :

- i) Γ_8 eigenstates of a quadratic Hamiltonian (like Eq. (105)) and infinite barrier heights [31],
- ii) Γ_6 in-plane dispersions for a Ben Daniel-Duke problem (see section II.3),
- iii) in-plane dispersions of the light particle states of a Kane Hamiltonian with the remote band effects neglected [4] but otherwise V_s, V_p steps of arbitrary height.

The numerical solutions of the 6×6 heterostructure Hamiltonian (i.e. including Γ_6 and Γ_8 bands and the remote band treated up to the second order in \mathbf{k}) have been extensively discussed by Altarelli *et al.* [2, 33, 34]. Under flat band

conditions an exact solution to the problem can be found [32] along the lines previously discussed. One simply expresses the fact that, in each kind of layer, the heterostructure state is a linear combination of propagating (or evanescent) bulk states. The energy ϵ and the in-plane wavevector k_{\perp} having been given, $k_z^{(A)}$ and $k_z^{(B)}$ can then be univoquely determined, together with the envelope function χ which is a 6×1 column vector. The unknown coefficients of the linear combination are then obtained by writing the continuity of χ and $\nabla \chi$ (Eq. (21)) across the A-B interfaces. In figures 21 and 22 we show the results of such calculations performed for a 70 Å - 70 Å and a 100 Å - 100 Å InAs-GaSb superlattice [32]. The first superlattice has simpler in-plane dispersions: there is no overlap

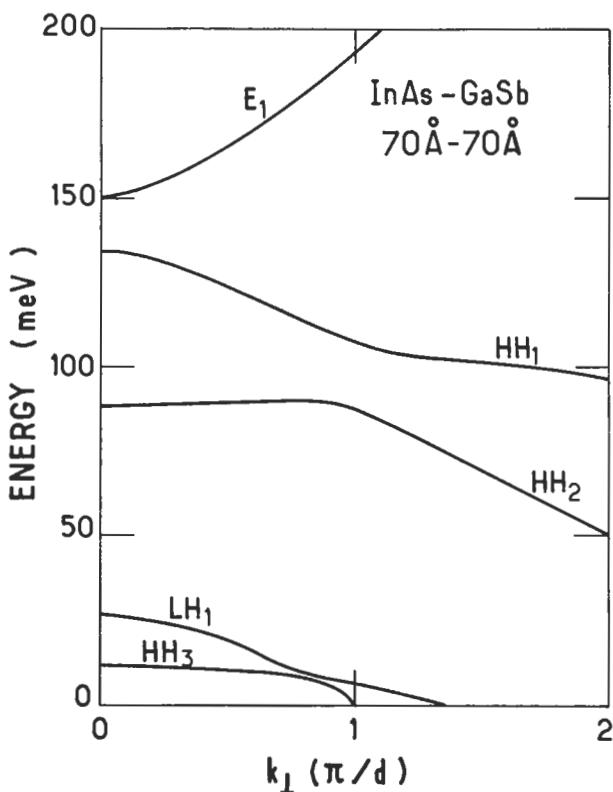


Fig. 21. — In-plane dispersion relations for a 70 Å-70 Å InAs-GaSb superlattice obtained in the 6 band model. $q = 0$; $\epsilon_A = 0.42$ eV; $m_{I_6^{(A)}} = 0.023 m_0$; $\epsilon_B = 0.81$ eV; $\gamma_1^{(A)} = \gamma_1^{(B)} = 3.7$; $\gamma_2^{(A)} = \gamma_2^{(B)} = \gamma_3^{(A)} = \gamma_3^{(B)} = 0.6$. The energy zero has been taken at the bottom of the bulk InAs conduction band. Courtesy of J.M. Berroir [32].

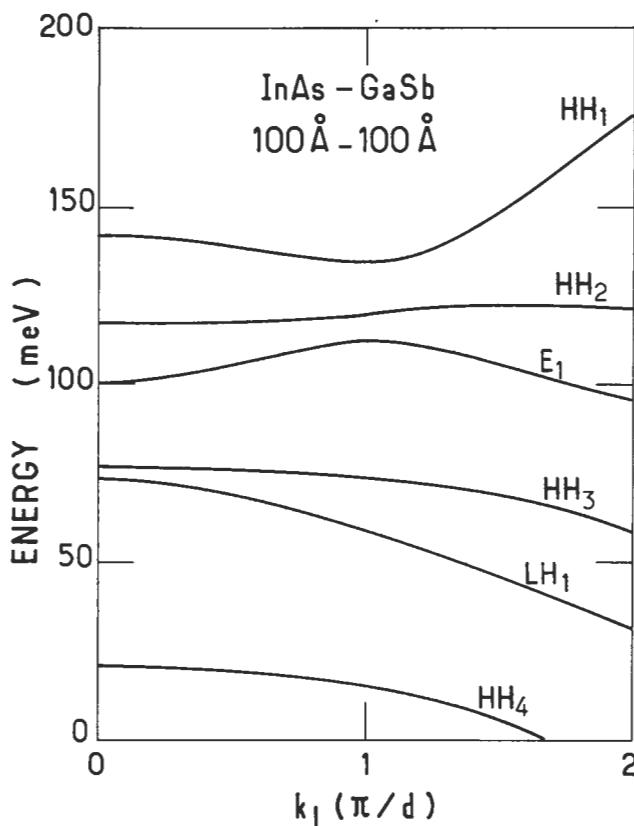


Fig. 22. — In-plane dispersion relations for a 100 Å - 100 Å InAs-GaSb superlattice obtained in the 6 band model. Flat band conditions have been assumed. The band parameters used in the calculations are the same as in figure 21. $q = 0$. Courtesy of J.M. Berroir [32].

between E_1 and HH_1 for any q or any \mathbf{k}_\perp . The 100 Å - 100 Å superlattice is such that at $q = 0$, $\mathbf{k}_\perp = \mathbf{0}$, $E_1 < HH_1$, HH_2 . However, as soon as $\mathbf{k}_\perp \neq \mathbf{0}$ a strong mixing occurs between E_1 and the heavy hole states. As a result, E_1 and HH_1 , HH_2 anti-cross thus preventing the formation of a true semi-metallic phase. Notice that the calculations are not self-consistent in the 100 Å - 100 Å case as the charge redistributions in a superlattice unit cell and their associated band bendings were not taken into account.

The in-plane dispersions of Γ_8 subbands in GaAs-Ga(Al)As heterostructures are the subject of active research [5, 32, 37]. In these materials the coupling with the Γ_6 bands is relatively weak and a parabolic description of the host bands is reasonable. Neglecting once more the inversion-asymmetry splitting, the valence

band Hamiltonian takes the form

$$\mathcal{H}_{\Gamma_8} = \begin{pmatrix} \left\langle \frac{3}{2}, \frac{3}{2} \right| & \left[\begin{array}{cccc} \mathcal{H}_{hh} & c & b & 0 \\ c^* & \mathcal{H}_{lh} & 0 & -b \\ b^* & 0 & \mathcal{H}_{lh} & c \\ 0 & -b^* & c^* & \mathcal{H}_{hh} \end{array} \right] \\ \left\langle \frac{3}{2}, -\frac{1}{2} \right| \\ \left\langle \frac{3}{2}, \frac{1}{2} \right| \\ \left\langle \frac{3}{2}, -\frac{3}{2} \right| \end{pmatrix} \quad (106)$$

where :

$$\mathcal{H}_{hh} = -\frac{1}{2m_0} p_z (\gamma_1 - 2\gamma_2) p_z + V_p(z) - \frac{\hbar^2 k_\perp^2}{2m_0} (\gamma_1 + \gamma_2) \quad (107)$$

$$\mathcal{H}_{lh} = -\frac{1}{2m_0} p_z (\gamma_1 + 2\gamma_2) p_z + V_p(z) - \frac{\hbar^2 k_\perp^2}{2m_0} (\gamma_1 - \gamma_2) \quad (108)$$

$$c(\mathbf{k}_\perp) = \frac{\hbar^2}{m_0} \frac{\sqrt{3}}{2} [\gamma_2(k_x^2 - k_y^2) - 2i\gamma_3 k_x k_y] \quad (109)$$

$$b(\mathbf{k}_\perp, p_z) = \frac{\sqrt{3}}{2} \frac{\hbar}{m_0} (k_x - ik_y)(\gamma_3 p_z + p_z \gamma_3) \quad (110)$$

In equations (107, 110) the γ 's are the Luttinger parameters which describe the coupling between Γ_8 and all the hosts' edges, including Γ_6 . These parameters are, in principle, position-dependent (since they are different in the A and B layers). Consequently the terms of the form γp_z have been symmetrized.

We notice that, as expected, the b and c terms vanish if $\mathbf{k}_\perp \neq \mathbf{0}$. If they were always negligible the Γ_8 Hamiltonian would split into two independent Ben Daniel-Duke Hamiltonians : one for the heavy holes ($m_J = \pm \frac{3}{2}$) and one for the light holes ($m_J = \pm \frac{1}{2}$). This zeroth order approximation, hereafter termed the diagonal approximation, is characterized by a mass reversal effect, i.e. the heavy hole effective mass is heavy along \hat{z} but light in the layer plane whereas the light hole effective mass is light along \hat{z} but heavy in the layer plane. This means that in the diagonal approximation the subbands HH_n and LH_n should cross. Such crossings are suppressed by the non-vanishing b and c terms and replaced by anti-crossings. In the following we shall keep the notations LH_n , HH_n to label the valence subbands, although it should be kept in mind that they retain little physical significance if $\mathbf{k}_\perp \neq \mathbf{0}$. The b and c terms very effectively mix the $\mathbf{k}_\perp \neq \mathbf{0}$ solutions.

The eigenstates of equation (106) are twice degenerate if $V_p(z)$ (and $-e\varphi(z)$ if there is a band bending) and the γ 's are even functions of z . This can be explicitly shown at small \mathbf{k}_\perp by eliminating the $\pm \frac{1}{2}$ components of the eigenstates of equation

(106) to the benefit of the $\pm \frac{3}{2}$ ones and by writing the effective 2×2 Hamiltonian which acts on the $\pm \frac{3}{2}$ components as :

$$\mathcal{H}_{\text{eff}} \begin{bmatrix} \psi_{3/2} \\ \psi_{-3/2} \end{bmatrix} = \begin{bmatrix} \mathcal{H}_+ & W \\ W^+ & \mathcal{H}_- \end{bmatrix} \begin{bmatrix} \psi_{3/2} \\ \psi_{-3/2} \end{bmatrix} = \varepsilon \begin{bmatrix} \psi_{3/2} \\ \psi_{-3/2} \end{bmatrix} \quad (111)$$

where :

$$\mathcal{H}_+ = \mathcal{H}_{hh} + c \frac{1}{\varepsilon - \mathcal{H}_{lh}} c^+ + b \frac{1}{\varepsilon - \mathcal{H}_{lh}} b^+ \quad (112)$$

$$\mathcal{H}_- = \mathcal{H}_{hh} + c^+ \frac{1}{\varepsilon - \mathcal{H}_{lh}} c + b^+ \frac{1}{\varepsilon - \mathcal{H}_{lh}} b \quad (113)$$

$$W = b \frac{1}{\varepsilon - \mathcal{H}_{lh}} c - c \frac{1}{\varepsilon - \mathcal{H}_{lh}} b \quad (114)$$

or more explicitly

$$W = \frac{3\hbar^3}{4m_0^2} (k_x - ik_y) \left\{ (\gamma_3 p_z + p_z \gamma_3) \frac{1}{\varepsilon - \mathcal{H}_{lh}} [\gamma_2 (k_x^2 - k_y^2) - 2i\gamma_3 k_x k_y] - \right. \\ \left. - [\gamma_2 (k_x^2 - k_y^2) - 2i\gamma_3 k_x k_y] \frac{1}{\varepsilon - \mathcal{H}_{lh}} (\gamma_3 p_z + p_z \gamma_3) \right\} \quad (115)$$

At $\mathbf{k}_\perp = \mathbf{0}$ the spectrum of \mathcal{H}_{eff} is twice degenerate. The eigenstates of equation (111) are :

$$\psi_{-3/2}^{(n)} = \begin{bmatrix} 0 \\ \chi_n(z) \end{bmatrix}; \quad \psi_{3/2}^{(n)} = \begin{bmatrix} \chi_n(z) \\ 0 \end{bmatrix} \quad (116)$$

where the χ_n 's are the $\mathbf{k}_\perp = \mathbf{0}$ -heavy hole eigenfunctions. If \mathbf{k}_\perp is small, the \mathbf{k}_\perp -dependent contributions to \mathcal{H}_+ and \mathcal{H}_- and the W term will also be small as they involve at least second and third powers of k_x , k_y respectively. We can therefore attempt a diagonalization of the \mathbf{k}_\perp -dependent contributions within the 2×2 subspace, spanned by the degenerate $\psi_{-3/2}^{(n)}$ and $\psi_{3/2}^{(n)}$ eigenstates, and neglect all the admixtures with the other $\psi_{-3/2}^{(m)}$, $\psi_{3/2}^{(m)}$ states. We readily obtain :

$$\varepsilon = \varepsilon_n \pm |\langle \chi_n | W | \chi_n \rangle| \quad (117)$$

where ε_n is the mean values of \mathcal{H}_+ or \mathcal{H}_- over χ_n . Equation (117) is in fact an implicit equation for ε , inasmuch as ε_n and W are energy dependent and involve an infinite number of terms which describe the indirect coupling between $\psi_{3/2}^{(n)}$ and $\psi_{-3/2}^{(n)}$ via all excursions to the \mathcal{H}_{lh} eigenstates.

The twofold degeneracy between $\psi_{3/2}^{(n)}$ and $\psi_{-3/2}^{(n)}$ will be lifted unless the matrix element $\langle \chi_n | W | \chi_n \rangle$ vanishes, which is exactly what happens when the heterostructure is symmetric with respect to the plane $z = 0$. In fact, under these circumstances

the γ 's, $V_p(z)$, $-e\varphi(z)$, \mathcal{H}_+ , \mathcal{H}_- and \mathcal{H}_{hh} are even in z while W is an odd function of z . As $\chi_n(z)$ may be chosen with a definite parity, it follows that $\langle \chi_n | W | \chi_n \rangle$ vanishes. Thus, at this lower order of perturbation, the twofold degeneracy is not lifted if the heterostructure is symmetrically designed. It is possible to check that this twofold degeneracy is preserved at higher orders of perturbation.

On the other hand, if the heterostructure lacks inversion symmetry (for examples of such heterostructures see Fig. 23), the twofold degeneracy at $\mathbf{k}_\perp = \mathbf{0}$ between $\psi_{3/2}^{(n)}$ and $\psi_{-3/2}^{(n)}$ is already lifted at the lowest order (see equation (117)) and consequently at any order in the coupling between the unperturbed $\psi^{(n)}$ doublet and the other $\psi^{(m)}$ doublets.

To summarize, the lifting of the twofold (Kramers) degeneracy requires that :

- i) $\mathbf{k}_\perp \neq \mathbf{0}$,
- ii) $V_p(z)$ and/or $-e\varphi_{sc}(z)$, $\gamma_i(z)$ lack of inversion symmetry,
- iii) the spin-orbit coupling is non vanishing.

The latter condition is only implicit in our calculations as the Γ spin-orbit energies Δ_A , Δ_B do not appear explicitly in equations (111-115). This is because they are

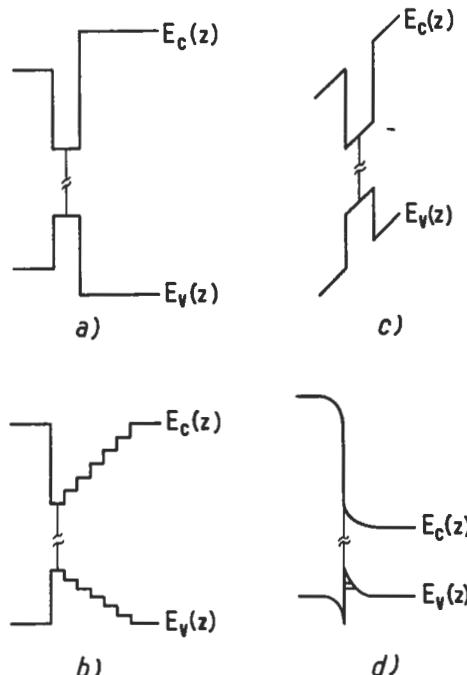


Fig. 23. — Band edges profiles of four different heterostructures lacking inversion-symmetry. a) asymmetric rectangular quantum well ; b) pseudo-sawtooth quantum well ; c) rectangular quantum well tilted by an external electric field ; d) modulation-doped p-type heterojunction.

assumed to be very large compared to the Γ_8 energies. On the other hand if both Δ_A , Δ_B vanish, the twofold degeneracy at finite \mathbf{k}_\perp is immediately recovered. It then coincides with the spin degeneracy ($\sigma_z = \pm \frac{1}{2}$) which obviously can not be lifted by scalar potentials like $V_p(z)$ and $-e\varphi_{sc}(z)$.

Let us finally remark that even if the heterostructure is symmetrically designed, lifting of the Kramers degeneracy should still take place due to the lack of inversion-symmetry of the host unit cells. This effect (linear \mathbf{k} terms in \mathcal{H}_{Γ_8}) is however small (and has been neglected in the previous analysis) compared with the lifting of the Kramers degeneracy obtained by asymmetrically designing the heterostructure. Notice also that asymmetrical situations should experimentally be the rule rather than the exception since growth processes often lead to some asymmetry between the two GaAs-Ga(Al)As interfaces.

Unless \mathbf{k}_\perp is very small, the effects of the b and c in \mathcal{H}_{Γ_8} cannot be treated by second order perturbation but require numerical diagonalization. Again an exact solution [32] is possible if flat band conditions prevail. Other treatments [5, 33-37] which are approximate, are capable of handling band bending effects. A variational scheme has been proposed by Altarelli [2, 34] which also incorporates the differences between the γ parameters in the A and B layers. If these parameters are assumed to be the same in both kinds of layers another computational scheme becomes possible. This approximation is reasonably justified in GaAs-Ga_{1-x}Al_xAs heterostructures. Let us denote by $\xi_m(z)$, $\varphi_n(z)$, the eigenfunctions of \mathcal{H}_{hh} and \mathcal{H}_{lh} at $\mathbf{k}_\perp = \mathbf{0}$ respectively. The eigenstates of \mathcal{H}_{Γ_8} at finite \mathbf{k}_\perp , can then be expanded [5, 35, 37] in the following manner

$$\psi = \begin{bmatrix} \sum_m \alpha_m^+ \xi_m(z) \\ \sum_n \beta_n^- \varphi_n(z) \\ \sum_n \beta_n^+ \varphi_n(z) \\ \sum_m \alpha_m^- \xi_m(z) \end{bmatrix} \quad (118)$$

If the summations in equation (118) run over all the $\mathbf{k}_\perp = \mathbf{0}$ -heavy and light hole states, ψ is an exact solution of the problem. In practice, M heavy hole states and N light hole states are retained and \mathcal{H}_{Γ_8} is diagonalized inside this $2(M + N)$ -dimensional subspace.

To illustrate the salient features of Γ_8 in-plane dispersions, we have presented in figures 24 to 26 the results of calculations using equation (118) as the eigenfunction at finite \mathbf{k}_\perp in GaAs-Ga_{0.7}Al_{0.3}As single quantum wells. The expansions over m and n have been restricted to the bound heavy and light states of the quantum wells. The following parameters were used : $V_p = -0.15$ eV ; $\gamma_1 = 0.85$; $\gamma_2 = 2.1$; $\gamma_3 = 2.9$.

In addition, the calculations were performed under the axial approximation [33, 34] which amounts to replacing γ_2 and γ_3 in equation (109) by their arithmetic average. This renders the in-plane dispersions isotropic, while retaining the correct confinement energies at $\mathbf{k}_\perp = \mathbf{0}$ [34].

Figure 24 displays the calculated valence subbands of two quantum wells ($L = 100 \text{ \AA}$ and $L = 150 \text{ \AA}$ respectively) under flat band conditions. As the heterostructure displays centro-symmetry with respect to the middle of the well each of the levels are twice degenerate. The dashed lines are the results obtained in the diagonal approximation ($b = c = 0$ in \mathcal{H}_{Γ_8}). The mass reversal effect causes HH_1 to cross LH_1 . The off-diagonal terms in \mathcal{H}_{Γ_8} replace this crossing by an anticrossing which is more pronounced at $L = 150 \text{ \AA}$ than at $L = 100 \text{ \AA}$. This is because the $\mathbf{k}_\perp = \mathbf{0}$ -levels are closer in the former situation than in the latter. The mixing between the $\mathbf{k}_\perp = \mathbf{0}$ -levels is very strong, resulting in highly non-parabolic subband dispersions. In particular, one notices that the LH_1 subband displays an **electron-like** curvature in the vicinity of $\mathbf{k}_\perp = \mathbf{0}$ [34]. This is due to the prevalent coupling between LH_1 and HH_2 (and other states of lower energies) over the repulsive coupling between HH_1 and LH_1 which pushes LH_1 towards lower energies. The electron-like mass of LH_1 is lighter in the 150 \AA than in the 100 \AA thick quantum well, a feature which again originates from the decreasing energy separations between the $\mathbf{k}_\perp = \mathbf{0}$ -states at larger L .

It is very difficult to predict general trends for the in-plane dispersion of a given heterostructure. The only model-independent conclusion is the increase in the

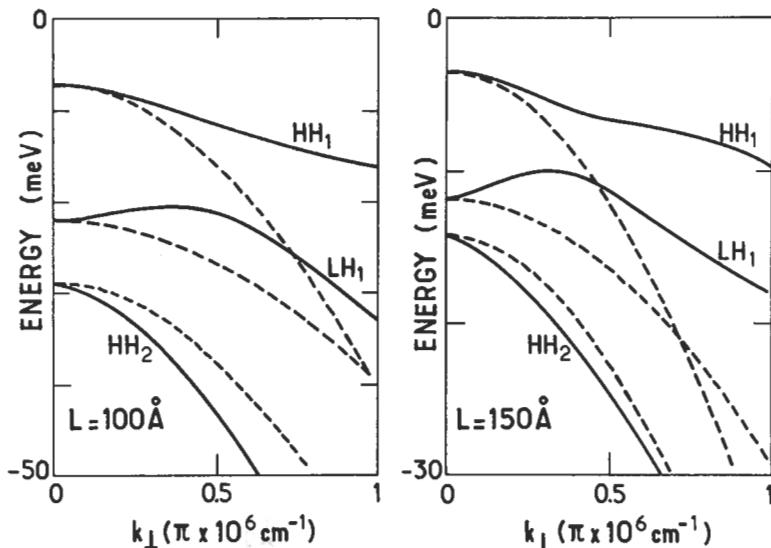


Fig. 24. — In-plane dispersion relations of the valence subbands of $\text{GaAs-Ga}_{0.7}\text{Al}_{0.3}\text{As}$ quantum wells ($L = 100 \text{ \AA}$ and $L = 150 \text{ \AA}$ respectively). The dashed lines are the subband dispersions obtained in the diagonal approximation.

HH_1 in-plane mass over the value given by the diagonal approximation: as HH_1 is the ground hole state the only result of off-diagonal perturbations is to lower its (hole) energy, thereby increasing its in-plane mass.

In figure 25 we have presented the L -dependence of the in-plane effective masses of the four lowest lying hole subbands (when existing) in GaAs-Ga_{0.7}Al_{0.3}As quantum wells in the vicinity of $k_{\perp} = 0$. These masses were obtained by fitting the calculated in-plane dispersion to the quadratic functions of k_{\perp}^2 . Notice that the validity of such a quadratic expansion narrows when L increases.

To illustrate the lifting of the Kramers degeneracy in asymmetric heterostructures, figure 26 shows the in-plane subbands of a GaAs-Ga_{0.7}Al_{0.3}As quantum well ($L = 100$ Å) tilted by an electric field F applied parallel to \hat{z} ($F = 10^5$ V/cm). Such field configurations and strengths are obtained, for instance, by inserting the quantum well into the intrinsic part of a reverse-biased p-i-n junction [38]. The centres of gravity of the two components of the various Kramer doublets follow trends which are similar to the ones already analysed in rectangular quantum wells, e.g. anticrossing between HH_1 and LH_1 , electron-like behaviour of LH_1 near $k_{\perp} = 0$... The splitting due to the externally controlled inversion-asymmetry is significant: 3-5 meV for HH_1 and LH_1 ... This is in fact comparable in magnitude to what was calculated for modulation-doped p-type GaAs-Ga(Al)As heterojunctions [33-37]. In the latter heterostructures, the inversion-asymmetry splitting is clearly evidenced by

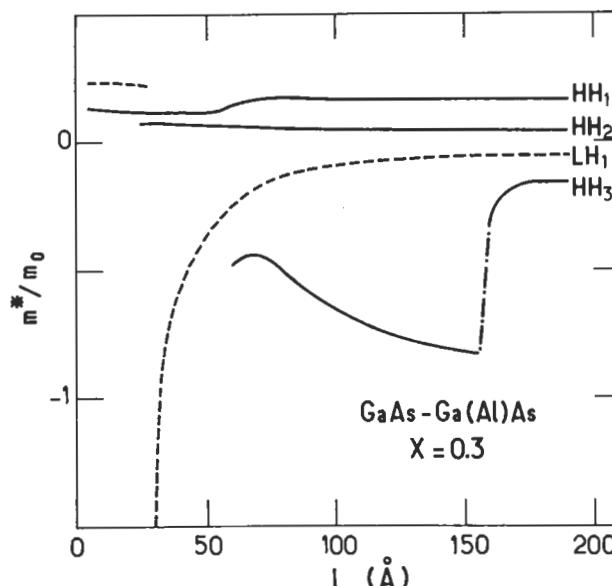


Fig. 25. — The in-plane hole effective masses near $k_{\perp} = 0$ of the four lowest lying hole subbands of GaAs-Ga_{0.7}Al_{0.3}As quantum wells are plotted *versus* the GaAs slab thickness L .

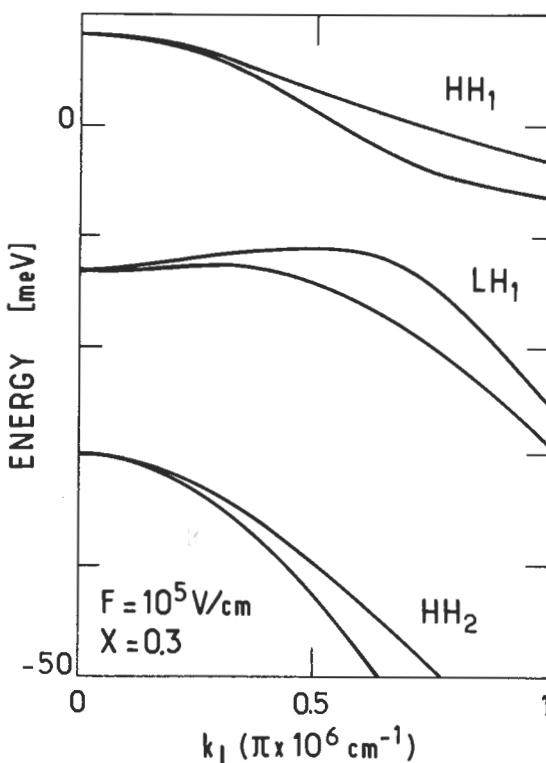


Fig. 26. — In-plane dispersions of the valence subbands of a GaAs-Ga_{0.7}Al_{0.3}As quantum well ($L = 100 \text{ \AA}$) subjected to an external electric field $F \parallel z$ ($F = 10^5 \text{ V/cm}$). The twofold Kramers degeneracy prevailing at $k_{\perp} = 0$ is lifted at $k_{\perp} \neq 0$.

beating effects in the Shunikov-de Haas oscillations as well as by the existence of two hole effective masses pertaining to the HH_1 doublet.

The study of the in-plane dispersion relations in heterostructures is presently being very actively pursued both experimentally and theoretically. We wish to point out that, apart from the numerical difficulties linked to the complicated nature of the valence Hamiltonian, the lack of knowledge of **bulk** valence parameters often makes it difficult to obtain precise results. Even in the most studied GaAs-Ga(Al)As system, the γ parameters of the Ga(Al)As alloys have not, to our knowledge, been measured accurately. *A fortiori*, in less studied systems like Ga_{0.47}In_{0.53}As-InP, GaSb-AlSb, Ga_{0.47}In_{0.53}As-Al_{0.48}In_{0.52}As, HgTe-CdTe, the knowledge of these parameters is even more questionable. This casts doubts on the quantitative reliability of valence subband calculations, although qualitatively the gross features are relatively well understood.

Appendix A.

Boundary conditions and stationary states.

Let us consider a Ben Daniel-Duke Hamiltonian for electrons

$$\mathcal{H} = \frac{1}{2} p_z \frac{1}{m(z)} p_z + V_s(z) + \frac{\hbar^2 k_\perp^2}{2m(z)} \quad (A1)$$

where :

$$m(z) = \begin{cases} m_A & z > 0 \\ m_B & z < 0 \end{cases} \quad V_s(z) = \begin{cases} 0 & z > 0 \\ V_s & z < 0 \end{cases} \quad (A2)$$

We are interested in checking the validity of the two boundary conditions for $\frac{\partial \psi}{\partial z}$

$$\frac{\partial \psi}{\partial z} \text{ continuous or } \frac{1}{m(z)} \frac{\partial \psi}{\partial z} \text{ continuous} \quad (A3)$$

given that the envelope function $\psi(z)$ is continuous everywhere.

Let us consider the time derivative of the probability amplitude

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial t} [\psi^*(z, t) \psi(z, t)] \quad (A4)$$

By using the Schrödinger equation

$$i\hbar \frac{\partial \psi(z, t)}{\partial t} = H\psi(z, t) \quad (A5)$$

we easily find that

$$\frac{\partial \rho}{\partial t} = \frac{i\hbar}{2} \frac{\partial}{\partial z} \left\{ \psi^*(z, t) \frac{1}{m(z)} \frac{\partial \psi}{\partial z}(z, t) - \psi(z, t) \frac{1}{m(z)} \frac{\partial \psi^*}{\partial z}(z, t) \right\} \quad (A6)$$

In a stationary state $\psi(z, t)$ factorizes into

$$\psi(z, t) = \psi(z) \exp \left(-\frac{i \varepsilon_n t}{\hbar} \right) \quad (A7)$$

where ε_n is the eigenenergy. Consequently $\frac{\partial \rho}{\partial t}$ vanishes, which leads to :

$$G(z) = \psi^*(z) \frac{1}{m(z)} \frac{\partial \psi}{\partial z}(z) - \psi(z) \frac{1}{m(z)} \frac{\partial \psi^*}{\partial z}(z) = \text{constant} \quad (A8)$$

Let us apply the conservation of $G(z)$ on each side of the A-B interface. We obtain :

$$\begin{aligned} \psi^*(0^-) \frac{1}{m_B} \frac{\partial \psi}{\partial z} (z = 0^-) - \psi(0^-) \frac{1}{m_B} \frac{\partial \psi}{\partial z} (z = 0^-) = \\ \psi^*(0^+) \frac{1}{m_A} \frac{\partial \psi}{\partial z} (z = 0^+) - \psi(0^+) \frac{1}{m_A} \frac{\partial \psi}{\partial z} (z = 0^+) \end{aligned} \quad (A9)$$

It can be seen that the boundary conditions

$$\psi(0^-) = \psi(0^+) ; \frac{1}{m_B} \frac{\partial \psi}{\partial z} (z = 0^-) = \frac{1}{m_A} \frac{\partial \psi}{\partial z} (z = 0^+) \quad (A10)$$

are compatible with (A9), while

$$\psi(0^-) = \psi(0^+) ; \frac{\partial \psi}{\partial z} (z = 0^-) = \frac{\partial \psi}{\partial z} (z = 0^+) \quad (A11)$$

do not ensure that (A9) is fulfilled.

Consequently the "traditional" boundary conditions (A11) have to be discarded for Ben Daniel-Duke problems since they do not ensure the conservation of $G(z)$ and ultimately the stationary nature of the eigenstates.

In the case of a multiband Kane Hamiltonian with remote band effects included up to the second order in k (Eq. (20)), the same kind of reasoning can be made. The only difference is that $\rho(t)$ should be written

$$\rho(t) = \sum_i \chi_i^*(z, t) \chi_i(z, t) \quad (A12)$$

where i labels the band edges which are included in the analysis and $\chi_i(z, t)$ the i^{th} component of the column vector envelope function $\chi(z, t)$. Proceeding as in (A6-A9) we find that the boundary conditions

$$\chi(z) , \frac{\partial \chi}{\partial z} \text{ continuous at the interfaces} \quad (A13)$$

are incompatible with the stationary nature of the states and thus should be discarded, whereas the conditions $\chi(z) , \mathcal{A}\chi$ where \mathcal{A} has been defined in equations (21, 22) are compatible with the stationary requirement.

Appendix B.

Coupling between light and heavy particle states due to inversion asymmetry splitting in bulk materials. Qualitative aspects.

The hosts' unit cells of actual heterostructures lack inversion-symmetry since the two atom basis of the Bravais lattice underlying the zinc-blende structure consists of two different elements (e.g. Ga and As in the case of bulk GaAs).

The effects associated with the inversion-asymmetry of the zinc-blende lattice have been thoroughly analysed in bulk materials (see e.g. [23, 39]). When combined with non zero spin-orbit coupling and the wavevector \mathbf{k} , they produce terms in the dispersion of Γ_8 bands which are linear in \mathbf{k} . When \mathbf{k} is parallel to [111] these terms displace the maximum of the valence band from the zone centre to $[k_0, k_0, k_0]$ and raise the valence band maximum by an energy δ . In addition, with the exception of the [001] direction, the inversion asymmetry splitting removes the Kramers degeneracy of the Γ_8 -related levels. The strength of the inversion-asymmetry effects is very small, more so in III-V than in II-VI compounds, to the point where a considerable scatter in the published values of δ and k_0 exists in all the III-V materials. A value of $\delta = 1$ meV seems to be an upper bound.

Our purpose in this Appendix is to discuss some of the effects of the host inversion asymmetry on the Γ_8 -related heterostructures states. A more complete analysis has already been given by Bychkov and Rashba [40] using the envelope function framework and by Schulman and Chang [6] using the tight-binding approach. For the sake of simplicity let us assume that parabolic Γ_8 bands exist in the hosts, that the growth direction is [001] and that $\mathbf{k}_\perp = (k_x, k_y) = \mathbf{0}$. We shall also neglect the eventual differences in the inversion-asymmetry constants of the different kinds of layers. Under these assumptions the Hamiltonian for Γ_8 -related levels becomes :

$$\mathcal{H}_{\Gamma_8} = \begin{bmatrix} \alpha p_z & 0 & 0 & 0 \\ -\frac{1}{2m_0} p_z (\gamma_1 - 2\gamma_2) p_z + V_p(z) & \alpha p_z & 0 & 0 \\ \alpha^* p_z & -\frac{1}{2m_0} p_z (\gamma_1 + 2\gamma_2) p_z + V_p(z) & 0 & 0 \\ 0 & 0 & -\frac{1}{2m_0} p_z (\gamma_1 + 2\gamma_2) p_z + V_p(z) & \alpha p_z \\ 0 & 0 & \alpha^* p_z & -\frac{1}{2m_0} p_z (\gamma_1 - 2\gamma_2) p_z + V_p(z) \end{bmatrix} \quad (B1)$$

In (B1) the energy zero has been taken at the Γ_8 point in the A material, γ_1 and γ_2 are Luttinger parameters which eventually take different values in A and B layers, and the terms in αp_z are the inversion-asymmetry contributions to the Γ_8 Hamiltonian [23, 39].

Despite our $\mathbf{k}_\perp = \mathbf{0}$ assumption, the heterostructure eigenstates are no longer eigenstates of J_z if α , i.e. the inversion-asymmetry, is nonvanishing. (B1) shows that the $m_J = \pm \frac{3}{2}$ components are in fact admixed with the $m_J = \mp \frac{1}{2}$ components. The form of the coupling term αp_z also indicates that the inversion-asymmetry contributions will admix light and heavy particle states, which at $\alpha = 0$ are of opposite parities. For instance, in the case of a single rectangular quantum well, the HH_2 state will be coupled to the LH_1 state.

Finally, we notice in (B1) that the eigenstates will remain twice degenerate (Kramers degeneracy) as the 4×4 Hamiltonian has been split into two identical and uncoupled blocks. This decoupling is no longer valid if $\mathbf{k}_\perp \neq \mathbf{0}$ and/or if the growth occurs in a symmetry direction lower than [001].

Exact solutions of (B1) exist if $V_p(z)$ is piecewise constant. The eigenstates of (B1) can be written as linear combinations of bulk solutions in the A and B layers. The matching conditions at the hetero-interfaces with $\alpha \neq 0$ are the same as those obtained with $\alpha = 0$ since, by integrating (B1) across an interface, the α -dependent contributions amount to re-writing the continuity of the envelope functions. By using the appropriate boundary conditions at large $|z|$ (Bloch theorem in the case of superlattices, exponential decay in the case of quantum wells), one ends up with a transcendental equation linking the energy ϵ to the corresponding wavevectors k_A, k_B in each kind of layer.

Bearing in mind a qualitative estimate of the inversion-asymmetry effects, let us consider a single GaAs-Ga(Al)As rectangular quantum well. At $\alpha = 0$, the eigenstates of \mathcal{H}_{Γ_8} are either the heavy hole states HH_n , each twice degenerate, or the light hole states LH_n , which are also each twice degenerate. Owing to the mass ratio $\frac{\gamma_1 - 2\gamma_2}{\gamma_1 + 2\gamma_2} > 4$ in GaAs it is easy to work out that at least two decoupled states

may be degenerate if $\alpha = 0$; namely HH_2 and LH_1 . In fact, $HH_2 < LH_1$ in thick quantum wells whereas in narrow wells ($L \rightarrow 0$) HH_2 fades away in the valence continuum while LH_1 remains bound in the well. Thus, for certain critical thickness L_c , HH_2 and LH_1 are degenerate at the zeroth order in α .

To qualitatively analyse the α -induced coupling at $L = L_c$, we diagonalize the inversion asymmetry contribution between HH_2 and LH_1 , while neglecting all the other α -induced couplings between these two states and LH_n, HH_m . We obtain a lifting of the degeneracy at $L = L_c$ which is given by

$$\Delta\epsilon = 2 |\langle HH_2 | \alpha p_z | LH_1 \rangle| \quad (B2)$$

For an order of magnitude estimate we use

$$\langle z | HH_2 \rangle = \sqrt{\frac{2}{L_c}} \sin \left(\frac{2\pi z}{L_c} \right) \quad (B3)$$

$$\langle z | LH_1 \rangle = \sqrt{\frac{2}{L_c}} \cos \left(\frac{\pi z}{L_c} \right) \quad (B4)$$

and obtain :

$$\Delta\epsilon = \frac{16}{3} \alpha \frac{\hbar}{L_c} \quad (B5)$$

The order of magnitude of α is now related to the quantity δ , which is the shift of the top of the valence band in bulk GaAs by

$$\alpha \sim \sqrt{\frac{2\delta}{(\gamma_1 - 2\gamma_2)m_0}} \quad (B6)$$

which gives

$$\Delta\epsilon \sim \frac{16}{3\pi} \sqrt{\delta HH_2} \quad (B7)$$

i.e. $\Delta\epsilon \sim$ few meV's.

Conclusions qualitatively similar to those obtained in this analysis were reached by Schulman and Chang [6] by means of empirical tight binding calculations. In such a model, the inversion-asymmetry splitting is automatically taken into account due to the difference in the atomic wavefunctions of the anions and cations from which the heterostructure states are built.

References

- [1] S.R. WHITE and L.J. SHAM, *Phys. Rev. Lett.* **47** (1981) 879.
- [2] M. ALTARELLI, *Phys. Rev.* **B28** (1983) 842 ; *Physica* **117 & 118 B** (1983) 747.
- [3] M.F.H. SCHUURMANS and G.W.'t HOOFT, *Phys. Rev.* **B31** (1985) 8041.
- [4] G. BASTARD, *Phys. Rev.* **B24** (1981) 5693 ; *Phys. Rev.* **B25** (1982) 7584.
- [5] G. BASTARD and J.A. BRUM in *Quantum Well Structures, Physics and Applications*, *IEEE J. Quant. Electr.* **QE22** (1986) 1625.
- [6] J.N. SCHULMAN and YIA-CHUNG CHANG, *Phys. Rev.* **B24** (1981) 4445 ; *Phys. Rev.* **B31** (1985) 2056.
- [7] YIA-CHUNG CHANG and J.N. SCHULMAN, *J. Vac. Sci. Technol.* **21** (1982) 540.
- [8] M. JAROS and K.B. WONG, *J. Phys.* **C17** (1984) L 765.
- [9] M. JAROS, K.B. WONG and M.A. GELL, *Phys. Rev.* **B31** (1985) 1205.
- [10] see e.g. M. QUILLEC, L. GOLDSTEIN, G. LE ROUX, J. BURGEAT and J. PRIMOT, *J. Appl. Phys.* **55** (1984) 2904.
- [11] see e.g. J.H. NEAVE, P.J. DOBSON, B.A. JOYCE and J. ZHANG, *Appl. Phys. Lett.* **47** (1985) 100.
- J.M. VAN HOVE, C.S. LENT, P.R. PUKITE and P.I. COHEN, *J. Vac. Sci. Technol.* **B1** (1983) 741.
- [12] G.C. OSBOURN, *J. Vac. Sci. Technol.* **B1** (1983) 379.
- [13] P. VOISIN in *Two dimensional Systems, Heterostructures and Superlattices* edited by G. Bauer, F. Kuchar and H. Heinrich, Springer Series in *Solid State Sci.* **53** (Springer Verlag, Berlin) 1984, p. 192.
- [14] J.Y. MARZIN in *Heterojunctions and Semiconductor Superlattices* edited by G. Allan, G. Bastard, N. Boccara, M. Lannoo and M. Voos (Springer Verlag, Berlin) 1986, p. 161.
- [15] J.N. SCHULMAN and T.C. MC GILL in *Synthetic Modulated Structure Materials* edited by L.L. Chang and B.C. Giessen (Academic Press, New York) 1984.
- [16] see e.g. E.J. FANTNER and G. BAUER in *Two Dimensional Systems, Heterostructures and Superlattices* edited by G. Bauer, F. Kuchar and H. Heinrich, Springer Series in *Solid State Sci.* **53**, p. 207 (Springer Verlag, Berlin) 1984.
- [17] W. HARRISON, *J. Vac. Sci. Technol.* **14** (1977) 1016 ; *J. Vac. Sci. Technol.* **B3** (1985) 1231.
- [18] W.R. FRENSLEY and H. KROEMER, *Phys. Rev.* **B16** (1977) 2642.
- [19] J. TERSOFF, *Phys. Rev.* **B30**, 4874 (1984) ; *Phys. Rev. Lett.* **52** (1984) 465 ; *Phys. Rev.* **B32** (1985) 6968.
- [20] G. DUGGAN, *J. Vac. Sci. Technol.* **B3** (1985) 1224.
- [21] H. KROEMER, *Surf. Sci.* **132** (1983) 543 ; *J. Vac. Sci. Technol.* **B2** (1984) 433 and in *Molecular Beam Epitaxy and Heterostructures* edited by L.L. Chang and K. Ploog (Martinus Nijhoff, Dordrecht) 1985, p. 331.

- [22] J.M. LUTTINGER, *Phys. Rev.* **102** (1956) 1030.
- [23] G.L. BIR and G.E. PIKUS *Symmetry and Strain-induced Effects in Semiconductors* (Wiley, New York) 1974.
- [24] G. DRESSELHAUS, A.F. KIP and C. KITTEL, *Phys. Rev.* **98** (1955) 368.
- [25] D.J. BEN DANIEL and C.B. DUKE, *Phys. Rev.* **152** (1966) 683.
- [26] see e.g. M.H. WEILER in *Semiconductor and Semimetals* vol. 16 edited by R.K. Willardson and A.C. Beer (Academic Press, New York) 1981.
- [27] Y. GULDNER, G. BASTARD, J.P. VIEREN, M. VOOS, J.P. FAURIE and A. MILLION, *Phys. Rev. Lett.* **51** (1983) 907.
- [28] YIA-CHUNG CHANG, J.N. SCHULMAN, G. BASTARD, Y. GULDNER, M. VOOS, *Phys. Rev. B* **31** (1985) 2557.
- [29] Y.R. LIN LIU and L.J. SHAM, *Phys. Rev. B* **32** (1985) 5561.
- [30] see e.g. L. ESAKI in *Narrow Gap Semiconductors. Physics and Applications Lecture Notes Phys.* **133** (Springer Verlag, Berlin) 1980.
- [31] S.S. NEDOREZOV, *Sov. Phys. Sol. State* **12** (1971) 1814.
- [32] J.M. BERROIR, 1985, unpublished
- [33] U. EKENBERG and M. ALTARELLI, *Phys. Rev. B* **30** (1984) 3569.
- [34] M. ALTARELLI in *Semiconductor Superlattices and Heterojunctions*, edited by G. Allan, G. Bastard, N. Boccara, M. Lannoo and M. Voos (Springer Verlag, Berlin) 1986, p. 12.
- [35] D.A. BRODO and L.J. SHAM, *Phys. Rev. B* **31** (1985) 888.
- [36] E. BANGERT and G. LANDWEHR, *Superlattices and Microstructures* **1** (1985) 363.
- [37] T. ANDO, *J. Phys. Soc. Japan* **54** (1985) 1528.
- [38] D.A.B. MILLER, D.S. CHEMLA, T.C. DAMEN, A.C. GOSSARD, W. WIEGMANN, T.H. WOOD and A.C. BURRUS, *Phys. Rev. B* **32** (1985) 1043.
- [39] E.O. KANE, *J. Phys. Chem. Sol.* **1** (1957) 249.
- [40] YU. A. BYCHKOV and E.I. RASHBA in the *Proceedings of the 17th International Conference on the Physics of Semiconductors*, San Francisco (1984) edited by J.D. Chadi and W.A. Harrison (Springer Verlag) 1985, p. 321.

CHAPTER IV

Coulombic bound states and interface defects in heterostructures

As with the bulk materials, the single or multiple heterojunctions contain various kinds of defects. In addition to interface defects, which are specific to heterostructures, one finds in these materials neutral or charged impurities which give rise to short range or long range potentials. In this chapter we will discuss the properties of the bound states supported on one hand by Coulombic potentials (impurity, exciton) and on the other, by the short-range interface defects.

COULOMBIC BOUND STATES

I. Qualitative aspects.

I.1 BULK HYDROGENIC IMPURITIES. — In an infinite crystalline material it is unnecessary to know precisely which unit cell hosts the impurity ion : two impurity sites \mathbf{R}_i and \mathbf{R}_j , which differ by a lattice vector are equivalent. The binding energy of the impurity located at \mathbf{R}_i will be the same as the binding energy of the impurity located at \mathbf{R}_j . Note that this is not the case if \mathbf{R}_i and \mathbf{R}_j are two unequivalent lattice sites of a given elementary cell (e.g. see the case of a substitutional Si atom in the GaAs lattice. The silicon gives rise to an acceptor level if it substitutes an As atom, or a donor level if it substitutes a Ga atom). For non-degenerate and isotropic bands with parabolic dispersion relations the envelope functions of the impurity states fulfill the equation [1] :

$$\left[\frac{p^2}{2m^*} - \frac{e^2}{\kappa r} \right] F(\mathbf{r}) = \varepsilon F(\mathbf{r}), \quad (1)$$

where m^* is the carrier effective mass and κ the relative dielectric constant of the material. The ground bound state is the 1S hydrogenic wavefunction :

$$F_{1S}(\mathbf{r}) = \frac{1}{(\pi a_0^3)^{1/2}} \exp(-r/a_0) \quad (2)$$

where a_0 is the effective three-dimensional Bohr radius :

$$a_0 = \frac{\kappa \hbar^2}{m^* e^2} = 0.53 \times \kappa \times \frac{m_0}{m^*} \text{ \AA.} \quad (2')$$

In GaAs $\kappa = 12.85$, $m^* = 0.067 m_0$ and $a_0 \sim 100 \text{ \AA}$. The 1S state has a binding energy R_0 given by

$$R_0 = \frac{1}{2} \frac{m^* e^4}{\kappa^2 \hbar^2} = \frac{13600}{\kappa^2} \times \frac{m^*}{m_0} \text{ meV} \quad (3)$$

In GaAs $R_0 \sim 5.5 \text{ meV}$.

I.2 IMPURITIES IN HETEROSTRUCTURES. — In contrast with bulk materials, the heterostructures are characterized by the lack of translational invariance along the growth axis. Thus, the impurity binding energy explicitly depends on the precise location of the impurity. It is important to know where the Si atom is placed in GaAs-Ga_{1-x}Al_xAs heterostructure, and the binding energy of the donor state associated with Si will depend on whether the Si atom sits at the center of the GaAs well, on-edge of this well or far away in the Ga_{1-x}Al_xAs barrier.

The second important feature of Coulombic problems in heterostructures is the variation of the impurity binding energy with the characteristic dimension of the well (e.g. L , the thickness of a rectangular quantum well). When the quantum well thickness decreases, the impurity binding energy increases, as long as the penetration of the unperturbed quantum well wavefunction $\chi_1(z)$ in the barrier remains small. This may seem surprising at first sight since we intuitively associate extra kinetic energy with the localization of a particle in a finite region of space. Moreover, it is actually true that the energy position of the ground bound state of the impurity $\epsilon_d(L, z_i)$ increases as L decreases, when it is measured from some fixed reference (e.g. the bottom of the well). However, the onset of the impurity continuum, which (for donors) coincides with the energy of the quantum well bound state $E_1(L)$, also moves up with decreasing L . The binding energy $E_1(L) - \epsilon_d(L, z_i)$ finally increases since, in the bound impurity states, the carrier is kept near to the attractive centre by the quantum well walls and thus experiences a larger potential energy than in the absence of a quantum well. On the other hand the onset of the impurity continuum $E_1(L)$ does not benefit from any extra potential energy gain, as the continuum states are hardly affected by the Coulombic potential. The previous considerations outlined for the impurity problem, also apply to the exciton in quantum wells : in excitons the electron and hole relative motion is described by a Hamiltonian which is similar to that of the impurity.

In figure 1 we have sketched the shape of the impurity wavefunctions, keeping the in-plane distance between the carrier and the attractive centre ρ ($\rho = \sqrt{x^2 + y^2}$) equal to zero. Several quantum well thicknesses (upper part of Fig. 1) and several impurity positions in a thick well lower part of Fig. 1) have been considered. For thick wells ($L \gg a_0$) the on-centre impurity wavefunction resembles that of the bulk 1S states. On the other hand, the on-edge impurity wavefunction approaches the shape of a $2p_z$ wavefunction if $\frac{L}{a_0} \gg 1$. This is because the barrier potential forces the impurity wavefunction to almost vanish at $z = \pm \frac{L}{2}$. If $\frac{L}{a_0} \ll 1$ and if the barrier height

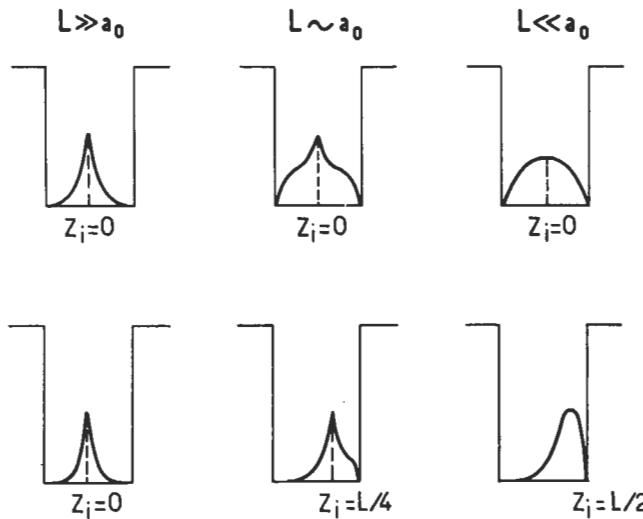


Fig. 1. — Schematic evolution of the shape of the impurity wavefunction in a quantum well of decreasing well thickness for an on-centre impurity (upper part) and with the impurity position in a thick well (lower part).

is very large, the electron z motion becomes forced by the quantum well potential. Thus, along the z axis the impurity wavefunction looks like the ground state wavefunction of the well $\chi_1(z)$.

II. Approximate solutions of the hydrogenic impurity problem.

II.1 FORMULATION OF THE PROBLEM. — In all that follows we will consider donor-like impurities unless otherwise specified. The conduction bands of both host materials of the quantum well are assumed to be isotropic and parabolic in \mathbf{k} . We neglect the effective mass jumps at the interfaces as well as the differences in the relative dielectric constants of the two host materials. The impurity envelope functions are the solutions of the effective Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + V_{\text{imp}} = \frac{p_z^2}{2m^*} + \frac{p_x^2 + p_y^2}{2m^*} + V_b Y \left[z^2 - \frac{L^2}{4} \right] - \frac{e^2}{\kappa \sqrt{\rho^2 + (z - z_i)^2}} \quad (4)$$

where V_b is the barrier height and $Y(x)$ the step function ($Y(x) = 1$ if $x > 0$; $Y(x) = 0$ if $x < 0$). The impurity position along the growth axis is z_i . The (x, y) origin is at the impurity site because all the impurity positions are equivalent in the layer plane. ρ is the projection of the electron position vector in the layer plane ($\rho = (x, y)$).

In the absence of the impurity the eigenstates of \mathcal{H}_0 are separable in (x, y) and z :

$$\mathcal{H}_0 |\nu, \mathbf{k}_\perp\rangle = \left(E_\nu + \frac{\hbar^2 k_\perp^2}{2m^*} \right) |\nu, \mathbf{k}_\perp\rangle, \quad (5)$$

where ν labels the quantum well eigenstates (energy E_ν), i.e. the quantum well bound ($E_\nu < V_b$) and unbound states ($E_\nu > V_b$) and $\mathbf{k}_\perp = (k_x, k_y)$. Since the $|\nu, \mathbf{k}_\perp\rangle$ basis is complete we may always expand the impurity wavefunction $|\psi_{\text{loc}}\rangle$ in the form :

$$|\psi_{\text{loc}}\rangle = \sum_{\nu, \mathbf{k}_\perp} c(\nu, \mathbf{k}_\perp) |\nu, \mathbf{k}_\perp\rangle \quad (6)$$

The Coulombic potential couples a given subband ν , as well as a given vector \mathbf{k}_\perp with all the others. The intersubband coupling (especially the one with the subbands of the quantum well continuum) is difficult to handle. In a quasi bi-dimensional situation we would like to set $c(\nu, \mathbf{k}_\perp) \equiv c_{\nu_0}(\mathbf{k}_\perp) \delta_{\nu, \nu_0}$, i.e. to neglect intersubband coupling [3]. This procedure is convenient as the impurity wavefunction displays a separable form :

$$\langle \mathbf{r} | \psi_{\text{loc}} \rangle = \chi_{\nu_0}(z) \varphi(\boldsymbol{\rho}). \quad (7)$$

The wavefunction $\varphi(\boldsymbol{\rho})$ is the solution of the two-dimensional Schrödinger equation

$$\left[\frac{p_x^2 + p_y^2}{2m^*} + V_{\text{eff}}(\rho) \right] \varphi(\boldsymbol{\rho}) = (\varepsilon - E_{\nu_0}) \varphi(\boldsymbol{\rho}) \quad (8)$$

where $V_{\text{eff}}(\rho)$ is the effective in-plane Coulombic potential :

$$V_{\text{eff}}(\rho) = \frac{-e^2}{\kappa} \int_{-\infty}^{+\infty} dz \chi_{\nu_0}^2(z) \frac{1}{\sqrt{\rho^2 + (z - z_i)^2}} \quad (9)$$

which is ν_0 and z_i dependent. A solution of equation (8) can be sought variationally, the simplest choice for the ground state being the nodeless one-parameter trial wavefunction

$$\varphi_0(\rho) = \frac{1}{\lambda} \sqrt{\frac{2}{\pi}} \exp(-\rho/\lambda) \quad (10)$$

where λ is the variational parameter. The bound state energy is obtained through the minimization of the function

$$\varepsilon_{\nu_0}(z_i, \lambda) = \frac{\hbar^2}{2m^* \lambda^2} - \frac{2e^2}{\kappa \lambda} \int_0^\infty x dx e^{-x} \int_{-\infty}^{+\infty} \frac{dz \chi_{\nu_0}^2(z)}{\sqrt{x^2 + \frac{4}{\lambda^2} (z - z_i)^2}} + E_{\nu_0} \quad (11)$$

and the binding energy is :

$$\eta_{\nu_0}(z_i) = E_{\nu_0} - \min_{\lambda} \varepsilon_{\nu_0}(z_i, \lambda) \quad (12)$$

One should be aware that the decoupling procedure (Eqs. (7-12)) runs into difficulties in the limits $L \rightarrow \infty$ and $L \rightarrow 0$. In the former case the energy difference $E_{\nu_0+1} - E_{\nu_0}$ becomes very small and many subbands become admixed by the Coulombic potential. Clearly, for infinite L one cannot describe the bulk 1S hydrogenic bound state by a separable wavefunction. In the latter case ($L \rightarrow 0$) one finds similar problems due to the energy proximity between the ground quantum well subband E_1 and the top of the well V_b . Consequently $\chi_1(z)$ leaks more and more heavily in the barrier. At $L = 0$ the only sensible result is to find a 1S bulk hydrogenic wavefunction corresponding to the barrier-acting material (in our model the latter has a binding energy equal to R_0). Once again this state cannot have a wavefunction like that of equation (7). If V_b is infinite the result at $L = 0$ is qualitatively different. The quantum well only has bound levels whose energy separation increases like L^{-2} when L decreases. The smaller the well thickness, the better the separable wavefunction becomes. At $L = 0$, one obtains a true bi-dimensional hydrogenic problem whose binding energy is $4R_0$ whereas $\lambda = \frac{1}{2}a_0$.

To circumvent the previous difficulties and obtain the exact limits at $L = 0$ and $L = \infty$ for any V_b one may use [4-6] :

$$\psi_{\text{loc}}(\mathbf{r}) = N \chi_1(z) \exp \left[\frac{-1}{\lambda} \sqrt{\rho^2 + (z - z_i)^2} \right] \quad (13)$$

where N is a normalization constant, λ the variational parameter and where attention is focussed on the ground bound state attached to the ground quantum well subband E_1 . Calculations are less simple than with the separable wavefunction (Eq. (7)). Comparing the binding energy deduced from equations (7, 13) one finds, for infinite V_b , that the separable wavefunction gives almost the same results as the non-separable wavefunction if $L/a_0 \leq 3$. This is the range where for most materials the quantum size effects are important.

Other variational calculations have been proposed [7, 8]. For example, instead of using a non-linear variational parameter one uses a finite basis set of fixed wavefunctions (often Gaussian ones) in which \mathcal{H} is numerically diagonalized. The numerical results obtained by using a single non linear variational parameter compare favourably with these very accurate treatments.

II.2 RESULTS FOR THE GROUND IMPURITY STATE ATTACHED TO THE GROUND SUBBAND. — Figures 2 to 4 give a sample of some calculated results for the ground impurity state attached to the ground subband. Two parameters control the binding energy :

i) Thickness dependence of the impurity binding energy : the dimensionless ratio L/a_0 indicates the amount of bi-dimensionality of the impurity state. If

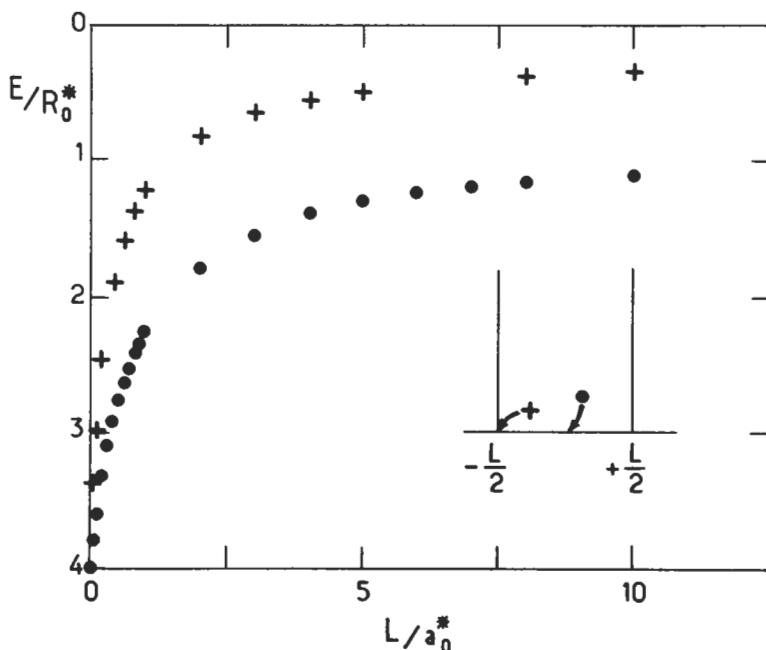


Fig. 2. — Calculated dependence of the on-edge (crosses) and on-centre (circles) hydrogenic donor binding energy *versus* the well thickness L in a quantum well with an infinite barrier height. R_0^* and a_0^* are the bulk effective Rydberg and Bohr radius respectively. After reference [4].

$\frac{L}{a_0} \geq 3$ or $\frac{L}{a_0} \leq 0.2$ and $V_b \sim 0.3$ eV in GaAs-Ga_{1-x}Al_xAs the problem is almost three-dimensional. This is either because the subbands are too close $\left(\frac{L}{a_0} \geq 3 \right)$ or because the quantum well continuum is too close $\left(\frac{L}{a_0} \leq 0.2 \right)$. The on-centre donor binding energy increases from R_0 ($L \rightarrow \infty$) to reach a maximum $\left(\frac{L}{a_0} \leq 1 \right)$ whose exact L -location and amplitude depend on V_b . Finally it decreases to the value R_0 at $L = 0$ [7, 8]. If V_b is infinite, the maximum is only reached at $L = 0$ and has a value of $4R_0$ [4].

ii) Position dependence of the impurity binding energy : the impurity binding energy monotonically decreases when the impurity location z_i moves from the centre to the edge of the well and finally deep into the barrier. In figure 5 it is shown that this decrease is rather slow for $z_i > \frac{L}{2}$; for instance, a donor placed 150 Å away from

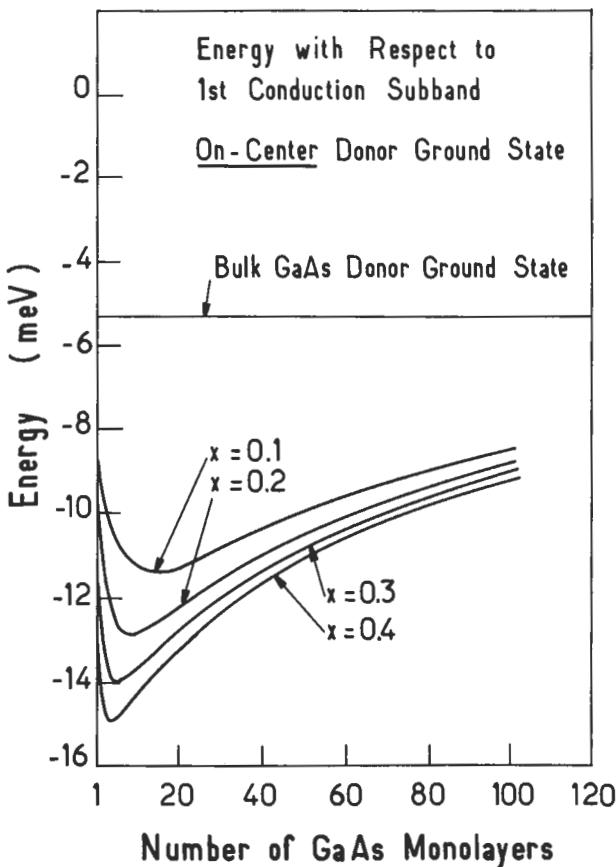


Fig. 3. — Calculated dependence of the on-centre hydrogenic donor binding energy in $\text{GaAs-Ga}_{1-x}\text{Al}_x\text{As}$ quantum wells *versus* the GaAs slab thickness L . $V_b(x) = 0.85 [\epsilon_g(\text{Ga}_{1-x}\text{Al}_x\text{As}) - \epsilon_g(\text{GaAs})]$. One GaAs monolayer is 2.83 \AA thick. After reference [7].

a $\text{GaAs-Ga}_{0.7}\text{Al}_{0.3}\text{As}$ quantum well ($L = 94.8 \text{ \AA}$) still binds a state by $\sim 0.5 R_0$ which is $\sim 2.5 \text{ meV}$ [5].

II.3 EXCITED SUBBANDS. CONTINUUM. — The procedure followed for the bound state attached to the E_1 subband can be generalized for excited subbands E_2, E_3, \dots as well as for the quantum well continuum.

However it becomes rapidly cumbersome since a correct variational procedure for excited states requires the trial wavefunctions to be orthogonal to all the states of lower energies. For separable wavefunctions (Eq. (7)) this requirement is automatically fulfilled and one may safely minimize $\epsilon_{\nu_0}(z_i, \lambda)$ to obtain a lower bound of

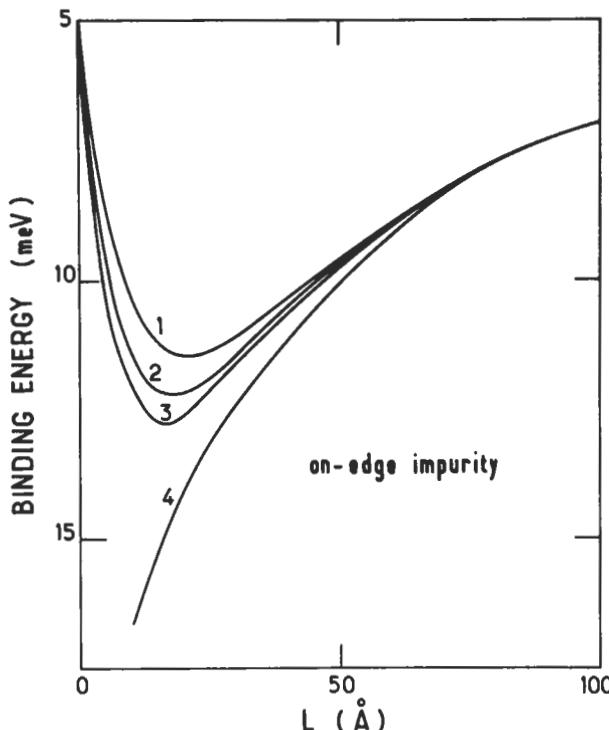


Fig. 4. — Calculated dependence of the on-edge hydrogenic donor binding energy in $\text{GaAs-Ga}_{1-x}\text{Al}_x\text{As}$ quantum wells *versus* the GaAs slab thickness L . $m^* = 0.067 m_0$, $\kappa = 13.1$. $V_b = 212$ meV (curve 1); 318 meV (curve 2); 424 meV (curve 3) and infinite (curve 4). After reference [12].

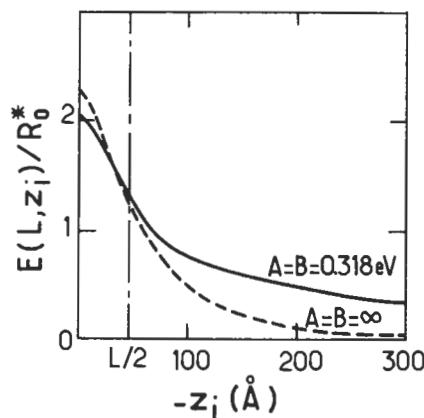


Fig. 5. — Calculated dependence of the hydrogenic donor binding energy in a quantum well *versus* the impurity position z_i a) in the case of a finite barrier well ($V_b = 0.318$ eV) and b) in the case of an infinite barrier well [4]. There is an interface at $-z_i = \frac{L}{2}$. $L = a_0 = 94.8$ Å. After reference [5].

$\eta_{\nu_0}(z_i)$. The new feature associated with the bound states attached to excited subbands is their finite lifetime. This is due to their degeneracy with the two-dimensional continua of the lower lying subbands. This effect however is not very large [3] since it arises from the intersubband coupling which is induced by the Coulombic potential : if the decoupling procedure is valid the lifetime of the quasi bound state calculated with equation (7) should be long.

For impurities located in the barriers, (an important practical topic with regard to the modulation-doping technique), we have seen that they weakly bind a state below the E_1 edge. There exists (at least) a second quasi-discrete level attached to the barrier edge, i.e. with an energy $\sim V_b - R_0$. This state is reminiscent of the hydrogenic ground state of the bulk barrier. Due to the presence of the quantum well slab it becomes a resonant state since it interferes with the two dimensional continua of the quantum well subbands ν ($E_\nu < V_b$). Its lifetime τ can be calculated using the expression

$$\frac{\hbar}{2\tau} = \frac{2\pi}{\hbar} \sum_{\nu, k_\perp} \left| \langle \tilde{\phi}(r, z_i) | \frac{-e^2}{\kappa \sqrt{\rho^2 + (z - z_i)^2}} | \nu k_\perp \rangle \right|^2 \times \delta \left[V_b - R_0 - E_\nu - \frac{\hbar^2 k_\perp^2}{2m^*} \right] \quad (14)$$

where $\tilde{\phi}$ is the 1S bulk hydrogenic wavefunction of the barrier. If the distance d separating the impurity from the quantum well edge is much larger than a_0 , equation (14) simplifies into

$$\frac{\hbar}{2\tau} \sim 16R_0 \sum_\nu P_\nu \left(\frac{R_0}{V_b - E_\nu} \right)^{3/2} \exp \left[\frac{-2d}{a_0} \sqrt{\frac{V_b - E_\nu}{R_0}} \right], \quad (15)$$

where P_ν is the total integrated probability of finding the carrier in the ν^{th} state (energy E_ν) in any of the two barriers of the quantum well. One sees from equation (15) that the lifetime τ is strongly dominated by the escape processes to the excited subband E_ν whose energy is nearest to $V_b - R_0$. As an example let us take a GaAs-Ga(Al)As quantum well with thickness $L = 50 \text{ \AA}$, $V_b = 0.2 \text{ eV}$ and assume that $d = 3 a_0$ (i.e. $d \sim 300 \text{ \AA}$). We then get $\tau \sim 3 \times 10^{-6} \text{ s}$. The quasi-bound state can thus be considered, to a reasonable approximation, as stationary $\left(\frac{\hbar}{\tau} \sim 4 \times 10^{-8} R_0 \right)$.

II.4 EXCITED IMPURITY LEVELS ATTACHED TO THE GROUND SUBBAND. — The Schrödinger equation (Eq. (4)) has several bound states below E_1 . Their binding energies have been calculated by several groups [7, 8]. The calculated energy difference between the ground on-centre donor state (quasi 1S) and the excited states (quasi $2P_x, 2P_y$) agrees with the far-infrared absorption and magneto-absorption data [9]. On-edge donor levels have also been investigated.

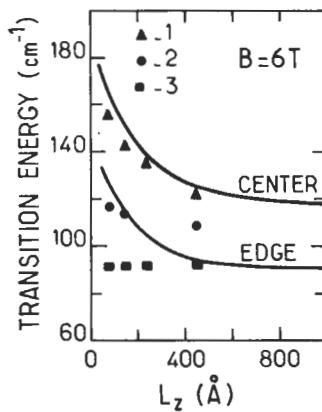


Fig. 6. — The transition energies of the peaks labelled 1, 2 and 3 observed in far infrared magneto-optical experiments are plotted *versus* the GaAs well width at $B = 6T$. The theoretically determined $1s \rightarrow 2p$ transition energies for donors at the centre and edge of the quantum well are indicated with solid lines. The Al mole fraction of the cladding barriers is 0.25. After reference [9].

II.5 ACCEPTOR LEVELS IN A QUANTUM WELL. — The problem of acceptor levels in semiconductor quantum wells is much more intricate than the equivalent donor problem. This is due to the degenerate nature of the valence bands in cubic semiconductors. In quantum wells this degeneracy is lifted (light and heavy holes have different confinement energies).

However the energy separation between the heavy and light hole subbands is seldom comparable to the bulk acceptor binding energies. Thus many subbands are coupled by the combined actions of the Coulombic potential and the quantum well confining barrier potential. No simple decoupling procedures appear manageable.

Masselink *et al.* [10] used variational calculations to estimate the binding energy of the acceptor level due to carbon (a well-known residual impurity in M.B.E. grown GaAs layers). Their results, shown in figure 7, agree remarkably with Miller *et al.*'s experiments [11]. One notices in figure 7 the same trends *versus* quantum well thickness as displayed by Coulombic donors (Figs. 3, 4). The binding energy first increases when L decreases (increasing tendency to bi-dimensional behaviour), then saturates and finally drops to the value of the acceptor binding energy of the bulk barrier. Finally, it should be noted that the relative increase in binding is smaller for acceptors than for donors. This is because the bulk acceptor Bohr radius is much smaller than that of the bulk donor.

III. Excitons.

When a semiconductor of direct bandgap ϵ_g is shone with near-bandgap light, electron-hole pairs (excitons) are created. If the electron and the hole were non-

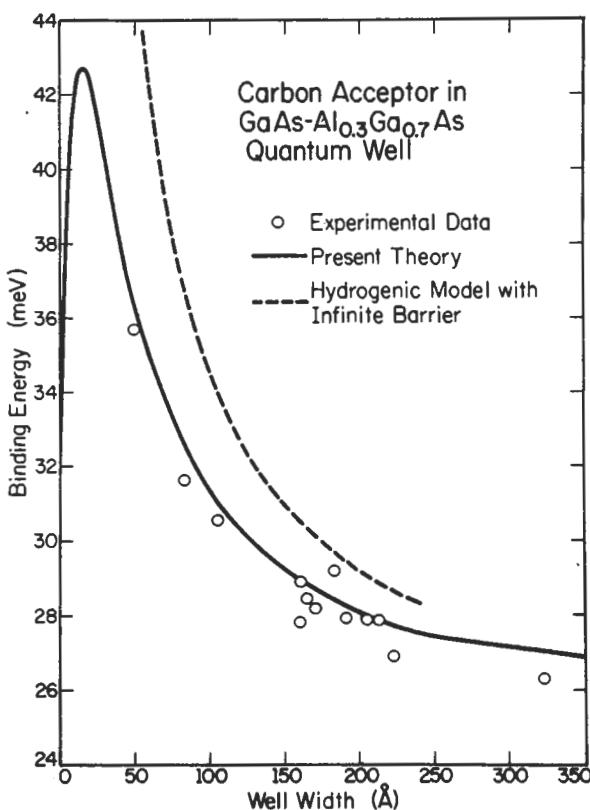


Fig. 7. — Calculated dependence of the on-centre carbon binding energy *versus* GaAs well width in GaAs-Ga_{1-x}Al_xAs quantum wells. $V_b(x) = 0.15 \times [\epsilon_g(\text{Ga}_{1-x}\text{Al}_x\text{As}) - \epsilon_g(\text{GaAs})]$ is the assumed hole confining barrier height. The open circles are the experimental values obtained by Miller *et al.* [11]. After reference [10].

interacting only photon energies $\hbar\omega > \epsilon_g$ would be absorbed and ϵ_g would be the absorption edge. The Coulombic electron-hole interaction greatly modifies this picture. The electron-hole attraction gives rise to bound states of the relative motion of the exciton [13]. The appearance of intense, narrow absorption lines below the fundamental absorption edge is the manifestation of these bound states.

The excitonic bound levels in quantum wells are, in many respects, analogous to the Coulombic impurity bound states discussed in sections I and II. In both cases one expects an increase in the binding energy due to the confinement of particles in the quantum well. There are however some genuine properties of the excitons which deserve some particular discussion.

- Two possibilities can be found for the respective locations of the electrons and

the holes. Either both electrons and holes are localized within the same layer (type I quantum wells, e.g. GaAs-Ga_{1-x}Al_xAs) or each layer of the quantum well locates one kind of particle (type II quantum wells, e.g. InP-Al_{0.52}In_{0.48}As where the electrons are mostly in the InP layers and the holes in Al_{0.53}In_{0.47}As, or InAs-GaSb where the electrons are mostly in the InAs layers and the holes are mostly in the GaSb layers). In type II quantum wells, we deal with interface excitons where the interacting particles are spatially separated, a situation which is reminiscent of that of the bound impurity states created by impurities placed in the barriers of the quantum well structure.

ii) The exciton problem involves both the dispersions of the conduction and valence subbands. Thus the complicated nature of the valence subbands will play some part in the magnitude of the excitonic binding. At the moment there is no general consensus (to the authors' knowledge) on how to evaluate the effects of the valence subband intricacies on the exciton binding energy. Thus only model calculations will be discussed hereafter.

We will proceed gradually starting with excitons in idealized bulk materials, then going on to the excitons in idealized type I and type II quantum wells [14]. Finally we will discuss Greene and Bajaj's model [16a, b], which is the most elaborate treatment of the "diagonal" exciton problem in quantum wells. Here, "diagonal" has the same meaning as in chapter III : the off-diagonal terms of the hole Hamiltonian have been discarded.

III.1 EXCITONS IN IDEALIZED BULK MATERIALS. — Consider a bulk semiconductor and assume that one can approximate its band structure by a single spherical conduction band with dispersion relation

$$\epsilon_c(\mathbf{k}) = \epsilon_g + \frac{\hbar^2 k^2}{2m_c^*}, \quad (16)$$

separated by the bandgap ϵ_g from a single spherical valence band with dispersion relation

$$\epsilon_v(\mathbf{k}) = \frac{-\hbar^2 k^2}{2m_v^*}. \quad (17)$$

The ground state of the semiconductor, taken as the zero of energy, consists of N electrons which occupy the N available places in the valence band, whereas all the conduction band levels are empty. An excited state of the crystal is built up by promoting a valence electron (wavevector \mathbf{k}_v) to the conduction band state corresponding to the wavevector \mathbf{k}_c and leaving the $N - 1$ other electrons in the valence band. The valence band with all but the $|\mathbf{k}_v\rangle$ states occupied can equally be viewed as a completely filled valence band (N occupied places) plus a hole with a wavevector $\mathbf{k}_h = -\mathbf{k}_v$. The hole is characterized by a positive mass ($+m_v^*$) and a positive charge ($+e$) [17]. If we neglect the Coulombic interaction among the electrons, the lowest lying excited state of the crystal corresponds to $\mathbf{k}_c = \mathbf{k}_v = \mathbf{0}$ and

has an energy of ε_g . However, due to the attractive Coulombic interaction between the conduction electron and the hole which is left in the valence band, it is possible to build up a crystalline excited state with a lower energy than ε_g . This is achieved by taking linear combinations of conduction and valence states which are such that they represent two clouds of particles of opposite signs orbiting around each other in a bound state. The spatial localization of the electron-hole pair carries an increase in kinetic energy over ε_g . This extra energy however is over-balanced due to the gain in electrostatic energy between the two clouds. The net result is a bound electron-hole pair which has a finite binding energy R_0^{ex} , compared with the unbound situation. This pair travels freely over the whole crystal. Because the electron-hole pair is weakly bound ($R_0^{\text{ex}} \ll \varepsilon_g$) and spreads over many lattice sites, the electron-hole interaction has to be screened by the static dielectric constant κ of the semiconductor [13]. Finally the exciton Hamiltonian can be written as :

$$\left[\frac{p_e^2}{2m_c^*} + \frac{p_h^2}{2m_v^*} - \frac{e^2}{\kappa |r_e - r_h|} \right] \psi(r_e, r_h) = (\varepsilon - \varepsilon_g) \psi(r_e, r_h). \quad (18)$$

Since the Coulombic term only affects the relative electron-hole coordinates, it is convenient to set

$$r = r_e - r_h; (m_c^* + m_v^*) \mathbf{R} = m_c^* r_e + m_v^* r_h \quad (19)$$

\mathbf{R} being the centre of mass of the electron-hole pair. The equation (18) is rewritten

$$\left[\frac{P^2}{2(m_c^* + m_v^*)} + \frac{p^2}{2\mu} - \frac{e^2}{\kappa r} \right] \psi(r, \mathbf{R}) = (\varepsilon - \varepsilon_g) \psi(r, \mathbf{R}) \quad (20)$$

where :

$$\mathbf{P} = -i\hbar \partial/\partial \mathbf{R}, \quad \mathbf{p} = -i\hbar \partial/\partial \mathbf{r} \quad (21)$$

and :

$$\mu = \frac{m_c^* m_v^*}{m_c^* + m_v^*} \quad (22)$$

is the reduced mass of the electron-hole pair. Since the Hamiltonian does not depend on \mathbf{R} , \mathbf{P} is a good quantum number with eigenvalue $\hbar\mathbf{K}$. Thus Ψ factorizes into :

$$\psi(r, \mathbf{R}) = \frac{1}{\sqrt{\Omega}} \exp(i\mathbf{K}\cdot\mathbf{R}) \varphi(r), \quad (23)$$

where Ω is the volume of the crystal.

By letting

$$\varepsilon = \varepsilon_g + \frac{\hbar^2 K^2}{2(m_c^* + m_v^*)} + \eta \quad (24)$$

we finally obtain

$$\left(\frac{p^2}{2\mu} - \frac{e^2}{\kappa r} \right) \varphi(r) = \eta \varphi(r) \quad (25)$$

which is the same as the Schrödinger equation found for the impurity problem (Eq. (1)) except that the reduced mass μ replaces the conduction mass m_c^* . The bound states correspond to $\eta < 0$ whereas positive η 's correspond to unbound electron-hole pairs. The ground state wavefunction is the 1S hydrogenic wavefunction

$$\varphi(\mathbf{r}) = \frac{1}{[\pi(a_0^{\text{ex}})^3]^{1/2}} \exp(-r/a_0^{\text{ex}}), \quad (26)$$

where a_0^{ex} is the exciton Bohr radius :

$$a_0^{\text{ex}} = \frac{\hbar^2 \kappa}{\mu e^2}. \quad (27)$$

The binding energy of the ground state ($\mathbf{K} = \mathbf{0}$) is the exciton Rydberg

$$R_0^{\text{ex}} = \frac{\mu e^4}{2\kappa^2 \hbar^2}. \quad (28)$$

Since $\mu < m_c^*$, the exciton binding energy is always smaller than that of the hydrogenic donor. Note finally that, although the relative motion of the electron-hole pair can be bound, the exciton is a delocalized excitation : the probability of finding the exciton within the volume d^3R around \mathbf{R} is equal to $\Omega^{-1} d^3R$ and is thus R -independent.

III.2 EXCITONS IN IDEALIZED QUANTUM WELL STRUCTURES. — Let us apply the previous considerations to a heterostructure which consists of a slab of A material inserted between two semi-infinite layers of B material. To keep the matter as simple as possible, let us assume that both materials have the same conduction m_c^* and valence m_v^* effective masses, as well as the same dielectric constant κ . Two cases have to be considered depending on whether the A material confines both the electron and the hole or only a single kind of particle (say the electrons).

i) A confines both electrons and holes (type I quantum wells) : let ϵ_g , V_e , V_h denote the bulk A bandgap and the barrier heights for the electron and hole respectively. Assuming, as in III.1 that the conduction and valence band dispersion relations in both materials are spherical and quadratic in \mathbf{k} , the Hamiltonian (Eq. (18)) can be generalized into

$$\mathcal{H} = \epsilon_g + \frac{p_e^2}{2m_c^*} + \frac{p_h^2}{2m_v^*} - \frac{e^2}{\kappa |\mathbf{r}_e - \mathbf{r}_h|} + V_e Y \left[z_e^2 - \frac{L^2}{4} \right] + V_h Y \left[z_h^2 - \frac{L^2}{4} \right] \quad (29)$$

where L is the quantum well thickness. In the absence of Coulombic interaction the eigenstates of \mathcal{H} are known :

$$\psi_0(\mathbf{r}_e, \mathbf{r}_h) = \frac{1}{S} \exp i [\mathbf{k}_{\perp e} \cdot \mathbf{p}_e + \mathbf{k}_{\perp h} \cdot \mathbf{p}_h] \chi_{n,c}(z_e) \chi_{m,h}(z_h) \quad (30)$$

where S is the sample area and $\chi_{n,c}(z_e)$, $\chi_{m,h}(z_h)$ are the eigenfunctions of the

conduction and valence quantum well problems respectively. The wavevectors $\mathbf{k}_{\perp e}$, $\mathbf{k}_{\perp h}$ are two-dimensional wavevectors which correspond to the free motions in the layer planes and $\boldsymbol{\rho}_e = (x_e, y_e)$; $\boldsymbol{\rho}_h = (x_h, y_h)$ are the position vectors in these planes.

The energy corresponding to $\psi_0(\mathbf{r}_e, \mathbf{r}_h)$ is :

$$\varepsilon(\mathbf{k}_{\perp e}, \mathbf{k}_{\perp h}, n, m) = \varepsilon_g + E_n + H_m + \frac{\hbar^2 k_{\perp e}^2}{2m_c^*} + \frac{\hbar^2 k_{\perp h}^2}{2m_v^*} \quad (31)$$

The lowest lying excited state of the heterostructure corresponds, in the absence of electron-hole interaction, to $\mathbf{k}_{\perp e} = \mathbf{k}_{\perp h} = \mathbf{0}$ and $n = m = 1$. Its energy exceeds the ground state energy of the heterostructure by $\varepsilon_g + E_1 + H_1$. Because the barrier potentials

$$V_e Y\left(z_e^2 - \frac{L}{4}\right), \quad V_h Y\left(z_h^2 - \frac{L}{4}\right)$$

break the translational invariance along the growth axis, one cannot use the transformations defined in equation (19) to simplify \mathcal{H} . However one can still define $\mathbf{R}_\perp, \boldsymbol{\rho}$ as :

$$(m_c^* + m_v^*) \mathbf{R}_\perp = m_c^* \boldsymbol{\rho}_e + m_v^* \boldsymbol{\rho}_h; \quad \boldsymbol{\rho} = \boldsymbol{\rho}_e - \boldsymbol{\rho}_h \quad (32)$$

to obtain :

$$\begin{aligned} \mathcal{H} = & \frac{P_\perp^2}{2(m_c^* + m_v^*)} + \frac{P_\perp^2}{2\mu} - \frac{e^2}{\kappa \sqrt{\rho^2 + (z_e - z_h)^2}} + \frac{P_{z_e}^2}{2m_c^*} + \frac{P_{z_h}^2}{2m_v^*} + \\ & V_e Y\left[z_e^2 - \frac{L^2}{4}\right] + V_h Y\left[z_h^2 - \frac{L^2}{4}\right] + \varepsilon_g \end{aligned} \quad (33)$$

which leads to the partial factorization

$$\psi(\mathbf{r}_e, \mathbf{r}_h) = \frac{1}{\sqrt{S}} \exp(i \mathbf{K}_\perp \cdot \mathbf{R}_\perp) \varphi(z_e, z_h, \boldsymbol{\rho}) \quad (34)$$

where $\hbar \mathbf{K}_\perp$, the centre of mass momentum, is the eigenvalue of \mathbf{P}_\perp and $\varphi(z_e, z_h, \boldsymbol{\rho})$ is the solution of a Hamiltonian similar to the hydrogenic donor Hamiltonian, except that

- i) the "carrier" has an anisotropic effective mass (m_c^* along z , μ in the layer plane)
- ii) the attractive centre can move along the z axis since it has a finite effective mass (m_v^*), whereas donor ions are immobile. This suggests that, qualitatively, the binding energy of the exciton is an average of the hydrogenic donor binding energy over the probability of finding the positive charge at the location z_h .

Since exciton and donor Hamiltonians are quite similar most of the considerations valid for the donors apply to the excitons. Two kinds of variational procedures have

been attempted to solve equation (33) : either by using Gaussian basis sets [16] or non linear variational parameters [14-16]. In the latter case advantage is taken of the prevailing effects of the confining barriers on the z_e , z_h motions. Thus we can write [18] :

$$\varphi(z_e, z_h, \rho) = N_1 \chi_{1,c}(z_e) \chi_{1,h}(z_h) \exp\left(\frac{-\rho}{\lambda}\right) \quad (35)$$

$$\text{or : } \varphi(z_e, z_h, \rho) = N_2 \chi_{1,e}(z_e) \chi_{1,h}(z_h) \exp\left[\frac{-1}{\lambda} \sqrt{\rho^2 + (z_e - z_h)^2}\right] \quad (36)$$

where N_1 and N_2 are two normalization constants, λ is the variational parameter and where the exciton ground bound state formed between an electron in the lowest conduction subband E_1 and a hole in the lowest hole subband H_1 have been considered.

At this point one should point out that there exists two kinds of excitons which, in our simplified approach, are completely decoupled. The first kind corresponds to the pairing between an electron and a heavy hole, which will be termed heavy hole excitons. The second kind corresponds to the pairing between an electron and a light hole which will be termed the light hole excitons.

Corresponding to each kind of exciton are scaling quantities which have the dimensions of a bulk effective Rydberg and a bulk effective Bohr radius respectively

$$R_{hh} = \mu_{hh} e^4 / 2\kappa^2 \hbar^2 ; R_{lh} = \mu_{lh} e^4 / 2\kappa^2 \hbar^2 \quad (37)$$

$$a_{hh} = \hbar^2 \kappa / \mu_{hh} e^2 ; a_{lh} = \hbar^2 \kappa / \mu_{lh} e^2 \quad (38)$$

where $hh(lh)$ denotes the heavy hole (light hole) and where

$$\mu_{hh} = m_{hh} m_c^* / (m_{hh} + m_c^*) ; \mu_{lh} = m_{lh} m_c^* / (m_{lh} + m_c^*) \quad (39)$$

are the reduced electron-hole masses in the layer plane (m_{hh} and m_{lh} are the effective masses of the heavy hole and the light hole respectively). In the simplified model used in this section the light hole exciton is always less bound than the heavy hole exciton because :

i) The light holes in LH_1 subband are spatially less confined than the heavy holes in HH_1 subband. Thus the wavefunction χ_{LH_1} is less localized than χ_{HH_1} which implies a weaker effective electron-hole interaction in the layer plane for the light hole exciton and consequently a weaker binding.

ii) The bulk light hole exciton Rydberg is smaller than the bulk heavy hole Rydberg since $m_{lh} < m_{hh}$.

↗ The exciton binding energy *versus* the quantum well thickness L , displays the same kind of behaviour as exhibited by hydrogenic donors. First the binding energy increases, reaches a maximum and then drops to the effective Rydberg of the barrier

material. If the confining barriers V_e, V_h are infinitely high the binding energy $R^*(L)$ increases monotonously to $4R_\infty$ where R_∞ is R_{hh} or R_{lh} , depending on whether one is dealing with the light hole exciton or the heavy hole exciton.

This is illustrated in figure 8 where the results obtained by using a separable (Eq. (35)) or a non-separable (Eq. (36)) exciton wavefunction in ρ and z_e, z_h , correspond to the curves labelled (1) and (2) respectively. In figure 8 the centre of mass energy is disregarded.

ii) Exciton in type II quantum wells : in these structures the electrons and holes are spatially separated. Also, because we are considering a double heterostructure and not a superlattice, the quantum states of one kind of particle will not be size-quantized along the growth axis. To be specific let us consider that the A material mostly confines the electrons and repels the holes whereas the cladding B layers repel the electrons and attract the holes (see Fig. 9). If we take the same simplifying assumptions which were made in i) concerning the band structure, the exciton Hamiltonian can be written :

$$\mathcal{H} = \frac{p_e^2}{2m_e^*} + V_e Y \left[z_e^2 - \frac{L^2}{4} \right] + \frac{p_h^2}{2m_h^*} - V_h Y \left[-\frac{L^2}{4} + z_h^2 \right] - \frac{e^2}{\kappa |r_e - r_h|} + \epsilon_A \quad (40)$$

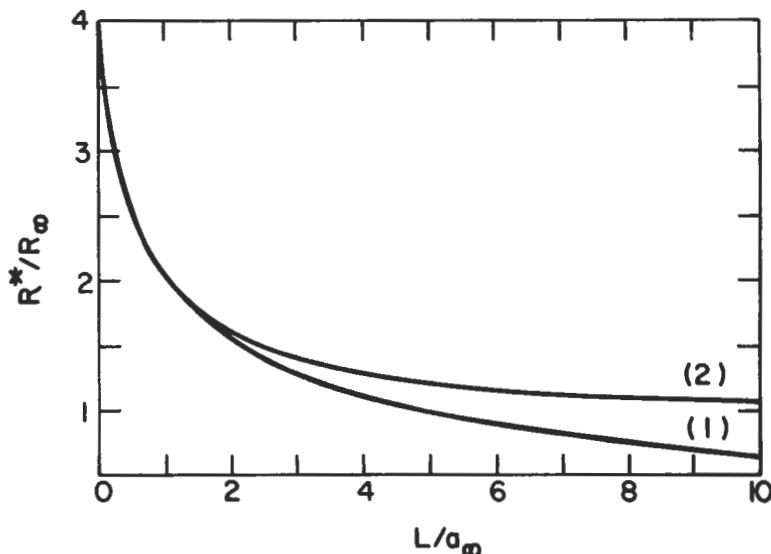


Fig. 8. — The exciton binding energy R^* in units of the effective Rydberg R_∞ is plotted against the dimensionless ratio of the type I quantum well thickness L to the effective Bohr radius a_∞ . The curves labelled (1) and (2) correspond respectively to an exciton wavefunction which is separable (Eq. (35)) or non separable (Eq. (36)) in ρ and z_e, z_h . After reference [14].

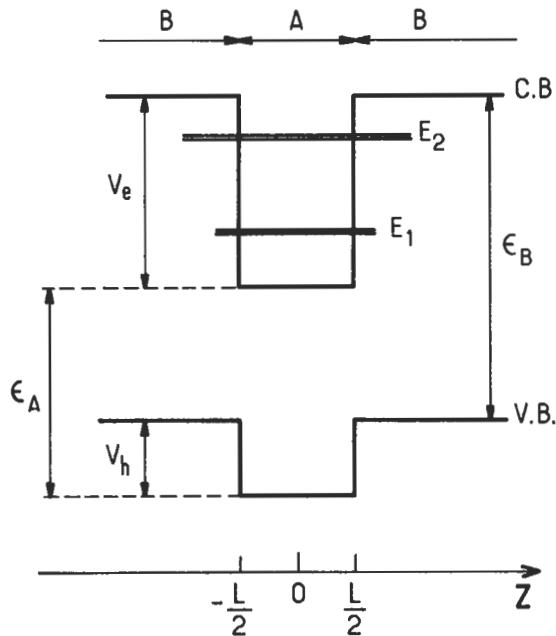


Fig. 9. — Conduction and valence band edge profiles in a type II quantum well.

In the absence of electron-hole interaction, the eigenstates of \mathcal{H} are written

$$\psi_0(r_e, r_h) = \frac{1}{S} \exp i [\mathbf{k}_{\perp e} \cdot \mathbf{p}_e + \mathbf{k}_{\perp h} \cdot \mathbf{p}_h] \chi_n(z_e) \xi(z_h) \quad (41)$$

where $\chi_n(z_e)$ is one of the electron quantum well bound states and $\xi(z_h)$ is an extended hole state labelled by a wavevector k_{z_h} which corresponds to a hole which stems from $z_h = -\infty$ and impinges on the A slab. This hole then tunnels through the A slab, and escapes to $z_h = +\infty$ in the right-hand size B layer. Another hole state corresponding to $-k_{z_h}$ has the same energy as $\xi(z_h)$. To build the exciton we are interested in the hole levels which are near the top of the valence band of the B layer. The characteristic penetration length of a heavy hole state ($m_{hh} = 0.4 m_0$) in the A layer is $l_{hh} \sim 10 \text{ \AA}$ (if $V_h = 0.1 \text{ eV}$). Thus, if the A layer is thick enough (say $L \geq 30 \text{ \AA}$) we can neglect the hole tunnelling and assume that the A slab is impenetrable for the holes. In this case

$$\begin{aligned} \xi_{\pm}(z_h) &= \sqrt{\frac{2}{\mathcal{L}}} \left\{ \sin k_{z_h} \left(z_h - \frac{L}{2} \right) Y \left(z_h - \frac{L}{2} \right) \pm \sin k_{z_h} \left(z_h + \frac{L}{2} \right) \times \right. \\ &\quad \left. \times Y \left(-z_h - \frac{L}{2} \right) \right\} \\ \xi_{\pm}(z_h) &= 0 \quad \text{if} \quad |z_h| < \frac{L}{2} \end{aligned} \quad (42)$$

where L is the A slab thickness and \mathcal{L} ($\gg L$) is the macroscopic length of the structure. In equation (42) $k_{z_h} > 0$ and \pm refers to the parity of the heavy hole state with respect to the centre of the A layer ($-$ corresponds to even states and $+$ to odd states).

We may use equations (32, 34) to eliminate the in-plane centre of mass coordinates and search for a variational solution of the relative motion of the electron-hole pair. We shall see later on that only parity-conserving transitions are optically allowed in symmetric heterostructures. A trial wavefunction for the lowest lying exciton state (photon energy $\leq \varepsilon_A - V_h + E_1$ in figure 9) is thus :

$$\varphi(\rho, z_e, z_h) = N \chi_1(z_e) \exp(-\rho/\lambda) \times \int_0^\infty dk_{z_h} \alpha(k_{z_h}) \xi_-(z_h) \quad (43)$$

where the summation over k_{z_h} is imposed by the continuous nature of the valence spectrum. In equation (43) N is a normalization constant, λ a variational parameter and $\alpha(k_{z_h})$ a function to be determined. To write down equation (43) we have assumed that the electron size-quantization is large enough to neglect the admixture between E_1 and $E_2, E_3 \dots$ due to the Coulombic interaction, thus forcing the z_e dependence for the exciton wavefunction to be identical to $\chi_1(z_e)$. On the other hand, because the hole spectrum is bulk-like, a hole wave packet concentrated near the interfaces ($\pm \frac{L}{2}$) is formed to maximize the electrostatic interaction with the electron charge distribution. It should also be noted that the excitons formed between the valence band and higher electron subbands $E_2, E_3 \dots$ are likely to be more tightly bound than the lowest lying exciton level, since χ_2, χ_3 leak more heavily in the B layers than χ_1 .

The exciton binding energy in GaSb-InAs-GaSb double heterostructures [14] has been calculated. These heterostructures are an extreme case of type II quantum wells inasmuch as the top of the GaSb valence lies above (by ~ 0.15 eV) the bottom of the InAs conduction band (Fig. 10). The electron confinement in the InAs slab was

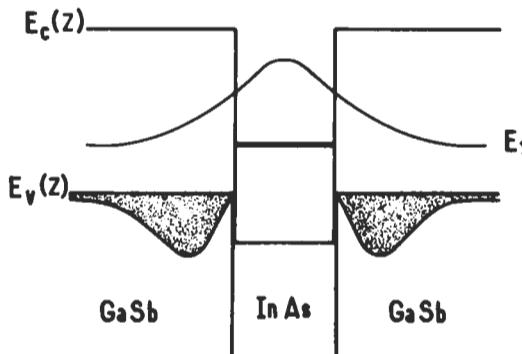


Fig. 10. — Conduction and valence band profiles in GaSb-InAs-GaSb double heterostructures. After reference [14]. The solid line enclosing the shaded area represents the deformed hole envelope function due to the electron-hole interaction.

assumed to be complete $\left(\chi_1(z_e) = \sqrt{\frac{2}{L}} \cos(\pi z_e/L) \right)$ and the z_h dependence of the exciton wavefunction was taken as a symmetric combination of Fang-Howard wavefunctions [18] :

$$\varphi(\rho, z_e, z_h) = N \sqrt{\frac{2}{L}} \cos\left[\frac{\pi z_e}{L}\right] \frac{b^{3/2}}{2} \left\{ \left(z_h - \frac{L}{2} \right) \exp\left[-\frac{b}{2}\left(z_h - \frac{L}{2}\right)\right] \times \right. \\ \left. \times Y\left[z_h - \frac{L}{2}\right] - \left(z_h + \frac{L}{2}\right) \exp\left[\frac{b}{2}\left(z_h + \frac{L}{2}\right)\right] Y\left(-z_h - \frac{L}{2}\right) \right\} \exp(-\rho/\lambda) \quad (44)$$

where b and λ are variational parameters. The mean distance from the hole to the A-B interface is $3/b$. Figure 11 shows the results of the variational calculation of the exciton binding energy in the GaSb-InAs-GaSb double heterostructures [14]. The

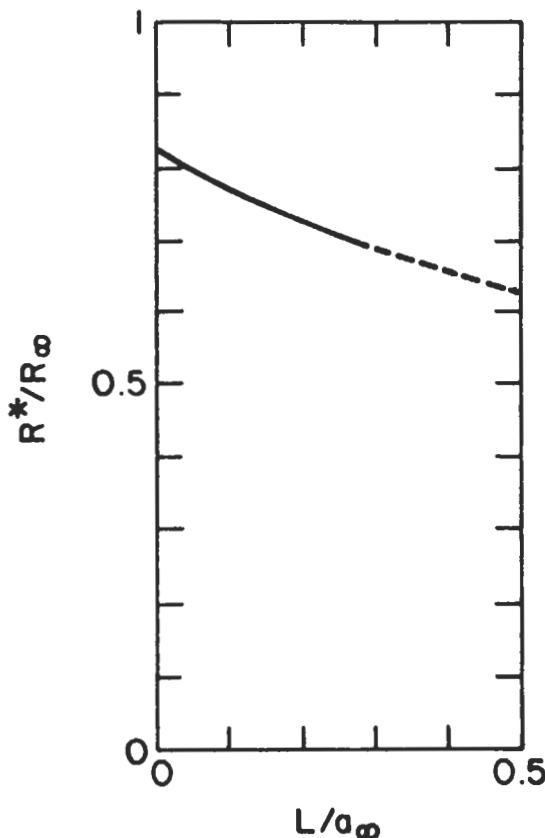


Fig. 11. — The exciton binding energy (in units of R_∞) is plotted against the InAs slab thickness (in units of a_∞) for GaSb-InAs-GaSb double heterostructures (solid line). The dashed line corresponds to InAs slab thicknesses such that the E_1 state lies below the top of the GaSb valence band. After reference [14].

quantities labelled R_∞, a_∞ are the bulk InAs exciton Rydberg and Bohr radius (1.3 meV and 370 Å respectively). We see clearly in figure 11 the strong influence of the spatial separation between the charges on the excitonic binding : the exciton is less bound in InAs-based type II quantum wells than in bulk InAs, despite the pronounced size quantization of the electrons in InAs slab.

From figure 11 it seems fair to conclude that excitonic effects are unimportant in type II quantum wells.

III.3 EXCITONS IN ACTUAL QUANTUM WELL STRUCTURES. — In diamond-like semiconductors the topmost valence band is fourfold degenerate at the zone centre (Γ_8 symmetry). In the spherical approximation the valence Hamiltonian can be written [19] :

$$\mathcal{H}_v(\mathbf{k}) = \frac{\hbar^2 k^2}{2m_1} \mathbf{1} + \frac{\hbar^2}{2m_2} (\mathbf{k} \cdot \mathbf{J})^2 \quad (45)$$

where $\mathbf{1}$ is the 4×4 identity matrix and \mathbf{J} a spin $3/2$ matrix. The effective masses m_1 and m_2 can be related to the heavy and light hole masses :

$$\frac{1}{m_1} = \frac{1}{8m_{hh}} - \frac{9}{8m_{lh}} ; \quad \frac{1}{m_2} = - \frac{1}{2m_{hh}} + \frac{1}{2m_{lh}} \quad (46)$$

To obtain equation (46) we have used the fact that for $\mathbf{k} // \mathbf{J}$ heavy hole states correspond to $m_J = \pm 3/2$ and light hole states to $m_J = \pm 1/2$.

The relative motion of the exciton is thus described by a Hamiltonian which is also a 4×4 matrix :

$$\mathcal{H}_{exc} = \left[\frac{p^2}{2m_c^*} - \frac{e^2}{\kappa r} \right] \mathbf{1} - \mathcal{H}_v \left(\frac{\mathbf{p}}{\hbar} \right) \quad (47)$$

Baldereschi and Lipari [20] have shown that \mathcal{H}_{exc} can be rewritten

$$\mathcal{H}_{exc} = P(\mathbf{r}, \mathbf{p}) \mathbf{1} + \begin{bmatrix} Q & L & M & 0 \\ L^* & -Q & 0 & M \\ M^* & 0 & -Q & -L \\ 0 & M^* & -L^* & Q \end{bmatrix} \quad (48)$$

where P, Q, L, M are functions of \mathbf{r} and \mathbf{p} which have S-symmetry (P) and D-symmetry (Q, L, M) respectively. The important point is that Q, L, M only involve valence band parameters whereas P also involves m_c^* and the Coulombic term $-e^2/\kappa r$. In fact :

$$P(\mathbf{r}, \mathbf{p}) = \frac{p^2}{2\mu_0} - \frac{e^2}{\kappa r} \quad (49)$$

where :

$$\frac{1}{\mu_0} = \frac{1}{m_c^*} + \frac{\gamma_1}{m_0} \quad (50)$$

γ_1 being one of the Luttinger parameters describing the Γ_8 hole kinematics. The diagonal term $P(\mathbf{r}, \mathbf{p})$ is thus much larger than the other ones. It represents the Coulombic binding between the conduction electron and the isotropic part of the Γ_8 hole. To the zeroth approximation in Q, L, M the exciton states formed between the Γ_6 electron and the Γ_8 hole are fourfold degenerate and can be calculated like the exciton states in idealized bulk materials (section III.1) except that the reduced mass of the exciton involves neither the heavy hole nor the light hole masses but an average of the two. As shown by Baldereschi and Lipari, the corrections to the zeroth approximation are very small under most circumstances.

To calculate exciton states in quantum well heterostructures one should add the barrier potentials for the electrons and the holes to equation (48). In type I structures this amounts to

$$V_{\text{barr}}(z_e, z_h) = V_e Y\left(z_e^2 - \frac{L^2}{4}\right) + V_h Y\left(z_h^2 - \frac{L^2}{4}\right) \quad (51)$$

where L is the quantum well thickness and V_e, V_h the barrier heights for the electrons and the holes. Once again we have neglected some band structure effects, e.g. the effective mass and dielectric mismatches between the host materials. For the L values where the size quantization is important, we expect the z_e, z_h motion to be forced by the quantum well effects. We thus need to keep the exact hole masses in order to evaluate the holes confinement energies correctly. This precludes the use of Baldereschi and Lipari's type of approach which is based on a suitable averaging of heavy and light hole masses. To the author's knowledge there does not exist at present any fully satisfactory treatment of the exciton binding in quantum well structures. Miller *et al.* [15] and Greene and Bajaj [16a, b] have approximated the exciton Hamiltonian by

$$\mathcal{H}_{\text{ex}} = (P + V_{\text{barr}}) \mathbf{1} + \begin{bmatrix} Q & & & 0 \\ & -Q & & \\ & & -Q & \\ 0 & & & Q \end{bmatrix} \quad (52)$$

Miller *et al.* took [15] V_e, V_h to be infinite, whereas Greene and Bajaj [16a, b] accounted for the finite barrier effects both groups were specifically interested in GaAs-Ga_{1-x}Al_xAs quantum wells. The advantage of including Q in \mathcal{H}_{ex} is that a correct evaluation of the size quantization of the holes can be obtained. The drawback of this approximation is the inclusion of terms (brought by Q) for the in plane exciton motion which are, in principle, as small as the leftover terms L, M (although the latter are off diagonal). If we remember that the L, M terms actually give rise to strong anticrossings between the hole subbands [21, 22], there is some possibility that they can significantly contribute to the excitonic binding itself.

As the exciton Hamiltonian (Eq. (52)) is a diagonal matrix, the excitons fall into two categories ; the heavy hole ($P + Q$) and the light hole ($P - Q$) excitons. The heavy hole exciton Hamiltonian corresponds to $m_J = \pm \frac{3}{2}$ and is written as

$$\mathcal{H}_{\text{ex}}^{\text{hh}} = \frac{p_{z_c}^2}{2m_c^*} + \frac{p_{z_h}^2}{2m_{hh}} - \frac{e^2}{\kappa |\mathbf{r}_e - \mathbf{r}_h|} + V_e Y \left[z_c^2 - \frac{L^2}{4} \right] + V_h Y \left[z_h^2 - \frac{L^2}{4} \right] + \frac{p_\perp^2}{2\mu_{hh}} \quad (53)$$

where m_{hh} and μ_{hh} are defined by

$$\frac{1}{m_{hh}} = \frac{1}{m_0} (\gamma_1 - 2\gamma_2); \quad \frac{1}{\mu_{hh}} = \frac{1}{m_c^*} + \frac{1}{m_0} (\gamma_1 + \gamma_2) \quad (54)$$

The light hole exciton Hamiltonian corresponds to $m_j = \pm \frac{1}{2}$ and is written as :

$$\mathcal{H}_{\text{ex}}^{\text{lh}} = \frac{p_{z_c}^2}{2m_c^*} + \frac{p_{z_h}^2}{2m_{lh}} - \frac{e^2}{\kappa |\mathbf{r}_e - \mathbf{r}_h|} + V_e Y \left[z_e^2 - \frac{L^2}{4} \right] + V_h Y \left[z_h^2 - \frac{L^2}{4} \right] + \frac{p_\perp^2}{2\mu_{lh}} \quad (55)$$

where m_{hh} and μ_{lh} are given by the expressions :

$$\frac{1}{m_{lh}} = \frac{1}{m_0} (\gamma_1 + 2\gamma_2); \quad \frac{1}{\mu_{lh}} = \frac{1}{m_c^*} + \frac{1}{m_0} (\gamma_1 - \gamma_2) \quad (56)$$

Thus, the heavy hole and light hole excitons once again resemble those obtained in idealized quantum well structures. However, in contrast with the results of section (III.2) μ_{lh} is not necessarily smaller than μ_{hh} . In fact in GaAs the opposite is true : $\mu_{lh} = 0.051 m_0$ and $\mu_{hh} = 0.04 m_0$ whereas $m_{lh} = 0.08 m_0$ and $m_{hh} = 0.45 m_0$. The inclusion of the Q term in \mathcal{H}_{exc} inverts the parts played by the heavy and light masses along and perpendicular to the z axis. The heavy hole exciton is indeed heavy along z but light in the layer plane and *vice versa*. Thus in GaAs-Ga_{1-x}Al_xAs quantum wells, the curves which represent the binding energies *versus* the GaAs slab thickness of the two kinds of excitons should cross. For large wells the light and heavy hole confinements (governed by m_{lh} and m_{hh} respectively) are almost complete and the light hole exciton is more tightly bound because its effective bulk Rydberg is larger than that of the heavy hole. On the other hand, for narrow GaAs wells, the light holes are less confined than the heavy holes. The Coulombic interaction between the electron and the hole in the light hole exciton is thus weaker than the one in the heavy hole exciton. Consequently the light hole exciton is less bound than the heavy hole exciton.

Figure 12 shows Greene *et al.*'s [16a, b] results concerning the ground bound exciton states in GaAs-Ga_{1-x}Al_xAs quantum wells for two aluminum mole fractions $x = 0.15$ and $x = 0.3$. In these curves, the Dingle's rule which states that the conduction band shares (85 %) of the total bandgap difference between GaAs and Ga_{1-x}Al_xAs, has been used.

Otherwise, the overall shapes of these curves look familiar : the exciton binding energies admit a maximum value *versus* the GaAs well thickness, whose location and amplitude depend on V_e , V_h and μ_{hh} , μ_{lh} .

To summarize, the Coulombic bound states in semiconductor heterostructures are qualitatively well understood. The effect of off-diagonal terms in the exciton Hamiltonian remains, however, an issue for quantitative understanding.

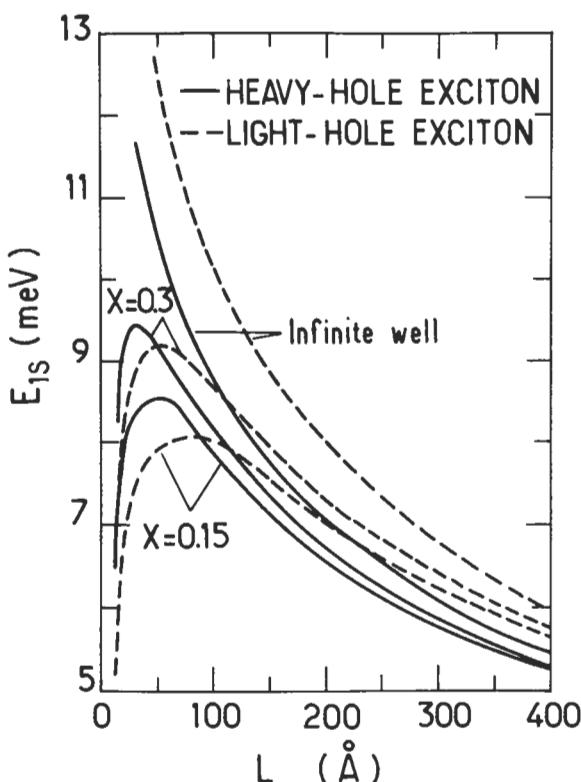


Fig. 12.— Variation of the binding energy of the ground state E_{1s} of the heavy hole exciton (solid lines) and the light hole exciton (dashed lines) as a function of the GaAs quantum well size (L) for aluminum concentration $x = 0.15$ and $x = 0.3$ and for infinite potential wells. The conduction band discontinuity is taken as 85 % of the total bandgap energy difference between GaAs and $\text{Ga}_{1-x}\text{Al}_x\text{As}$. After reference [16b].

INTERFACE DEFECTS

Up to now we have neglected the real structure of the interfaces which separate the two materials A and B. In fact within the envelope function scheme these interfaces, if perfect, can be reduced to planes perpendicular to the z axis. However, we know that on the atomic scale such an idealized picture is wrong. In this sub-chapter we will not attempt the difficult problem of providing an atomic description of the interfaces, but rather we will deal, (within the envelope function scheme) with the fact that interfaces are never perfect [23].

I. Graded interfaces.

The interface can best be viewed as a transition region where parameters continuously evolve from those characteristic of the A material to those characteristic

of the B material. Consider, for instance, the ideal case of two perfect, lattice-matched binary semiconductors, NP and N'P' (Fig. 13). In the bulks, bonds are found between the N and P atoms on the one hand, and between N' and P' on the other hand. Near the interface, one finds instead a transition layer of thickness $2b$ where hybrid bonds (PNP') necessarily occur before the crystalline order of the bulk hosts is established. With respect to the idealized interface $z = \text{constant}$, the situation depicted in figure 13 will be termed a graded interface.

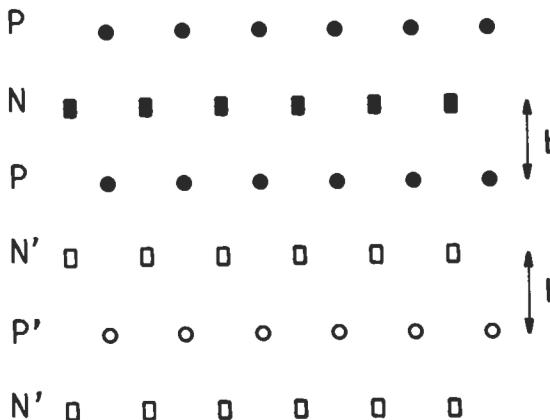


Fig. 13. — Schematic representation of a perfect interface between two lattice-matched semiconductors NP and N'P'.

A perfect but graded interface modifies the energy levels of the heterostructures with respect to those calculated if we assume the interface to be ideal. This effect is however small [24] if the characteristic thickness of the heterostructure remains large compared to the thickness of the transition regions. Moreover, we show in Appendix A that, under the same conditions, the bound states of a quantum well with narrow graded interfaces are identical to those of a *square* quantum well with idealized interfaces. The latter has a thickness slightly larger than the nominal thickness. It is important to stress that a perfect graded interface is *not* a defect since it preserves the translational invariance of the heterostructure in the layer planes. Thus it cannot limit the in-plane electrical conductivity.

II. Quantum well interface defects.

What does affect the electrical conductivity and can even give rise to bound states for electrons, holes or excitons, is a local protrusion of the A material into the B material or *vice versa*. Clearly the potential energy associated with such a protrusion depends on z but also on x and y . The interface defects were detected by X ray imaging and electron microscopy. In good GaAs-Ga_{1-x}Al_xAs materials they are one monolayer in depth (2.83 Å) (because As is common to both GaAs and Ga_{1-x}Al_xAs) and a few

hundred Angströms wide in the layer planes [25]. Note that because $\text{Ga}_{1-x}\text{Al}_x\text{As}$ is an alloy, a protrusion of GaAs in $\text{Ga}_{1-x}\text{Al}_x\text{As}$ layers can also be viewed as originating from a local fluctuation in the Al content (alloy clustering [29]). The extreme case of the alloy clustering is a GaAs bridge "piercing" the $\text{Ga}_{1-x}\text{Al}_x\text{As}$ barrier which separates two GaAs wells. This possibility was extensively discussed in the early eighties but is now thought to be improbable in high quality materials.

Because the sizes of the interface defects are unknown and probably depend on the growth conditions, one is forced to describe these defects very simply. A given defect is characterized by a depth b (along z) and one (or two) lateral dimensions a (or a and c). To be specific let us analyse the bound states created by interface defects in a single quantum well. Let $z = 0$ represent the position of one of the perfect interfaces between the well-acting (A) and the barrier-acting (B) materials, and let V_b be the barrier height, say for the electrons, of the quantum well. Attractive defects correspond to the protrusions of the A material in the B material. Simple models for interface defects which still retain some physical insights are :

i) the shoe-box defects

$$V_{\text{def}}(x, y, z) = -V_b Y(-z) Y(b+z) Y\left[\frac{a}{2} - |x|\right] Y\left[\frac{c}{2} - |y|\right] \quad (57)$$

a limiting case being the trench-like defect ($c \rightarrow \infty$)

ii) the pill-box defects (cylindrical symmetry)

$$V_{\text{def}}(\rho, z) = -V_b Y(-z) Y(b+z) Y(a - \rho) \quad (58)$$

where :

$$\rho^2 = x^2 + y^2; \quad (59)$$

iii) the semi-Gaussian defect (cylindrical symmetry)

$$V_{\text{def}}(\rho, z) = -V_b Y(-z) \exp\left[-\frac{\rho^2}{2a^2}\right] \exp\left[-\frac{z^2}{2b^2}\right]. \quad (60)$$

Note that in $\text{GaAs-Ga}_{1-x}\text{Al}_x\text{As}$ (or more generally in type I structures) a defect attracting the electron will also attract the hole. The converse is true in type II structures where attractive defects for electrons will repel the holes.

To calculate the bound states associated with a given interface defect we can proceed in the same way as we did for Coulombic impurities. The Hamiltonian which includes the defect potential is :

$$\mathcal{H} = \mathcal{H}_0 + \frac{p_x^2 + p_y^2}{2m^*} + V_{\text{def}}(\mathbf{r}) \quad (61)$$

where \mathcal{H}_0 is the one-dimensional Hamiltonian for the quantum well with perfect interfaces. The general solution can be expanded in the form

$$\psi(\mathbf{r}) = \sum_n \chi_n(z) f_n(x, y) \quad (62)$$

where $\chi_n(z)$ is the n^{th} eigenfunction of the unperturbed quantum well problem (energy E_n). If we assume that $V_{\text{def}}(\mathbf{r})$ induces negligible couplings between the unperturbed quantum well subbands, we find that the function $f_n(x, y)$ is the eigensolution of the effective bi-dimensional Schrödinger equation

$$\left[\frac{p_x^2 + p_y^2}{2m^*} + V_{\text{eff}}(x, y) \right] f_n(x, y) = (\varepsilon - E_n) f_n(x, y) \quad (63)$$

where :

$$V_{\text{eff}}(x, y) = \int_{-\infty}^{+\infty} dz \chi_n^2(z) V_{\text{def}}(\mathbf{r}). \quad (64)$$

We solve equation (63) in the specific case of the pill-box defect. We restrict our considerations to the ground defect state attached to the ground ($n = 1$) quantum well subband and try the variational solution :

$$f_1(\rho) = \frac{1}{\lambda} \sqrt{\frac{2}{\pi}} \exp[-\rho/\lambda]. \quad (65)$$

Then, we have to minimize

$$E(\lambda) = \frac{\hbar^2}{2m^* \lambda^2} - V_b \left\{ 1 - \left(1 + \frac{2a}{\lambda} \right) \exp\left[-\frac{2a}{\lambda}\right] \right\} \times \int_{-b}^0 \chi_1^2(z) dz + E_1. \quad (66)$$

Let P_b (respectively κ_b) represents the integrated probability of finding the carrier in the barriers (respectively the characteristic wavevector of the evanescent wave in the barrier) for the E_1 quantum well state. Then $E(\lambda)$ admits a minimum for :

$$\lambda_{\min} = 2a/\text{Ln} \left\{ \frac{2m^* a^2}{\hbar^2} V_b P_b [1 - \exp(-2b\kappa_b)] \right\}. \quad (67)$$

In order to make sense, equation (67) should be such that $\lambda_{\min} > 0$. Thus the variational solution is such that there exists a minimum lateral extension a_c of the defect below which there should be no bound states created by the defect. Remember however that the exact solution of the Schrödinger equation is lower in energy than $E(\lambda_{\min})$. Thus, even if we find that $E(\lambda_{\min}) = E_1$, i.e. that the state is unbound, this may only be a sign that the variational solution (Eq. (65)) is inaccurate.

A typical value for a_c is 140 Å (60 Å) for $m^* = 0.067 m_0$ ($m^* = 0.4 m_0$) if $V_b P_b = 3$ meV. These values are still smaller than the defect sizes found by X ray imaging. For the statistically dominant defects, the variational procedure is expected to lead to a reasonable, qualitative, description of the bound states.

In the limit of infinitely large defects, the binding energy η ($\eta = E_1 - E_1(\lambda_{\min})$) extrapolates to the value

$$\eta(a = \infty) = \frac{1}{2} V_b P_b [1 - \exp(-2b\kappa_b)] \quad (68)$$

This value is close to the exact solution which is $E_1(L) - E_1(L + b)$. In fact, equation (68) coincides with the first order energy shift of the E_1 state in a quantum well of width L , perturbed by the potential $-V_b Y(-z) Y(z + b)$, i.e. that of the extra layer of width b .

The calculations with infinite V_b should proceed differently. Exact numerical solutions may be found by expanding the wavefunctions in terms of the unperturbed solutions of perfect quantum wells with thicknesses L and $L + b$ respectively, and by matching the wavefunction and its derivative at the defect boundaries. The results of this type of calculation are presented in figure 14 for trench-like defects [27, 28]. This analysis has also been done for pill-box defects [26, 28]. An interesting result of these calculations is the rapidity with which the bound state converges towards the limiting value $E_1 - \eta(a = \infty)$, which corresponds to the confinement energy of a well of thickness $L + b$.

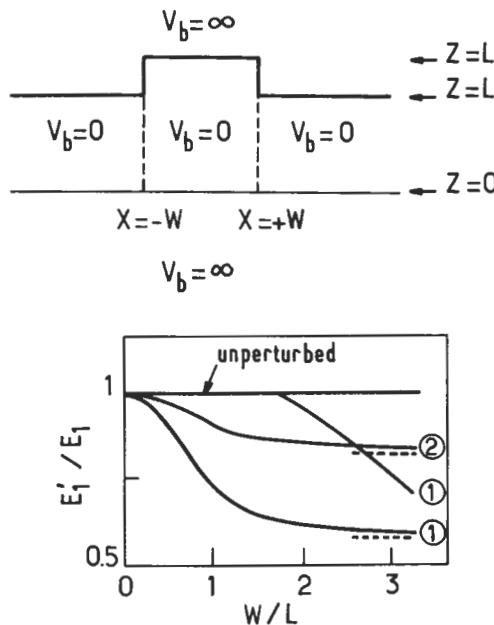


Fig. 14. — Exact bound states of the trench-like defect model with $V_b = \infty$. The bound state energies are plotted against the dimensionless lateral defect size W/L for two values of the ratio L'/L . 1) $\frac{L'}{L} = 1.314$; 2) $\frac{L'}{L} = 1.105$. After reference [27].

The interface defects can also bind an exciton. In the perfect well, the exciton formed between an electron in the E_1 subband and a hole in the HH_1 subband is described by the envelope function

$$\varphi_{\text{exc}}(\mathbf{r}_e, \mathbf{r}_h) = \frac{1}{\sqrt{S}} \exp(i\mathbf{K}_\perp \cdot \mathbf{R}_\perp) \chi_{E_1}(z_e) \chi_{HH_1}(z_h) f_\nu(\boldsymbol{\rho}) \quad (69)$$

where we have made the assumption that decoupled hole subbands exist, and where f_ν is the function describing the relative motion of the electron-hole pair in the layer plane for the ν^{th} exciton state. If the defect potential for the pair can be considered as small compared with the energy distance between say, the "1S" and the "2S" state of the exciton, it will only affect the degree of freedom of the exciton's centre of mass. Then a trial wavefunction for the exciton bound to the defect is

$$\psi_{\text{exc}}(\mathbf{r}_e, \mathbf{r}_h) = g_{\text{loc}}(\mathbf{R}_\perp) \chi_{E_1}(z_e) \chi_{H_1}(z_h) f_{1S}(\rho). \quad (70)$$

This procedure has been applied to semi-Gaussian defects by taking a Gaussian shape for $g_{\text{loc}}(\mathbf{R}_\perp)$. Figure 15 presents the calculated values of the trapped exciton

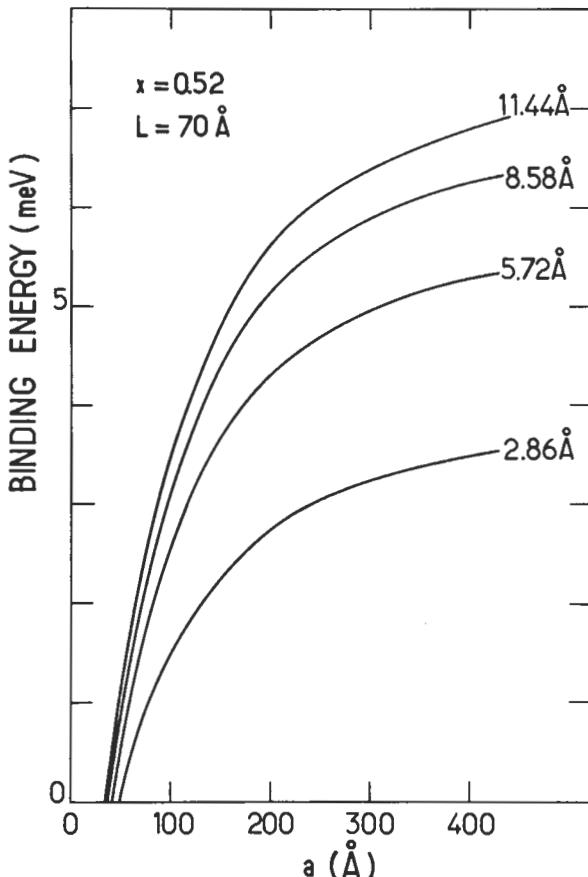


Fig. 15. — The binding energy of an exciton trapped on a semi-Gaussian interface defect, is plotted against the lateral extension of the defect a . This is done for several values of the characteristic defect depth b in a GaAs-Ga_{0.48}Al_{0.52}As quantum well of thickness $L = 70 \text{ \AA}$. The conduction band discontinuity is taken as 551 meV. After reference [30].

binding energy η versus the lateral defect size and several values of defect depth b . Figure 16 presents the a dependence of η , keeping b fixed, for different values of the quantum well thickness L .

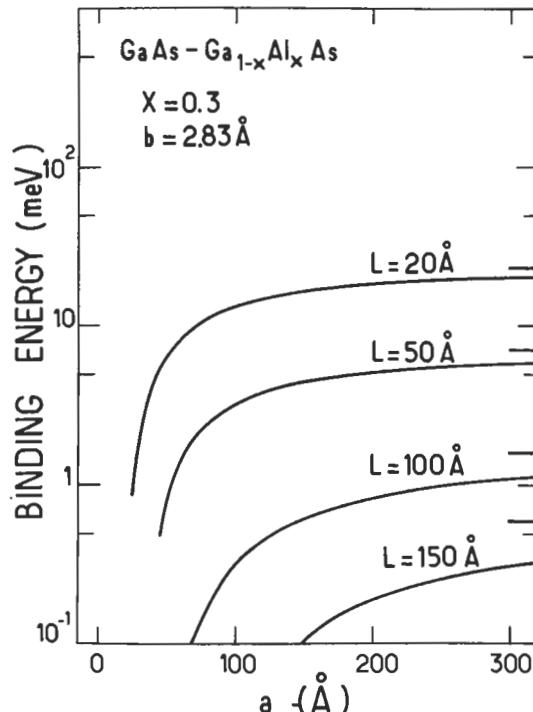


Fig. 16. — The binding energy of an exciton trapped on a semi-Gaussian interface defect is plotted against the lateral extension of the defect a , keeping the characteristic defect depth b fixed to 2.83 Å and varying the quantum well thickness L in GaAs-Ga_{0.7}Al_{0.3}As quantum wells. The conduction band discontinuity is taken as 318 meV. The horizontal bars which appear on the right hand side vertical axis correspond to the binding energy at infinite a .

An important physical feature of interface defects is their unavoidably random lateral size and depth. The statistical distribution $P(a, b)$ of sizes a and b implies that the binding energies of the carriers or excitons on these defects are in general distributed according to some unknown law $P(\eta)$ which is most likely sample-dependent. This complicates the analysis of the experimental data to a considerable extent and it is hard to expect more than qualitative insights from such analysis. As an example, we have presented in figure 17 the calculated trapped exciton density of states $P(\eta)$ which was obtained by assuming a rather broad Gaussian distribution of sizes a and b of semi-Gaussian interface defects :

$$P(\eta) = \sum_{a, b} P(a, b) \delta [\eta - \eta(a, b)]. \quad (71)$$

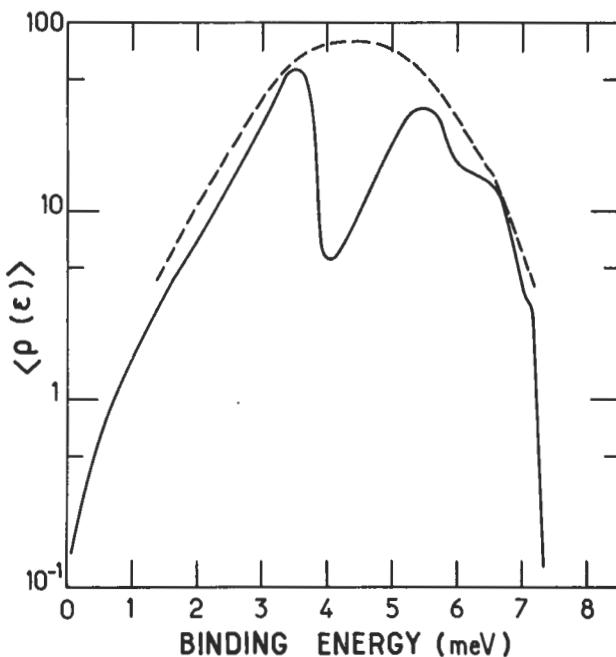


Fig. 17. — Average density of states of the trapped exciton bands in a GaAs-Ga_{0.48}Al_{0.52}As quantum well of thickness $L = 70$ Å. The solid line is obtained by assuming that the defect depth parameter b is discrete ($b = n \times 2.86$ Å, $n = 1, 2, \dots$). The dashed line is the expected curve if one let b vary continuously. After reference [30].

A completely different type of problem appears when we consider coupling between defects. What do the interfaces really look like? Are there so few interface defects that we can safely describe them as independent, like the independent impurities in lightly-doped bulk materials? Or is the interface merely a succession of attractive and repulsive interface defects? One of the interests of the quantum well structures is the possibility of optically studying the effects of a weak disorder experienced by neutral particles (the excitons). This weak disorder is caused by the neutral, and thus rather short-ranged, interface defects. The GaAs-Ga_{1-x}Al_xAs structures are at present the best candidates for helping to answer, at least partially, the above questions.

III. Superlattice defect states.

In addition to the defects at each interface, which separate the A and the B layers of a binary superlattice (period $d = l_A + l_B$, where l_A, l_B are the thicknesses of the A and B layers) and which give rise to bound states whose wavefunctions decay to zero in any direction away from the defect, there exist defects which are specific to the

superlattice structures. They consist of a layer (say A) whose thickness L_A is uniformly larger or smaller than the nominal layer thickness l_A [31, 32].

If $L_A > l_A$ the defect is attractive, say for electrons, and gives rise to bound levels (at least one) whose energy ϵ_D is smaller than the onset of the first conduction superlattice subband E_1 . The wavefunction associated with ϵ_D decays to zero when $z \rightarrow \pm \infty$. On the other hand the carrier is free to move in the layer plane. This gives rise to a step-like density of states associated with the defect. Figure 18 presents the binding energy of an electron on the InAs layer which is thicker than the nominal (30 Å) InAs thickness in a 30 Å-50 Å InAs-GaSb superlattice [32]. For such a type II system, the defect which is attractive for the electrons is repulsive for the holes, and *vice versa*.

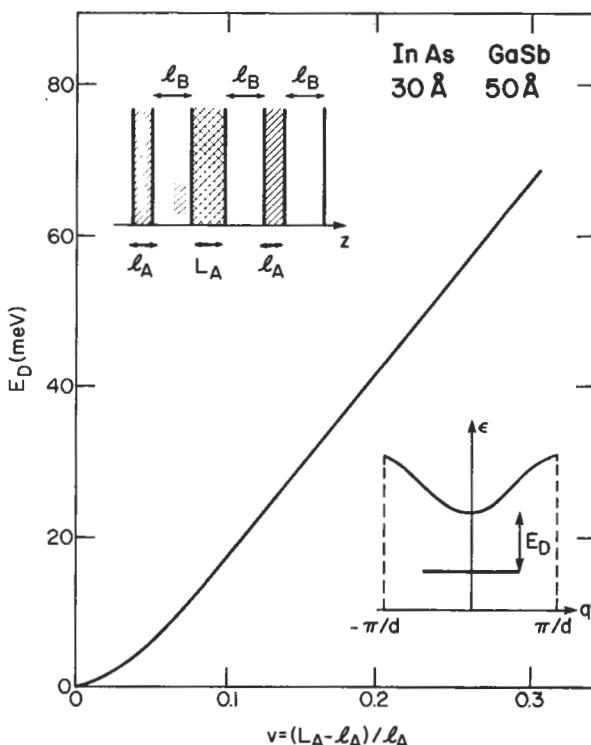


Fig. 18. — The binding energy of an electron in a InAs layer (which is thicker than the nominal InAs layer thickness), is plotted against $v = (L_A - l_A)/l_A$ for a (30 Å - 50 Å) InAs-GaSb superlattice. After reference [32].

On the other hand, in type I structures such as GaAs-Ga_{1-x}Al_xAs superlattices, an enlarged GaAs well attract both the electrons and the holes. Thus, by deliberately introducing such enlarged wells into a superlattice, one may use them as probes for studying the carrier dynamics in the superlattice. One can also obtain, by optical

studies, information on the various relaxation times which characterize [33] the electron motion along the superlattice growth axis.

Appendix.

Energy levels in a quantum well with narrow graded interfaces.

We consider the semiconductor quantum well whose conduction band edge profile is shown in figure 19. The electron experiences a potential energy $V_b(z)$ which is equal to V_b if $z \geq L + b'$ or $z \leq -b$. $V_b(z)$ vanishes between the points $z = +b$ and $z = L - b'$. We assume the existence of smoothly graded interfaces $[|z| < b$ and $|z - L| < b']$. $V_b(z)$ has thus an unknown z dependence, apart from the fact that it is smooth in the graded regions. The electron effective mass is m_B when $V_b(z) = V_b$, m_A when $V_b(z) = 0$ and varies smoothly with z in the graded regions.

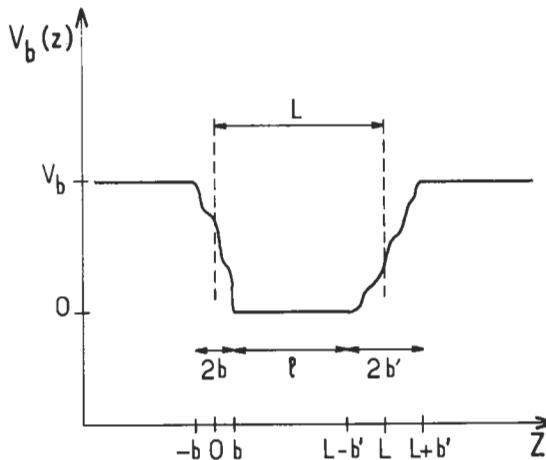


Fig. 19. — Potential energy profile in a quantum well with graded interfaces. l is the nominal quantum well thickness.

The stationary solutions of the Schrödinger equation which corresponds to the energy $\epsilon < V_b$ can be written :

$$\begin{cases} \chi(z) = A \exp[\kappa_b(z + b)] & z \leq -b \\ \chi(z) \text{ unknown} & |z| \leq b \\ \chi(z) = B \cos[k_0(z - b)] + C \sin[k_0(z - b)] & b \leq z \leq L - b' \\ \chi(z) \text{ unknown} & |z - L| < b' \\ \chi(z) = D \exp[-\kappa_b(z - L - b')] & z \geq L + b' \end{cases} \quad (A1)$$

where :

$$\frac{\hbar^2 \kappa_b^2}{2m_B} = V_b - \epsilon ; \quad \frac{\hbar^2 k_0^2}{2m_A} = \epsilon . \quad (A2)$$

The calculation consists in expressing the boundary conditions $[\chi(z)$ and $\frac{1}{m(z)} \frac{d\chi}{dz}$ continuous] at $z = 0$ and $z = L$ and in evaluating, by a first order expansion in b or b' , the values of $\chi(z)$ and $\frac{1}{m(z)} \frac{d\chi}{dz}$ at $z = 0$ [respectively $z = L$] as a function of $\chi(z = \pm b)$, $\frac{d\chi}{dz}(z = \pm b)$, [respectively $\chi(z = L \pm b')$, $\frac{d\chi}{dz}(z = L \pm b')$]. The validity of such linear interpolations requires that the transition regions be thin. More precisely $b\kappa_b$, bk_0 should be much smaller than 1. This is certainly true if

$$b \sqrt{\frac{2m_A V_b}{\hbar^2}} \ll 1 \quad \text{and} \quad b \sqrt{\frac{2m_B V_b}{\hbar^2}} \ll 1 \quad (\text{A3})$$

and similarly for b' .

For GaAs-Ga_{1-x}Al_xAs $\sqrt{\frac{\hbar^2}{2m_A V_b}} \sim 14 \text{ \AA}$ if $m_A \sim 0.067 m_0$ and $V_b = 0.3 \text{ eV}$.

Thus for transition regions extending over one monolayer (2.83 Å) the inequality (A3) will reasonably be fulfilled.

After calculations, one obtains the equation fulfilled by the energy levels in the quantum well with graded interfaces :

$$\begin{aligned} \eta^2 \sin k_0 \ell + 2\eta \cos k_0 \ell - \sin k_0 \ell + (b + b') k_0 \times \\ [-2\eta \sin k_0 \ell + (\eta^2 - 1) \cos k_0 \ell] + 0(bb') = 0 \quad (\text{A4}) \end{aligned}$$

where : $\eta = \frac{\kappa_b m_A}{k_0 m_B}$. (A5)

Equation (A4) can also be rewritten in the form

$$F(k_0 \ell, \varepsilon) + (b + b') k_0 \left[\frac{\partial F}{\partial (k_0 \ell)} \right]_{\varepsilon \text{ fixed}} + 0(bb') = 0 \quad (\text{A6})$$

where :

$$F(k_0 \ell, \varepsilon) = \eta^2 \sin k_0 \ell + 2\eta \cos k_0 \ell - \sin k_0 \ell \quad (\text{A7})$$

is nothing but the function whose zeros give the energy positions of the bound state in the square well of thickness ℓ .

To the linear approximation in b , b' , equation (A6) coincides with

$$F[k_0(\ell + b + b'), \varepsilon] \equiv F[k_0 L, \varepsilon] = 0. \quad (\text{A8})$$

Thus, irrespective of the exact spatial variations of $V_b(z)$ and $m^*(z)$ in the graded interface regions, the bound states equation of a graded quantum well coincides with that of a square quantum well whose effective thickness L is slightly larger (one half

of the sum of the transition regions thickness) than that of the nominal quantum well thickness ℓ .

References

- [1] see, e.g., W. KOHN, *Solid State Phys.* **5** (1957) 257.
- [2] J.S. BLAKEMORE, *J. Appl. Phys.* **53** (1982) R 123.
- [3] C. PRIESTER, G. ALLAN, M. LANNOO, *Phys. Rev.* **B28** (1983) 7194 and *Phys. Rev.* **B29** (1984) 3408.
- [4] G. BASTARD, *Phys. Rev.* **B24** (1981) 4714.
- [5] K. TANAKA, M. NAGAOKA and T. YAMABE, *Phys. Rev.* **B28** (1983) 7068.
- [6] S. CHAUDHURI, *Phys. Rev.* **B28** (1983) 4480.
- [7] C. MAILHIOT, YIA-CHUNG CHANG and T.C. MCGILL, *Phys. Rev.* **B26** (1982) 4449.
- [8] R.L. GREENE and K.K. BAJAJ, *Solid State Commun.* **45** (1983) 825.
- [9] R.J. WAGNER, B.V. SHANABROOK, J.E. FURNEAUX and J. COMAS, N.C. JAROSIK and B.D. McCOMBE in *GaAs and related compounds 1984* edited by B. de Crémoux, Institute of Physics Conference Series 74, (Adam Hilger Ltd, Bristol) 1985.
- [10] W.T. MASSELINK, YIA-CHUNG CHANG and H. MORKOC, *Phys. Rev.* **B28** (1983) 7373.
- [11] R.C. MILLER, A.C. GOSSARD, W.T. TSANG and O. MUNTEANU, *Phys. Rev.* **B25** (1982) 3871.
- [12] C. PRIESTER, G. BASTARD, G. ALLAN, M. LANNOO, *Phys. Rev.* **B30** (1984) 6029.
- [13] For a detailed theoretical treatment of excitons in bulk semiconductors see R.S. Knox *Theory of Excitons, Solid State Phys.*, Supplement 3 (Academic Press, New York) 1963.
- [14] G. BASTARD, E.E. MENDEZ, L.L. CHANG and L. ESAKI, *Phys. Rev.* **B26** (1982) 1974.
- [15] R.C. MILLER, D.A. KLEINMAN, W.T. TSANG and A.C. GOSSARD, *Phys. Rev.* **B24** (1981) 1134.
- [16a] R.L. GREENE, and K.K. BAJAJ, *Solid State Commun.* **45** (1983) 831.
- [16b] R.L. GREENE, K.K. BAJAJ and D.E. PHELPS, *Phys. Rev.* **B29** (1984) 1807.
- [17] see, e.g., N.W. ASHCROFT and N.D. MERMIN *Solid State Phys.* (Holt, Rinehart and Winston, New York) 1976.
- [18] F.F. FANG and W.E. HOWARD, *Phys. Rev. Lett.* **16** (1966) 797.
- [19] J.M. LUTTINGER, *Phys. Rev.* **102** (1956) 1030.
- [20] A. BALDERESCHI and N.O. LIPARI, *Phys. Rev.* **B3** (1971) 439.
- [21] M. ALTARELLI, *Phys. Rev.* **B28** (1983) 842.
- [22] A. FASOLINO and M. ALTARELLI in *Two-Dimensional Systems, Heterostructures and Superlattices* edited by G. Bauer, F. Kuchar and H. Heinrich, Springer Series in *Solid State Sci.* **43** (1984) 176.
- [23] P.J. PRICE and F. STERN, *Surf. Sci.* **132** (1983) 577.
- [24] F. STERN and J.N. SCHULMAN, *Superlattices and Microstructures* **1** (1985) 303.
- [25] P.M. PETROFF., *J. Vac. Sci. Technol.* **14** (1977) 973.
- [26] P.M. FRILINK, 1984, unpublished.
- [27] P. VOISIN, G. BASTARD, C.E.T. GONCALVES DA SILVA, M. VOOS, L.L. CHANG and L. ESAKI, *Solid State Commun.* **39** (1981) 79.
- [28] C.E.T. GONCALVES DA SILVA and P. FULCO, *Rev. Bras. Fis.* **12** (1982) 325.
- [29] N. HOLONYAK Jr., W.B. LAIDIG, B.A. VOJAK, K. HESS, J.J. COLEMAN, P.D. DAPKUS and J. BARDEEN, *Phys. Rev. Lett.* **45** (1980) 1703.

- [30] G. BASTARD, C. DELALANDE, M.H. MEYNADIER, P.M. FRIJLINK and M. VOOS, *Phys. Rev.* **B29** (1984) 7042.
- [31] M. COMBESCOT and C. BENOIT A LA GUILLAUME, *Solid State Commun.* **39** (1981) 651.
- [32] G. BASTARD, *Phys. Rev.* **B25** (1982) 7584.
- [33] A. CHOMETTE, B. DEVEAUD, J.Y. EMERY, A. REGRENY and B. LAMBERT in *Proceedings of the 17th International Conference on the Physics of Semiconductors*, San Francisco 1984, edited by J.D. Chadi and W.A. Harrison (Springer Verlag, New York) 1985, p. 441.

CHAPTER V

Energy levels in modulation-doped heterostructures.

In this chapter we describe the energy levels of heterostructures containing free carriers. The realization of a high-mobility quasi-bidimensional electron gas has been made possible thanks to the modulation-doping technique, which was first applied by Dingle *et al.* and Störmer [1] to GaAs-Ga_{1-x}Al_xAs heterostructures. We shall first discuss its physical origin. Then, once we have assumed the existence of quasi bidimensional carriers, we shall attempt to answer the practical question "How many transferred charges are there, if the heterostructure parameters are given ?".

This will lead us to a discussion of the size-quantization of the electron motion in a potential whose shape depends on the presence of free carriers. The latter feeds back on the energy levels of these carriers. The self-consistency requirement is an aspect of the energy level calculations in doped heterostructure, which was absent in undoped ones, where the confining potentials are *a priori* known and fixed.

The single heterojunctions quickly proved to be far superior to the double or multiple heterostructures, as far as the achievement of high mobility quasi bidimensional electron gas is concerned. However, the modulation doped quantum wells are interesting for their own sake and their quality could be greatly improved. We will also discuss their energy levels which present genuine features compared with those of single heterojunctions.

In all that follows, band structure effects, such as non-parabolicity, will be neglected. This is a fair approximation for Ga_{1-x}Al_xAs-GaAs heterojunctions and a reasonable one for InP-Ga_{0.47}In_{0.53}As. For heterostructures involving narrow gap materials (e.g. InSb or Hg_{1-x}Cd_xTe) band non-parabolicity should be taken into account [2].

For accumulation or inversion layers involving holes, the bulk valence band degeneracy precludes the use of any simple treatment. This topic, not yet settled in III-V heterostructures, will not be discussed here.

I. The modulation doping of heterojunctions. Qualitative aspects.

The idea of selective doping of heterostructures is very simple and has proved to be very powerful. During the growth of the epitaxial layer a dopant is added (e.g. Si in Ga_{1-x}Al_xAs for n type doping) at some place in the heterostructure. As the impurity levels depend on the impurity position in a medium which lacks for translational invariance, very different physical situations arise, depending on the relative positions of the doped part of the heterostructure and the interface. Let us examine how things change when going from an undoped heterojunction to a doped one. In the latter case, the impurity position will dramatically affects the resulting physical situation.

I.1 THE UNPERTURBED HETEROJUNCTION (FLAT BAND CASE). — Let us consider an abrupt heterojunction (Fig. 1) between a barrier-acting material B and a well-acting material A. In practice, one often finds GaAs, $\text{In}_{1-x}\text{Ga}_x\text{As}$, GaSb in place of A and $\text{Ga}_{1-x}\text{Al}_x\text{As}$, InP, AlSb in place of B. We consider for simplicity n-type doping and assume that the conduction bands in the A and B materials are nondegenerate, spherical and display quadratic dispersion relations upon the wavevector characterized by effective masses m_A , m_B . We denote by V_b the conduction band discontinuity $E_c^{(B)} - E_c^{(A)}$ at the interface $z = 0$. The envelope functions are the eigenstates of a Ben Daniel-Duke type of Hamiltonian [3] :

$$\left\{ \frac{1}{2m(z)} (p_x^2 + p_y^2) + \frac{1}{2} p_z \frac{1}{m(z)} p_z + V_b Y(-z) \right\} \psi(\mathbf{r}) = \varepsilon \psi(\mathbf{r}) \quad (1)$$

with the boundary conditions :

$$\frac{1}{m_A} \frac{\partial \psi}{\partial z} \Big|_{z=0^+} = \frac{1}{m_B} \frac{\partial \psi}{\partial z} \Big|_{z=0^-} \quad (2)$$

$$\psi(z=0^+) = \psi(z=0^-) \quad (3)$$

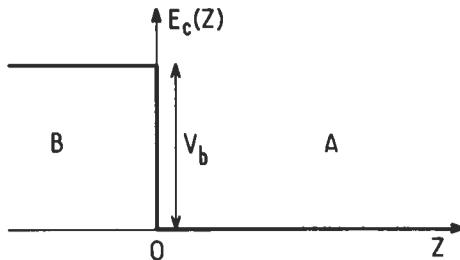


Fig. 1. — Conduction band profile of an abrupt heterojunction under flat band conditions.

If we denote the area of the heterostructure by $L_x L_y$, the exact solutions of equation (1) factorize into

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{L_x L_y}} \exp[i(k_x x + k_y y)] \chi(z) \quad (4)$$

For energies ε such that

$$\varepsilon - \frac{\hbar^2}{2m_B} (k_x^2 + k_y^2) < V_b \quad (5)$$

the wavefunction should decay exponentially in the barrier. Thus

$$\chi(z) = \begin{cases} B \exp \kappa_b z & z \leq 0 \\ A \sin(k_{\parallel} z + \delta) & z \geq 0 \end{cases} \quad (6)$$

with :

$$\frac{\hbar^2 \kappa_B^2}{2m_B} = V_B - \varepsilon + \frac{\hbar^2}{2m_B} (k_x^2 + k_y^2) \quad (7)$$

$$\frac{\hbar^2 \kappa_A^2}{2m_A} = \varepsilon - \frac{\hbar^2}{2m_A} (k_x^2 + k_y^2) \quad (8)$$

The boundary conditions at $z = 0$ only determine δ :

$$\tan \delta = \frac{k_{\parallel}}{\kappa_B} \frac{m_B}{m_A} \quad (9)$$

Thus, the allowed energy states form a three-dimensional continuum (labelled by k_{\parallel}, k_x, k_y) and the density of states $\rho(\varepsilon)$ is bulk-like : $\rho(\varepsilon) \sim \varepsilon^{1/2}$.

The single, undoped, heterojunction is of little practical use :

- i) it has no free carrier and thus cannot conduct the current ;
- ii) it displays a bulk-like density of states, which lacks for the energy shifts associated with size quantization which are so useful in quantum well systems.

I.2. SINGLE HETEROJUNCTION CONTAINING DILUTED COULOMBIC IMPURITIES. — As a simplifying assumption we assume that the dielectric permittivities of the A and B materials are identical. To the Hamiltonian equation (1) we have to add the Coulombic term

$$V(r) = \sum_i \frac{-e^2}{\kappa \sqrt{(\rho - \rho_i)^2 + (z - z_i)^2}} \quad (10)$$

where z_i is the impurity position along the growth axis, κ the static dielectric constant and ρ, ρ_i the two-dimensional position-vectors of the electron and the impurities in the layer plane. Two situations of unequal importance are met : either $z_i \geq 0$ or $z_i < 0$

- i) $z_i \geq 0$

When we consider the impurities as isolated, we notice that each of them supports at least one bound state. This is obvious if $z_i \gg a^*$, where a^* is the three-dimensional effective Bohr radius : the impurity state resembles that of a 1S hydrogenic level, with a binding energy R^* . Even if $z_i = 0$, a quasi $2P_z$ state is an approximate solution of the problem with a lower, yet finite, binding energy ($\sim R^*/4$). At very low temperatures ($k_B T \ll R^*$), the extra carriers remain frozen onto the impurity sites and the heterostructure is insulating (only a faint hopping conduction remains possible). When $k_B T \sim R^*$, a carrier detrapping takes place. The extended states of the heterojunction ($\varepsilon \geq 0$) which have been outlined in 1.1 become thermally populated. The electrical conduction in these extended states is possible, the carriers undergoing scattering by the positive ions. In fact, the situation is much the same as is found in lightly-doped bulk materials.

ii) $z_1 < 0$

When the impurities are placed in the barrier, one may repeat the previous reasoning. However, the bound states (quasi 1S if $|z_1| \geq a^*$) are now bound relative to the conduction band edge of the B material. (i.e. their energies are $\sim V_b - R^*$ above the onset of the heterojunction continuum discussed in 1.1). Clearly, the situation where all the electrons are frozen at $T = 0$ K on the impurity sites is an unstable one, for it does not ensure the equality of the chemical potential μ in both sides of the heterojunctions : on the one hand, μ should coincide with the donor levels and on the other hand μ should be negative since there are no free electrons in the A material. To establish thermal equilibrium, some of the carriers, assumed to be trapped at $t = 0$ onto the donor sites, escape by tunnel (or thermoionic) effect in the A material (Fig. 2). Once this has occurred, the electrons quickly loose energy by emitting phonons and relax to $\epsilon = 0$ (in a time scale of $\sim 10^{-12}$ s). This process takes place as long as available states exist below $\sim V_b - R^*$. The reverse process, i.e. recapture by the impurity, requires the electron to absorb a great many phonons in order to match the energy difference $\sim V_b - R^*$ (otherwise the tunnelling process will correspond to the capture of an electron moving quickly in the layer plane, which is both inefficient and unprobable).

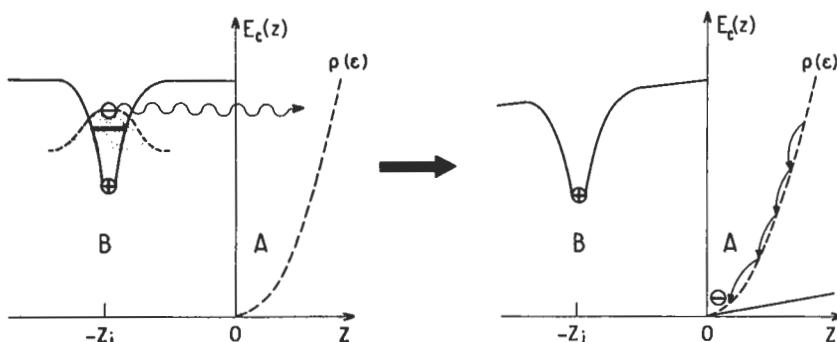


Fig. 2. — Composite sketch of the charge transfer mechanism in a single heterojunction whose barrier B has been selectively doped with donors. The electron Θ , assumed to be placed at $t = 0$ in a quasi bound state near the donor ion \oplus , can tunnel in the A material [density of states $\rho(\epsilon)$]. Once in the A layer, the electron quickly loses energy by emitting phonons and relaxes towards $\epsilon = 0$. As a result of the charge transfer, flat band conditions in the A layer are no longer ensured.

In fact, this is impossible at $T = 0$ K and difficult at room temperature since $V_b - R^* \sim 0.2$ eV and $k_B T \leq 25$ meV. Thus, by doping only the barrier-acting material, a spontaneous and irreversible charge transfer to the well-acting material is induced. The fact that the electrons and their parent donors are spatially separated has two important consequences :

i) A band bending takes place due to the dipole formed between the plus (ionized

donors) and the minus (electrons) charges. Averaging the donor distribution in the layer plane, this band bending only depends on z . The resulting potential shape felt by an electron is thus quasi triangular (see Fig. 3) near the interface, and is limited by the barrier at $z = 0$ and the z -dependent conduction band edge $E_c(z)$ at $z > 0$. As a result, bound states $E_1, E_2 \dots$ are formed for the z motion and, if the energy spacings $E_3 - E_2, E_2 - E_1 \dots$ between levels is much larger than the thermal or collisional broadenings, the carrier motion becomes effectively quasi bi-dimensional.

MODULATION DOPING

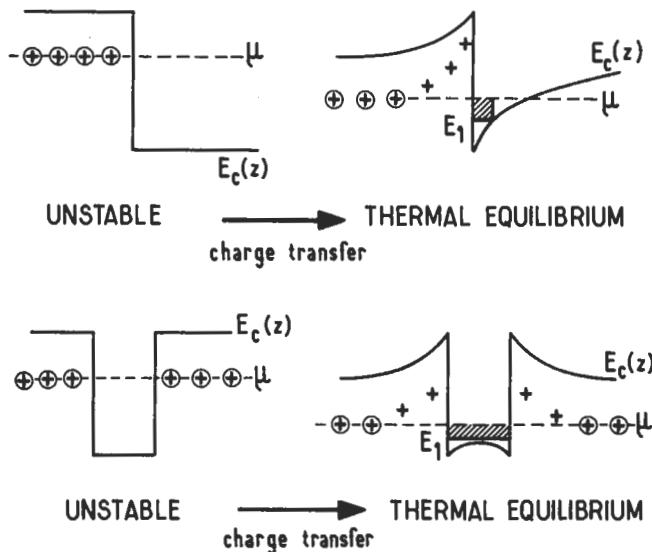


Fig. 3. — Charge transfer effects in modulation-doped single and double heterostructures. At thermal equilibrium, band bending takes place. The resulting potential felt by the electron is quasi-triangular near the interfaces.

ii) Due to the presence of free carriers, electrical conduction can take place, even at $T = 0$ K, in the layer plane. On the other hand, the current flow along the z axis is very difficult due to the size quantization of the electron z -motion. The remarkable advantage of the modulation-doping, as far as impurity scattering is concerned, is the tremendous improvement in the mobility brought about by the spatial separation between the carriers and their parent donors. This spatial separation can be further enhanced by inserting a spacer, which is a nominally undoped part of the barrier, between the donors and the quasi bi-dimensional electron gas. The low-temperature mobility of quasi bi-dimensional electron gases in heterostructures will be discussed in the following chapter.

I.3 ACTUAL HETEROSTRUCTURES. — In reality, semiconductor heterojunctions contain residual impurities. N_d donors per cubic centimeter have been deliberately introduced in the barrier. However, assuming a p-type residual doping for both kinds of layers, one should also deal with N_A ($N_{b,A}$) unwanted acceptors in the well-acting (hereafter abbreviated as channel) and the barrier-acting materials respectively (Fig. 4). Apart from their harmful effect on the electron mobility, these residual impurities influence the energy levels E_1 , E_2 ... and thus the amount of charge transfer which one can obtain in the heterostructure.

The interface grading, that is to say the fact that there exists a transition region of finite thickness between the barrier and the channel also affects the energy levels E_i . However, this effect is modest in GaAs-Ga_{1-x}Al_xAs heterostructures as shown by Stern *et al.* [4, 5]. For other pairs of materials, the grading may be much more

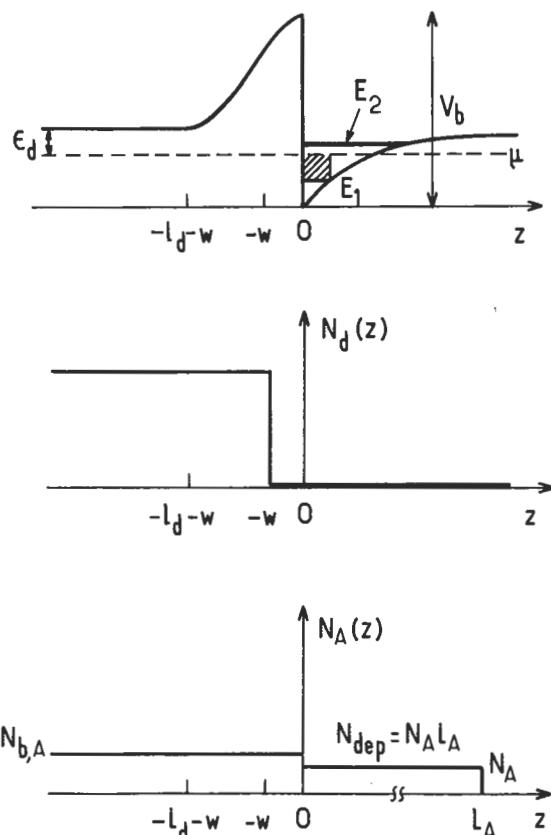


Fig. 4. — Assumed band bending and impurity distribution in a single modulation doped heterojunction. Upper panel: conduction band edge profile near the interface. Middle panel: donor profile in the heterojunction. Lower panel: ionized acceptor profile in the heterojunction.

severe, essentially due to an imperfect control of the interdiffusion between the two materials.

The barrier is of finite thickness. Due to technological reasons, the barrier is often grown on top of the channel. The interface between the barrier and the vacuum (or the ambient) contains several defects which act as deep traps. As a result, the chemical potential is pinned to these deep surface traps. This induces another band bending which, if the barrier is too thin, may overlap with that due to the charge transfer. In all that follows we shall admit that the barrier is thick enough to be able to neglect the effect of surface depletion on charge transfer.

Finally, we shall only consider ungated heterostructures. In gated ones, e.g. the Two-Dimensional Electron Gas Field Effect Transistors (T.E.G.F.E.T.), the charge transfer in the channel is controlled by means of a Schottky contact placed on top of the barrier. For a quantum analysis of gated heterostructures the reader is referred to Vinter [6].

II. Self-consistent calculations of energy levels and charge transfer in single heterojunctions.

The fact that the electronic z motion is bound at large z by the electrostatic potential created by residual impurities and the electrons themselves, means that the charge distribution reacts on the confining potential, which itself determines the charge distributions. This imposes a stringent self consistency requirement on the energy level calculations.

Here, we shall deal with the electron-electron interaction, a many body problem, within the simplest scheme : the Hartree approximation. It amounts to replacing the exact many-electrons potential by an average one. Each electron is assumed to move in a self-consistent potential $\Phi_{s.c.}(\mathbf{r})$ and we have to solve the coupled Poisson and Schrödinger equations

$$\nabla^2 \Phi_{s.c.}(\mathbf{r}) = \frac{4\pi e}{\kappa} \left[\sum_{\alpha \text{ occupied}} |\psi_{\alpha}(\mathbf{r})|^2 - \sum_{\mathbf{R}_{\text{don}}} \delta(\mathbf{r} - \mathbf{R}_{\text{don}}) + \sum_{\mathbf{R}_{\text{acc}}} \delta(\mathbf{r} - \mathbf{R}_{\text{acc}}) \right] \quad (11)$$

$$\left[-\frac{\hbar^2}{2m(z)} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) - \frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{m(z)} \frac{\partial}{\partial z} + V_b(z) - e\Phi_{s.c.}(\mathbf{r}) \right] \psi_{\alpha}(\mathbf{r}) = \varepsilon_{\alpha} \psi_{\alpha}(\mathbf{r}) \quad (12)$$

In equations (11, 12), κ is the relative dielectric constant of the heterostructure. \mathbf{R}_{acc} and \mathbf{R}_{don} label the sites effectively occupied by an ionized acceptor or donor. These sites are distributed at random in the (x, y) plane, but not along the z axis on a macroscopic scale (selective doping). $m(z)$ is the position dependent effective mass and $V_b(z)$ the conduction band profile along the z axis in the absence of impurities and free carriers. If there is no grading at the interface, taken as the origin of the z coordinate, $V_b(z)$ reduces to :

$$V_b(z) = V_b Y(-z) \quad (13)$$

where $Y(x)$ is the step function.

We now make the usual approximation of doped semiconductors, which consists of splitting the impurity contributions to equations (11, 12) into two parts. The first part is the spatial average of the impurity potentials $N_d^+(z)$, $N_A^-(z)$

$$\begin{aligned} N_d^+(z) &= \frac{1}{\omega} \int_{\omega} \delta(\mathbf{r} - \mathbf{R}_{\text{don}}) d^3 R_{\text{don}}; \\ N_A^-(z) &= \frac{1}{\omega} \int_{\omega} \delta(\mathbf{r} - \mathbf{R}_{\text{acc}}) d^3 R_{\text{acc}} \end{aligned} \quad (14)$$

where ω is a volume which is macroscopic (i.e. contains many impurities). The second part is the local deviation from this average. The first term, which still depends on the z coordinate due to the presence of the heterostructure and to selective doping, will be retained in equations (11, 12). The second one will be left over for the energy level calculations. Due to a lack of translational invariance in the layer plane, it will scatter the charge carriers. Its effects will be more thoroughly investigated in the next chapter.

Once the translational invariance in the layer plane has been restored, a number of simplifications of equations (11, 12) follow. $\psi_a(\mathbf{r})$ can be factorized :

$$\psi_a(\mathbf{r}) = \frac{1}{\sqrt{S}} \exp i [(k_x x + k_y y)] \chi_i(z) \quad (15)$$

where $\chi_i(z)$ is the solution of the one-dimensional Schrödinger equation :

$$\begin{aligned} \left[-\frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{m(z)} \frac{\partial}{\partial z} + \frac{\hbar^2}{2m(z)} (k_x^2 + k_y^2) + V_b(z) - e\Phi_{\text{s.c.}}(z) \right] \chi_i(z) = \\ = \varepsilon_i(k_x, k_y) \chi_i(z) \end{aligned} \quad (16)$$

$\Phi_{\text{s.c.}}(z)$ being the solution of a one-dimensional Poisson equation. In equation (16), the position-dependence of the effective mass induces a coupling between the electron motions along and perpendicular to the z axis. However, this effect is usually small [7] and hardly distinguishable from the band non parabolicity of the host materials. We can get rid of this inconvenient coupling in the following way. We replace $m^{-1}(z)$ in front of $(k_x^2 + k_y^2)$ by $m_i^{-1} + (m^{-1}(z) - m_i^{-1})$, where m_i is a suitable in-plane effective mass of the electron while in the i^{th} subband. The zeroth order term in $(m^{-1}(z) - m_i^{-1})$ gives us the subband dispersion relation :

$$\varepsilon_i(k_x, k_y) = E_i + \frac{\hbar^2}{2m_i} (k_x^2 + k_y^2) \quad (17)$$

where E_i is the eigenvalue of equation (16) at $k_x = k_y = 0$. The first order correction in $m^{-1}(z) - m_i^{-1}$ will vanish if we define m_i as

$$\frac{1}{m_i} = \int dz \chi_i^2(z) \frac{1}{m(z)} \quad (18)$$

The second order corrections will involve fourth order terms in k_x, k_y , which we shall discard.

The one-dimensional Poisson equation which has to be solved, together with the Schrödinger equation is :

$$\frac{d^2\Phi_{s.c.}}{dz^2} = \frac{4\pi e}{\kappa} \left[\sum_{i \text{ occupied}} n_i \chi_i^2(z) - N_d^+(z) + N_A^-(z) \right] \quad (19)$$

where n_i is the areal concentration of electrons which occupy the i^{th} subband. At finite temperature T , the chemical potential of the electrons μ and the quantities n_i, E_i, m_i are related by :

$$n_i = m_i \frac{k_B T}{\pi \hbar^2} \ln \left[1 + \exp \left(\frac{\mu - E_i}{k_B T} \right) \right] \quad (20)$$

where k_B is the Boltzmann constant. At $T = 0$ K, the equation (20) reduces to the familiar expression

$$n_i = \frac{m_i}{\pi \hbar^2} (\mu - E_i) Y(\mu - E_i) \quad (21)$$

The set of equations (16, 21) needs to be completed by boundary conditions. As for the Schrödinger equation (16), the bound state envelope function $\chi_i(z)$ should go to zero when $z \rightarrow \pm \infty$, and, like $m^{-1}(z) \frac{d\chi_i}{dz}$, should be continuous everywhere.

As for the Poisson equation (19), we require the heterojunction to be in electrical equilibrium. This implies the charge balance

$$\sum_{i \text{ occupied}} n_i + \int_{-\infty}^{+\infty} dz [N_A^-(z) - N_d^+(z)] = 0 \quad (22)$$

or equivalently that the self-consistent electric fields, $-\Phi'_{s.c.}(z)$ are the same at $z = \pm \infty$.

In addition, the heterostructure is in thermodynamical equilibrium. This requires the chemical potential to be constant. For a residual p-type doping of the channel, μ should coincide with the acceptor levels far away in the bulk. On the other hand, when far in the barrier, μ should coincide with the donor levels. Thus :

$$E_v(z = +\infty) + \varepsilon_A = E_c(z = -\infty) - \varepsilon_d \quad (23)$$

where $\varepsilon_d(\varepsilon_A)$ is the binding energy of the donor (acceptor) in the barrier (channel) and $E_c(E_v)$ are the conduction (valence) band edges (see Fig. 5).

At $T = 0$ K one can introduce further simplifications linked to the sharpness of the Fermi-Dirac distribution function. In particular, the notion of acceptor and donor depletion lengths becomes unambiguous. The acceptor depletion length l_A is computed by writing that all the channel acceptors are ionized between $z = 0$ and $z = l_A$, whereas beyond $z = l_A$ they are neutral (or only compensate minority

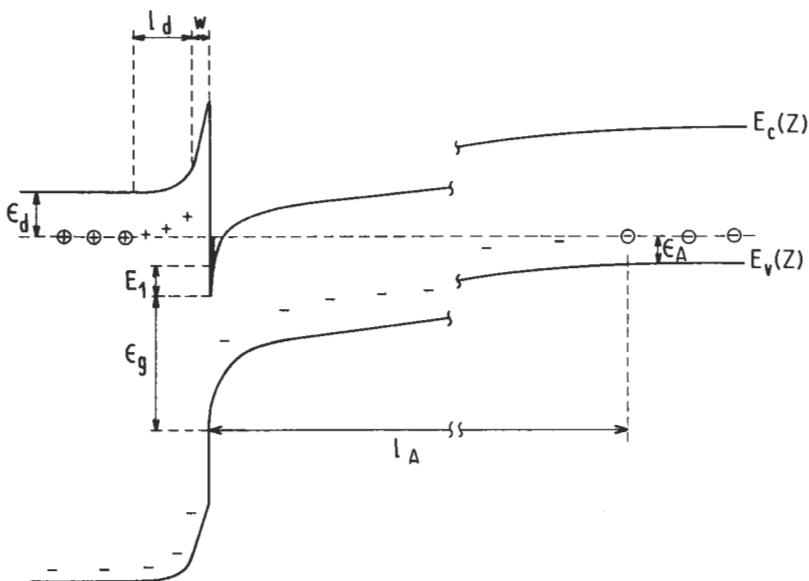


Fig. 5. — Position dependence of the conduction and valence band edges in a modulation-doped heterojunction whose channel has a residual p-type doping. The acceptors are negatively charged for $z < l_A$. Beyond $z = l_A$ the bands are flat. w is the spacer thickness and l_d the donor depletion length. For $z < -l_d - w$ the bands are flat.

donors). Thus, if we only retain that part $\delta\Phi_A(z)$ of $\Phi_{s.c.}(z)$, which is due to acceptors, we have $\delta\Phi'_A(l_A) = 0$. Taking the zero of the electrostatic potential at $z = 0$, one then obtains

$$\delta\Phi_A(z) = \frac{2\pi e}{\kappa} N_A^- z(z - 2l_A) \quad (24)$$

where we have made use of the fact that the electric field vanishes at $z = l_A$. The acceptor depletion length is determined by writing

$$-e\delta\Phi_A(l_A) = \frac{2\pi e^2 N_A^- l_A^2}{\kappa} = \epsilon_g - \epsilon_A \quad (25)$$

or:

$$l_A = \sqrt{\frac{\kappa(\epsilon_g - \epsilon_A)}{2\pi e^2 N_A^-}} \quad (26)$$

The depletion length varies like $(N_A^-)^{-1/2}$. For GaAs channels containing 10^{14} cm^{-3} acceptors, one finds $l_A \sim 4.6 \mu\text{m}$. The acceptor depletion length is thus considerably larger than the spatial extension of the χ_i 's. Therefore, to solve the Schrödinger equation, one may safely approximate $\delta\Phi_A(z)$ by the linear law

$$\delta\Phi_A(z) = -\frac{4\pi e}{\kappa} N_{\text{dep}} z \quad (27)$$

where :

$$N_{\text{dep}} = N_{\text{A}}^- l_{\text{A}} \quad (28)$$

is the equivalent areal concentration of the depletion charges. For GaAs channels containing 10^{14} cm^{-3} acceptors, $N_{\text{dep}} \approx 0.46 \times 10^{11} \text{ cm}^{-2}$. We may proceed in the same way with the donor depletion length l_{d} . The doped part of the barrier is assumed to extend from $z = -\infty$ to $z = -w$, where w is the spacer thickness. From $z = -\infty$ to $z = -l_{\text{d}} - w$ some donors are ionized to compensate for the residual barrier acceptors (volume concentration N_{bA}) but the net ionized donor concentration is zero. The electric field which is zero at $z = -\infty$ should also be zero at $z = -l_{\text{d}} - w$. In between $z = -l_{\text{d}} - w$ and $z = -w$ all the donors are ionized. Thus, the $\delta\Phi_{\text{d}}(z)$ part of $\Phi_{\text{s.c.}}$ which is due to the ionized donors, varies in the following way

$$\delta\Phi_{\text{d}}(z) = -\frac{2\pi e}{\kappa} (N_{\text{d}} - N_{\text{b,A}}) (z + l_{\text{d}} + w)^2 + \Phi_0 \quad (29)$$

where Φ_0 is a constant which has to be determined by continuity of the electrostatic potential at $z = -w$.

In the depletion length approximation the charge balance equation can be rewritten as

$$N_{\text{d}} l_{\text{d}} = \sum_i n_i + N_{\text{dep}} + N_{\text{b,A}} (w + l_{\text{d}}) \quad (30)$$

Note that l_{d} has still to be determined self-consistently from equations (16-21, 23, 30).

The major simplification, which renders the self consistent treatment of the heterostructures almost analytically tractable, is the separation between the Poisson and Schrödinger equations on the one hand and the equilibrium conditions on the other.

Such a separation is made possible thanks to the negligible penetration in the barrier of the electron wavefunction $\chi_1(z)$. If this condition is fulfilled, the details of the electrostatic potential in the barrier (its charge state, the actual value of the donor depletion length etc..) do not significantly influence χ_i and E_i . As a result, we may split the self consistent loop into two parts.

i) We assume that the n_i 's are given in order to solve the Poisson and Schrödinger equations. We then obtain $E_i(n_i, N_{\text{dep}})$.

ii) We insert the functionals $E_i(n_i, N_{\text{dep}})$ into the two equilibrium relationships and obtain the equilibrium value of the transferred charge $\sum_i n_i = n_{\text{e}}$. This value will

depend on several parameters : V_{b} , ϵ_{d} , w , N_{d} , $N_{\text{b,A}}$ and it may be of interest to ascertain whether the measured charge transfer can be reasonably interpreted.

II.1 ENERGY LEVEL CALCULATIONS. — The most elaborate energy level calculations in $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterostructures are the numerical calculations by Stern and Das Sarma [4] and Ando [7]. These authors went beyond the Hartree approximation by including the exchange and correlation effects within the local density approxi-

mation scheme. It turns out that for conduction subbands, the exchange and correlation terms are small compared to the Hartree terms. Thus, we can reasonably limit ourselves to the Hartree approximation.

In the Hartree approximation, a convenient and transparent solution of the coupled Poisson and Schrödinger equations is obtained using a self-consistent variational procedure [8, 9]. In many circumstances, only a single (E_1) subband is populated at $T = 0$ K (Electric Quantum Limit). Thus, one only has to imagine a plausible shape for χ_1 in order to determine $\Phi_{s.c.}(z)$ and in turn E_1 by a minimization procedure.

Let us suppose that the trial wavefunction χ_1 depends on the variational parameters $\beta_1, \beta_2, \dots, \beta_n$. Then, by applying the variational principle, one has to minimize the *total* energy of the n_e electrons with respect to $\beta_1, \beta_2, \dots, \beta_n$. Since all the electrons are assumed to be in the lowest quantum states for their z motion, the extremalization procedure reduces to a minimization of the *modified* one-electron energy \tilde{E}_1 , in which that part of the self-consistent potential which arises from the interaction of a given electron with the others, is divided by 2. In such a way, we avoid double-counting the electron-electron interaction. Once \tilde{E}_1 has been minimized with respect to $\beta_1, \beta_2, \dots, \beta_n$, the best trial wavefunction $\chi_1(z, \beta_1^{\min}, \beta_2^{\min}, \dots, \beta_n^{\min})$ is known, and the one-electron energy E_1 is obtained as the expectation value of the full one-electron Hamiltonian over $\chi_1(z, \beta_1^{\min}, \beta_2^{\min}, \dots, \beta_n^{\min})$.

Long ago, Fang and Howard [10] when dealing with the inversion layers of Si, proposed the trial wavefunction

$$\begin{cases} \chi_1(z) = 0 & z \leq 0 \\ \chi_1(z) = \sqrt{\frac{b^3}{2}} z \exp\left(-\frac{bz}{2}\right) & z \geq 0 \end{cases} \quad (31)$$

where b is the trial parameter. The related algebra has been fully developed by Stern [11]. In equation (31) it is assumed that the wavefunction vanishes in the barrier. This assumption is certainly much worse in the III-V heterostructures ($V_b \sim 0.3$ eV) than in the Si-SiO₂ system ($V_b \sim 2$ eV). Thus, we could improve the Fang-Howard solution by allowing χ_1 to leak into the barrier. This is readily achieved by choosing a trial wavefunction of the form

$$\begin{cases} \chi_1(z) = M \exp(\kappa_b z/2) & z \leq 0 \\ \chi_1(z) = N(z + z_0) \exp(-bz/2) & z \geq 0 \end{cases} \quad (32)$$

By expressing the continuity conditions at $z = 0$ and the wavefunction normalization, it is possible to calculate M, N, z_0 in terms of the two independent variational parameters b and κ_b . Actually, if our decoupling procedure is going to work well, E_1 should depend much less on κ_b than on b . In fact, one may fix κ_b at the value

$$\kappa_b = 2 \sqrt{\frac{2m_b V_b}{\hbar^2}} \quad (33)$$

and use b as the single variational parameter.

A shortcoming of the trial solution (Eq. (32)) is that it does not properly decay to zero in the region $\frac{3}{b} \ll z \ll l_A$ where $-e\Phi_{s.c.}(z)$ reduces to the depletion term $-e\delta\Phi_A(z)$ due to the ionized acceptors in the channel. For such a linearly varying potential, it is known that the exact asymptotic form for $\chi_1(z)$ is $\exp(-az^{3/2})$. Instead, $\chi_1(z)$ decreases in the same way as $z \exp(-bz/2)$ (see Eq. (32)). This introduces a few meV errors for E_1 .

We present in figures (6-11) some numerical results obtained for E_1 in $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ and $\text{InP}-\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ heterojunctions.

For $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions $V_b \sim 1060x$ meV if one adopts Dingle's rule [12] ($Q_c = \frac{\Delta E_c(x)}{\Delta E_g(x)} = 85\%$ where $\Delta E_g(x)$ is the bandgap energy difference

between $\text{Ga}_{1-x}\text{Al}_x\text{As}$ and GaAs , and $\Delta E_c(x)$ is the conduction band offset between the two materials, i.e. V_b). Using the recently proposed value of $Q_c = 60\%$ [13, 14], V_b varies like $\sim 750x$ meV. In tables I and II, are listed some of the characteristics of the energy levels and wavefunctions for a $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}-\text{GaAs}$ heterojunction. This aluminum percentage corresponds to $V_b = 0.318$ eV (Tab. I) or $V_b = 0.225$ eV (Tab. II), depending on which Q_c value is chosen.

The conduction band offset of the $\text{InP}-\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ heterojunction is not very well known, the experimental determination varying from 0.2 eV to 0.53 eV. In

Table I. — List of some of the average values of physical quantities which characterize the quasi bi-dimensional electron gas located near the interface ($z = 0$) of a $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ modulation doped heterojunction. The notation $\langle A \rangle$ stands for $\int \chi_1^2(z) A(z) dz$ where χ_1 is the self-consistent variational solution of the heterojunction problem. The Electric Quantum Limit is assumed. $N_{\text{dep}} = 0.46 \times 10^{11} \text{ cm}^{-2}$; $m_A = 0.07 m_0$; $m_B = 0.088 m_0$; $V_b = 318$ meV i.e. $x = 0.3$ if $Q_c = 85\%$. The notation $\langle A \rangle^{\text{Fang Howard}}$ means that V_b has been taken as infinite.

$n_c (10^{11} \text{ cm}^{-2})$	2	4	6	8	10
E_1 (meV)	32.00	45.6	56.71	66.30	74.68
$\langle -e\Phi_{s.c.}(z) Y(z) \rangle$ (meV)	17.85	29.05	37.76	45.07	51.13
$\langle -e\Phi_{s.c.}(z) Y(-z) \rangle$ (meV)	-0.015	-0.044	-0.092	-0.16	-0.25
$\langle V_b Y(-z) \rangle$ (meV)	2.22	3.54	4.86	6.2	7.62
$P_b = \langle Y(-z) \rangle$ (%)	0.7	1.11	1.53	1.95	2.40
z_0 (Å)	12.97	12.91	12.91	12.94	12.99
$\langle z \rangle$ (Å)	100.03	82.56	72.43	65.55	60.14
$\langle z \rangle_1^{\text{Fang Howard}}$ (Å)	116.09	99.27	89.28	82.39	77.23
$E_1^{\text{Fang Howard}}$ (meV)	38.63	56.60	72.1	86.09	99.04
$\langle -e\delta\Phi_A(z) \rangle$ (meV)	5.09	4.04	3.44	3.03	2.72

Table II. — Same as table I except that $V_b = 225$ meV, i.e. $x = 0.3$ if $Q_c = 60\%$.

$n_c (10^{11} \text{ cm}^{-2})$	2	4	6	8	10
$E_1 (\text{meV})$	30.9	43.71	53.87	62.4	69.87
$\langle -e\Phi_{s.c.}(z) Y(z) \rangle (\text{meV})$	17.06	27.51	35.2	41.33	46.58
$\langle -e\Phi_{s.c.}(z) Y(-z) \rangle (\text{meV})$	-0.03	-0.1	-0.21	-0.38	-0.62
$\langle V_b Y(-z) \rangle (\text{meV})$	2.59	4.17	5.84	7.64	9.5
$P_b = \langle Y(-z) \rangle (\%)$	1.15	1.85	2.6	3.4	4.23
$z_0 (\text{\AA})$	15.34	15.38	15.48	15.64	15.85
$\langle z \rangle (\text{\AA})$	96.34	79.24	68.84	61.70	56.62
$\langle z \rangle_1^{\text{Fang Howard}} (\text{\AA})$	116.09	99.27	89.28	82.39	77.23
$E_1^{\text{Fang Howard}} (\text{meV})$	38.63	56.60	72.1	86.09	99.04
$\langle -e\delta\Phi_A(z) \rangle (\text{meV})$	4.73	3.71	3.1	2.69	2.39

addition, the InP-Ga_{0.47}In_{0.53}As heterostructures, which have mostly been grown by low pressure Metal-Organic-Chemical-Vapor-Deposition, display a n-type residual doping of the Ga_{0.47}In_{0.53}As channel. Thus, in this system one deals with accumulation layers instead of inversion layers. In the case of weak accumulation, the chemical potential should coincide at large z with the shallow donors of Ga_{0.47}In_{0.53}As. Thus, from $z = 0$ to $z = l_d$ the majority donors have captured one electron per site and what remains are the minority acceptors which contribute to $-e\Phi_{s.c.}(z)$. Weak accumulation or inversion layers can thus be handled with the same formalism, except that the equivalent two-dimensional depletion charge N_{dep} is smaller in the former case than is the latter.

Several features are noticeable in figures 6 to 11 and tables I and II.

Firstly, the central assumption underlying the decoupling procedure, which is the faint penetration of the electron wavefunction in the barrier, is very well justified in both kinds of heterojunctions. This is witnessed by the small integrated probability P_b of finding the electron in the barrier (few %) and, concomitantly, by the very small effect of the self-consistent potential in the barrier on the energy levels.

Despite the small leakage of the χ_1 wavefunction in the barrier, it should be noticed that the confinement energy E_1 , is much smaller than the one obtained by letting V_b be infinite (labelled $E_1^{\text{Fang Howard}}$ in Tabs. I, II). The wavefunction at finite V_b resembles that at infinite V_b except that it has been shifted by some $\sim 10 \text{ \AA}$ towards the interface (Fig. 10).

The energy E_1 increases sublinearly with n_c at large n_c . When $n_c \gg N_{\text{dep}}$, E_1 varies roughly like $n_c^{2/3}$, which is reminiscent of the E_1 variation versus the slope of a potential which increases linearly with z in the channel (see Appendix B). At vanishing n_c , E_1 extrapolates to a finite value. The latter coincides with the confinement energy of an electron bound in the exactly triangular potential formed by the depletion potential (Eq. (27)) and the barrier potential $V_b Y(-z)$.

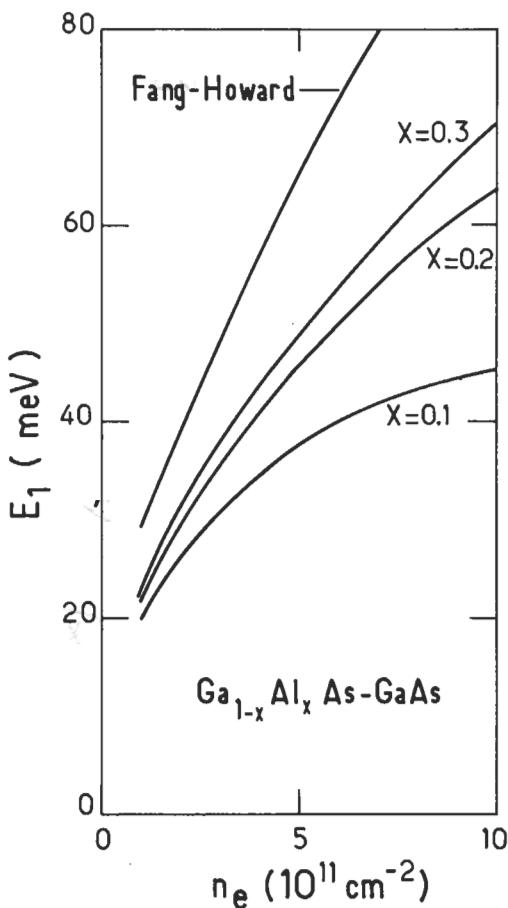


Fig. 6. — The confinement E_1 of the ground subband is plotted *versus* the two-dimensional electron concentration n_e for several Al concentrations x in $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ single heterojunctions. In the calculations, the following parameters have been used, $m_A = 0.07 m_0$, $m_B = (0.07 + 0.06 x) m_0$, where m_0 is the free electron mass. The conduction band offset is taken as $V_b(x) = 748.2 x$ meV. $N_{\text{dep}} = 4.6 \times 10^{11} \text{ cm}^{-2}$. The curve labelled Fang-Howard corresponds to infinite V_b .

The confinement energy E_1 is sensitively affected by the concentration of the residual acceptors in the channel (see Fig. 11 for $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterojunctions). Good GaAs channels are believed to contain $\sim 10^{14} \text{ cm}^{-3}$ acceptors. This affects E_1 by several meV's as n_e varies. At low n_e the residual doping contributes to a significant fraction of E_1 . The latter increases by more than 30 % when going from the quasi-accumulation regime ($N_A = 10^8 \text{ cm}^{-3}$, $N_{\text{dep}} = 0.46 \times 10^8 \text{ cm}^{-2}$) to $N_A = 9 \times 10^{14} \text{ cm}^{-3}$ ($N_{\text{dep}} = 1.38 \times 10^{11} \text{ cm}^{-2}$).

The residual doping is an even more prevalent factor in the determination of the

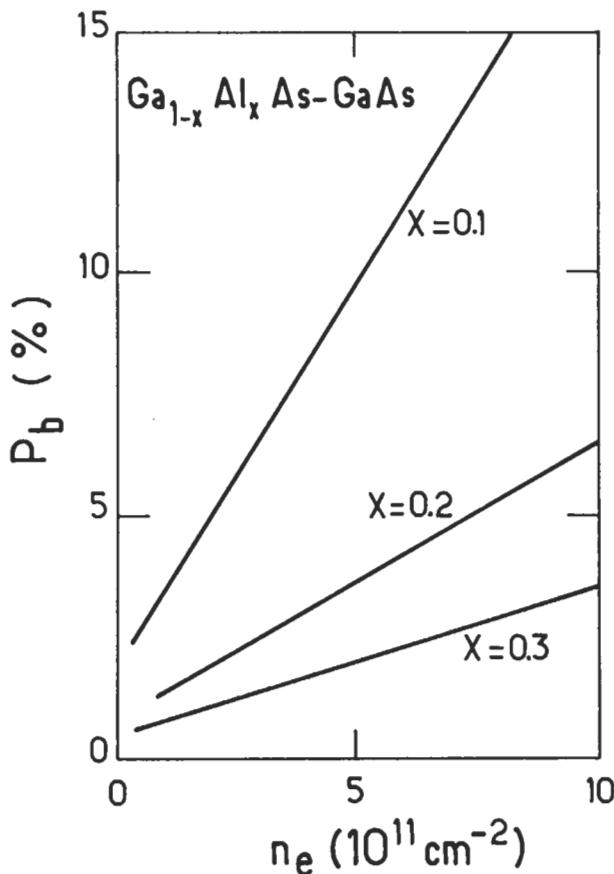


Fig. 7. — The integrated probability P_b of finding the electron in the barrier for the E_1 state is plotted versus n_e , the two-dimensional electron concentration, for several Al concentrations x in $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions. The material parameters are the same as those used in figure 6.

confinement energy E_2 of the second subband. By a straightforward generalization of equation (32), the wavefunction χ_2 can be sought in the form:

$$\chi_2(z) = \begin{cases} M_2 \exp(\kappa_b z/2) \\ N_2(z + \alpha)(z + \beta) \exp(-cz/2) \end{cases} \quad (34)$$

where the parameters M_2 , N_2 , α , β can be expressed in terms of c and κ_b by means of the continuity relations at $z = 0$, the wavefunction normalization and the required orthogonality between χ_1 and χ_2 . As in the E_1 case, E_2 is predominantly dependent on c rather than on κ_b . In the Electric Quantum Limit, χ_2 does not contribute to

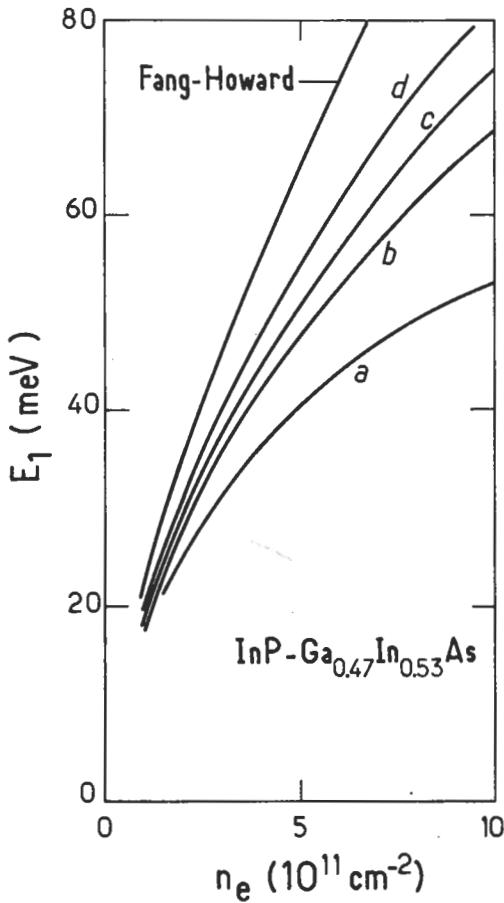


Fig. 8. — The confinement energy E_1 of the ground electron subband is plotted *versus* the two-dimensional electron concentration n_e in $\text{InP-Ga}_{0.47}\text{In}_{0.53}\text{As}$ heterojunctions for several values of the assumed conduction band discontinuity V_b . The curves, labelled a, b, c, d, correspond to $V_b = 0.1$ eV, 0.2 eV, 0.3 eV and 0.53 eV respectively. In the calculations, quasi-accumulation conditions have been assumed: $N_{\text{dep}} = 10^7 \text{ cm}^{-2}$; $m_B = 0.08 m_0$; $m_A = 0.047 m_0$.

$\Phi_{\text{s.c.}}(z)$. Thus, the variational procedure is a conventional one and to obtain E_2 we only have to minimize $\langle \chi_2 | \mathcal{H} | \chi_2 \rangle$ with respect to c .

Since χ_2 is orthogonal to χ_1 , it is quite small (and in fact has a node) in the region where χ_1 and thus the electron-electron part of $-e\Phi_{\text{s.c.}}(z)$ are important. The binding of the E_2 state is thus essentially provided by the depletion potential, which explains the marked sensitivity of the E_2 confinement energy to N_{dep} . This has an important consequence on the eventual population of the E_2 level.

The energy difference $\mu - E_1$ increases linearly with n_e . On the other hand, the subband spacing $E_2 - E_1$ increases sublinearly with n_e . Thus, there should exist some

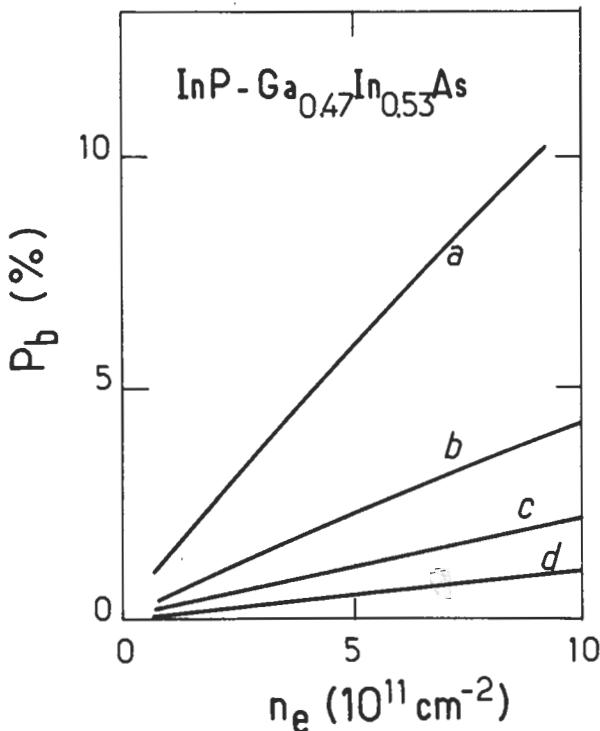


Fig. 9. — The integrated probability P_b of finding the electron in the InP barrier is plotted versus the two-dimensional electron concentration n_e for several values of the assumed barrier height V_b in InP-Ga_{0.47}In_{0.53}As heterojunctions. The material parameters are the same as used in figure 8.

critical concentration n_e^c beyond which E_2 becomes populated. In figure 12 we present the calculated variation of n_e^c upon the depletion charge N_{dep} , together with the results of Ando [7] and Stern and Das Sarma [4]. All the calculations agree in demonstrating a strong increase of n_e^c with N_{dep} . For 10^{14} cm^{-3} residual acceptors ($N_{dep} = 0.46 \times 10^{11} \text{ cm}^{-2}$) the second subband should become populated near $n_e^c = 7 \times 10^{11} \text{ cm}^2$ at $T = 0 \text{ K}$ if $V_b = 0.3 \text{ eV}$. The relatively good agreement between the self consistent variational treatment performed in the Hartree approximation and the exact numerical calculations, which include many-body corrections, is probably a result of cancellations between the approximations inherent to the former type of calculations. Nevertheless, one may conclude from figure 12 that despite shortcomings, the approximate treatment remains useful.

II.2 CHARGE TRANSFER IN SINGLE HETEROJUNCTIONS. — Once the functional dependences $E_i(n_e)$ have been established, we can solve the two equilibrium

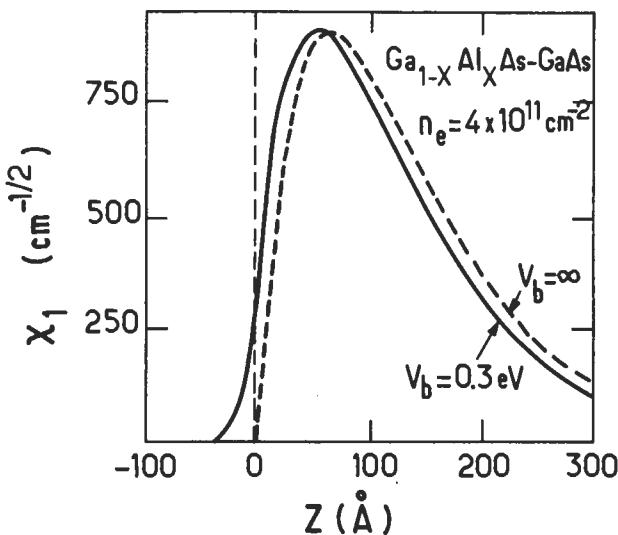


Fig. 10. — An example of the spatial dependence of the wavefunction amplitude of the E_1 state in modulation-doped $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterojunctions. $V_b = 0.3 \text{ eV}$; $n_e = 4 \times 10^{11} \text{ cm}^{-2}$; $N_{\text{dep}} = 4.6 \times 10^{10} \text{ cm}^{-2}$.

equations (23, 30). Expressing the donor depletion length l_d in terms of n_e by means of equation (30), we finally obtain the equation which governs the transferred charge at equilibrium :

$$E_1 + \frac{\pi \hbar^2 n_e}{m_1} = V_b - \varepsilon_d - \frac{4\pi e^2}{\kappa} n_e \int_{-\infty}^0 z \chi_1^2(z) dz - \frac{2\pi e^2}{\kappa} \times \{ (N_d - N_{b,A}) l_d (l_d + 2w) - N_{b,A} w^2 \} \quad (35)$$

In the right-hand side of equation (35) the term involving $\chi_1(z)$ is usually very small, a direct consequence of the validity of the decoupling procedure. Several factors influence the charge transfer but to quite different extents. This is illustrated in figures 13 to 15.

- i) the dominant effect is the monotonic decrease of the charge transfer as the spacer thickness w increases.
- ii) the donor concentration N_d which has been deliberately introduced into the barrier hardly influences the charge transfer, at least for usual concentrations ($N_d > 10^{17} \text{ cm}^{-3}$) (see Fig. 13) and $w \geq 100 \text{ \AA}$.
- iii) the barrier compensation ($N_{b,A}$) plays some part, especially at large w where the residual barrier acceptors capture a significant number of electrons released by the donors (see Fig. 14). Notice that the slope of the curve n_e versus w becomes steeper as $N_{b,A}$ increases.

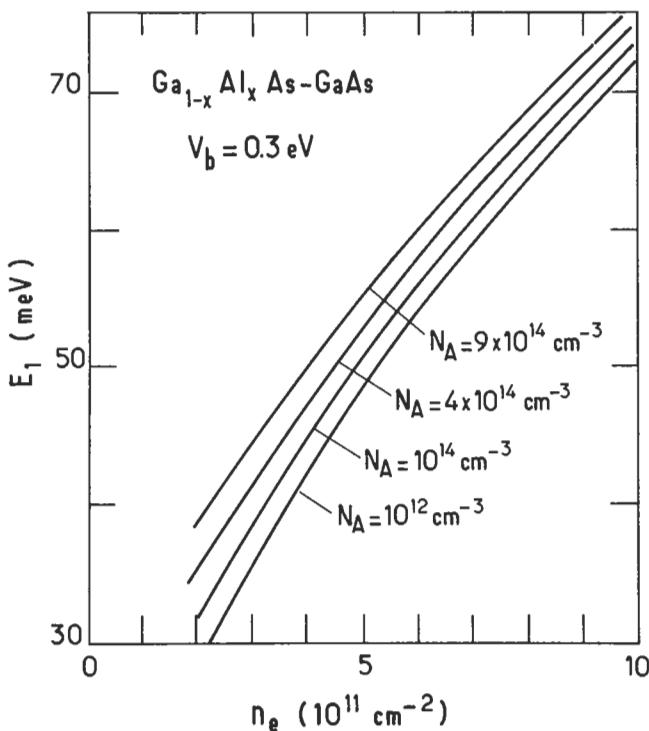


Fig. 11. — The confinement energy E_1 of the ground subband is plotted *versus* the two dimensional electron concentration n_e in $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterojunctions for several values of the volume concentration of ionized acceptors N_A in the channel, $N_A = 10^{12} \text{ cm}^{-3}$ corresponds to the quasi-accumulation conditions. $V_b = 0.3 \text{ eV}$; $m_A = 0.07 m_0$.

iv) the energy difference $V_b - \varepsilon_d$ significantly influences the charge transfer (Fig. 15). It may be noticed that by decreasing $V_b - \varepsilon_d$ all the n_e *versus* w curves retain roughly the same slope. In addition, the results are slightly different depending on whether one varies ε_d while keeping V_b fixed or *vice versa*. This is due to the implicit dependence of E_1 upon V_b .

The charge transfer in modulation-doped n-type $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterojunctions has been widely investigated in different laboratories. It has been known for a long time that fewer electrons are transferred than is theoretically predicted, but no generally accepted explanation of this phenomenon has been provided.

The compensation of $\text{Ga}_{1-x}\text{Al}_x\text{As}$ cannot be the sole cause of the discrepancy between the measurements and the calculations. Unreasonably large N_{bA} have to be postulated to account for some measured transferred charges. If these large N_{bA} were actual, the channel mobility could barely exceed $\sim 10^5 \text{ cm}^2/\text{Vs}$, whereas mobilities as large as $\sim 5 \times 10^6 \text{ cm}^2/\text{Vs}$ have been reported. In addition, the slope of

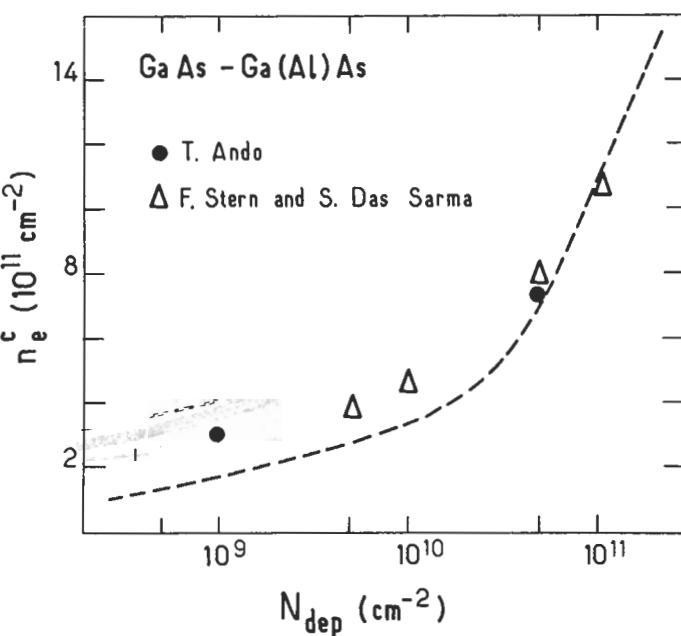


Fig. 12. — The critical concentration n_e^c beyond which the second electron subband E_2 becomes populated in $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions ($V_b = 0.3$ eV), is plotted *versus* N_{dep} , the areal concentration of depletion charges in the channel. The two circles are Ando's results [7] and the triangles are some of the values obtained by Stern and Das Sarma [4].

the measured transferred charge does not agree with that which was theoretically calculated, assuming N_{bA} to be large.

We have already stated that the barrier height V_b between GaAs and $\text{Ga}_{1-x}\text{Al}_x\text{As}$ is still a matter of controversy. Depending on which value of Q_c (85 % or 60 %) is retained, V_b is changed by nearly 0.1 eV if $x = 0.3$. This sensitively affects the calculated charge transfer (see Fig. 15).

In addition, the Si donor binding energy ε_d in $\text{Ga}_{1-x}\text{Al}_x\text{As}$ has been shown to vary with x in a strange fashion [15]. If one interprets the temperature dependences of the electrical conductivity and the Hall effect in bulk $\text{Ga}_{1-x}\text{Al}_x\text{As}$ layers (grown by M.B.E.) in terms of the activation of electrons from a single type of donor, one finds a nearly hydrogenic donor up to $x \simeq 0.2$ (i.e. $\varepsilon_d \sim 8$ meV). For $x > 0.2$, ε_d increases markedly reaching ~ 140 meV for $x \sim 0.4$ and decreases again at larger x (see Fig. 16).

However, the apparent increase of ε_d could be the result of the incorrect assumption that a single type of donor controls the electron population in bulk $\text{Ga}_{1-x}\text{Al}_x\text{As}$. Schubert and Ploog [16] have in fact shown that the $\text{Ga}_{1-x}\text{Al}_x\text{As}$ electrical measurements can be fitted if one assumes that there exists two kinds of donor species which are associated with the Si dopant. Their binding energies are fairly constant with x . One is shallow ($\varepsilon_{d1} \sim 10$ meV) and the other is deep

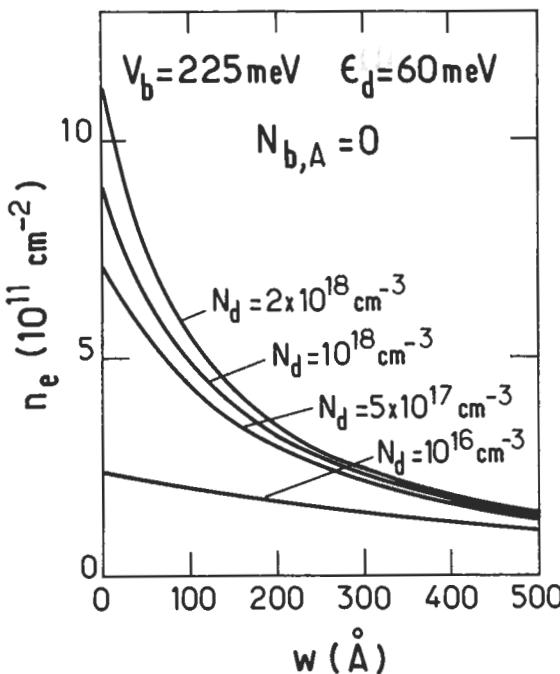


Fig. 13. — The areal concentration n_e of transferred charges is plotted *versus* the spacer thickness w in $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As-GaAs}$ heterojunction for several N_d values of the volume concentration of donors in the barrier. The following material parameters have been used in the calculations : $V_b = 225$ meV, $N_{\text{dep}} = 4.6 \times 10^{10} \text{ cm}^{-2}$, $N_{bA} = 0$, $\epsilon_d = 60$ meV.

($\epsilon_{d_2} \sim 120$ meV). What varies with x is their relative amounts. For $x \leq 0.2$ the shallow donors are more numerous, whereas for $x \geq 0.2$ the deep donors progressively become the dominant species (Fig. 17).

The fact that deep traps exist in $\text{Ga}_{1-x}\text{Al}_x\text{As}$ is certain. These traps may involve a native defect or a complex between a native defect and a foreign impurity. Their activation energy depends on their charge state due to a strong interaction with the lattice. They are probably responsible for the phenomenon called "persistent photoconductivity" [17-19]. It consists of the irreversible increase in the carrier concentration in bulk $\text{Ga}_{1-x}\text{Al}_x\text{As}$ or in the transferred charge in $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterojunctions at low temperature after light exposure. The initial situation can be restored by heating up the sample. Thus, this parasitic effect can be used to vary the carrier concentration in a given sample. However, at the same time it should be borne in mind that the number of charged impurities in the barrier varies in a rather unknown way. The persistent photoconductivity effect is much less pronounced in p-type materials.

Despite all the uncertainties on the material parameters, one can attempt to interpret the charge transfer in $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ n-type heterojunctions. In figure

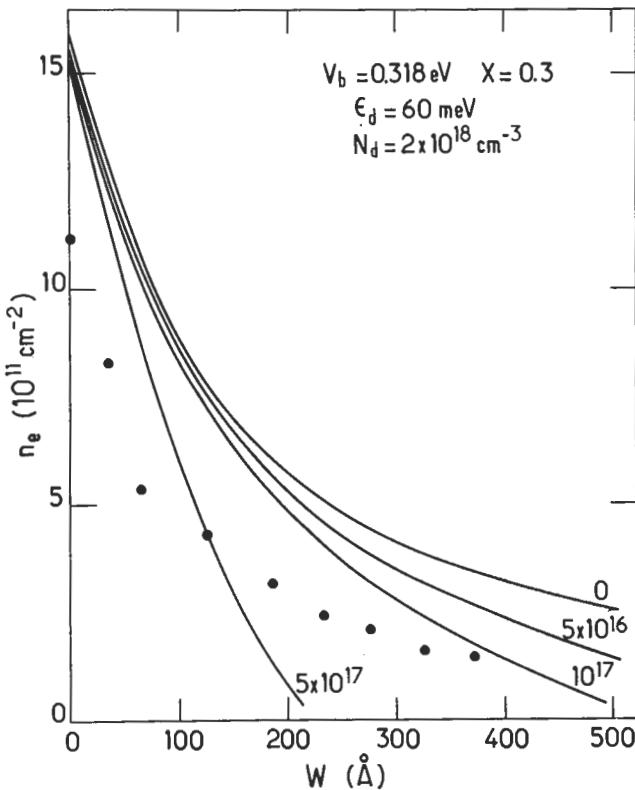


Fig. 14. — Influence of the volume concentration N_{bA} of compensating acceptors in the barriers on the spacer thickness dependence of the transferred charge n_e in $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions. $N_d = 2 \times 10^{18} \text{ cm}^{-3}$, $N_{\text{dep}} = 0.46 \times 10^{11} \text{ cm}^{-2}$, $\epsilon_d = 60 \text{ meV}$, $V_b = 0.318 \text{ eV}$. The symbols are Hwang *et al.*'s experimental results [20].

18, we show a comparison between the measured [20] and calculated n_e versus w for $x = 0.3$. The parameters which have been used in the calculations are $\epsilon_d = 60 \text{ meV}$, $Q_c = 60 \%$ (i.e. $V_b = 225 \text{ meV}$), $N_{\text{dep}} = 0.46 \times 10^{11} \text{ cm}^{-2}$ and $N_{b,A} = 0$. The $E_1(n_e)$ dependence has been calculated using the self-consistent variational procedure outlined in section II.1. It is seen that there exists a good agreement between the calculated and measured transferred charge. Also, one notices in figure 18 that using $Q_c = 85 \%$ and keeping the same ϵ_d sensitively degrades the agreement. However, no definitive conclusions can be drawn from figure 18 because, for example, the surface band bending has not been included in the calculation and, more importantly, because the measured transferred charges on nominally identical heterostructures scatter significantly.

Recently, studies of the hole transfer (Fig. 19) in modulation-doped p-type $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions have been reported [14]. They show

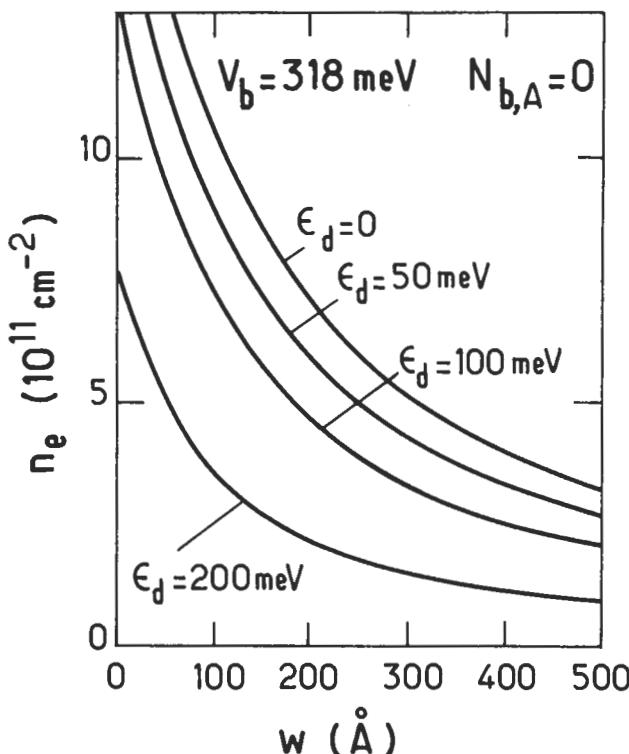


Fig. 15. — Influence of the energy difference $V_b - \epsilon_d$ on the spacer thickness dependence of the transferred charge n_e in $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions. $N_d = 2 \times 10^{18} \text{ cm}^{-3}$, $N_{\text{dep}} = 0.46 \times 10^{11} \text{ cm}^{-2}$, $N_{bA} = 0$. In these curves V_b is kept fixed (0.318 eV) while ϵ_d is varied.

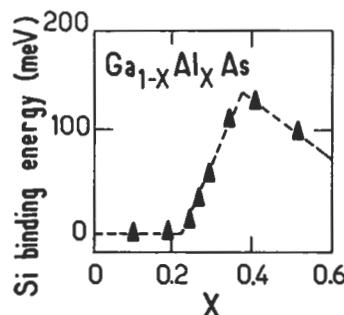


Fig. 16. — The Si donor binding energy ϵ_d in bulk $\text{Ga}_{1-x}\text{Al}_x\text{As}$ alloys is plotted against the Al mole fraction x . After reference [15].

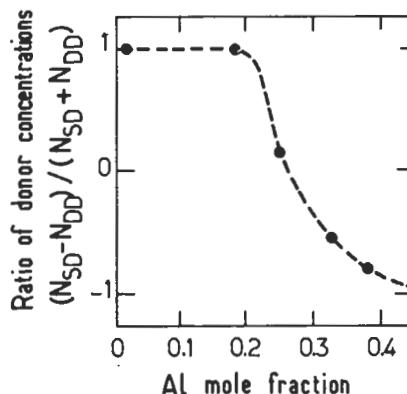


Fig. 17. — The relative amount of shallow and deep Si donor levels in bulk $\text{Ga}_{1-x}\text{Al}_x\text{As}$ alloys is plotted against the Al mole fraction x . At small x the Si concentration N_{Si} is equal to the shallow donor concentration N_{SD} . At large x the deep donor concentration N_{DD} is approximately equal to N_{Si} . After reference [16].

convincingly that the valence offset between $\text{Ga}_{1-x}\text{Al}_x\text{As}$ and GaAs ought to be larger than the one deduced from Dingle's rule, the measured charge transfer being much larger than theoretically predicted using $Q_c = 85\%$. For $x = 0.5$, Wang *et al.* [14] deduced from their experiments that Q_c should be closer to 60 % rather than 85 % and thus $Q_v = \frac{\Delta E_v}{\Delta E_g}$ should be 40 % instead of 15 % as was previously thought.

The hole subbands in p-type heterojunctions are harder to calculate than the electron ones. Firstly, the valence band is fourfold degenerate at the Γ point in the bulk materials. This leads to a complex mixing of the four kinds of bulk solutions due to the heterojunction electric field. Secondly, the Hartree approximation is probably not sufficient enough to account for the Coulombic interaction between the carriers, and the other many-body terms (exchange and correlations) may be quite large, as found in other heterostructures involving carriers with large effective masses (Si-SiO₂). The only calculations known to us of hole subbands in $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterojunctions have either been performed in the Hartree approximation, with account for the valence band complexities [21-23], or have had exchange and correlation effects taken into account but the intricacies of the hole kinematics [14] neglected.

Despite the difficulties of evaluating the hole confinement energies *versus* the carrier concentration, we may infer that they should be much smaller than the corresponding electron ones due to the larger hole masses. This, in turn, implies that in the hole transfer equation the dominant term is $V_b - \epsilon_A$. In addition, since the binding energy of the Be acceptor level smoothly varies with x in $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$, it is conceivable that the hole transfer measurements might give more reliable information on the valence band offsets than the electron transfer measurements provide on the conduction band offset.

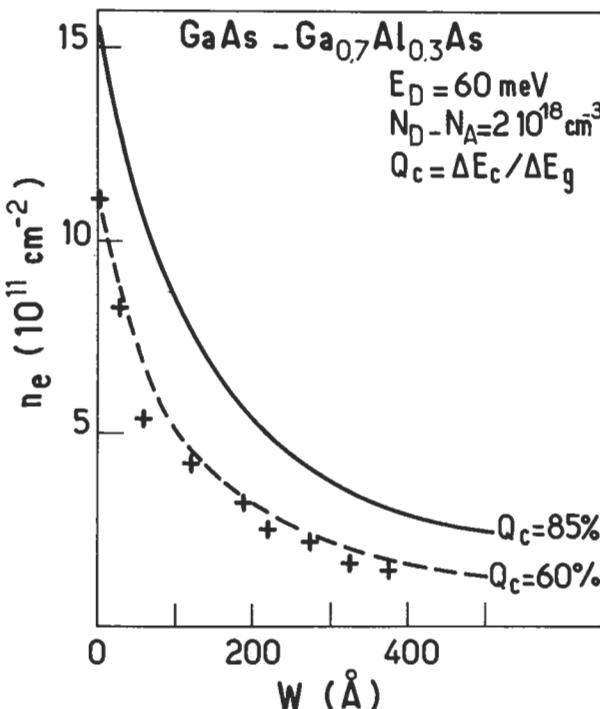


Fig. 18. — Comparison between the calculated and measured transferred charge in $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions. $x = 0.3$, $N_d = 2 \times 10^{18} \text{ cm}^{-3}$, $N_{\text{dep}} = 0.46 \times 10^{11} \text{ cm}^{-2}$, $E_d = 60 \text{ meV}$, $N_{b,A} = 0$. Two values of the parameter Q_c have been used: $Q_c = 60\% : V_b = 225 \text{ meV}$ (dashed line) and $Q_c = 85\% : V_b = 318 \text{ meV}$ (solid line). The symbols correspond to Hwang *et al.*'s data [20].

III. Energy levels in modulation-doped quantum wells.

The modulation-doped quantum wells have received much less attention than the single heterojunctions, even though one should conceivably be able to transfer more carriers in double heterojunctions than in single ones and consequently obtain larger conductivities, which is one of the desirable features of high performance TEGFET's [24].

The reason for this lack of interest may be due to the mediocre mobilities which have been achieved in modulation-doped quantum wells (the low temperature mobilities have seldom exceeded $10^5 \text{ cm}^2/\text{Vs}$ in these structures, which is more than one order of magnitude lower than in single heterojunctions). These relatively low mobilities are probably related to the growth of one inverted $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ interface (i.e. GaAs on top of $\text{Ga}_{1-x}\text{Al}_x\text{As}$) which has long been known to be difficult. The precise reasons for the bad quality of the inverted interface have not yet been fully elucidated. It may be due to an impurity segregation near the interface

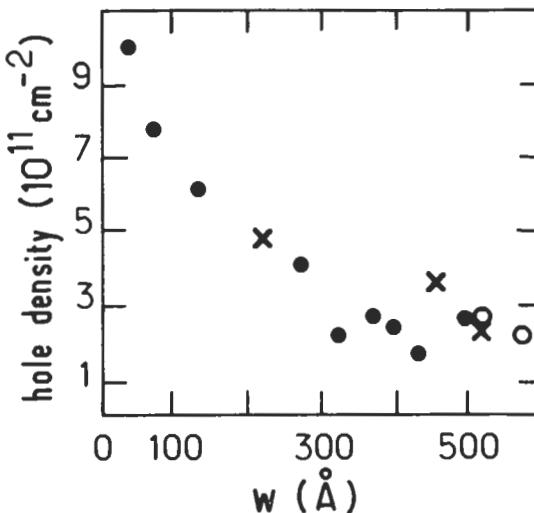


Fig. 19. — The areal concentration of transferred holes in modulation-doped p-type $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions is plotted *versus* the spacer thickness w . The solid dots correspond to $x = 0.5$ and the other symbols to slightly larger Al mole fractions. After reference [14].

once the Al flux is terminated, or to the difficulty of growing thick $\text{Ga}_{1-x}\text{Al}_x\text{As}$ layers with a smooth surface, or to the misfit strains, which arise due to the slight lattice mismatch between GaAs and $\text{Ga}_{1-x}\text{Al}_x\text{As}$. Notice that the problem of the non symmetrical growth of two materials on top of each other (i.e. A on B *versus* B on A) is probably not restricted to the $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ pair but may also exist in other systems (e.g. InAs-GaSb, InP-Ga_{0.47}In_{0.53}As etc...). Recently Drummond *et al.* [25] succeeded in greatly improving the quality of the inverted $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunctions by growing a thin superlattice prior to the heterojunction itself. With this method, mobilities of $\sim 2 \times 10^5 \text{ cm}^2/\text{Vs}$ were obtained in inverted heterojunctions at low temperature. It is not yet clear however if actual modulation-doped quantum wells are characterized by equivalent interfaces.

The energy level models [24, 26-29] in modulation-doped quantum wells have, to our knowledge, invariably assumed a perfect symmetry of the heterostructure with respect to the centre of the quantum well. In the following, we shall retain this approximation, because it greatly simplifies the algebra. It has to be kept in mind however, that great care has to be exercised when comparing the energy levels of such idealized systems with the experiments.

The salient feature of the energy levels in modulation-doped quantum wells is that the size-quantization (i.e. the formation of bound states for the z motion) is less critically dependent upon the band bending than in single heterojunctions. In quantum wells, bound states exist even in the absence of carriers and fixed charges

whereas in single heterojunctions, bound states are formed only due to the nonvanishing band bending.

This leads to the opportunity of using the perturbative approach [28] to account for the band bending contributions to $E_1, E_2 \dots$ states. This approach is useless in single heterojunctions.

The energy levels in modulation-doped quantum wells should extrapolate to the single heterojunction bound states when the quantum well thickness L becomes much larger than the characteristic extension of the carrier wavefunction in a single heterojunction.

Self consistent calculations of energy levels in modulation-doped quantum wells have been performed either numerically [24, 26, 29] or using a variational method [27] or a perturbative approach [28]. Here, we shall outline the perturbative approach, originally derived by Fishman [28], since the variational calculations are a straightforward (but tedious) generalization of those performed in single heterojunctions.

We want to solve the Schrödinger equation

$$\left[p_z \frac{1}{m(z)} p_z + V_b Y \left[z^2 - \frac{L^2}{4} \right] - e\Phi_{s.c.}(z) \right] \chi_i(z) = [H_0 - e\Phi_{s.c.}(z)] \chi_i(z) = E_i \chi_i(z) \quad (36)$$

in conjunction with the Poisson equation

$$\Phi''_{s.c.}(z) = \frac{4\pi e}{\kappa} \left\{ \sum_{i \text{ occupied}} n_i \chi_i^2(z) - N_d^+(z) + N_A^-(z) \right\} \quad (37)$$

where all the notations have the same definition as in paragraph II.1. The boundary conditions for the Poisson equation are such that the electric field vanishes both at the centre of the quantum well ($z = 0$) and at $z = \pm \left(\frac{L}{2} + l_d + w \right)$ where

l_d is the donor depletion length and w the spacer thickness (Fig. 20). It may be remarked that, normally, the residual doping of the channel barely influences the energy levels of modulation-doped quantum wells: their areal concentration $N_{\text{dep}} \leq 10^9 \text{ cm}^{-2}$ being much smaller than in single heterojunctions.

Even in the absence of band bending, equation (36) already admits bound states. Let us denote the eigenstate of H_0 by $\chi_i^{(0)}(z)$ and their associated eigenvalues by $E_i^{(0)}$. Roughly speaking, the level spacing varies like L^{-2} whereas the band bending contributions $\langle \chi_i^{(0)} | -e\Phi_{s.c.} | \chi_j^{(0)} \rangle$ are linear in n_c and in L . Thus, if L is small enough, it should be possible to treat $-e\Phi_{s.c.}(z)$ as a small perturbation of H_0 .

We focus our attention on the lowest lying states E_i . The self consistent potential will admix the unperturbed bound state $E_i^{(0)}$ not only with the bound states $E_j^{(0)}$, but also with the quantum well continuum states ($E_j^{(0)} > V_b$). This latter admixture will be small because the energy denominators in the perturbation expansion will be large and the matrix element will be small (since the excited states

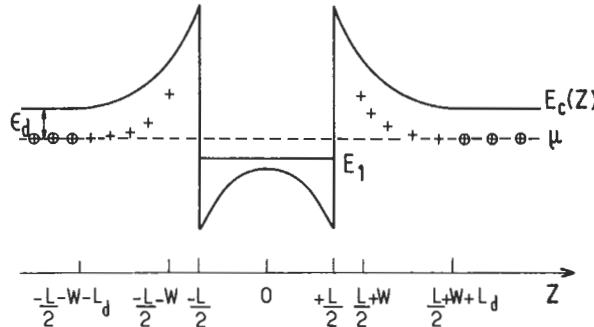


Fig. 20. — Conduction band profile in a symmetrically modulation-doped single quantum well.

are delocalized). Thus, we shall only retain the couplings between the bound states of the unperturbed quantum well.

Since the heterostructure is assumed to be perfectly symmetrical, $-e\Phi_{\text{s.c.}}(z)$ should be an even function of z . The unperturbed solutions $\chi_i^{(0)}(z)$ also have a definite parity. Thus, $-e\Phi_{\text{s.c.}}(z)$ can only couple unperturbed levels of the same parity, i.e. $\langle \chi_i^{(0)} | -e\Phi_{\text{s.c.}} | \chi_j^{(0)} \rangle$ is non vanishing only if $i + j$ is even.

To be more specific, let us take an example and assume that the unperturbed quantum well admits four bound states. This means that

$$3\pi < \sqrt{\frac{2m_A V_b L^2}{\hbar^2}} < 4\pi. \quad (38)$$

We project $H_0 - e\Phi_{\text{s.c.}}(z)$ on the four unperturbed states $\chi_i^{(0)}(z)$. Then, by exploiting the symmetry properties of $\Phi_{\text{s.c.}}(z)$ and $\chi_i^{(0)}(z)$, we end up with a 4×4 matrix Hamiltonian which is block-diagonal :

$$\begin{array}{c|cccc}
 & \chi_1^{(0)} & \chi_3^{(0)} & \chi_2^{(0)} & \chi_4^{(0)} \\ \hline
 \chi_1^{(0)} & E_1^{(0)} + \delta E_1 & V_{13} & 0 & 0 \\
 \chi_3^{(0)} & V_{31} & E_3^{(0)} + \delta E_3 & 0 & 0 \\
 \chi_2^{(0)} & 0 & 0 & E_2^{(0)} + \delta E_2 & V_{24} \\
 \chi_4^{(0)} & 0 & 0 & V_{42} & E_4^{(0)} + \delta E_4
 \end{array} \quad (39)$$

where :

$$\delta E_i = \langle \chi_i^{(0)} | -e\Phi_{\text{s.c.}} | \chi_i^{(0)} \rangle \quad (40)$$

$$V_{ij} = V_{ji} = \langle \chi_i^{(0)} | -e\Phi_{\text{s.c.}} | \chi_j^{(0)} \rangle \quad (41)$$

The eigenenergies in the presence of band bending can thus be written :

$$E_{\text{even}}^{\pm} = \frac{E_1^{(0)} + E_3^{(0)} + \delta E_1 + \delta E_3}{2} \pm \sqrt{\left[\frac{E_3^{(0)} + \delta E_3 - E_1^{(0)} - \delta E_1}{2} \right]^2 + V_{13}^2} \quad (42)$$

$$E_{\text{odd}}^{\pm} = \frac{E_2^{(0)} + \delta E_2 + E_4^{(0)} + \delta E_4}{2} \pm \sqrt{\left[\frac{E_4^{(0)} + \delta E_4 - E_2^{(0)} - \delta E_2}{2} \right]^2 + V_{24}^2} \quad (43)$$

for states which are even and odd in z respectively. In equations (42, 43), the $+$ ($-$) sign corresponds to the upper (lower) level of each pair of coupled states. The normalized wavefunctions which correspond to the eigenenergies (42, 43) are :

$$\chi_{\text{even}}^{\pm}(z) = \sqrt{\frac{(E_1^{(0)} + \delta E_1 - E_{\text{even}}^{\pm})^2}{(E_1^{(0)} + \delta E_1 - E_{\text{even}}^{\pm})^2 + V_{13}^2}} \times \left\{ \frac{V_{13}}{E_1^0 + \delta E_1 - E_{\text{even}}^{\pm}} \chi_1^{(0)}(z) - \chi_3^{(0)}(z) \right\} \quad (44)$$

$$\chi_{\text{odd}}^{\pm}(z) = \sqrt{\frac{(E_2^{(0)} + \delta E_2 - E_{\text{odd}}^{\pm})^2}{(E_2^0 + \delta E_2 - E_{\text{odd}}^{\pm})^2 + V_{24}^2}} \times \left\{ \frac{V_{24}}{E_2^0 + \delta E_2 - E_{\text{odd}}^{\pm}} \chi_2^{(0)}(z) - \chi_4^{(0)}(z) \right\} \quad (45)$$

The self consistency now requires that in the Poisson equation the charge distribution of each subband is given by the squares of equations (44, 45). Since the admixture coefficients in equations (44, 45) depends on the perturbed energies E^{\pm} , an iterative process has to be designed. At the input we use the unperturbed solutions χ_i^0 to calculate the band bending. Then, we compute the δE_i 's and V_{ij} 's and obtain the modified energy E^{\pm} and modified χ_i 's which are injected in the Poisson equation etc...until the convergence is reached.

The main advantages of the self-consistent perturbative approach are that it is physically transparent and analytically tractable. This method also provides explicit proof that for symmetrical modulation doped quantum wells of moderate thickness (i.e. binding few levels), the band bending effects are quite small since they arise from the coupling between states which are not adjacent but next nearest neighbours. These qualitative considerations are more easily appreciated in an example. Let us evaluate the excited level admixture into the ground one which occurs due to band bending [28]. This is summarized in table III for a GaAs-Ga_{1-x}Al_xAs quantum well with $L = 204$ Å, $V_b = 190$ meV, $N_{\text{dep}} = 0$, $m_A = m_B = 0.068 m_0$ and $n_e = 4.2 \times 10^{11} \text{ cm}^{-2}$.

When the quantum well thickness decreases, the matrix elements of $-e\Phi_{\text{s.c.}}$

Table III. — *Admixture coefficients C_i ($1 \leq i \leq 4$) in a modulation-doped quantum well. $L = 204 \text{ \AA}$, $V_b = 190 \text{ meV}$; $n_e = 4.1 \times 10^{11} \text{ cm}^{-2}$. The wavefunctions are written as: $\chi_{\text{even}}(z) = c_1 \chi_1^{(0)}(z) + c_3 \chi_3^{(0)}(z)$, $\chi_{\text{odd}}(z) = c_2 \chi_2^{(0)}(z) + c_4 \chi_4^{(0)}(z)$, respectively. After reference [28].*

$n_e (10^{11} \text{ cm}^{-2})$	0	4.2
$C_1 (\%)$	100	99.8
$C_3 (\%)$	0	-7
$C_2 (\%)$	100	99.9
$C_4 (\%)$	0	-5.4

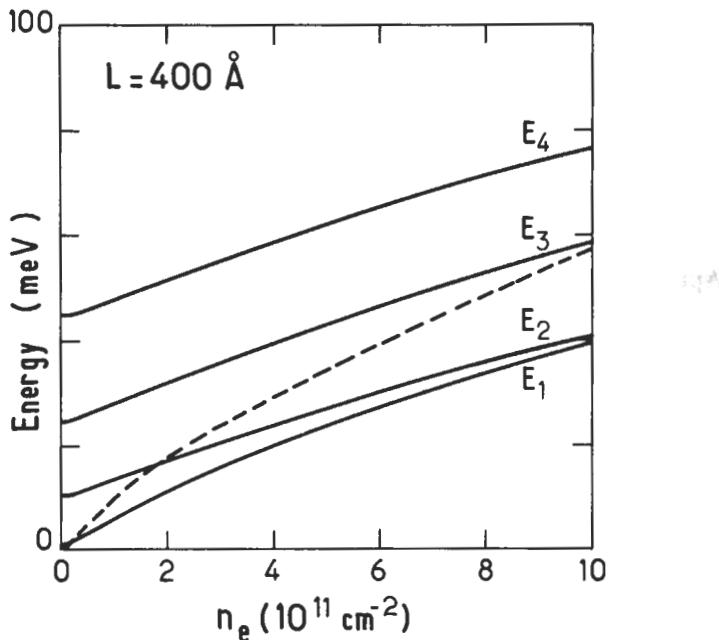


Fig. 21. — Numerical self-consistent calculations of energy levels in a symmetrical modulation-doped single $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ quantum well ($L = 400 \text{ \AA}$). The exchange and correlation effects are responsible for the small dip at low n_e . The dashed curve corresponds to the Fermi energy. After reference [29].

decreases almost linearly with L , whereas the unperturbed level spacing increases. The perturbative approach thus becomes more and more justified. Notice, however, that the number of unperturbed bound states decreases as L decreases. For instance,

when there are only three bound levels $E_1^{(0)}, E_2^{(0)}, E_3^{(0)}$, the second level $E_2^{(0)}$ is not coupled to any other by $-e\Phi_{s.c.}(z)$ and consequently

$$E_2 = E_2^{(0)} + \langle \chi_2^{(0)} | -e\Phi_{s.c.} | \chi_2^{(0)} \rangle \quad (46)$$

When a further decrease in L occurs, only two levels are bound by the unperturbed quantum well. These levels are shifted by terms which are linear in $\Phi_{s.c.}(z)$

$$E_1 = E_1^{(0)} + \langle \chi_1^{(0)} | -e\Phi_{s.c.} | \chi_1^{(0)} \rangle \quad (47)$$

$$E_2 = E_2^{(0)} + \langle \chi_2^{(0)} | -e\Phi_{s.c.} | \chi_2^{(0)} \rangle \quad (48)$$

and the associated wavefunctions are identical to $\chi_1^{(0)}$ and $\chi_2^{(0)}$ respectively.

When the quantum well becomes thick enough to bind many levels, the self-consistent perturbative approach loses its simplicity : the ground state wavefunction resembles more and more the symmetric combination of two single heterojunction χ_1 wavefunctions and thus clearly requires a large number of unperturbed quantum well eigenfunctions $\chi_i^{(0)}(z)$ to be correctly approximated. One should then resort to self consistent variational treatments or numerical procedures to find the energy levels and wavefunctions of thick modulation-doped quantum well. An example of such a numerical calculation [29] is shown in figure 21 for a GaAs-Ga_{0.6}Al_{0.4}As modulation-doped quantum well with $L = 400$ Å.

Appendix A.

The algebra of the modified Fang Howard wavefunction.

In this appendix, we give some of the algebraical formula related to the self-consistent calculations of the confinement energy E_1 in a heterostructure. The trial wavefunction $\chi_1(z)$ is assumed to have the form

$$\chi_1(z) = \begin{cases} M \exp(\kappa_b z/2) & z \leq 0 \\ N(z + z_0) \exp(-bz/2) & z \geq 0 \end{cases} \quad (A1)$$

By writing that $\chi_1(z)$ and $\frac{1}{m(z)} \frac{d\chi_1}{dz}$ are continuous at the interface $z = 0$ and by accounting for the normalization condition we obtain :

$$z_0 = \frac{2}{b + \kappa_b \frac{m_A}{m_B}} \quad (A2)$$

$$N = \sqrt{\frac{b^3}{2}} \times \frac{1}{\left[1 + bz_0 + \frac{1}{2} b^2 z_0^2 \left(1 + \frac{b}{\kappa_b} \right) \right]^{1/2}} \quad (A3)$$

$$M = Nz_0 \quad (A4)$$

The expressions (A2-A4) completely determine M, N, z_0 in terms of b and κ_b . The self-consistent electrostatic potential $\Phi_{s.c.}(z)$ is the sum of contributions arising from the electrons and the ionized acceptors in the channel (if we assume a zero compensation of the barrier and neglect the ionized donors term).

$$\Phi_{s.c.}(z) = \delta \Phi_A(z) + \delta \Phi_{el-el.} \quad (A5)$$

with :

$$\delta \Phi_A(z \geq 0) = - \frac{4\pi e}{\kappa} N_{dep} z \quad (A6)$$

$$\delta \Phi_{el-el.}(z \geq 0) = \frac{4\pi e}{\kappa} n_e N^2 [(\alpha z^2 + \beta z + \gamma) e^{-bz} - \gamma] \quad (A7)$$

where $\Phi_{s.c.}(0)$ is taken as equal to zero and

$$\alpha = \frac{1}{b^2}; \quad \beta = \frac{2z_0}{b^2} + \frac{4}{b^3}; \quad \gamma = \frac{z_0^2}{b^2} + \frac{4z_0}{b^3} + \frac{6}{b^4} \quad (A8)$$

The self-consistent potential in the spacer part of the barrier is conveniently written as

$$\delta \Phi_{s.c.}(z) = - F_{s.c.} z \quad (A9)$$

with

$$F_{s.c.} = - \left[\frac{d}{dz} \Phi_{s.c.}(z \geq 0) \right]_{z=0} = - \frac{4\pi e}{\kappa} [N^2 n_e (\beta - b\gamma) - N_{dep}] \quad (A10)$$

The average value of the kinetic energy T is equal to

$$T = \left\langle -\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{m(z)} \frac{d}{dz} \right\rangle = -\frac{\hbar^2}{8m_B} M^2 \kappa_b + \frac{\hbar^2 N^2}{2m_A} \left\{ \frac{1}{2b} + \frac{1}{2} z_0 - \frac{1}{4} b z_0^2 \right\} \quad (A11)$$

whereas

$$\langle -e \delta \Phi_A(z) \rangle = \frac{4\pi e^2}{\kappa} N_{dep} N^2 \left\{ \frac{6}{b^4} \left[1 + \frac{2}{3} b z_0 + \frac{1}{6} b^2 z_0^2 \right] \right\} \quad (A12)$$

$$\langle -e \delta \Phi_{el-el.} Y(z) \rangle = \frac{4\pi e^2}{\kappa} n_e \frac{N^4}{b^7} \times \left\{ \frac{33}{4} + \frac{25}{2} b z_0 + \frac{17}{2} b^2 z_0^2 + 3b^3 z_0^3 + \frac{1}{2} b^4 z_0^4 \right\} \quad (A13)$$

$$\langle -e \Phi_{s.c.} Y(-z) \rangle = -e F_{s.c.} \frac{N^2 z_0^2}{\kappa_b^2} \quad (A14)$$

Finally the integrated probability of finding the particle in the barrier is equal to

$$P_b = \langle Y(-z) \rangle = \frac{N^2 z_0^2}{\kappa_b} \quad (A15)$$

and the mean distance separating the electron gas from the interface is equal to

$$\langle z \rangle = \frac{6N^2}{b^4} \left[1 + \frac{2b}{3} z_0 + \frac{1}{6} b^2 z_0^2 - \frac{z_0^2}{6\kappa_b^2} b^4 \right] \quad (\text{A16})$$

The energy $\tilde{E}_1(b)$ which has to be minimized is

$$\tilde{E}_1(b) = T + \frac{1}{2} \langle -e\delta\Phi_{\text{el.el}}(z) \rangle + \langle -e\delta\Phi_A(z) \rangle$$

and the single particle energy is :

$$E_1(b) = [T + \langle -e\Phi_{\text{s.c.}}(z) \rangle]_{b=b_{\min}}$$

Appendix B.

Heterojunction energy levels : qualitative aspects.

In a qualitative approach to the Schrödinger equation which describes the z motion of an electron in a self consistent potential $-e\Phi_{\text{s.c.}}(z) + V_b Y(-z)$, we can expand $-e\Phi_{\text{s.c.}}(z)$ in the vicinity of $z=0$ to obtain

$$\left[\frac{p_z^2}{2m^*} + V_b Y(-z) + eFz \right] \chi_1(z) = E_1 \chi_1(z) \quad (\text{B1})$$

Let us once again use the Heisenberg inequality to evaluate E_1 . Let α be the characteristic spatial extension of the χ_1 wavefunction (assumed to vanishingly penetrate the barrier). We then know that the kinetic energy will be of the order of $\frac{\hbar^2}{2m^*\alpha^2}$ and the potential energy of the order $eF\alpha$. Thus :

$$E_1(\alpha) \sim \frac{\hbar^2}{2m^*\alpha^2} + eF\alpha \quad (\text{B2})$$

Considered as a function of α , $E_1(\alpha)$ admits a minimum for :

$$\alpha_{\min} = \left[\frac{\hbar^2}{m^*eF} \right]^{1/3} \quad (\text{B3})$$

the minimum value of the confinement energy $E_1(\alpha_{\min})$ being equal to :

$$E_1(\alpha_{\min}) \sim \frac{3}{2} \left[\frac{\hbar^2}{m^*} \right]^{1/3} (eF)^{2/3} \quad (\text{B4})$$

Since we know from the Poisson law that F is related to the areal electron concentration n_e by

$$F = \frac{4\pi e}{\kappa} n_e \quad (\text{B5})$$

we see that the spatial extension varies like $n_e^{-1/3}$ and the confinement energy like $n_e^{2/3}$. Let us take GaAs as an example. We find that for conduction electrons $\alpha \sim 10^2 \text{ \AA}$ and $E_1 \sim 25 \text{ meV}$ if $n_e = 10^{11} \text{ cm}^{-2}$.

Although the scaling laws $\alpha \sim n_e^{-1/3}$, $E_1 \sim n_e^{2/3}$ are derived from a very crude model, they are useful for obtaining rough estimates of the electron confinement energies in potentials which do not depart substantially from a linear variation upon z .

In fact, the same scaling laws are also obtained if a Fang Howard trial wavefunction [10] is used to evaluate E_1 by a variational procedure. Let us search for a $\chi_1(z)$ wavefunction in the form

$$\chi_1(z) = \sqrt{\frac{b^3}{2}} z \exp(-bz/2), \quad z \geq 0 \quad (B6)$$

$$\chi_1(z) = 0, \quad z \leq 0 \quad (B7)$$

where b is the variational parameter. We then easily obtain

$$\tilde{E}_1(b) = \frac{\hbar^2 b^2}{8m^*} + \frac{1}{2} \frac{3eF}{b} \quad (B8)$$

$$E_1(b) = \frac{\hbar^2 b^2}{8m^*} + \frac{3eF}{b} \quad (B9)$$

where $\tilde{E}_1(b)$ is the function which has to be minimized and $E_1(b)$ is the single-particle confinement energy. $\tilde{E}_1(b)$ has a minimum for :

$$b_{\min} = \left[\frac{6m^* e F}{\hbar^2} \right]^{1/3} \quad (B10)$$

and thus the one-electron confinement energy is equal to :

$$E_1(b_{\min}) = \left[\frac{6^{2/3}}{8} + \frac{3}{6^{1/3}} \right] (eF)^{2/3} \left[\frac{\hbar^2}{m^*} \right]^{1/3} \quad (B11)$$

E_1 has the same dependence upon e , F , \hbar , m^* as the expression (B4). Only the numerical coefficient in front of $(eF)^{2/3} \left[\frac{\hbar^2}{m^*} \right]^{1/3}$ is ~ 2.06 instead of 1.5.

The scaling laws are in fact a direct consequence of the assumed linear dependence of the potential upon z and of the non-penetration of the χ_1 wavefunction in the barrier : the Schrödinger equation

$$\left[\frac{p_z^2}{2m^*} + eFz \right] \chi_n(z) = E_n \chi_n(z) \quad (B12)$$

with the boundary conditions

$$\chi_n(0) = \chi_n(\infty) = 0 \quad (B13)$$

can be rewritten in terms of dimensionless variables. By choosing

$$z = \left(\frac{\hbar^2}{m^* e F} \right)^{1/3} \xi ; \quad E_n = \varepsilon_n \left[\frac{\hbar^2 e^2 F^2}{m^*} \right]^{1/3} \quad (B14)$$

the equation (B12) is transformed into :

$$-\frac{1}{2} \frac{d^2 \chi_n(\xi)}{d\xi^2} + \xi \chi_n(\xi) = \varepsilon_n \chi_n(\xi) \quad (B15)$$

whose solutions are the Airy functions of argument $2^{1/3} (\varepsilon_n - \xi)$.

In practice $-e\Phi_{s.c.}(z)$, after a linear increase upon z (whose slope depends on n_c and the areal density of ionized acceptors in the channel N_{dep}), asymptotically approaches $\frac{4\pi e}{\kappa} N_{dep} z$ when z becomes large in comparison with the spatial extension of the χ_1 wavefunction. If one wishes to work out the confinement energies of the excited levels $E_2, E_3 \dots$, which are less spatially localized than E_1 , one should take an effective F , which decreases with the level index to approach $\frac{4\pi e}{\kappa} N_{dep}$ for very excited subbands. This procedure however, becomes very imprecise.

References

Two general references concerning the energy levels in heterostructures are :

- i) T. ANDO, A.B. FOWLER and F. STERN, *Rev. Mod. Phys.* **54** (1982) 437.
- ii) F. STERN, in the Proceedings of the Les Houches Winterschool "Heterojunctions and Semiconductor Superlattices" edited by G. Allan, G. Bastard, N. Boccara, M. Lannoo and M. Voos (Springer Verlag, Berlin) 1986.
- [1] R. DINGLE, H. L. STORMER, A.C. GOSSARD and W. WIEGMANN, *Appl. Phys. Lett.* **33** (1978) 665 ;
H.L. STORMER, *J. Phys. Soc. Japan Suppl. A* **49** (1980) 1013.
- [2] Y. TAKADA, K. ARAI, N. UCHIMURA and Y. UEMURA, *J. Phys. Soc. Japan* **49** (1980) 1851.
- [3] D.J. BEN DANIEL and C.B. DUKE, *Phys. Rev.* **152** (1966) 683.
- [4] F. STERN and S. DAS SARMA, *Phys. Rev.* **B30** (1984) 840.
- [5] P.J. PRICE and F. STERN, *Surf. Sci.* **132** (1983) 577.
- [6] B. VINTER, *Solid State Commun.* **48** (1983) 151.
- [7] T. ANDO, *J. Phys. Soc. Japan* **51** (1982) 3893.
- [8] G. BASTARD, *Surf. Sci.* **142** (1984) 284.
- [9] T. ANDO, *J. Phys. Soc. Japan* **51** (1982) 3900.
- [10] F.F. FANG and W.E. HOWARD, *Phys. Rev. Lett.* **16** (1966) 797.
- [11] F. STERN, *Phys. Rev.* **B5** (1972) 4891.
- [12] R. DINGLE, in *Festkörperprobleme XV (Adv. Solid State Phys.)* edited by H.J. Queisser (Pergamon Vieweg, Braunschweig) 1975, p 21
- [13] R.C. MILLER, A.C. GOSSARD, D.A. KLEINMAN and O. MUNTEANU, *Phys. Rev.* **B29** (1984) 3740.

- [14] W.I. WANG, E.E. MENDEZ and F. STERN, *Appl. Phys. Lett.* **45** (1984) 639.
 - [15] see e.g. T. ISHIKAWA, J. SAITO, S. SASA and S. HIYAMIZU, *Japan. J. Appl. Phys.* **21** (1982) L 675.
 - [16] E.F. SCHUBERT and K. PLOOG, *Phys. Rev.* **B30** (1984) 7021.
 - [17] H.L. STORMER, A.C. GOSSARD, W. WIEGMANN and K. BALDWIN, *Appl. Phys. Lett.* **39** (1981) 912.
 - [18] J. KLEM, W.T. MASSELINK, D. ARNOLD, R. FISHER, T.J. DRUMMOND, H. MORKOC, K. LEE and M.S. SHUR, *J. Appl. Phys.* **54** (1983) 5214.
 - [19] A. KASTALSKY and J.C.M. HWANG, *Solid State Commun.* **51** (1984) 317.
 - [20] J.C.M. HWANG, A. KASTALSKY, H.L. STORMER and V.G. KERAMIDAS, *Appl. Phys. Lett.* **44** (1984) 802.
 - [21] A. BRODO and L.J. SHAM, *Phys. Rev.* **B31** (1985) 888.
 - [22] U. EKENBERG and M. ALTARELLI, *Phys. Rev.* **B30** (1984) 3569.
 - [23] E. BANGERT and G. LANDWEHR, *Superlattices and Microstructures* **1** (1985) 363.
 - [24] K. INOUE, H. SAKAKI and J. YOSHINO, *Japan. J. Appl. Phys.* **23** (1984) L 767.
 - [25] T.J. DRUMMOND, J. KLEM, D. ARNOLD, R. FISCHER, R.E. THORNE, W.G. LYONS and H. MORKOC, *Appl. Phys. Lett.* **42** (1983) 615.
 - [26] T. ANDO and S. MORI, *J. Phys. Soc. Japan* **47** (1979) 1518.
 - [27] G. BASTARD, E.E. MENDEZ, L.L. CHANG and L. ESAKI, *J. Vac. Sci. Technol.* **21** (1982) 531.
 - [28] G. FISHMAN, *Phys. Rev.* **B27** (1983) 7611.
 - [29] F. STERN and J.N. SCHULMAN, *Superlattices and Microstructures* **1** (1985) 303.
-

CHAPTER VI

Electrical conductivity of quasi bi-dimensional electron gases

In the previous chapter we studied the energy levels of quasi bidimensional electron gases at thermal and electrical equilibrium. We found that the carrier motion along the growth axis is bound, whereas its in-plane motion is free.

Under equilibrium conditions there is no permanent electrical current but transient local electrical currents do exist (since the electron have a non-zero in-plane momentum). However, the unavoidable electron scatterings by the various defects and imperfections randomize the momentum directions on a time scale $\sim 10^{-12}$ s. As a result, no macroscopic permanent current flows in the layer plane. As for electrical transport along the z axis, it is impossible since the matrix element $\langle \chi_1 | p_z | \chi_1 \rangle$ vanishes for a bound state.

Suppose that a weak, constant and homogeneous electric field $\mathbf{F} = F \hat{x}$ has been applied along the \hat{x} axis. In addition to the scattering events, the carrier now experiences an acceleration along \mathbf{F} , resulting in a non-zero drift velocity. In the steady state situation, and for weak electric fields, the areal density of the electrical current $\mathbf{J} = J \hat{x}$ will be proportional to \mathbf{F} : the system has reacted to a weak disturbance by a flux which is proportional to the disturbance (ohmic regime). Since we will only consider systems which are isotropic in the layer plane, the proportionality between the electrical current and the external field will not depend on the field orientation and thus will be a c -constant :

$$\mathbf{J} = \sigma \mathbf{F} \quad (1)$$

where σ is the two-dimensional ohmic conductivity of the electron gas.

In this chapter we shall study how to calculate σ at low temperatures and how to relate σ to the material parameters. The electron scatterings by the imperfections will be assumed to be elastic. Such an assumption makes it possible to define a relaxation time for each of the quasi bi-dimensional subbands, which significantly simplifies the algebra. Coulombic scattering by impurities, interface roughness and alloy scatterings are elastic mechanisms. The phonon scattering, which arises from the existence of a non vanishing electron-phonon interaction and which is so efficient at room temperature, is inelastic. Consequently, to be fully analysed, it requires a numerical integration of the Boltzmann equation. The phonon scattering is discussed in references [1-4].

In multiple quantum well structures, or in superlattices, the levels of the isolated wells hybridize. There is thus a possibility of electrical conduction along the growth axis. In fact, in their pioneering work, Esaki and Tsu [5] advocated the possibility of

negative differential resistance occurring in superlattices. In comparison with the in-plane transport, very little effort has been devoted to the transport along the growth axis (also called vertical transport). We shall discuss some of the features of this longitudinal transport, which appears to be so crucially dependent on the quality of the structure.

I. Static conductivity of a quasi bi-dimensional electron gas.

We shall retain the same assumptions as in the previous chapter. Thus, we shall hereafter assume that the carriers can be considered as independent and that the (x, y) and z components of their motion can be decoupled. Apart from spin, a quantum state is labelled by n , the subband index (E_1, E_2, \dots) and $\mathbf{k}_\perp = (k_x, k_y)$ the two-dimensional wavevector which characterizes the in-plane motion. The energy and wavefunction associated with $|n\mathbf{k}_\perp\rangle$ are thus written as

$$\epsilon_n(\mathbf{k}_\perp) = E_n + \frac{\hbar^2 k_\perp^2}{2 m_n} \quad (2)$$

$$\langle \mathbf{r} | n\mathbf{k}_\perp \rangle = \chi_n(z) \frac{1}{\sqrt{S}} \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp) \quad (3)$$

where $\mathbf{r}_\perp = (x, y)$, S is the sample area, and m_n the effective mass for the in-plane motion of the n^{th} subband.

At thermal and electrical equilibrium, the distribution function of the states $|n\mathbf{k}_\perp\rangle$ (which are the diagonal elements of the one-electron density matrix) are the Fermi Dirac functions

$$f_n^{(0)}(\mathbf{k}_\perp) = \{1 + \exp[\beta(\epsilon_n(\mathbf{k}_\perp) - \mu)]\}^{-1} \quad (4)$$

$$\text{where :} \quad \beta = \frac{1}{k_B T} \quad (5)$$

and k_B is the Boltzmann constant, T the electron temperature and μ the chemical potential.

We are now interested in finding the linear electrical response of the quasi bi-dimensional electron gas to a weak, constant and homogeneous electric field \mathbf{F} applied parallel to the x axis.

$$\mathbf{F} = F \hat{\mathbf{x}} \quad (6)$$

The one-electron Hamiltonian, defect included is :

$$\mathcal{H} = \mathcal{H}_0 + eFx + \mathcal{H}_{\text{def}} \quad (7)$$

where \mathcal{H}_0 is the unperturbed Hamiltonian whose eigenstates are $|n\mathbf{k}_\perp\rangle$. \mathcal{H}_{def} includes the local fluctuations of the electrostatic potentials due to the defects (notice that the spatial averages of these potentials are included in \mathcal{H}_0 and give rise to band-bending).

\mathcal{H}_{def} is not translationally invariant in the layer plane. Thus, it broadens the $|n\mathbf{k}_\perp\rangle$ state and degrades the momentum of the carriers. \mathcal{H}_{def} has the property of being the sum of the contributions arising from randomly distributed scatterers located at \mathbf{R}_i

$$\mathcal{H}_{\text{def}} = \sum_{\mathbf{R}_i} \mathcal{H}_{\text{def}}(\mathbf{r} - \mathbf{R}_i) \quad (8)$$

where \mathbf{R}_i are random variables.

A long time ago Kohn and Luttinger [6] showed how to rigorously make a connection between a fully quantum treatment of the electrical transport (i.e. based on the solution of the Liouville equation for the density matrix) and the more familiar approach based on the solution of the Boltzmann equation [7]. We shall follow the latter approach here.

Due to the presence of the $e\mathbf{F}x$ and \mathcal{H}_{def} terms in equation (7), the distribution functions of the $|n\mathbf{k}_\perp\rangle$ states are no longer the $f_n^{(0)}$'s. Instead, each of the f_n 's is permanently decreased (increased) by transitions from (to) the state $|n\mathbf{k}_\perp\rangle$ which arise from the other states $|n'\mathbf{k}'_\perp\rangle$. As a result, in the permanent regime $\left(\frac{\partial f_n}{\partial t} = 0\right)$, the $f_n(\mathbf{k}_\perp)$ are the solutions of the coupled Boltzmann equations :

$$-\frac{e}{m_n} \hbar \mathbf{k}_\perp \cdot \mathbf{F} \frac{\partial f_n}{\partial \varepsilon_n(\mathbf{k}_\perp)} = \sum_{n', \mathbf{k}'_\perp} W_{n' \mathbf{k}'_\perp \rightarrow n \mathbf{k}_\perp} f_{n'}(\mathbf{k}'_\perp) [1 - f_n(\mathbf{k}_\perp)] - W_{n \mathbf{k}_\perp \rightarrow n' \mathbf{k}'_\perp} f_n(\mathbf{k}_\perp) [1 - f_{n'}(\mathbf{k}'_\perp)] \quad (9)$$

where $W_{\alpha \rightarrow \beta}$ is the transition probability per unit time that a transition induced by \mathcal{H}_{def} takes place from the state $|\alpha\rangle$ to the state $|\beta\rangle$. From microreversibility we know that [8] :

$$W_{\alpha \rightarrow \beta} = W_{\beta \rightarrow \alpha} \quad (10)$$

The rates $W_{\alpha \rightarrow \beta}$ will be evaluated in the Born approximation. Thus

$$W_{\alpha \rightarrow \beta} = \frac{2\pi}{\hbar} \delta(\varepsilon_\alpha - \varepsilon_\beta) |\langle \alpha | \mathcal{H}_{\text{def}} | \beta \rangle|^2 \quad (11)$$

When the scatterers are very efficient the Born approximation may not be sufficient and a more refined treatment, the self-consistent Born approximation [9], may be necessary. In the latter scheme the collision broadening is also included in the energy levels ε_α , resulting in the replacement of the delta function by a Lorentzian function whose width and level shift have to be determined self-consistently.

Retaining equation (11) for simplicity, the Boltzmann equation is rewritten in the form

$$-\frac{e}{m_n} \hbar \mathbf{k}_\perp \cdot \mathbf{F} \frac{\partial f_n}{\partial \varepsilon_n(\mathbf{k}_\perp)} = \frac{2\pi}{\hbar} \sum_{n', \mathbf{k}'_\perp} \langle |\langle n\mathbf{k}_\perp | \mathcal{H}_{\text{def}} | n'\mathbf{k}'_\perp \rangle|^2 \rangle_{\text{average}} \times [f_{n'}(\mathbf{k}'_\perp) - f_n(\mathbf{k}_\perp)] \times \delta[\varepsilon_n(\mathbf{k}'_\perp) - \varepsilon_n(\mathbf{k}_\perp)] \quad (12)$$

The bracket $\langle \dots \rangle_{\text{average}}$ means that an average over the \mathbf{R}_i 's has been performed :

$$\langle A(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \rangle_{\text{average}} = \int A(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) P(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \times \text{d}^3R_1 \text{d}^3R_2 \dots \text{d}^3R_N \quad (13)$$

where $P(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)$ is the probability density of finding a scatterer at \mathbf{R}_1 , another at \mathbf{R}_2 etc...

In the limit of diluted impurities the scattering sites become uncorrelated. Thus $P(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)$ factorizes :

$$P(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = \prod_i P(\mathbf{R}_i) \quad (14)$$

and $P(\mathbf{R}_i)$ reduces to Ω_μ^{-1} where Ω_μ is the macroscopic volume of the heterostructure which is occupied by the μ^{th} kind of scatterers, e.g. $\Omega_\mu = Sl_d$ in the case of the ionized donors in a single modulation-doped heterojunction. Assuming uncorrelated scattering sites, the average over the impurity sites of a double sum like

$$S_{\mu\mu} = \sum_{\mathbf{R}_i^\mu} \sum_{\mathbf{R}_j^\mu} \langle n\mathbf{k}_\perp | \mathcal{H}_{\text{def}}(\mathbf{R}_i^\mu) | n'\mathbf{k}'_\perp \rangle \langle n' \mathbf{k}'_\perp | \mathcal{H}_{\text{def}}(\mathbf{R}_j^\mu) | n\mathbf{k}_\perp \rangle \quad (15)$$

reduces to a sum of the diagonal elements in i and j :

$$S_{\mu\mu} = \frac{N_\mu}{\Omega_\mu} \int \text{d}^3R_\mu \left| \langle n\mathbf{k}_\perp | \mathcal{H}_{\text{def}}(\mathbf{r} - \mathbf{R}_\mu) | n'\mathbf{k}'_\perp \rangle \right|^2 \quad (16)$$

The averaged sums $S_{\mu\nu}$ of crossed terms involving scatterers of different species vanish. Equation (16) represents a considerable simplification since we can already anticipate that each of the defect species will contribute separately to the right-hand side of equation (12).

The assumption of diluted impurities is almost always used in the analysis of conductivity phenomena. For the specific case of modulation-doped heterojunctions it is probably an excellent approximation for the residual impurities, whose volume concentration is usually very low ($\sim 10^{14} \text{ cm}^{-3}$). Its validity, as far as the deliberately introduced donors are concerned, is opened to question since the donor concentration is usually quite large ($\sim 10^{18} \text{ cm}^{-3}$). The validity of the assumption of diluted impurities has not, to knowledge, been critically examined.

We look for a solution to equation (12) in the form

$$f_j(\mathbf{k}_\perp) = f_j^{(0)}(\mathbf{k}_\perp) + \frac{e}{m_j} \hbar \mathbf{k}_\perp \cdot \mathbf{F} \frac{\partial f_j^0}{\partial \epsilon_j(\mathbf{k}_\perp)} \tau_j[\epsilon_j(\mathbf{k}_\perp)] \quad (17)$$

and linearize equation (12) with respect to the field. This means that the τ_j 's as well as the conductivity σ , will be field independent. Siggia and Kwok [10] and

Ando and Mori [11] have shown that the ansatz (Eq. (17)) is a solution of equation (12), provided that the τ_j 's fulfil the linear relations

$$-\sum_{\mathbf{k}_\perp\sigma}\frac{k_\perp^2}{m_i}\frac{\partial f_i^{(0)}}{\partial \varepsilon_{\mathbf{k}_\perp}} = \sum_j K_{ij} \tau_j \quad (18)$$

where :

$$\begin{aligned} K_{ij} = & \frac{2\pi}{\hbar} \sum_{\substack{\mathbf{k}_\perp, \mathbf{k}'_\perp \\ \sigma, \sigma'}} \delta_{ij} \left\{ \frac{\mathbf{k}_\perp}{m_i} \cdot (\mathbf{k}'_\perp - \mathbf{k}_\perp) \delta[\varepsilon_i(\mathbf{k}_\perp) - \varepsilon_i(\mathbf{k}'_\perp)] \times \right. \\ & \left\langle \left| \langle i \mathbf{k}_\perp \sigma' | \mathcal{H}_{\text{def}} | i \mathbf{k}'_\perp \sigma' \rangle \right|^2 \right\rangle_{\text{average}} \frac{\partial f_i^{(0)}}{\partial \varepsilon_{\mathbf{k}_\perp}} - \frac{k_\perp^2}{m_i} \frac{\partial f_i^{(0)}}{\partial \varepsilon_{\mathbf{k}_\perp}} \sum_l (1 - \delta_{il}) \times \\ & \delta[\varepsilon_i(\mathbf{k}_\perp) - \varepsilon_l(\mathbf{k}'_\perp)] \left\langle \left| \langle i \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | l \mathbf{k}'_\perp \sigma' \rangle \right|^2 \right\rangle_{\text{average}} \} + \\ & + (1 - \delta_{ij}) \frac{\mathbf{k}_\perp \cdot \mathbf{k}'_\perp}{m_j} \delta(\varepsilon_i(\mathbf{k}_\perp) - \varepsilon_j(\mathbf{k}'_\perp)) \times \left\langle \left| \langle i \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | l \mathbf{k}'_\perp \sigma' \rangle \right|^2 \right\rangle_{\text{average}} \frac{\partial f_i^{(0)}}{\partial \varepsilon_{\mathbf{k}_\perp}} \end{aligned} \quad (19)$$

and where the spin variables σ, σ' have been explicitly introduced. At $T = 0$ K equations (18, 19) can be further simplified. In this limit, the chemical potential μ coincides with the Fermi energy ε_F and :

$$\frac{\partial f_i^{(0)}}{\partial \varepsilon_{\mathbf{k}_\perp}} = -\delta[\varepsilon_F - \varepsilon_i(\mathbf{k}_\perp)] \quad (20)$$

$$\sum_{\mathbf{k}_\perp\sigma} \hbar^2 \frac{k_\perp^2}{m_i} \frac{\partial f_i^{(0)}}{\partial \varepsilon_{\mathbf{k}_\perp}} = -\frac{2m_i S}{\pi\hbar^2} [\varepsilon_F - E_i] \quad (21)$$

Thus, equations (18, 19) can be rewritten :

$$\varepsilon_F - E_i = \sum_j K_{ij}(\varepsilon_F) \tau_j(\varepsilon_F) \quad (22)$$

with :

$$\begin{aligned} K_{ij}(T = 0) = & \frac{\pi^2 \hbar^3}{m_i S} \sum_{\substack{\mathbf{k}_\perp, \mathbf{k}'_\perp \\ \sigma, \sigma'}} \left[\delta_{ij} \left\{ \frac{\mathbf{k}_\perp}{m_i} \cdot (\mathbf{k}'_\perp - \mathbf{k}_\perp) \delta[\varepsilon_F - \varepsilon_i(\mathbf{k}'_\perp)] \times \right. \right. \\ & \delta[\varepsilon_F - \varepsilon_i(\mathbf{k}_\perp)] \left\langle \left| \langle i \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | i \mathbf{k}'_\perp \sigma' \rangle \right|^2 \right\rangle_{\text{average}} + \frac{k_\perp^2}{m_i} \delta[\varepsilon_F - \varepsilon_i(\mathbf{k}_\perp)] \times \\ & \sum_l (1 - \delta_{il}) \delta[\varepsilon_F - \varepsilon_l(\mathbf{k}'_\perp)] \left\langle \left| \langle i \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | l \mathbf{k}'_\perp \sigma' \rangle \right|^2 \right\rangle_{\text{average}} \} - \\ & - (1 - \delta_{ij}) \frac{\mathbf{k}_\perp \cdot \mathbf{k}'_\perp}{m_j} \delta[\varepsilon_F - \varepsilon_i(\mathbf{k}_\perp)] \delta[\varepsilon_F - \varepsilon_j(\mathbf{k}'_\perp)] \left\langle \left| \langle i \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | j \mathbf{k}'_\perp \sigma' \rangle \right|^2 \right\rangle_{\text{average}} \right] \end{aligned} \quad (23)$$

The areal density of the electrical current averaged over the distribution function f_n is equal to

$$\langle J \rangle = -\frac{e}{S} \sum_{i, \mathbf{k}_\perp, \sigma} \frac{\hbar \mathbf{k}_\perp}{m_i} f_i[\varepsilon_i(\mathbf{k}_\perp)] \quad (24)$$

As expected the zeroth order terms $f_i^{(0)}$ do not contribute to $\langle J \rangle$ and we are left with

$$\langle J \rangle = \frac{e^2 \hbar^2}{S} \sum_{i, \mathbf{k}_\perp, \sigma} \mathbf{k}_\perp \frac{\mathbf{k}_\perp \cdot \mathbf{F}}{(m_i)^2} \tau_i(\varepsilon_F) \delta[\varepsilon_F - \varepsilon_i(\mathbf{k}_\perp)] \quad (25)$$

at $T = 0$ K. Exploiting the in-plane isotropy of the dispersion relations $\varepsilon_i(\mathbf{k}_\perp)$, we can re-express the formula (25) in terms of the areal densities of carriers $n_e^{(i)}$ in the i^{th} subband :

$$n_e^{(i)} = \frac{m_i}{\pi \hbar^2} (\varepsilon_F - E_i) \quad (26)$$

We find the Ohm law

$$\langle J \rangle = \sigma \mathbf{F} \quad (27)$$

where :

$$\sigma = \sum_i \sigma_i \quad (28)$$

and

$$\sigma_i = \frac{n_e^{(i)} e^2 \tau_i(\varepsilon_F)}{m_i} \quad (29)$$

The carrier mobility of the i^{th} subband at zero temperature is given by :

$$\mu_i = \frac{e}{m_i} \tau_i(\varepsilon_F) \quad (30)$$

In terms of the total areal density of carriers n_e

$$n_e = \sum_i n_e^{(i)} \quad (31)$$

one can define an effective mobility $\bar{\mu}$ which is related to σ by

$$\sigma = n_e e \bar{\mu} \quad (32)$$

Thus :

$$\bar{\mu} = \sum_i \mu_i \frac{n_e^{(i)}}{n_e} \quad (33)$$

The very fact that $\tau_i(\varepsilon_F)$ is not simply proportional to $\varepsilon_F - E_i$ but should be derived by inverting the K_{ij} matrix, clearly shows that elastic inter-subband transitions contribute to $\langle J \rangle$. It is only in the Electric Quantum Limit that equation (22) can be simplified to display a more transparent structure.

I.1 ELECTRICAL CONDUCTIVITY IN THE ELECTRIC QUANTUM LIMIT. — When only the E_1 subband is populated, the only non-vanishing K_{ij} term is K_{11} . Let us define \mathbf{q}_\perp by

$$\mathbf{k}'_\perp - \mathbf{k}_\perp = \mathbf{q}_\perp \quad (34)$$

$\hbar \mathbf{q}_\perp$ is the momentum gained by the carrier during a collision against defects (see Fig. 1). Assume for simplicity that \mathcal{H}_{def} is spin-conserving (i.e. does not involve spin-orbit couplings). K_{11} can then be rewritten

$$K_{11} = -\frac{2\pi^2\hbar^3}{(m_1)^2 S} \sum_{\mathbf{k}_\perp, \mathbf{q}_\perp} \mathbf{k}_\perp \cdot \mathbf{q}_\perp \delta[\epsilon_F - \epsilon_1(\mathbf{k}_\perp + \mathbf{q}_\perp)] \delta[\epsilon_F - \epsilon_1(\mathbf{k}_\perp)]$$

$$\langle |\langle 1 \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | 1 \mathbf{k}'_\perp \sigma \rangle|^2 \rangle_{\text{average}} \quad (35)$$

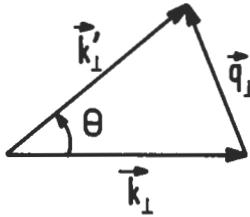


Fig. 1.

However from figure 1

$$\mathbf{k}_\perp \cdot \mathbf{q}_\perp = -2 k_\perp^2 \sin^2 \frac{\theta}{2}$$

Thus, by performing the summation over \mathbf{k}_\perp and \mathbf{k}'_\perp , we obtain

$$K_{11}(T=0) = \frac{m_1 S}{2\pi\hbar^3} \times [\epsilon_F - E_1] \times \int_0^{2\pi} d\theta (1 - \cos \theta) \times$$

$$\langle |\langle 1 \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | 1 \mathbf{k}_\perp + \mathbf{q}_\perp \sigma \rangle|^2 \rangle_{\text{average}}^{q_\perp = 2k_F \left| \sin \frac{\theta}{2} \right|} \quad (36)$$

and

$$\frac{1}{\tau_1(\epsilon_F)} = \frac{m_1 S}{2\pi\hbar^3} \int_0^{2\pi} d\theta (1 - \cos \theta) \times$$

$$\langle |\langle 1 \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | 1 \mathbf{k}_\perp + \mathbf{q}_\perp \sigma \rangle|^2 \rangle_{\text{average}}^{q_\perp = 2k_F \left| \sin \frac{\theta}{2} \right|} \quad (37)$$

where k_F is the Fermi wavevector

$$\epsilon_F - E_1 = \frac{\hbar^2 k_F^2}{2m_1} \quad (38)$$

The expression (37) shows that the velocity relaxation frequency $\frac{1}{\tau_1}$ is the angular average over all possible elastic collisions which occur on the Fermi circle of the product of the transition probability for an elastic scattering to change the carrier velocity orientation by θ , by the relative velocity decrement $k_F(1 - \cos \theta)/k_F$.

I.2 INTERSUBBAND SCATTERING. — Before evaluating \mathcal{H}_{def} for specific defect potentials, let us qualitatively describe the mobility anomalies which take place when the Fermi energy coincides with the onset of the first excited subband E_2 . Experimentally, two techniques have been used to modify ε_F in Ga(Al)As-GaAs heterojunctions. n_e has been changed by using the persistent photoconductivity effect or by biasing the heterostructure *via*, for instance, a Cu electrode placed on the GaAs side of the heterojunctions [12].

We assume that the E_2 density of states is a perfect staircase and that $T = 0$ K. When $\varepsilon_F = E_2 - \eta$, $\eta \rightarrow 0$ we have :

$$\mu(\varepsilon_F) = \frac{e\tau_1}{m_1} \quad \tau_1 = \frac{(\varepsilon_F - E_1)}{K_{11}} \quad (39)$$

On the other hand, when $\varepsilon_F = E_2 + \eta$, $\eta \rightarrow 0$, we obtain

$$\mu(\varepsilon_F) = \frac{e\tau_1}{m_1} \quad \tau_1 = \frac{\varepsilon_F - E_1}{K_{11} - K_{21}K_{12}/K_{22}} \quad (40)$$

The second subband does not yet contribute to the conductivity (because $n_2 \rightarrow 0$) but the τ_1 expression has changed compared to equation (39) due to the non-vanishing K_{12} and the change in the K_{11} expression. Let us examine K_{12} and K_{11} in some details. From equation (23) :

$$K_{12} = -\frac{2\pi^2\hbar^3}{Sm_1m_2} \sum_{\mathbf{k}_\perp, \mathbf{k}'_\perp} \mathbf{k}_\perp \cdot \mathbf{k}'_\perp \delta\left[\varepsilon_F - E_1 - \frac{\hbar^2 k_\perp^2}{2m_1}\right] \delta\left[\varepsilon_F - E_2 - \frac{\hbar^2 k'_\perp^2}{2m_2}\right] \times \langle | \langle 1 \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | 2 \mathbf{k}'_\perp \sigma \rangle |^2 \rangle_{\text{average}} \quad (41)$$

Since $\varepsilon_F \sim E_2$, $\mathbf{k}'_\perp \sim 0$. The \mathcal{H}_{def} matrix element is finite. Thus $K_{12} \sim 0$ because of the factor \mathbf{k}'_\perp sitting in front of the delta functions.

As for K_{11} , we see from inspection of equation (23) that it can be rewritten as

$$K_{11}(\varepsilon_F = E_2 + \eta) = K_{11}(\varepsilon_F = E_2 - \eta) + \frac{2\pi^2\hbar^3}{m_1 S} \sum_{\mathbf{k}_\perp} \frac{k_\perp^2}{m_1} \times \left[\delta\left(\varepsilon_F - E_1 - \frac{\hbar^2 k_\perp^2}{2m_1}\right) \right] \sum_{\mathbf{k}'_\perp} \delta\left(\varepsilon_F - E_2 - \frac{\hbar^2 k'_\perp^2}{2m_2}\right) \times \langle | \langle 1 \mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | 2 \mathbf{k}'_\perp \sigma \rangle |^2 \rangle_{\text{average}} \quad (42)$$

The important difference between equations (41, 42) is the absence in equation (42) of a \mathbf{k}'_1 term which multiplies the delta function. Thus the sum over \mathbf{k}'_1 is non vanishing. It amounts to the density of states of the E_2 subband, which is constant, if the \mathcal{H}_{def} matrix elements are q independent. Therefore, at the onset of the population of the second subband, $\tau_1(\varepsilon_F)$ and thus $\mu(\varepsilon_F)$ have a discontinuity (a drop since $\delta K_{11} > 0$) (see Fig. 2) which is entirely due to the peculiar shape of the two-dimensional density of states [11, 13].

Notice that for bulk materials (whose density of states vanishes at the band extrema) the same kind of reasoning would lead to a continuous $\mu(\varepsilon_F)$; only the derivative $d\mu/d\varepsilon_F$ would be singular. For more singular densities of states (e.g. those of quasi uni-dimensional materials), K_{11} would diverge and thus $\mu(\varepsilon_F)$ would vanish

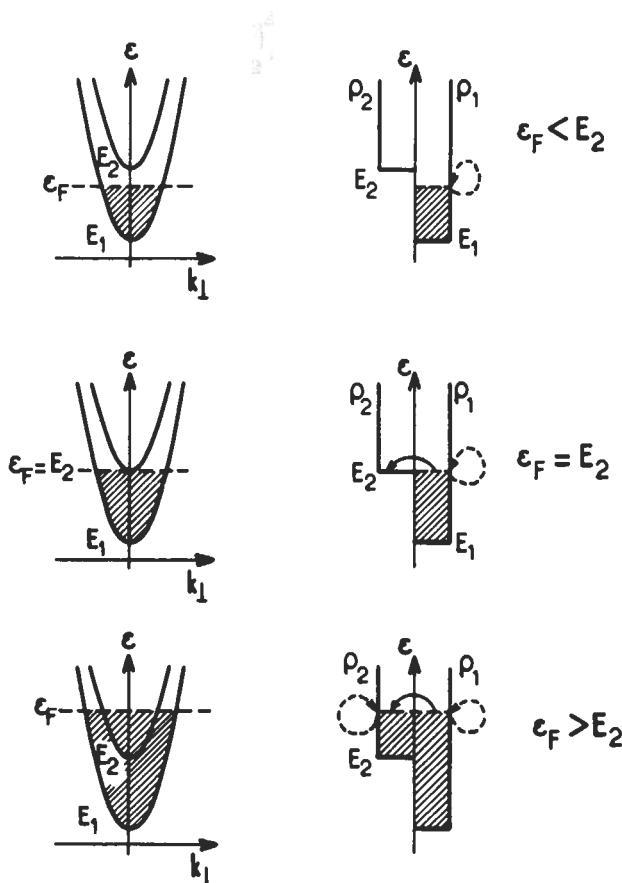


Fig. 2. — Schematic representation of interband scattering processes in a quasi bi-dimensional electron gas at $T = 0$ K. Three situations have been sketched : $\varepsilon_F < E_2$; $\varepsilon_F = E_2$; $\varepsilon_F > E_2$. The mobility exhibits a drop when $\varepsilon_F = E_2$ due to the finite value of the density of states $\rho_2(E_F)$ at $\varepsilon_F = E_2$.

at the onset of the occupation of a one-dimensional subband (Fig. 3). However, near such singularities the simple Born approximation which was used is no longer justified (the scattering is no longer weak) and self-consistent treatments are required to provide sensible results.

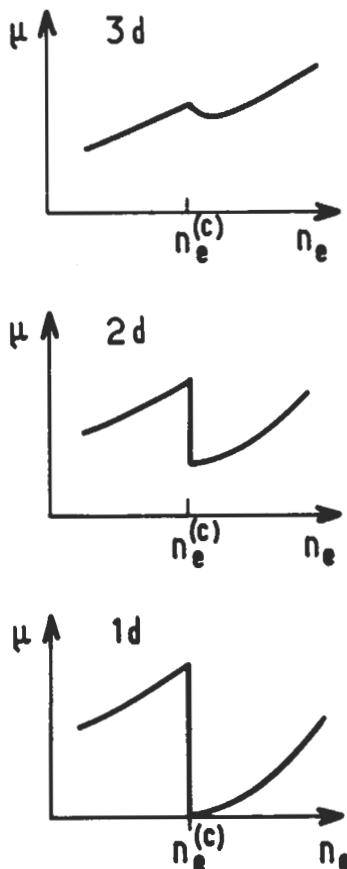


Fig. 3. — Schematic comparison between the mobility anomalies which take place at the onset of occupancy of an excited subband (or band) in three-dimensional (upper panel), quasi bidimensional (middle panel) and quasi uni-dimensional (lower panel) materials respectively. Zero temperature and Born approximation are assumed in these sketches.

In practice, the mobility drop can never be achieved in quasi bidimensional materials due to thermal and collisional broadenings which blur the E_2 onset. Nevertheless, pronounced decreases in mobility have been observed in GaAs-Ga(Al)As heterojunctions (see Fig. 4) and rather detailed calculations of this mobility drop have been performed at $T = 0$ K (see Figs. 5, 6).

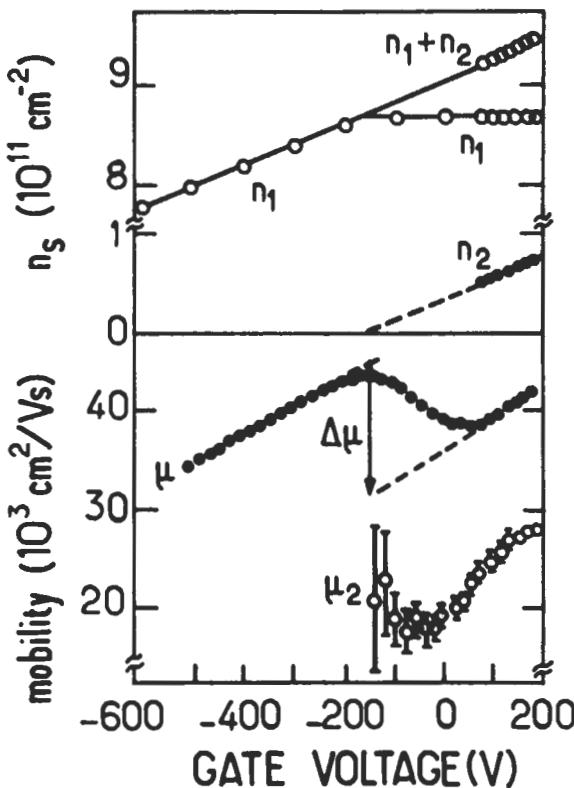


Fig. 4. — Experimental mobility drop in GaAs-Ga(Al)As heterojunctions which takes place at the onset of occupancy of the E_2 subband. $T = 4.2$ K. After reference [12].

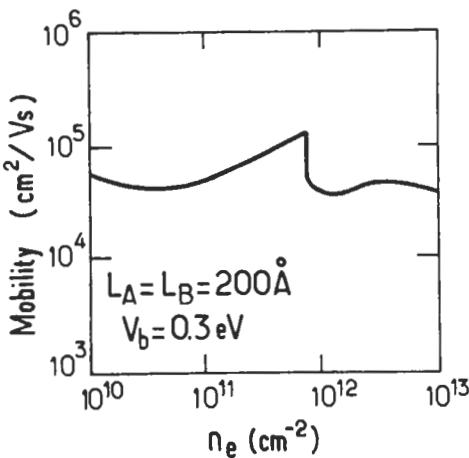


Fig. 5. — Calculated mobility drop which occurs in a GaAs-Ga(Al)As superlattice when the second electron subband becomes occupied. The electrons are assumed to be scattered by screened coulombic impurities. After reference [11].

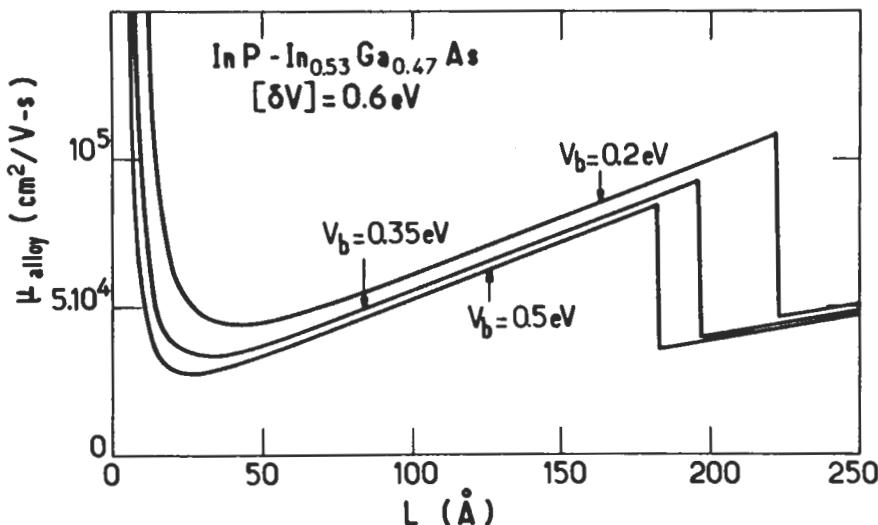


Fig. 6. — Calculated mobility drop which occurs in InP-Ga_{0.47}In_{0.53}As quantum wells when the second subband becomes occupied. The electrons are assumed to be scattered by short-range alloy fluctuations. After reference [42].

I.3 SCREENING IN A QUASI BI-DIMENSIONAL ELECTRON GAS. APPLICATION TO COULOMBIC IMPURITIES. — In modulation-doped heterostructures, the two dimensional electron gas is heavily degenerate ($\epsilon_F - E_1)/k_B T \gg 1$) at low temperatures. The interaction between a given electron and an impurity is strongly modified by the presence of the other electrons. In fact, the impurity induces a spatially dependent charge density in the electron gas and the actual electron-impurity interaction should account for this spatial rearrangement of the electron gas. This modification of the electron-impurity interaction with respect to that of the bare one is called the screening. In the following, we shall adopt a linear response formalism to calculate the screening effects. This means that the electron-impurity potential will be assumed weak enough to modify the electron wavefunctions only to the first order in V_{imp} . Thus, our derivation will closely follow Lindhart's in bulk materials [14]. An additional assumption will be that only a single subband is occupied (Electric Quantum Limit). This facilitates the algebra.

Suppose then that an impurity gives rise to a bare electrostatic potential $\phi^{\text{ext}}(\mathbf{r})$. The latter induces a charge distribution $\rho^{\text{ind}}(\mathbf{r})$, which itself gives rise to the electrostatic potential $\phi^{\text{ind}}(\mathbf{r})$. The net electrostatic potential is $\phi^{\text{ext}}(\mathbf{r}) + \phi^{\text{ind}}(\mathbf{r})$, and an electron of charge $-e$ ($e > 0$), instead of experiencing only the self-consistent potential due to the heterojunction band bending, (translationally invariant in the layer plane), has an Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + V(\mathbf{r}) \quad (43)$$

$$V(\mathbf{r}) = -e [\phi^{\text{ext}}(\mathbf{r}) + \phi^{\text{ind}}(\mathbf{r})] = -e \phi^{\text{tot}}(\mathbf{r}) \quad (44)$$

where \mathcal{H}_0 is the unperturbed, one-electron Hamiltonian whose solutions have been studied in the previous chapter. In the Electric Quantum Limit only the E_1 subband is occupied. We are interested in evaluating the induced charge density due to $V(\mathbf{r})$ by a perturbative scheme. $V(\mathbf{r})$ induces two kinds of mixing : the intra-subband one, which is equivalent to the one found in bulk metals and which can occur with an arbitrarily small energy (excitation on the Fermi circle) ; and the intersubband one, which is more like the one found in bulk semiconductors, and which is a virtual, non-resonant, mechanism, since the energy difference separating the occupied E_1 states from the E_2, \dots empty states is at least $E_2 - \epsilon_F$.

If $V(\mathbf{r})$ is small, we can only retain the intra-subband terms and neglect the others. The perturbed wavefunctions of the E_1 states are thus

$$\bar{\psi}_{1\sigma}(\mathbf{r}) = \chi_1(z) \frac{1}{\sqrt{S}} \exp i \mathbf{k}_\perp \cdot \mathbf{r}_\perp | \sigma \rangle + \sum_{\mathbf{k}'_\perp \sigma'} \frac{\langle E_1 \mathbf{k}'_\perp \sigma' | V(\mathbf{r}) | E_1 \mathbf{k}_\perp \sigma \rangle}{\frac{\hbar^2}{2m_1} (k_\perp^2 - k'^2_\perp)} \times \chi_1(z) \frac{1}{\sqrt{S}} \exp i \mathbf{k}'_\perp \cdot \mathbf{r}_\perp | \sigma' \rangle \quad (45)$$

where σ is the spin quantum number.

The induced charge density is

$$\rho^{\text{ind}}(\mathbf{r}) = -e \sum_{\sigma \mathbf{k}_\perp} \left\{ |\psi_{1\sigma}(\mathbf{r})|^2 - \frac{1}{S} |\chi_1(z)|^2 \right\} f_1^{(0)}(\mathbf{k}_\perp) \quad (46)$$

where $f_1^{(0)}$ is the Fermi Dirac distribution function of the E_1 states.

Retaining only the terms which are linear in V we obtain after some manipulations

$$\rho^{\text{ind}}(\mathbf{r}) = e^2 \sum_{\mathbf{k}_\perp, \mathbf{q}_\perp, \sigma} \frac{[f_1^{(0)}(\mathbf{k}_\perp) - f_1^{(0)}(\mathbf{k}_\perp + \mathbf{q}_\perp)]}{\frac{\hbar^2}{2m_1} [k_\perp^2 - (\mathbf{k}_\perp + \mathbf{q}_\perp)^2]} \times \langle E_1 \mathbf{k}_\perp + \mathbf{q}_\perp \sigma | \phi^{\text{ext}}(\mathbf{r}) + \phi^{\text{ind}}(\mathbf{r}) | E_1 \mathbf{k}_\perp \sigma \rangle \times \frac{1}{S} \chi_1^2(z) \exp(i \mathbf{q}_\perp \cdot \mathbf{r}_\perp) \quad (47)$$

where $\phi_{\text{ext}}(\mathbf{r})$ and $\phi_{\text{tot}}(\mathbf{r})$ have been explicitly assumed to be spin independent and the spin variables dropped where convenient.

The induced charge density can be rewritten

$$\rho^{\text{ind}}(\mathbf{r}) = e^2 \sum_{\mathbf{q}_\perp} e^{i \mathbf{q}_\perp \cdot \mathbf{r}_\perp} \chi_1^2(z) \phi_{11}^{\text{total}}(\mathbf{q}_\perp) A_{11}(\mathbf{q}_\perp) \quad (48)$$

where :

$$A_{11}(\mathbf{q}_\perp) = \frac{2}{S} \sum_{\mathbf{k}_\perp} \frac{f_1^{(0)}(\mathbf{k}_\perp + \mathbf{q}_\perp) - f_1^{(0)}(\mathbf{k}_\perp)}{\frac{\hbar^2}{2m_1} [(\mathbf{k}_\perp + \mathbf{q}_\perp)^2 - k_\perp^2]} \quad (49)$$

Notice that $A_{11}(\mathbf{q})$ is independent of the details of the heterostructure (shape of the χ_1 wavefunction etc...). It is however strongly influenced by the bi-dimensionality of the carrier motion (see Appendix A for a comparison at $T = 0$ K between the A_{11} 's evaluated for bulk, bi-dimensional and uni-dimensional carrier motions).

Equation (48) is simply the two-dimensional Fourier expansion of $\rho^{\text{ind}}(\mathbf{r})$:

$$\begin{aligned}\rho^{\text{ind}}(\mathbf{r}) &= \sum_{\mathbf{q}_\perp} \rho^{\text{ind}}(\mathbf{q}_\perp, z) e^{i\mathbf{q}_\perp \cdot \mathbf{r}_\perp}; \\ \rho^{\text{ind}}(\mathbf{q}_\perp, z) &= e^2 \phi_{11}^{\text{total}}(\mathbf{q}_\perp) A_{11}(\mathbf{q}_\perp) \chi_1^2(z)\end{aligned}\quad (50)$$

On the other hand $\phi^{\text{ind}}(\mathbf{r})$ is equal to

$$\phi^{\text{ind}}(\mathbf{r}) = \frac{1}{\kappa} \int \frac{d^2 r'_\perp dz' \rho^{\text{ind}}(\mathbf{r}')}{\sqrt{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 + (z - z')^2}} \quad (51)$$

where κ is the static permittivity of the heterostructures. Thus $\phi^{\text{ind}}(\mathbf{r})$ also admits a two-dimensional Fourier expansion

$$\phi^{\text{ind}}(\mathbf{r}) = \sum_{\mathbf{q}_\perp} \phi^{\text{ind}}(\mathbf{q}_\perp, z) e^{i\mathbf{q}_\perp \cdot \mathbf{r}_\perp} \quad (52)$$

where :

$$\phi^{\text{ind}}(\mathbf{q}_\perp, z) = \frac{1}{\kappa} \int \frac{d^2 r'_\perp dz' e^{i\mathbf{q}_\perp \cdot \mathbf{r}'_\perp} \rho^{\text{ind}}(\mathbf{q}_\perp, z')}{\sqrt{r'^2_\perp + (z - z')^2}} \quad (53)$$

Let us define the matrix element

$$\phi_{11}^{\text{ind}}(\mathbf{q}_\perp) = \int dz \chi_1^2(z) \phi^{\text{ind}}(\mathbf{q}_\perp, z) \quad (54)$$

Then $\phi_{11}^{\text{ind}}(\mathbf{q}_\perp)$ and $\phi_{11}^{\text{total}}(\mathbf{q}_\perp)$ fulfil the linear relationship

$$\phi_{11}^{\text{ind}}(\mathbf{q}_\perp) = A_{11}(\mathbf{q}_\perp) \phi_{11}^{\text{total}}(\mathbf{q}_\perp) T_{11}^{11}(\mathbf{q}_\perp) \quad (55)$$

where : $T_{11}^{11}(\mathbf{q}_\perp) = \frac{e^2}{\kappa} \int dz dz' \chi_1^2(z) \chi_1^2(z') \int \frac{d^2 r'_\perp e^{i\mathbf{q}_\perp \cdot \mathbf{r}'_\perp}}{\sqrt{r'^2_\perp + (z - z')^2}}$ (56)

which can be simplified into

$$T_{11}^{11}(\mathbf{q}_\perp) = \frac{2\pi e^2}{\kappa} \frac{1}{q_\perp} \int dz dz' \chi_1^2(z) \chi_1^2(z') e^{-q_\perp |z - z'|} \quad (57)$$

if one uses the bi-dimensional Fourier expansion of $|\mathbf{r} - \mathbf{r}'|^{-1}$:

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{2\pi}{S} \sum_{\mathbf{K}_\perp} \frac{1}{K_\perp} \exp[i \mathbf{K}_\perp \cdot (\mathbf{r}_\perp - \mathbf{r}'_\perp)] \exp[-K_\perp |z - z'|] \quad (58)$$

We are close to the end since the linear relation equation (55) can be inverted to provide the explicit relationship between the total and external ϕ_{11} 's respectively :

$$\phi_{11}^{\text{tot}}(\mathbf{q}_\perp) = \frac{1}{\epsilon(\mathbf{q}_\perp)} \phi_{11}^{\text{ext}}(\mathbf{q}_\perp) \quad (59)$$

where :

$$\epsilon(\mathbf{q}_\perp) = 1 - A_{11}(\mathbf{q}_\perp) T_{11}^{11}(\mathbf{q}_\perp) \quad (60)$$

is the wavevector dependent dielectric constant of the quasi bi-dimensional electron gas [15].

We have only partly succeeded in evaluating the screening action of the electron gas on the electron-external potential interaction. Only the averages of the external and total potentials over the $\chi_1^2(z)$ charge distributions were easily obtained. In a bulk material, we would have gone further and obtained a direct relationship between the three-dimensional Fourier coefficients of $\phi^{\text{tot}}(\mathbf{r})$ and $\phi^{\text{ext}}(\mathbf{r})$. By inverting the Fourier expansion, the explicit \mathbf{r} -dependence of $\phi^{\text{tot}}(\mathbf{r})$ would have resulted. For instance, when $\epsilon(\mathbf{q})$ is approximated by its long wavelength limit ($\epsilon(\mathbf{q}) = 1 + \lambda^2/q^2$), the potential takes the familiar Yukawa form :

$$\phi^{\text{tot}}(\mathbf{r}) = \frac{e}{\kappa r} \exp(-\lambda r) \quad (61)$$

whereas

$$\phi^{\text{ext}}(\mathbf{r}) = \frac{e}{\kappa r} \quad (62)$$

In quasi bi-dimensional systems, there is a strong anisotropy between the responses of the electron gas to an external perturbation along and perpendicular to the growth axis. The in-plane response is similar to the bulk one, but along the z axis the carrier is not liable to distort its wavefunction. Thus the formal analogy between the dielectric constants $\epsilon(\mathbf{q}_\perp)$ in bulk and quasi bi-dimensional systems can only be partial. It is only in the exactly bi-dimensional limit (which never exists in reality) that a complete analogy is recovered. In the latter situation, the electrons move on a plane located at $z = z_0$. Thus $\chi_1^2(z) = \delta(z - z_0)$ and T_{11}^{11} reduces to $2\pi e^2/\kappa q_\perp$ which is the bidimensional equivalent of $4\pi e^2/\kappa q^2$ in bulk materials, i.e. the Fourier component of the bare Coulombic interaction.

I.4 DISCUSSION OF THE DIELECTRIC FUNCTION. — We rewrite equation (60) in the form

$$\epsilon(\mathbf{q}_\perp) = 1 - \frac{2\pi e^2}{\kappa q_\perp} g_s(\mathbf{q}_\perp) A_{11}(\mathbf{q}_\perp) \quad (63)$$

where $g_s(\mathbf{q}_\perp)$ is the screening form factor arising from the finite extension of the electron wavefunctions along the z axis :

$$g_s(\mathbf{q}_\perp) = \int \chi_1^2(z) \chi_1^2(z') \exp(-q_\perp |z - z'|) dz dz' \quad (64)$$

It should be realized that, as g_s is smaller than 1, the screening action is diminished compared with the exactly bi-dimensional case (where $g_s = 1$). In fact at large wavevectors we obtain :

$$g_s(q_\perp) \xrightarrow[q_\perp \rightarrow \infty]{} \frac{2}{q_\perp} \int_{-\infty}^{+\infty} \chi_1^4(z) dz \quad (65)$$

whereas in the opposite limit

$$g_s(q_\perp) \xrightarrow[q_\perp \rightarrow 0]{} 1 \quad (66)$$

These limiting behaviour have led Price [16] to propose that $g_s(q_\perp)$ could be fitted to a simple expression like

$$g_s(q_\perp) = (1 + bq_\perp)^{-1} \quad (67)$$

where b is a c -constant, indicative of the spatial spread of the quasi bi-dimensional electron gas along the z axis. Figure 7 shows an example of calculated $g_s(q_\perp)$ for the case of quantum wells with infinite depth.

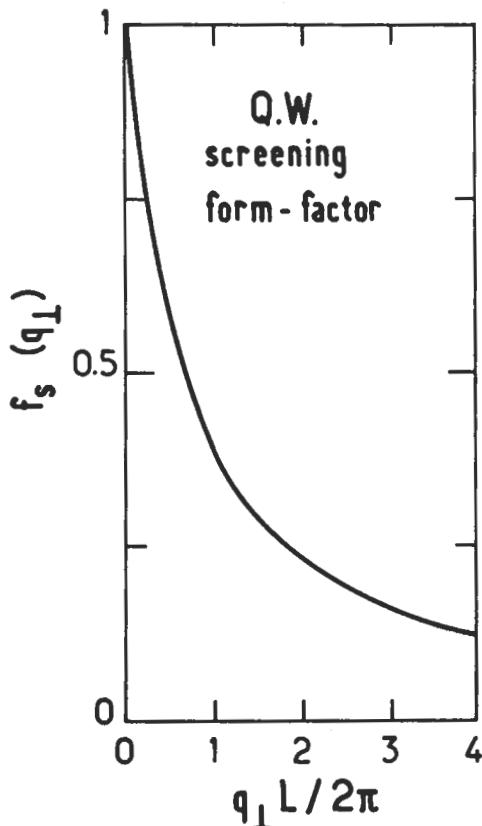


Fig. 7. — The calculated screening form factor of a quantum well clad between infinite barriers is plotted *versus* the dimensionless product $q_\perp L / 2\pi$.

The second factor which influences the screening action is $A_{11}(\mathbf{q}_\perp)$. As noticed above, it does not depend on the exact shape of the carrier wavefunction, but only on the bi-dimensional nature of the carrier motion. At $T = 0$ K, the Fermi surface is sharp and the Fermi-Dirac distribution functions reduce to step functions. As a result, $A_{11}(\mathbf{q}_\perp)$ can be obtained in closed forms if collision broadenings are neglected :

$$A_{11}(\mathbf{q}) = -\frac{m_1}{\pi\hbar^2} \left\{ 1 - \sqrt{1 - \left(\frac{2k_F}{q_\perp} \right)^2} Y(q_\perp - 2k_F) \right\} \quad (68)$$

where k_F is the Fermi wavevector ($k_F = \sqrt{2\pi n_e}$). It is remarkable that $A_{11}(\mathbf{q}_\perp)$ has a kink at $q_\perp = 2k_F$. Such a singularity is associated with the occurrence of elastic transitions which involve initial and final states which differ by one Fermi diameter (see Eq. (49)). The same kind of transitions leads, in bulk materials, to much weaker singularities (see Appendix A, where it is shown that only the derivative of A_{11} has a logarithmic singularity at $q = 2k_F$). We once again see that the reduction of the dimensionality of carrier motions emphasizes the critical behaviours.

By collecting all the results, we can rewrite $\varepsilon(\mathbf{q}_\perp)$ in the final form

$$\varepsilon(\mathbf{q}_\perp) = 1 + \frac{q_0}{q_\perp} g_s(\mathbf{q}_\perp) \alpha(\mathbf{q}_\perp) \quad (69)$$

with :

$$q_0 = \frac{2m_1 e^2}{\kappa\hbar^2} = \frac{2}{a_0^*} \quad (70)$$

where a_0^* is the bulk effective Bohr radius for the carriers under consideration and

$$\alpha(\mathbf{q}_\perp) = 1 - \sqrt{1 - \left(\frac{2k_F}{q_\perp} \right)^2} Y(q_\perp - 2k_F) \quad (71)$$

The long wavelength limit of $\varepsilon(\mathbf{q}_\perp)$ is particularly interesting, as we should be able to recover the two-dimensional analogue of the Thomas Fermi analysis of the electron gas response to a slowly varying disturbance. In the limit $q_\perp \rightarrow 0$ equation (69) leads to

$$\varepsilon(\mathbf{q}_\perp) \underset{q_\perp \rightarrow 0}{\approx} 1 + \frac{q_0}{q_\perp} \quad (72)$$

that is to say a screening effect which is independent of the carrier concentration. This result is in marked contrast to the bulk behaviour where we know that $\varepsilon(\mathbf{q}_\perp)$ increases with the bulk carrier concentration N_e .

The reason for such a great difference between the bulk and quasi bidimensional systems can be traced back to the different energy dependences of their respective density of states. In fact, let us rederive equation (72) using semiclassical arguments. Firstly the z motion should be discarded since the level quantization along the z axis can not be described in terms of a semiclassical analysis of the carrier motion. Thus, we assume a truly bi-dimensional situation. A slowly varying bi-dimensional

potential $\phi^{\text{ext}}(\mathbf{r}_\perp)$ perturbs an electron gas whose areal density is n_e in the absence of a perturbing potential. It induces a charge density $-e\delta n(\mathbf{r}_\perp)$, which itself gives rise to an induced potential $\phi^{\text{ind}}(\mathbf{r}_\perp)$. In the presence of the total potential $\phi^{\text{tot}}(\mathbf{r}_\perp)$ the semiclassical analysis [17] tells us that the electron energy is :

$$\epsilon_1(\mathbf{k}_\perp, \mathbf{r}_\perp) = E_1 + \frac{\hbar^2 k_\perp^2}{2m_1} - e\phi^{\text{tot}}(\mathbf{r}_\perp) \quad (73)$$

The electron density is no longer constant in space, but depends on \mathbf{r}_\perp through the implicit relation

$$n_e(\mathbf{r}_\perp) = \frac{2}{S} \sum_{\mathbf{k}_\perp} \left\{ 1 + \exp \left[\beta \left(E_1 + \frac{\hbar^2 k_\perp^2}{2m_1} - e\phi^{\text{tot}}(\mathbf{r}_\perp) - \mu \right) \right] \right\}^{-1} \quad (74)$$

The induced charge density is therefore :

$$-e\delta n(\mathbf{r}_\perp) = -e[n_e(\mu + e\phi^{\text{tot}}(\mathbf{r}_\perp)) - n_e(\mu)] \quad (75)$$

which after linearization gives :

$$-e\delta n(\mathbf{r}_\perp) = -e^2 \frac{\partial n_e}{\partial \mu} \phi^{\text{tot}}(\mathbf{r}_\perp) \quad (76)$$

The induced potential is thus :

$$\phi^{\text{ind}}(\mathbf{r}_\perp) = -\frac{e^2}{\kappa} \frac{\partial n_e}{\partial \mu} \int \frac{d^2 r'_\perp \phi^{\text{tot}}(\mathbf{r}'_\perp)}{|\mathbf{r}_\perp - \mathbf{r}'_\perp|} \quad (77)$$

If we Fourier-expand the Coulombic term $|\mathbf{r}_\perp - \mathbf{r}'_\perp|^{-1}$ we obtain from equation (77) a linear relationship between the bidimensional Fourier components $\phi^{\text{ind}}(\mathbf{q}_\perp)$ and $\phi^{\text{tot}}(\mathbf{q}_\perp)$ of the induced and total electrostatic potentials respectively :

$$\phi^{\text{ind}}(\mathbf{q}_\perp) = -\frac{2\pi e^2}{\kappa q_\perp} \frac{\partial n_e}{\partial \mu} \phi^{\text{tot}}(\mathbf{q}_\perp) \quad (78)$$

Finally equation (78) leads to the relation

$$\phi^{\text{tot}}(\mathbf{q}_\perp) = \frac{1}{\epsilon(\mathbf{q}_\perp)} \phi^{\text{ext}}(\mathbf{q}_\perp) \quad (79a)$$

where :

$$\epsilon(\mathbf{q}_\perp) = 1 + \frac{q_0}{q_\perp} \quad (79b)$$

$$q_0 = \frac{2\pi e^2}{\kappa} \frac{\partial n_e}{\partial \mu} \quad (79c)$$

As in bulk materials, we find that a semiclassical analysis provides the same result as the long wavelength limit of a Lindhard-type calculation (Eq. (72)). The new feature is that the screening wavevector q_0 is, at $T = 0$ K, concentration-independent, $\frac{\partial n_e}{\partial \mu}$ being equal to the density of states at the Fermi energy. This would imply

that a single electron is sufficient to completely screen any external potential, which is unphysical. As pointed out by Ando *et al.* [9] the $n_e \rightarrow 0$ limit cannot be taken within a Thomas Fermi scheme since, if n_e is small, k_F is also small and the q_\perp dependence of $\alpha(q_\perp)$ can no longer be neglected. Nevertheless, it is true that in the large n_e limit (still compatible with the Electric Quantum Limit assumption) the screening effects in quasi bi-dimensional electron gas saturate instead of continuing to increase as they do in bulk systems.

The calculations we have presented can be improved in several ways. The thermal and collisional broadenings can be included in the calculations. A recent self-consistent treatment of $\epsilon(q_\perp)$, including electron-impurity interactions, has been performed by Ando [18]. He found that the $2k_F$ singularity in $\alpha(q_\perp)$ is rounded off in somewhat the same manner as is produced by finite temperature (see Fig. 8).

Finally, when several subbands are occupied, the screening effects become complicated due to the interplay of intra-subband and intersubband contributions to $\epsilon(q)$.

The best we can do is to relate the matrix elements

$$\phi_{nj}^{\text{ind}}(q_\perp) = \int dz \chi_n(z) \chi_j(z) \phi^{\text{ind}}(q_\perp, z) \quad (80a)$$

$$\phi_{nj}^{\text{tot}}(q_\perp) = \int dz \chi_n(z) \chi_j(z) \phi^{\text{tot}}(q_\perp, z) \quad (80b)$$

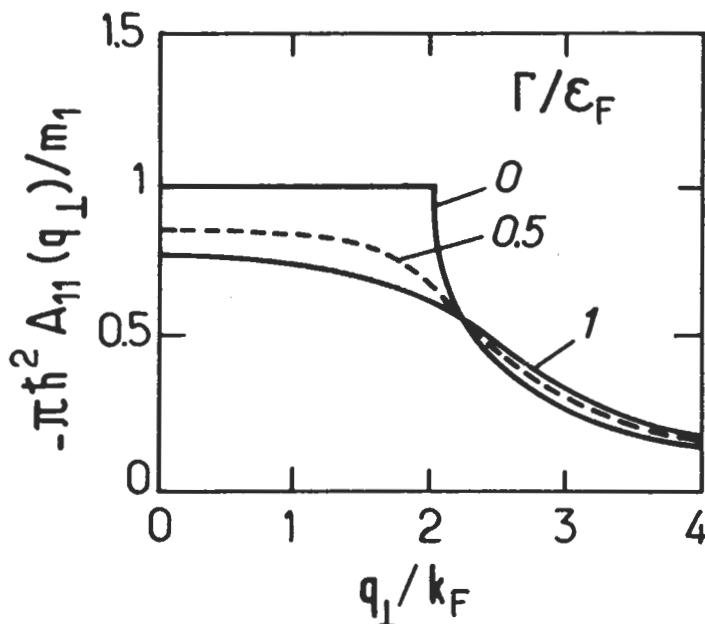


Fig. 8. — Calculated wavevector dependence of $A_{11}(q_\perp)$ in a quasi bi-dimensional electron gas perturbed by short-range scatterers for three different values of the broadening parameter Γ/ϵ_F . After reference [18].

(where $\phi(\mathbf{q}_\perp, z)$ has been defined in Eq. (52)) by a linear relationship in the form :

$$\phi_{nj}^{\text{ind}}(\mathbf{q}_\perp) = \sum_{l,m} A_{lm}(\mathbf{q}_\perp) T_{lm}^{jn}(\mathbf{q}_\perp) \phi_{lm}^{\text{tot}}(\mathbf{q}_\perp) \quad (81a)$$

where :

$$T_{lm}^{jn}(\mathbf{q}_\perp) = \frac{2\pi e^2}{\kappa q_\perp} \iint dz dz' \chi_j(z) \chi_n(z) \chi_l(z') \chi_m(z') \exp[-q_\perp |z - z'|] \quad (81b)$$

and

$$A_{lm}(\mathbf{q}_\perp) = \frac{2}{S} \sum_{\mathbf{k}_\perp} \frac{f_l^{(0)}(\mathbf{k}_\perp) - f_m^{(0)}(\mathbf{k}_\perp + \mathbf{q}_\perp)}{\epsilon_l(\mathbf{k}_\perp) - \epsilon_m(\mathbf{k}_\perp + \mathbf{q}_\perp)} \quad (81c)$$

II. Evaluation of some scattering mechanisms.

II.1 MOBILITY LIMITED BY COULOMBIC SCATTERING. — In quasi bi-dimensional electron gases, the low temperature mobility is sensitively affected by coulombic scatterings due to charged impurities. These are of two kinds in modulation-doped heterostructures.

i) The impurities which have been deliberately introduced in the barrier-acting layers and which have transferred their electrons into the channel.

ii) The residual impurities which take place either in the channel or in the barrier.

We make use of the formalism derived in section I to calculate the mobility limited by the Coulombic scatterings experienced by the quasi bi-dimensional electrons on charged impurities.

At $T = 0$ K and in the Electric Quantum Limit :

$$\mu(\epsilon_F) = \frac{e \tau_1}{m_1} \quad (82)$$

$$\frac{1}{\tau_1} = \frac{m_1 S}{2\pi\hbar^3} \int_0^{2\pi} d\theta (1 - \cos \theta) \left\langle \left| \langle 1\mathbf{k}_\perp \sigma | \mathcal{H}_{\text{def}} | 1\mathbf{k}_\perp + \mathbf{q}_\perp \sigma \rangle \right|^2 \right\rangle_{\text{average}}^{q_\perp = 2k_F \left| \sin \frac{\theta}{2} \right|} \quad (83)$$

where \mathcal{H}_{def} is the electron-impurity potential, or more exactly, the electron-impurity potential minus its spatial average in the layer plane. The latter term has already been included in energy level calculations (band bending). This spatial average being, by definition, translationally invariant in the layer-plane, has a zero matrix element between \mathbf{k}_\perp and $\mathbf{k}_\perp + \mathbf{q}_\perp$ plane wave states and can thus be dropped. The impurities are assumed to be randomly placed at the sites \mathbf{R}_i and the bare electron-impurity interaction can be Fourier analysed :

$$\mathcal{H}_{\text{def}}^{\text{bare}}(\mathbf{r}) = -\frac{2\pi e}{\kappa S} \sum_{\mathbf{R}_i} Z_i e_i \sum_{\mathbf{Q}_\perp} \frac{1}{Q_\perp} \exp[-Q_\perp |z - z_i|] \times \exp[i \mathbf{Q}_\perp \cdot (\mathbf{r}_\perp - \mathbf{R}_{i\perp})] \quad (84)$$

where $Z_i e_i$ is the algebraic charge of the impurity sitting at the site \mathbf{R}_i , $\mathbf{R}_i = (\mathbf{R}_{i\perp}, z_i)$. By Fourier expanding $\mathcal{H}_{\text{def}}^{\text{bare}}$ and \mathcal{H}_{def} in the following manner

$$\mathcal{H}_{\text{def}}^{\text{bare}}(\mathbf{r}) = \sum_{\mathbf{Q}_\perp} \mathcal{H}_{\text{def}}^{\text{bare}}(\mathbf{Q}_\perp, z) \exp(i \mathbf{Q}_\perp \cdot \mathbf{r}_\perp) \quad (85)$$

$$\mathcal{H}_{\text{def}}(\mathbf{r}) = \sum_{\mathbf{Q}_\perp} \mathcal{H}_{\text{def}}(\mathbf{Q}_\perp, z) \exp(i \mathbf{Q}_\perp \cdot \mathbf{r}_\perp) \quad (86)$$

and by taking the averages over the $\chi_1(z)$ wavefunction :

$$\mathcal{H}_{\text{def}}^{\text{bare}}(\mathbf{Q}_\perp) = \int_{-\infty}^{+\infty} \chi_1^2(z) \mathcal{H}_{\text{def}}^{\text{bare}}(\mathbf{Q}_\perp, z) dz \quad (87)$$

$$\mathcal{H}_{\text{def}}(\mathbf{Q}_\perp) = \int_{-\infty}^{+\infty} \chi_1^2(z) \mathcal{H}_{\text{def}}(\mathbf{Q}_\perp, z) dz \quad (88)$$

we readily obtain :

$$\langle 1 \mathbf{k}_\perp + \mathbf{q}_\perp \sigma | \mathcal{H}_{\text{def}} | 1 \mathbf{k}_\perp \sigma \rangle = \frac{1}{\varepsilon(\mathbf{q}_\perp)} \mathcal{H}_{\text{def}}^{\text{bare}}(\mathbf{q}_\perp) \quad (89)$$

$$= \sum_{\mathbf{R}_i} \frac{1}{\varepsilon(\mathbf{q}_\perp)} \left(\frac{-2\pi e e_i Z_i}{\kappa q_\perp S} \right) \exp[-i \mathbf{q}_\perp \cdot \mathbf{R}_{i\perp}] \int_{-\infty}^{+\infty} \exp[-q_\perp |z - z_i|] \chi_1^2(z) dz \quad (90)$$

Equation (90) can finally be rewritten as

$$\langle 1 \mathbf{k}_\perp + \mathbf{q}_\perp \sigma | \mathcal{H}_{\text{def}} | 1 \mathbf{k}_\perp \sigma \rangle = \frac{-2\pi e}{\kappa S [q_\perp + q_0 g_s(q_\perp)]} \sum_{\mathbf{R}_i} Z_i e_i e^{-i \mathbf{q}_\perp \cdot \mathbf{R}_{i\perp}} g_{\text{imp}}(q_\perp, z_i) \quad (91)$$

Thus, in spite of our incomplete evaluation of the screened impurity potential (only its average over $\chi_1(z)$ having been explicitly evaluated), the matrix elements of the coulombic scattering can be derived in closed forms. Notice that in equation (91), the $\alpha(\mathbf{q}_\perp)$ term which appears in equation (69) does not show up. This is due to our $T = 0$ K approximation, which forces q_\perp to be smaller or equal to $2 k_F$.

In equation (91) a new form factor appears : $g_{\text{imp}}(q_\perp, z_i)$. It is of fundamental importance to the understanding of the beneficial action of modulation-doping [19, 20], for it depends explicitly on the relative positions of the impurities and the electron gas :

$$g_{\text{imp}}(q_\perp, z_i) = \int_{-\infty}^{+\infty} \chi_1^2(z) \exp[-q_\perp |z - z_i|] \quad (92)$$

The q_\perp -dependence of g_{imp} is similar to that of g_s , being equal to one at vanishing q_\perp and dropping to zero like q_\perp^{-1} at large q_\perp . When the impurity is located far away from the electron gas, the latter can be assimilated to a sheet of negative charges. Therefore

$$g_{\text{imp}}(q_\perp, z_i) \sim \exp[-q_\perp d] \quad (93)$$

where d is the distance separating the impurity from the electrons. Equation (93) shows that, for remote impurities, the forward scattering events (i.e. $q_{\perp} \rightarrow 0$, or equivalently $\theta \rightarrow 0$) are strongly favoured over the backward scattering events. But this means that the carrier motion is barely affected by the scatterers (owing to the $1 - \cos \theta$ factor appearing in Eq. (83)) and, in consequence, the carrier mobility is high.

Due to the afore-mentioned q_{\perp} -dependence of the impurity form factor, one may anticipate that residual impurities sitting near the conducting channel may be more efficient scatterers than the remote impurities deliberately introduced into the barriers, although the latter are far more numerous than the former. In fact, for the Ga(Al)As-GaAs heterojunctions which display high mobilities, the scattering by impurities from the doped part of the barrier (volume concentration $\sim 2 \times 10^{18} \text{ cm}^{-3}$) is comparable to that of background impurities (volume concentration $\sim 10^{14} \text{ cm}^{-3}$) [21].

To achieve our task of evaluating $1/\tau(\varepsilon_F)$, we now make use of the random nature of the impurity sites. Using equations (15, 16), denoting the volume concentration of the μ^{th} impurity species (charge e_{μ}) by c_{μ} and the length over which this impurity species is found in the heterostructure by \mathcal{L}_{μ} ($\mathcal{L}_{\mu} = z_{\max} - z_{\min}$), $\hbar/\tau(\varepsilon_F)$ can finally be rewritten as :

$$\frac{\hbar}{\tau(\varepsilon_F)} = \frac{m_1}{\pi \hbar^2} \sum_{\mu} c_{\mu} \int_0^{\pi} d\theta (1 - \cos \theta) \left[\frac{2\pi e Z_{\mu} e_{\mu}}{\kappa \left[2k_F \sin \frac{\theta}{2} + q_0 g_s (2k_F \sin \frac{\theta}{2}) \right]} \right]^2 \times \\ \times \int_{z_{\min}}^{z_{\max}} dz_i g_{\text{imp}}^2 \left[2k_F \sin \left(\frac{\theta}{2} \right), z_i \right] \quad (94)$$

The expression (94) makes it possible to summarize the main trends of the Coulombic scattering in heterostructures :

i) \hbar/τ is proportional to $\frac{m_1}{\pi \hbar^2}$ which is the two-dimensional density of states. Thus, comparing electron and hole mobilities in heterostructures characterized by the same angular integral, one would expect to find a mobility ratio μ_e/μ_h proportional to $\left(\frac{m_h}{m_e} \right)^2$, and not to $\frac{m_h}{m_e}$ as is so often claimed in literature. In practice, such "mirror-like" heterostructures do not exist and the ratio μ_e/μ_h is of the order of 10 when the best electron and hole mobilities are compared ($\mu_e \sim 2 \times 10^6 \text{ cm}^2/\text{Vs}$, $\mu_h \sim 2 \times 10^5 \text{ cm}^2/\text{Vs}$)

ii) \hbar/τ is proportional to $c_{\mu} Z_{\mu}^2 e_{\mu}^2$. Each species scatters electrons independently and each impurity site is decorrelated from the others. We recall that this result is valid only for impurities which are diluted enough, a condition which is not always fulfilled, e.g. for the highly concentrated remote impurities.

- iii) \hbar/τ is an average over the scattering angles of the relative momentum loss $1 - \cos\theta$ weighed by the square of the scattering matrix elements. At small scattering angles (forward scattering), the screened coulombic scattering matrix element is finite, in contrast with the unscreened one.
- iv) The proportionality between $\hbar/\tau(\varepsilon_F)$ and $Z_\mu^2 e_\mu^2 e^2$ is characteristic of the Born approximation : attractive and repulsive impurities scatter in a similar way.
- v) This is finally the relative position of the scatterers with respect to the electron gas which more directly influences the carrier mobility through the impurity form factor.

Although the integral which appears in equation (94) has to be numerically worked out to obtain accurate results, an analytical model can be derived using simplifying assumptions. Let us then consider a very idealized modulation-doped heterojunction which has been spike-doped at z_0 with donors (areal concentration N_d), i.e. all the donors sit at $|z_0|$ from the electron gas. The latter is in the Electric Quantum Limit and we assume for simplicity that we can assimilate $\chi_1^2(z)$ to $\delta(z)$. We consider that all the donors are ionized. Thus, the areal concentration of electrons n_e is equal to N_d . With these assumptions $g_s(q_\perp)$ is unity and moreover

$$c_\mu \int_{z_{\min}}^{z_{\max}} dz_i g_{\text{imp}}^2 \left[2k_F \sin \frac{\theta}{2}, z_i \right] = N_d \exp \left[-4k_F |z_0| \sin \frac{\theta}{2} \right] \quad (95)$$

Consequently

$$\frac{\hbar}{\tau(\varepsilon_F)} = \frac{m_1 N_d}{\pi \hbar^2} \left(\frac{2\pi e^2}{\kappa} \right)^2 \int_0^\pi d\theta \frac{(1 - \cos \theta)}{\left[q_0 + 2k_F \sin \frac{\theta}{2} \right]^2} \times \exp \left[-4k_F |z_0| \sin \frac{\theta}{2} \right] \quad (96)$$

If $k_F z_0 \gg 1$, we may replace in equation (96) $\sin \frac{\theta}{2}$ by $\frac{\theta}{2}$, $1 - \cos \theta$ by $\frac{\theta^2}{2}$ and the denominator by q_0^2 to obtain

$$\mu(\varepsilon_F) = 16 |z_0|^3 \frac{e}{\hbar} \sqrt{2\pi n_e} \quad (97)$$

Thus, μ should vary like z_0^3 and $n_e^{\frac{1}{2}}$. Note that the latter dependence partially results from the neutrality condition $n_e = N_d$. If the scatterers were different from the donating impurities, we would have obtained, instead of equation (97), the expression :

$$\mu(\varepsilon_F) = 16 |z_0|^3 \frac{e}{\hbar} \frac{n_e}{N_d} \sqrt{2\pi n_e} \quad (98)$$

i.e. μ would have varied like $n_e^{3/2}$. Notice also that μ is effective mass independent, a result of a cancellation between the explicit m_1 dependence and the implicit one (in q_0). For $n_e = 10^{12} \text{ cm}^{-2}$, $|z_0| = 150 \text{ \AA}$ one finds $\mu \sim 2 \times 10^5 \text{ cm}^2/\text{Vs}$ from

equation (97), which is a reasonable order of magnitude. The n_e and z_0 dependences cannot be compared with the experiments since mobility figures in spike-doped heterojunctions have not yet been reported. For actual modulation-doped GaAs-Ga(Al)As heterojunctions, one often finds $\mu \sim n_e^\alpha$, $1 \leq \alpha \leq 2$ at low temperatures [22].

Various other simplified approaches of impurity-limited mobility in heterojunctions have been reported (see e.g. [2]). As in the previous analysis, they cannot be expected to give more than qualitative descriptions. A more quantitative approach to the n_e -dependence of mobility has been proposed by Stern [21], who took a Fang-Howard wavefunction to describe $\chi_1(z)$, but otherwise proceeded with the evaluation of all the form factors and angular integration (see Fig. 9). The electron concentration was varied (in a gedanken experiment) by changing the spacer thickness, which alters the charge transfer (see previous chapter). Stern was able to show that, at low n_e , the residual impurities in the channel and in the (nominally) undoped spacer layer play a very significant part. He also obtained peak mobility $\sim 10^6 \text{ cm}^2/\text{Vs}$ which compares rather favourably with experiments. Ando [22] fully analysed the low temperature data obtained by Hiyamizu *et al.* [23] and Tsui *et al.* [24]. He was able to find a good agreement between his very complete mobility

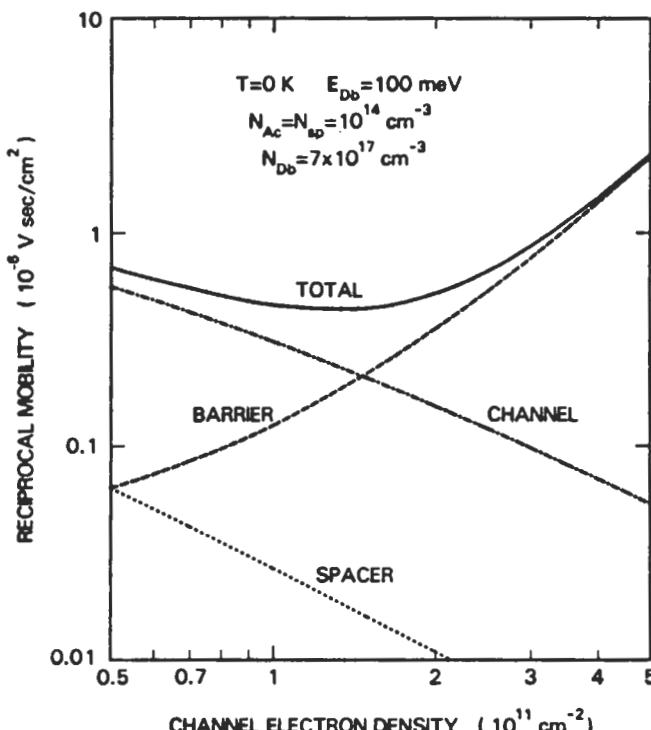


Fig. 9. — The reciprocal mobility of a quasi bi-dimensional electron gas in Ga(Al)As-GaAs heterostructures is plotted *versus* the areal electron concentration n_e . After reference [21].

calculations and the experiments by adjusting the effective number of ionized donors and by neglecting any residual impurities in the undoped spacer layer (but not in the channel).

Perhaps the clearest proof that residual impurities decisively influence the mobility of some high quality GaAs-Ga(Al)As heterojunctions is the complete lack of correlation between the mobility and spacer thickness reported by Hwang *et al.* [25] for n-type heterojunctions. Similar trends were reported by Wang *et al.* for p-type heterojunctions [26].

The $\text{Ga}_{1-x}\text{Al}_x\text{As}$ -GaAs modulation-doped heterojunctions have been very extensively studied. Their low temperature top mobility has steadily increased with time, starting from $\sim 3 \times 10^4 \text{ cm}^2/\text{Vs}$ to reach at present $\sim 5 \times 10^6 \text{ cm}^2/\text{Vs}$ ($x = 0.3$) [29b]. The latter figure is close to an inherent limit, which is the mobility limited by acoustical phonon scattering. Thus, there is little hope of future improvements. The increase of the top mobilities is largely due to improved control of the growth conditions of Molecular Beam Epitaxy. The achievement of high mobilities requires :

i) a fairly low concentration of residual charged impurities in the channel and in the adjacent spacer layer. In the Molecular Beam Epitaxy growth of GaAs epilayers,

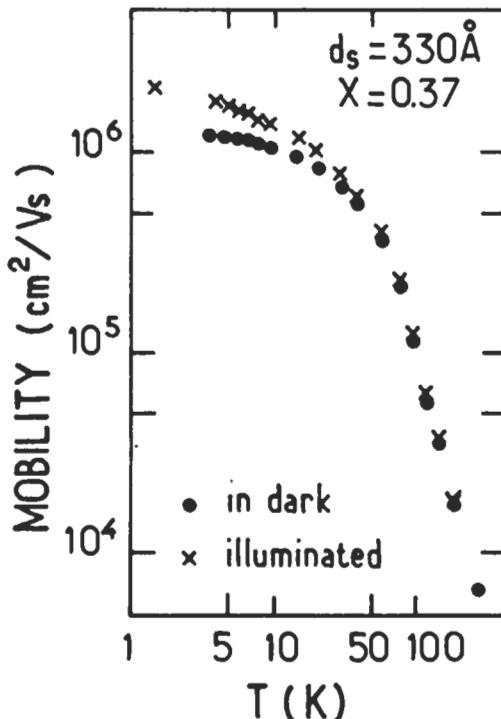


Fig. 10. — Temperature dependence of the electron mobility in a high quality Ga(Al)As-GaAs heterostructure. After reference [29a].

one often gets residual p-doping, possibility due to carbon contamination. Concentrations of residual acceptors $N_A^- \sim 10^{14} \text{ cm}^{-3}$ are often published.

ii) a thick spacer layer ($w \geq 200 \text{ \AA}$) to widely separate the two-dimensional electron gas from its parent impurities and thus diminish the scattering by remote impurities. This, in turn, implies (see the previous chapter) that high mobility heterojunctions display low carrier concentrations ($n_e \sim 2 \times 10^{11} \text{ cm}^{-2}$), i.e. that their electrical conductivities are not very large.

It may be remarked that condition ii) is far from what is required for the successful room-temperature operation of Two-Dimensional-Field-Effect-Transistors (TEG-FET). These devices consist of a modulation-doped heterojunction topped by a gate in order to achieve an external control of the charge density in the channel. Because one wants low Joule losses under operating conditions, and since room-temperatures mobilities are essentially controlled by phonon scattering which is insensitive to w , people have searched for the highest channel conductance. They have thus used thin spacer layers ($w \sim 50 \text{ \AA}$) to get a substantial charge transfer.

Up to this point, the reader might think that the low temperature mobilities in $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterostructures are more or less understood. This is probably true for the "normal" heterojunctions, that is to say, those in which the growth sequence ends up with the ternary layer grown on top of the GaAs layer. The "inverted" heterojunctions, i.e. those obtained by evaporating i) a thick undoped $\text{Ga}_{1-x}\text{Al}_x\text{As}$ layer on top of a semi-insulating GaAs substrate, ii) a doped $\text{Ga}_{1-x}\text{Al}_x\text{As}$ part, iii) a spacer layer and finally iv) the GaAs layer, have been known for a long time to be of poor quality. The charge transfer was somewhat inefficient and channel mobility could barely exceed $\sim 4 \times 10^4 \text{ cm}^2/\text{Vs}$ at low temperatures. The exact physical reason for such a discrepancy between "normal" and "inverted" heterojunctions is still rather obscure. Difficulty in growing high quality thick $\text{Ga}_{1-x}\text{Al}_x\text{As}$ layers, the surface roughness of $\text{Ga}_{1-x}\text{Al}_x\text{As}$, the impurity segregation in GaAs when the Al flux is terminated, may be some of the reasons for the failure to successfully grow "inverted" heterostructures. The reported low mobility figures call for the existence of numerous charged defects located in (or near) the two-dimensional electron gas. Recently, a group at Illinois University (U.S.A.) succeeded in greatly improving the quality of the "inverted" heterojunctions. Drummond *et al.* [27], by purposefully growing a thin $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ superlattice prior to the growth of the main GaAs layer, reported low temperature mobilities of $\sim 2 \times 10^5 \text{ cm}^2/\text{Vs}$ in an inverted $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ heterojunction. The beneficial action of the thin superlattice may involve trapping the impurities which float on the $\text{Ga}_{1-x}\text{Al}_x\text{As}$ growing front, preventing their incorporation in the GaAs channel. The thin superlattice may also relieve misfit strains and improve the interface morphology. In any event, the adjunction of a thin superlattice prior to the GaAs growth has been a major breakthrough in the realization of good "inverted" interfaces. Nowadays it is currently used in many laboratories.

The difficulties connected with the "inverted" $\text{Ga}_{1-x}\text{Al}_x\text{As}$ interface are probably the reason for the poor mobilities obtained in modulation-doped quantum wells or multiple quantum wells. In fact the low temperature mobilities seldom exceed

$10^5 \text{ cm}^2/\text{Vs}$ in these heterostructures and show a systematic drop as quantum well thicknesses decrease. This explains why little effort has been devoted to the modulation-doped quantum wells although from the device (TEGFET) point of view, they should be quite interesting due to their inherent capability for storing more charges than single heterojunctions [28]. With the recent improvement in the inverted heterojunctions, one hopes that the quality of modulation-doped quantum wells will significantly increase.

II.2 ALLOY SCATTERING. — When a semiconductor is an alloy, e.g. $\text{Ga}_{1-x}\text{Al}_x\text{As}$, one of the constituting atoms (As) occupies the sites of one of the face-centred cubic lattices of the zinc blende crystal. The other kind of atoms (Ga or Al) are randomly distributed on the sites of the second face-centred cubic lattice. Thus, although the underlying lattice is periodic, the crystal potential is actually non-periodic. In principle, there are no Bloch waves, dispersion relations etc... The simplest way to overcome this difficulty is to use the virtual crystal approximation. This amounts to replacing the periodic potential due to Ga or Al by a periodic one which is produced by an "average" atom. For an alloy $\text{A}_x\text{B}_{1-x}\text{C}$ the actual potential is replaced by $xV_A + (1-x)V_B + V_C$ which is periodic. With this periodic potential, one may define Bloch waves, gaps... The difference between the actual and the virtual crystal potentials scatters the Bloch waves, which gives rise to band tailing in the energy spectrum and limits the electrical conductivity (alloy scattering). This very naive approach works rather well in many cases. It may be improved by using, e.g., the coherent potential approximation which is more sophisticated (see e.g. reference [30]). Here we shall keep the virtual crystal description.

The physical characteristic of alloy disorder is that it is produced by a sum of *short range* scatterers. In fact, since the A and B atoms are isovalent, the scattering potential is essentially effective within the unit cell of an A (or B) atom. This means that on the scale of envelope functions, slowly varying on the atomic scale, the scattering potential is a sum of delta-like potentials.

With respect to the alloy scattering, one may classify the heterojunctions into three categories :

- i) the electron gas is essentially localized in the binary material and faintly experiences the alloy scattering in the ternary barrier ($\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$)
- ii) the electron gas is essentially localized in the ternary alloy and weakly penetrates the barrier ($\text{InP}-\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$)
- iii) both well-acting and barrier-acting materials are alloys ($\text{Al}_{1-x}\text{In}_x\text{As}-\text{Ga}_{1-x}\text{In}_x\text{As}$).

It is clear that the alloy scattering plays a negligible part in materials i) but may be important in materials ii) and iii). In fact, the low temperature top mobilities of the $\text{InP}-\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ or $\text{Al}_{1-x}\text{In}_x\text{As}-\text{Ga}_{1-x}\text{In}_x\text{As}$ modulation-doped heterojunctions are one order of magnitude ($\mu_{\text{top}} \sim 1.5 \times 10^5 \text{ cm}^2/\text{Vs}$) smaller than the $\text{Al}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ top ones, despite the lighter conduction band effective mass of $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ compared to that of GaAs.

Let us now proceed with more detailed calculations. In the ternary material the actual potential is :

$$V(\mathbf{r}) = \sum_{\mathbf{R}_A} V_A(\mathbf{r} - \mathbf{R}_A) + \sum_{\mathbf{R}_B} V_B(\mathbf{r} - \mathbf{R}_B) + \sum_{\mathbf{R}_C} V_C(\mathbf{r} - \mathbf{R}_C) \quad (99)$$

We split $V(\mathbf{r})$ into a virtual crystal component $W(\mathbf{r})$ and a fluctuating potential $V(\mathbf{r}) - W(\mathbf{r})$:

$$W(\mathbf{r}) = \sum_{\mathbf{R}_i = \mathbf{R}_A, \mathbf{R}_B} x V_A(\mathbf{r} - \mathbf{R}_i) + (1 - x) V_B(\mathbf{r} - \mathbf{R}_i) + \sum_{\mathbf{R}_C} V_C(\mathbf{r} - \mathbf{R}_C) \quad (100)$$

$$V(\mathbf{r}) - W(\mathbf{r}) = \sum_{\mathbf{R}_A} (1 - x) \delta V(\mathbf{r} - \mathbf{R}_A) + \sum_{\mathbf{R}_B} -x \delta V(\mathbf{r} - \mathbf{R}_B) \quad (101)$$

where :

$$\delta V(\mathbf{r}) = V_A(\mathbf{r}) - V_B(\mathbf{r}) \quad (102)$$

The fluctuating potential is the sum of strongly localized perturbations which have different strengths and opposite signs on the randomly distributed sites of a face-centred cubic lattice. If one performs an average of $V(\mathbf{r}) - W(\mathbf{r})$ on a volume which is large compared with a unit cell, one obtains zero.

Let us now evaluate the intraband matrix element of $V(\mathbf{r}) - W(\mathbf{r})$ between two bulk states. For simplicity, we assume that the bulk band under consideration is non degenerate and that for relevant \mathbf{k} , $\mathbf{k} + \mathbf{q}$ wavectors one can write

$$\langle \mathbf{r} | c\mathbf{k} \rangle = \Omega^{-1/2} \exp(i\mathbf{k}\cdot\mathbf{r}) u_{co}(\mathbf{r}) \quad (103)$$

we obtain

$$\begin{aligned} \langle c\mathbf{k} | V(\mathbf{r}) - W(\mathbf{r}) | c\mathbf{k} + \mathbf{q} \rangle &= \Omega_0 \delta V \times \frac{1}{\Omega} \left[\sum_{\mathbf{R}_A} (1 - x) \exp(i\mathbf{q}\cdot\mathbf{R}_A) - \right. \\ &\quad \left. \sum_{\mathbf{R}_B} x \exp(i\mathbf{q}\cdot\mathbf{R}_B) \right] \end{aligned} \quad (104)$$

where Ω_0 is the volume of the unit cell and

$$\delta V = \frac{1}{\Omega_0} \int_{\Omega_0} \delta V(\mathbf{r}) u_{co}^*(\mathbf{r}) u_{co}(\mathbf{r}) d^3r \quad (105)$$

By inspecting equation (105) we see that $V(\mathbf{r}) - W(\mathbf{r})$ can be replaced by the effective operator

$$\delta V_{\text{eff}}(\mathbf{r}) = \Omega_0 \delta V \left[\sum_{\mathbf{R}_A} (1 - x) \delta(\mathbf{r} - \mathbf{R}_A) - x \sum_{\mathbf{R}_B} \delta(\mathbf{r} - \mathbf{R}_B) \right] \quad (106)$$

which acts in the plane waves envelope functions $\Omega^{-1/2} \exp(i\mathbf{k}\mathbf{r})$. Notice that δV , $\delta V_{\text{eff}}(\mathbf{r})$ both have the dimension of an energy, the Ω_0 term cancelling the (volume) $^{-1}$ term implicit in the delta functions.

We are now in a position to apply the previous considerations to an heterojunction whose envelope eigenstates can be written as

$$\langle \mathbf{r} | n\mathbf{k}_\perp \rangle = \frac{1}{\sqrt{S}} \chi_n(z) \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp) \quad (107)$$

where \mathbf{k}_\perp , \mathbf{r}_\perp are two-dimensional vectors. To evaluate the velocity relaxation time in the Electric Quantum Limit we calculate

$$\begin{aligned} \langle 1\mathbf{k}_\perp | \delta V_{\text{eff}}(\mathbf{r}) | 1\mathbf{k}_\perp + \mathbf{q}_\perp \rangle = \Omega_0 \frac{\delta V}{S} & \left\{ \sum_{\mathbf{R}_A} (1-x) \chi_1^2(z_A) \exp(i\mathbf{q}_\perp \cdot \mathbf{R}_{A\perp}) - \right. \\ & \left. - \sum_{\mathbf{R}_B} x \chi_1^2(z_B) \exp(i\mathbf{q}_\perp \cdot \mathbf{R}_{B\perp}) \right\} \end{aligned} \quad (108)$$

We need to calculate the average of the square modulus of this matrix element. Each of the \mathbf{R}_A or \mathbf{R}_B sites can be found in a volume $\omega = \mathcal{L}S$ (where \mathcal{L} is the macroscopic length of the ternary material) and are assumed to be decorrelated from the other sites. Each site has a probability x of being occupied by an A atom and a probability $1-x$ of being occupied by a B atom. Then we obtain

$$\langle | \langle 1\mathbf{k}_\perp | \delta V_{\text{eff}}(\mathbf{r}) | 1\mathbf{k}_\perp + \mathbf{q}_\perp \rangle |^2 \rangle_{\text{average}} = \frac{\Omega_0}{S} x(1-x) [\delta V]^2 \int_{\mathcal{L}} dz \chi_1^4(z) \quad (109)$$

The averaged scattering matrix element is proportional to $[\delta V]^2$, as expected from a Born-type calculation. Interestingly it is q_\perp -independent. This is a result of the short-range nature of $\delta V(\mathbf{r})$. A delta-like potential (or a sum of uncorrelated delta-like potentials) can only scatter the S-like part of the plane waves, all the other parts vanishing at the scattering sites. Finally, the averaged scattering matrix element is proportional to the integral of $\chi_1^4(z)$ over the part of the heterostructure where the alloy scattering takes place. This integral leads to a strong asymmetry between the alloy scattering which occurs in $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ and $\text{InP-Ga}_{0.47}\text{In}_{0.53}\text{As}$ [31, 32]. By inserting equation (109) into equation (37) and by neglecting the screening effects (the scatterers are short-ranged) we obtain

$$\mu_{\text{alloy}} = \frac{e\tau_{\text{alloy}}(\varepsilon_F)}{m_1}; \quad \frac{1}{\tau_{\text{alloy}}(\varepsilon_F)} = m_1 \frac{\Omega_0}{\hbar^3} [\delta V]^2 x(1-x) \int_{\mathcal{L}} dz \chi_1^4(z) \quad (110)$$

Since the averaged scattering matrix element is q_\perp -independent, the velocity relaxation time is exactly proportional to the two-dimensional density of states. It may also be remarked that $\tau_{\text{alloy}}(\varepsilon_F)$ coincides with the lifetime of the $|1\mathbf{k}_\perp\rangle$ level, a peculiarity of short range potentials. Finally, τ^{-1} is proportional to $x(1-x)$, a rule first established by Nordheim for bulk metallic solid solutions [43].

The only difficulties consist in evaluating $[\delta V]$ and the integral which appears in equation (110).

As for $[\delta V]$, it should be close to the conduction band offset between AlAs and GaAs in the $\text{Ga}_{1-x}\text{Al}_x\text{As}$ -GaAs heterostructures, i.e. $[\delta V] \sim 970 \text{ meV}$ (respectively $[\delta V] \sim 1370 \text{ meV}$), depending on whether the conduction band offset is 60 % (respectively 85 %) of the bandgap difference between GaAs and $\text{Ga}_{1-x}\text{Al}_x\text{As}$.

Similar estimates are harder in $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ because GaAs and InAs are strongly lattice-mismatched. Extensive mobility data have been obtained in bulk $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ alloys and a value $[\delta V] \sim 0.6 \text{ eV}$ has been extracted from these data [33].

The dependence of μ_{alloy} upon the two-dimensional electron concentration n_e is governed by the integral which appears in equation (110). For $\text{InP-Ga}_{0.47}\text{In}_{0.53}\text{As}$ we saw in the previous chapter that the E_1 state was heavily localized in the ternary material. Thus, in order to roughly evaluate μ_{alloy} , a Fang Howard $\chi_1(z)$ wavefunction may be used :

$$\begin{cases} \chi_1(z) = \sqrt{\frac{b^3}{2}} z \exp\left(-\frac{bz}{2}\right) & z \geq 0 \\ \chi_1(z) = 0 & z \leq 0 \end{cases} \quad (111)$$

and we obtain

$$\mu_{\text{alloy}} = \frac{e}{(m_1)^2} \frac{\hbar^3}{\Omega_0 [\delta V]^2 x(1-x)} \frac{16}{3b} \quad (112)$$

When $n_e \gg N_{\text{dep}}$, $b \sim n_e^{1/3}$ and therefore μ_{alloy} decreases like $n_e^{1/3}$ when n_e increases. Such behaviour contrasts with the impurity scattering limited mobility where the larger n_e , the larger the Fermi velocity and thus the larger μ (see Eq. (97)). In the present case another mechanism is operative : the larger n_e , the more concentrated the $\chi_1(z)$ wavefunction and thus the more sensitive becomes the carrier to alloy fluctuations which occur at a fixed scale (the unit cell Ω_0).

More quantitative results can be obtained by using a modified Fang Howard wave function. These are shown in figure 11. We see that the alloy scattering is rather effective, limiting the mobility to less than $2 \times 10^5 \text{ cm}^2/\text{Vs}$ when n_e is larger than $3 \times 10^{11} \text{ cm}^{-2}$.

In $\text{Ga}_{1-x}\text{Al}_x\text{As}$ -GaAs heterojunctions we need to use a modified Fang Howard wavefunction to evaluate μ_{alloy} since the alloy scattering takes place in the barrier. With

$$\chi_1(z) = N \exp \kappa_b \frac{z}{2} \quad z \leq 0 ; \quad \kappa_b = 2 \left[\frac{2m_1}{\hbar^2} (V_b - E_1) \right]^{1/2} \quad (113)$$

we obtain :

$$\mu_{\text{alloy}} = \frac{e \hbar^3}{(m_1)^2 \Omega_0 [\delta V]^2 x(1-x)} \times \frac{2}{\kappa_b P_b^2} \quad (114)$$

where P_b is the integrated probability of finding the electron in the barrier. In the previous chapter, we saw that $P_b \sim \text{few \%}$. This explains why for most

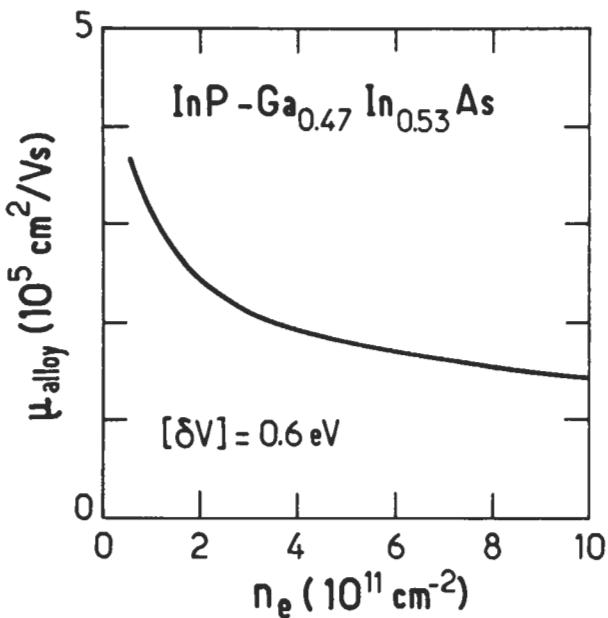


Fig. 11. — The calculated alloy scattering-limited mobility of the quasi bi-dimensional electron gas in $\text{InP-Ga}_{0.47}\text{In}_{0.53}\text{As}$ heterojunctions is plotted *versus* n_e the areal concentration of transferred charges. After reference [32].

$\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterojunctions, the alloy scattering is unimportant ($\mu_{\text{alloy}} > 10^7 \text{ cm}^2/\text{Vs}$) [34].

II.3 INTERFACE ROUGHNESS SCATTERING. — The interface separating two even lattice-matched semiconductors A and B is never perfect. Some A material protrudes into the B material and *vice versa*. In chapter IV, we have discussed the bound electron or exciton states supported by such defects. Here, we are concerned by their effects on the low-temperature mobility of charge carriers. Very little is known on the microscopic structure of such defects and we are forced to use very simplified descriptions. A defect will be characterized by its depth h along the growth axis and by its lateral (in-plane) extensions l_x, l_y . The ideal interface will be taken at the $z = 0$ plane. A convenient model of a repulsive (i.e. barrier protruding in the well) defect is given by

$$V_{\text{def}}^{\text{bare}}(\mathbf{r}, \boldsymbol{\rho}_i) = V_b Y[z(h - z)] Y\left[\frac{l_x^2}{4} - (x - x_i)^2\right] Y\left[\frac{l_y^2}{4} - (y - y_i)^2\right] \quad (115)$$

for a defect centred at $\boldsymbol{\rho}_i = (x_i - y_i)$ in the layer plane. V_b is the barrier height and $Y(x)$ the step function. In equation (115) we have assumed that the defects have sharp boundaries. This is an approximation which can be easily relaxed, if necessary.

At $T = 0$ K, and in the Electric Quantum Limit the velocity relaxation time τ at the Fermi energy ($k_{\perp} = k_F$) is given by :

$$\frac{\hbar}{2\pi \tau_{\text{rough}}(\varepsilon_F)} = \sum_{k_{\perp}'} \delta[\varepsilon_1(k_{\perp}') - \varepsilon_F] \frac{(1 - \cos \theta)}{\varepsilon^2(q_{\perp})} \times \left\langle \left| \langle 1k_{\perp} | \sum_{\rho_i} V_{\text{def}}^{\text{bare}} | 1, k_{\perp} + q_{\perp} \rangle \right|^2 \right\rangle_{\text{average}} \quad (116)$$

where $\varepsilon(q_{\perp})$ has been defined in equation (69), θ in figure 1 and the average has to be performed over the defect positions ρ_i as well as on their sizes.

The matrix element which appears in equation (116) is readily evaluated using equations (3, 115) :

$$\langle 1k_{\perp} | V_{\text{def}}^{\text{bare}}(\mathbf{r}, \rho_i) | 1k_{\perp} + q_{\perp} \rangle = \frac{4 V_b}{S} \exp[i q_{\perp} \cdot \rho_i] \frac{\sin\left(\frac{1}{2} q_x l_x\right)}{q_x} \times \frac{\sin\left(\frac{1}{2} q_y l_y\right)}{q_y} \int_0^h \chi_1^2(z) dz \quad (117)$$

If the defect depth h is small compared to the characteristic length of variation of the carrier wavefunction near $z = 0$, equation (117) can be simplified into :

$$\langle 1k_{\perp} | V_{\text{def}}^{\text{bare}}(\mathbf{r}, \rho_i) | 1k_{\perp} + q_{\perp} \rangle = \frac{4 V_b h}{S} \exp[i q_{\perp} \cdot \rho_i] \frac{\sin\left(\frac{1}{2} q_x l_x\right)}{q_x} \times \frac{\sin\left(\frac{1}{2} q_y l_y\right)}{q_y} \chi_1^2(0) \quad (118)$$

This simplification allows us to simultaneously treat the repulsive and attractive defects, the latter case corresponding to negative h 's in equation (118). In "good" heterojunctions such as the $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$ ones, the interface extends over \sim one monolayer, i.e. 2.83 Å. This is smaller than the characteristic length of the variation of the envelope function near $z = 0$ ($\sim \left[\frac{2m_1(V_b - E_1)}{\hbar^2} \right]^{-1/2}$) and thus the approximation leading from equation (117) to equation (118) is not too unreasonable.

The average over the defect positions ρ_i can be easily performed. One obtains :

$$\left\langle \left| \langle 1k_{\perp} | \sum_{\rho_i} V_{\text{def}}^{\text{bare}}(\mathbf{r}, \rho_i) | 1k_{\perp} + q_{\perp} \rangle \right|^2 \right\rangle_{\text{average}} = \frac{16h^2 V_b^2}{S} N_{\text{def}} \chi_1^4(0) \left[\frac{\sin\left(\frac{1}{2} q_x l_x\right) \sin\left(\frac{1}{2} q_y l_y\right)}{q_x q_y} \right]^2 \quad (119)$$

where N_{def} is the areal concentration of the defects. Finally, performing the integration over q_{\perp} , we are left with :

$$\frac{\hbar}{\tau_{\text{roug}}(\varepsilon_F)} = 16 \hbar^2 V_b^2 N_{\text{def}} \chi_1^4(0) \frac{m_1}{\pi \hbar^2} \times \int_0^\pi \frac{d\theta}{\varepsilon^2 \left(2k_F \sin \frac{\theta}{2} \right)} \times \frac{\sin^2 \left[\frac{k_F l_x}{2} (\cos \theta - 1) \right]}{k_F^4 \sin^2 \theta (1 - \cos \theta)} \times \sin^2 \left[\frac{k_F l_y}{2} \sin \theta \right] \quad (120)$$

where k_F is the Fermi wavevector ($k_F = (2\pi n_e)^{1/2}$). An average over the statistical distribution of defect sizes h, l_x, l_y remains to be performed.

Since we have used the Born approximation and assumed that the scatterers are independent, the condition $N_{\text{def}} l_x l_y \ll 1$ should be fulfilled to ensure incoherence between consecutive scattering events.

The interface roughness scattering displays several trends already found in the alloy scattering. Firstly, the term $\chi_1^4(0)$ in equation (120) recalls the integral of $\chi_1^4(z)$ which appears in the alloy scattering limited mobility (Eq. (110)). Both arise from the short range nature of the scattering potentials : the alloy scattering involves $\delta(\mathbf{r} - \mathbf{r}_i)$ functions whereas the interface roughness has been assumed to be a localized perturbation near the interface. Secondly, one may check from equation (120) that, in single heterojunction, $\tau_{\text{roug}}(\varepsilon_F)$ decreases with increasing areal concentration of carriers n_e , at least because $\chi_1^4(0)$ increases with n_e . A similar decrease in the velocity relaxation time was found in the case of alloy scattering. In both situations, the scatterer ranges are fixed but the $\chi_1(z)$ wavefunction becomes more and more localized with increasing n_e .

It may be noticed in equation (120) that, despite the explicit V_b^2 term, τ_{roug}^{-1} is actually almost independent of V_b . This is because $\chi_1^4(0)$ is proportional to V_b^{-2} if $V_b - E_1$ is large (see Appendix A of the previous chapter).

All these features are in agreement with Ando's calculations who used a more refined treatment of the interface roughness scattering [22].

Quantitatively, the interface roughness scattering is difficult to assess since the distribution of sizes, h, l_x, l_y is unknown. Using parameters adapted to $\text{SiO}_2\text{-Si}$ interfaces, which should over-estimate the actual roughness in good $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterostructures, Ando [22] showed that the interface roughness scattering is not very important in these heterostructures : $\mu_{\text{roug}} = e\tau_{\text{roug}}/m_1$ remains larger than $5 \times 10^5 \text{ cm}^2/\text{Vs}$ if $n_e < 10^{12} \text{ cm}^{-2}$.

III. Vertical transport.

The idea of vertical transport, i.e. transport occurring along the growth axis of a microstructure, was the reason why Esaki and Tsu put forward the concept of

semiconductor superlattices [5]. Let us consider a one-dimensional superlattice and assume that the carrier dispersion relations can be written :

$$\varepsilon_n(q, \mathbf{k}_\perp) = \frac{\hbar^2 k_\perp^2}{2m_1} + \varepsilon_n(q) \quad (121)$$

with :

$$\varepsilon_n\left(q + \frac{2\pi}{d}\right) = \varepsilon_n(q) \quad (122)$$

A static electric field \mathbf{F} , of small intensity, is applied parallel to the growth axis. If we assume that the field does not induce intersubband transitions and thus restrict our attention to the lowest miniband, the semi-classical equations of motion become exact. They are :

$$\hbar \frac{d\mathbf{k}}{dt} = -e\mathbf{F} \quad (123)$$

$$\mathbf{v} = \frac{1}{\hbar} \frac{\partial \varepsilon_1(\mathbf{k})}{\partial \mathbf{k}} \quad (124)$$

The x, y motions are readily calculated :

$$\mathbf{k}_\perp = \mathbf{k}_0; \quad \mathbf{v}_\perp = \frac{\hbar \mathbf{k}_0}{m_1} \quad (125)$$

The z motion is more interesting, for it leads to an oscillatory behaviour of the velocity component v_z upon time (Bloch oscillator), despite the fact that the electric field is *static* [17]

$$q = q_0 - \frac{eFt}{\hbar} \quad (126)$$

$$v_z(t) = -\frac{1}{eF} \frac{\partial \varepsilon_1}{\partial t} \left[q_0 - \frac{eFt}{\hbar}, \mathbf{k}_0 \right] \quad (127)$$

The velocity is periodic upon time with a frequency

$$\nu = \frac{eFd}{2\pi\hbar} \quad (128)$$

For a field strength of 10^4 V/cm and a superlattice period of 200 Å, one obtains $\nu \sim 5 \times 10^{12}$ Hz. It is quite remarkable that the v_z periodicity upon time arises from equation (122), i.e. from the spatial periodicity of the superlattice. The realization of a Bloch oscillator would considerably broaden the available spectrum of millimeter devices. However, no Bloch oscillators have yet been demonstrated, although the appearance of a negative differential resistance in GaAs-Ga(Al)As superlattices has been reported [35].

The point is that to exhibit oscillations the carrier must not experience any scattering event during the time elapse $(2\pi\nu)^{-1}$. Thus, the scattering time

τ_s should be such that $2\pi\nu\tau_s > 1$. The mobility relaxation time τ_μ is an upper bound to τ_s . Bloch oscillations would thus become unobservable if $2\pi\nu\tau_\mu \leq 1$. For GaAs-Ga(Al)As superlattices this means a mobility smaller or equal to $\sim 8 \times 10^2 \text{ cm}^2/\text{Vs}$. This mobility, although relatively low, is still 80 times larger than the one recently reported by Palmier *et al.* [36] (see Fig. 12).

Another physical reason which may lead to a failure of the Bloch oscillator is the field-induced intersubband scattering which was explicitly neglected in equations (123, 124). To our knowledge, this effect has not yet been investigated in great detail in semiconductor superlattices.

The vertical transport in GaAs-Ga(Al)As superlattices has been interpreted [36] in terms of phonon-assisted hopping between the localized levels in one well to those in the neighbouring one. The occurrence of such a hopping type of transport (instead of the standard conduction in minibands) has been proved by two observations.

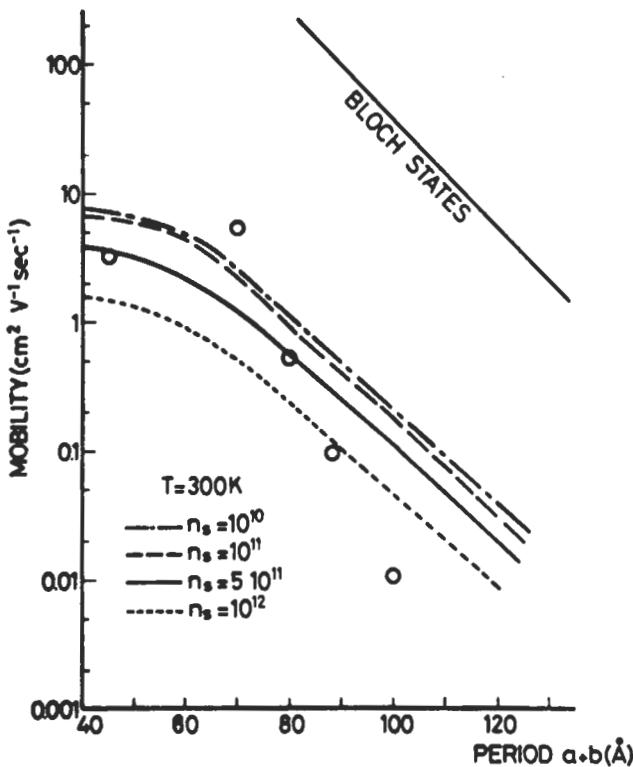


Fig. 12. — Calculated effective perpendicular mobility as a function of the superlattice period in Ga(Al)As-GaAs superlattices. Two conduction mechanisms have been investigated: phonon-limited mobility of Bloch states or phonon-assisted hopping. The various symbols correspond to different samples and different techniques of extracting the mobility. After reference [36].

- i) The order of magnitude of the experimental mobility which was found to be much lower than predicted by assuming a conduction through the delocalized Bloch states.
- ii) The experimental mobility increase with temperature which dismisses any interpretation of the data in terms of *phonon-limited* relaxation time but strongly favours an interpretation in terms of *phonon-assisted* tunneling between wells. If the conduction were to occur *via* the extended Bloch states, the electron-phonon scattering would increase the relaxation frequency with increasing T as more phonons are available at higher temperatures. In turn, the mobility μ , which is proportional to τ_μ would decrease with increasing T , as is invariably observed in the in-plane transport. The hopping type of transport describes the electrical conductivity as arising from phonon-assisted hopping between wells. Then, by the Fermi Golden Rule, the conductivity is proportional to the scattering frequency and increases with increasing temperature.

In other GaAs-Ga(Al)As superlattices Chomette *et al.* have reported optical evidence of vertical transport at low temperatures [37]. They purposedly introduced GaAs enlarged wells into short period GaAs-Ga(Al)As superlattices and studied at low temperatures the excitonic photoluminescence which arises from both the superlattice and the enlarged wells. Some of their results are presented in figure 13. If the coupling between wells were negligible, the intensities of both types of luminiscence would roughly be in proportion to the number of enlarged wells to superlattices periods. These results are far from what was measured (see Fig. 13). The observed data were explained by means of miniband conduction. Rate equations involving the trapping time from the superlattice to the enlarged wells and the carrier lifetimes in the superlattice and the enlarged wells have been used to calculate the photoluminescence intensities. A reasonable agreement was found between the calculated trapping time and the one deduced from the experiments.

IV. Resonant tunnelling.

Another vertical transport mechanism, which is physically different from both defect-limited conduction in the miniband of a superlattice and the phonon-assisted hopping between wells, is the transport *via* resonant tunnelling. As is the case of the Bloch oscillator, there is no contribution from perturbative agents (impurities, defects...) at the lowest order in the resonant tunnelling. The latter is a quantum effect of intrinsic origin which should be best observed in perfect heterostructures. The physics of resonant tunnelling, is still a matter of active research, especially in its time-dependent aspect [38].

Here we only intend to summarize some basic considerations.

First we shall assume that the carrier motions parallel and perpendicular to the growth axis are decoupled. Thus the total carrier energy can be split into :

$$\varepsilon = \varepsilon_{\parallel} + \frac{\hbar^2 k_{\perp}^2}{2m_1} \quad (129)$$

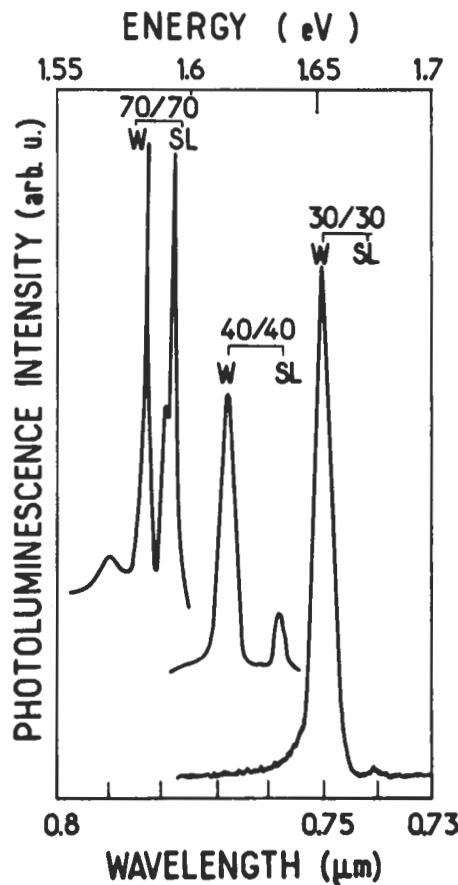


Fig. 13. — Comparison of the luminescence spectra obtained under the same excitation condition in three different Ga(Al)As-GaAs superlattices containing enlarged wells. The peaks labelled W(SL) are due to the enlarged wells (the superlattices). $T = 1.7$ K. After reference [37].

where ϵ_{\parallel} is the carrier energy associated with the longitudinal motion. Under this assumption, the tunnelling problem becomes quasi unidimensional. Let us consider the simplest device structure aimed at exploiting the resonant tunnelling effect. It consists of a quantum well, in principle of arbitrary shape (but in the following assumed to be rectangular), flanked by two finite barriers (see Fig. 14). Electrical contacts are placed on each side of the structure.

For $\epsilon_{\parallel} \leq V_b$, the transmission coefficient of a single barrier is an exponentially decaying function of the barrier thickness b and of the barrier height V_b :

$$T_{(1)}(\epsilon_{\parallel}) = |t_{(1)}(\epsilon_{\parallel})|^2 = \left[1 + \frac{1}{4} \left(\xi + \frac{1}{\xi} \right)^2 \sin h^2(\kappa_b b) \right]^{-1}; \quad \epsilon_{\parallel} \leq V_b \quad (130)$$

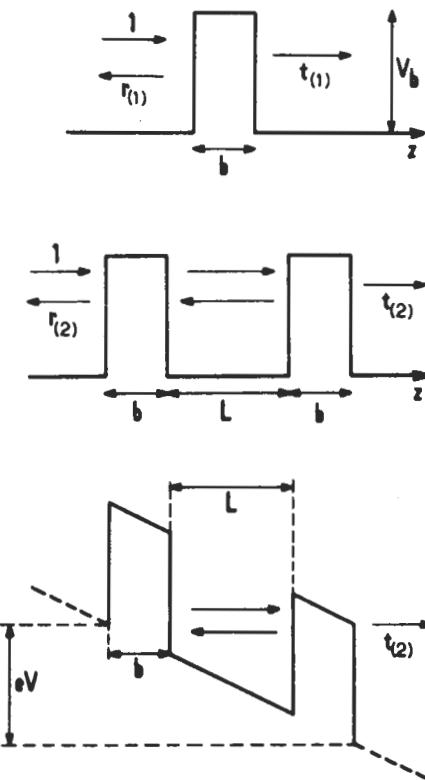


Fig. 14. — Conduction band profile of an unbiased single (upper panel) and double (middle panel) barrier structure. When a voltage V is applied between the left and right hand side electrodes, the double barrier structure is no longer symmetrical (lower panel).

where :

$$\xi = \frac{k_w}{m_w} \times \frac{m_b}{\kappa_b} \quad (131)$$

and :

$$k_w = \sqrt{\frac{2m_w}{\hbar^2} \epsilon_{\parallel}} ; \quad \kappa_b = \sqrt{\frac{2m_b}{\hbar^2} (V_b - \epsilon_{\parallel})} \quad (132)$$

In the case of a double symmetrical barrier, the transmission $T_{(2)}(\epsilon_{\parallel})$ will be even lower for most ϵ_{\parallel} . However, $T_{(2)}(\epsilon_{\parallel})$ will reach unity for certain discrete energies $\epsilon_{i\parallel}$ (see Fig. 15). If L denotes the quantum well thickness, these energies are the roots of the equation

$$\cos k_w L - \frac{1}{2} \left[\xi - \frac{1}{\xi} \right] \sin k_w L = 0 \quad (133)$$

i.e. the $\epsilon_{i\parallel}$'s coincide with the bound states of a quantum well of thickness L and barrier height V_b (see chapter I)

For these particular energies, the transmitted and reflected electron beams interfere constructively inside the quantum well and build up a virtual bound state. The state is only virtually bound due to the finite thickness of the cladding barriers. As in our previous discussion (chapter I) of the virtual bound states which occur in the continuum of a single quantum well, the virtual bound states of the double barrier structure can equally be viewed as transmission resonances (see again what happens in a laser). The width $\delta \epsilon_i$ of the i^{th} resonance $\epsilon_{i\parallel}$ are decreasing functions of the barrier thickness b (see Fig. 15). Thus, the contrast between the maximum and minimum values of $T_{(2)}(\epsilon)$ increases with increasing b .

These transmission resonances can be exploited to produce sharp peaks in the current-voltage characteristic of a biased double barrier (see Fig. 14). Whenever the Fermi level ϵ_F of the metallic electrode coincides with one of the $\epsilon_{i\parallel}$, one may expect a current peak. For other bias the current is very small. This implies that, by tuning the bias voltage V between the two electrodes, one finds negative differential resistance regions once a virtual bound state has passed through ϵ_F : a further increase in V leads to a sharp drop in the current.

The exact bias value at which a current peak develops is difficult to assess simply from the calculation of transmission resonances in unbiased double barrier structures.

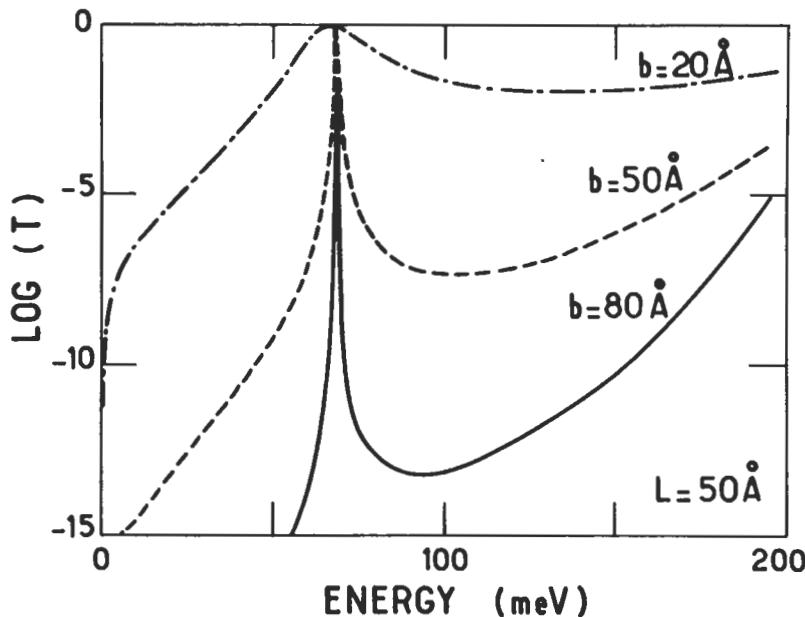


Fig. 15. — The natural logarithm of the transmission of a symmetrical double barrier structure is plotted *versus* the energy ϵ_{\parallel} of the longitudinal motion. $V_b = 0.2 \text{ eV}$; $m_w = 0.07 m_0$; $m_b = 0.088 m_0$; $L = 50 \text{ \AA}$. Three barrier thicknesses have been considered: $b = 20 \text{ \AA}$; 50 \AA ; 80 \AA . The transmission peaks coincide with the only bound state of the isolated well ($E_1 \approx 68 \text{ meV}$).

If the bias is vanishingly small (which seldom happens in practice) the current peaks take place when

$$eV = 2 \epsilon_{i\parallel} \quad (134)$$

However, the biased structure in general

- i) is no longer symmetrical
- ii) can no longer be described in terms of field-independent virtual bound states.

As pointed out by Ricco and Azbel [38] the piling up of charges inside the well under the resonance condition produces a band bending. In turn, this shifts the resonant state energy, making the achievement of resonances a complicated self-consistent, time-dependent problem.

The first clear-cut resonant tunnelling effects in GaAs-Ga(Al)As double barriers were demonstrated by Chang *et al.* (see Fig. 16) in 1974 [39]. Since then, improved MBE growth has made possible the observation of narrower peaks in the current voltage characteristics. Very recently, the resonant tunnelling of holes in high quality AlAs-GaAs-AlAs heterostructures was reported by Mendez *et al.* [40] (see Fig. 17). Although one should expect to see peaks associated with GaAs heavy and light hole virtual bound states, the actual peaks in the current-voltage characteristics did not fit these assumptions. This calls for detailed calculations of the resonant tunnelling of carriers with complicated subband structures.

The resonant tunnelling effects in multiple barriers were calculated by Tsu and Esaki [41]. If N is the number of identical barriers and $(N - 1)$ thus the number of identical wells, the double barrier virtual bound states which fulfil equation (133),

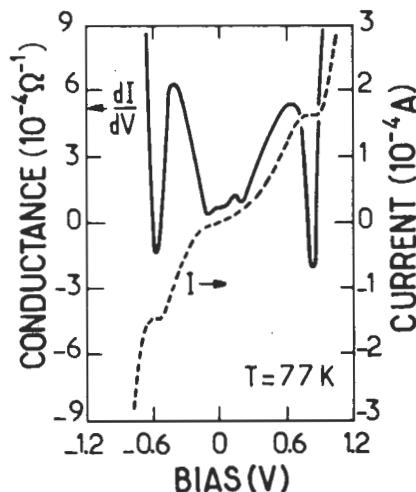


Fig. 16. — Current-voltage and conductance-voltage characteristics of a double barrier structure of GaAs between two $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ barriers sandwiched between n^+ -GaAs regions. $L = 50 \text{ \AA}$; $b = 80 \text{ \AA}$; $T = 77 \text{ K}$. After reference [39].

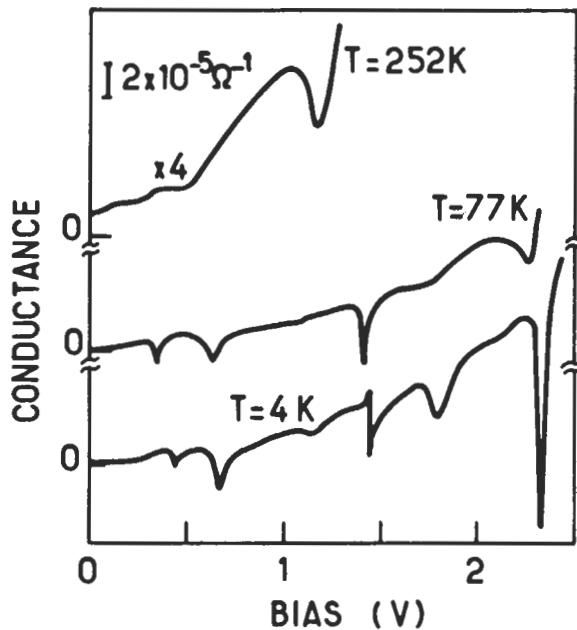


Fig. 17. — Conductance *versus* voltage bias, for representative temperatures, of an undoped AlAs-GaAs-AlAs double barrier structure sandwiched between p^+ -GaAs regions. $L = 50 \text{ \AA}$; $b = 80 \text{ \AA}$. After reference [40].

split into $(N - 1)$ states where the transmission is unity. At the limit $N \rightarrow \infty$, one recovers the superlattice minibands, where the transmission is unity for energies corresponding to the allowed superlattice states and zero for energies corresponding to the superlattice bandgaps.

Appendix.

Dimensional dependence of the screening effects.

In this Appendix we shall evaluate the quantity

$$A(\mathbf{q}) = \frac{2}{\Omega_d} \sum_{\mathbf{k}} \frac{f_1^{(0)}(\mathbf{k} + \mathbf{q}) - f_1^{(0)}(\mathbf{q})}{\frac{\hbar^2}{2m_1} [(\mathbf{k} + \mathbf{q})^2 - \mathbf{k}^2]} \quad (\text{A1})$$

at $T = 0 \text{ K}$. In (A1), Ω_d denotes the d -dimensional volume, i.e. $\Omega_d = L_x L_y L_z$, $L_x L_y$, L_x for $d = 3, 2$ and 1 respectively. The vectors \mathbf{k} and \mathbf{q} are also to be

understood as d -dimensional wavevectors. The summation over \mathbf{k} is converted into an integration following the usual rule

$$\frac{1}{\Omega_d} \sum_{\mathbf{k}} \rightarrow \frac{1}{(2\pi)^d} \int d^d k \quad (A2)$$

where $d = 1, 2, 3$. The results are the following :

$$A_{(1)}(q_z) = -\frac{2m_1}{\pi\hbar^2 k_F} \times \frac{k_F}{q_z} \ln \left| \frac{1 + \frac{q_z}{2k_F}}{1 - \frac{q_z}{2k_F}} \right| ; \quad d = 1 \quad (A3)$$

$$A_{(2)}(q_{\perp}) = -\frac{m_1}{\pi\hbar^2} \times \left\{ 1 - \sqrt{1 - \left(\frac{2k_F}{q_{\perp}} \right)^2} Y[q_{\perp} - 2k_F] \right\} ; \quad d = 2 \quad (A4)$$

$$A_{(3)}(q) = -\frac{m_1 k_F}{\pi^2 \hbar^2} \times \left\{ \frac{1}{2} + \frac{k_F}{2q} \left(1 - \frac{q^2}{4k_F^2} \right) \ln \left| \frac{1 + \frac{q}{2k_F}}{1 - \frac{q}{2k_F}} \right| \right\} ; \quad d = 3 \quad (A5)$$

Each of the A functions can be re-expressed in the form :

$$A_{(d)}(q) = -\rho_{(d)}(\varepsilon_F) f_{(d)}\left[\frac{q}{2k_F}\right] ; \quad d = 1, 2, 3 \quad (A6)$$

where $\rho_{(d)}(\varepsilon_F)$ is the value of the d -dimensional density of states per unit d -dimensional volume at the Fermi energy and $f_{(d)}\left(\frac{q}{2k_F}\right)$ is a dimensionless function

which approaches unity when q tends to zero. The long wavelength limits ($q \rightarrow 0$) of $A_{(d)}$ coincide with the semiclassical screening theory (Debye-Hückel's laws). However, when the wavevector approaches the Fermi diameter $2k_F$, the finite size of the Fermi surface can no longer be neglected and singularities appear in the $A_{(d)}$ functions. The singularities are more and more pronounced by decreasing d (see Fig. 18). On a plot of $f_{(d)}$ versus $q/2k_F$ the $q = 2k_F$ singularity is barely visible if $d = 3$. If $d = 2$, $f_{(2)}(q)$ has a kink at $q = 2k_F$, whereas $f_{(1)}(q)$ logarithmically diverges near $q = 2k_F$.

Notice finally that the semiclassical screening theory, i.e. $A_{(d)}\left(\frac{q}{2k_F}\right) = -\rho_d(\varepsilon_F)$, leads to unphysical results in the limit of zero electron concentration ($k_F = 0$) if $d = 1, 2$. As both $\rho_{(1)}(\varepsilon)$ and $\rho_{(2)}(\varepsilon)$ do not vanish at the subband edges, the semiclassical result would imply that a single electron completely screens an external potential. Such unphysical behaviour is, however, suppressed if the finite size of the Fermi surface is taken into account.

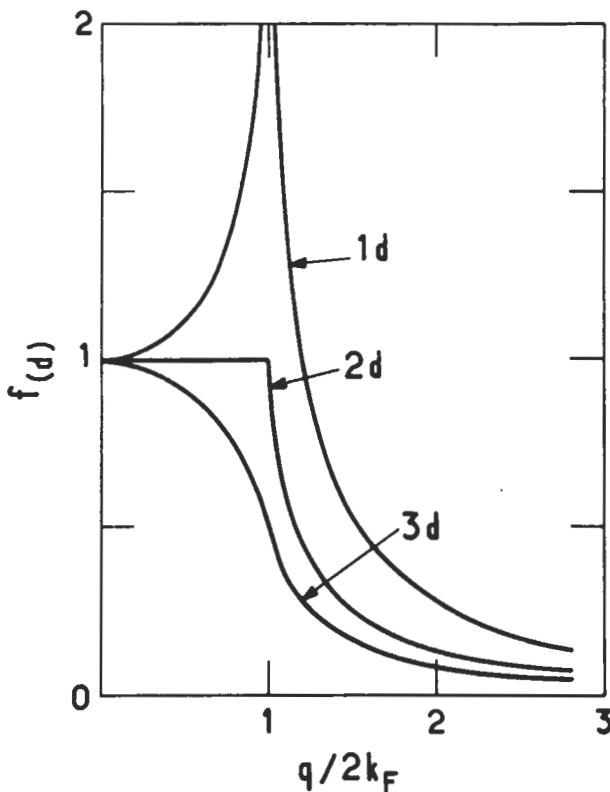


Fig. 18. — The dimensionless functions $f_{(d)}$ are plotted versus $q/2 k_F$ for $d = 1, 2, 3$.

References

- [1] P.J. PRICE, *Surf. Sci.* **113** (1982) 199, and in *Proceedings of the 17th International Conference on the Physics of Semiconductors, San Francisco 1984* edited by J.D. Chadi and W.A. Harrison, Springer Verlag, New York, 1985, p. 433.
- [2] K. HESS, *Appl. Phys. Lett.* **35** (1979) 484.
- [3] J.F. PALMIER and A. CHOMETTE, *J. Phys. France* **43** (1982) 381.
- [4] B. VINTER, *Appl. Phys. Lett.* **45** (1984) 581.
- [5] L. ESAKI and R. TSU, *IBM J. Res. Dev.* **14** (1970) 61.
- [6] W. KOHN and J.M. LUTTINGER, *Phys. Rev.* **108** (1957) 590.
J.M. LUTTINGER and W. KOHN, *Phys. Rev.* **109** (1958) 1892.
- [7] see, e.g., D. CALECKI and J. TAVERNIER, *Introduction aux phénomènes de transport linéaires dans les semiconducteurs*, Masson Paris (1969).
- [8] see, e.g., A. MESSIAH, *Mécanique Quantique*, Dunod Paris (1959).
- [9] see, e.g., T. ANDO, A.B. FOWLER and F. STERN, *Rev. Mod. Phys.* **54** (1982) 437.
- [10] E.D. SIGGIA and P.C. KWOK, *Phys. Rev.* **B2** (1970) 1024.
- [11] S. MORI and T. ANDO, *J. Phys. Soc. Japan* **48** (1980) 865.

- [12] H.L. STORMER, A.C. GOSSARD and W. WIEGMANN, *Solid State Commun.* **41** (1982) 707.
- [13] G. FISHMAN and D. CALECKI, *Phys. Rev.* **B29** (1984) 5778.
- [14] J. LINDHARD, K., *Dan. Vidensk. Selsk. Mat. Fys. Medd.* **28** (1954) 1.
- [15] F. STERN, *Phys. Rev. Lett.* **18** (1967) 546.
- [16] P.J. PRICE, *J. Vac. Sci. Technol.* **19** (1981) 599.
- [17] see, e.g., N.W. ASHCROFT and N.D. MERMIN, *Solid State Physics*, Holt, Rinehart and Winston New York (1976).
- [18] T. ANDO, *J. Phys. Soc. Japan* **51** (1982) 3215.
- [19] R. DINGLE, H.L. STORMER, A.C. GOSSARD and W. WIEGMANN, *Appl. Phys. Lett.* **33** (1978) 665.
- [20] H.L. STORMER, *J. Phys. Soc. Japan Suppl. A* **49** (1980) 1013.
- [21] F. STERN, *Appl. Phys. Lett.* **43** (1983) 974.
- [22] T. ANDO, *J. Phys. Soc. Japan* **51** (1982) 3900.
- [23] S. HIYAMIZU, J. SAITO, K. NANBU, T. ISHIKAWA, *Japan. J. Appl. Phys.* **22** (1983) L 609.
- [24] D.C. TSUI, A.C. GOSSARD, G. KAMINSKI and W. WIEGMANN, *Appl. Phys. Lett.* **39** (1981) 712.
- [25] J.C.M. HWANG, A. KASTALSKY, H.L. STORMER and V.G. KERAMIDAS, *Appl. Phys. Lett.* **44** (1984) 802.
- [26] W.I. WANG, E.E. MENDEZ and F. STERN, *Appl. Phys. Lett.* **45** (1984) 639.
- [27] T.J. DRUMMOND, J. KLEM, D. ARNOLD, R. FISCHER, R.E. THORNE, W.G. LYONS and H. MORKOC, *Appl. Phys. Lett.* **42** (1983) 615.
- [28] K. INOUE, H. SAKAKI and J. YOSHINO, *Japan. J. Appl. Phys.* **23** (1984) L 767.
- [29a] G. WEIMANN and W. SCHLAPP, *Appl. Phys. Lett.* **46** (1985) 411.
- [29b] J.H. ENGLISH, A.C. GOSSARD, H.L. STORMER and K.W. BALDWIN, *Appl. Phys. Lett.* **50** (1987) 1826.
- [30] P. SOVEN, *Phys. Rev.* **156** (1967) 809 ;
B. VELICKY, S. KIRKPATRICK and H. EHRENREICH, *Phys. Rev.* **175** (1968) 747.
- [31] P.K. BASU and B.R. NAG, *Appl. Phys. Lett.* **43** (1983) 689.
- [32] G. BASTARD, *Appl. Phys. Lett.* **43** (1983) 591 ; *Surf. Sci.* **142** (1984) 284.
- [33] J.R. HAYES, A.R. ADAMS and P.D. GREENE in *GaInAsP Alloy Semiconductors*, edited by T.P. Pearsall, Wiley New York (1982) p. 189.
- [34] W. WALUKIEWICZ, H.E. RUDA, J. LAGOWSKI and H.C. GATOS, *Phys. Rev.* **B29** (1984) 4818, and *Phys. Rev.* **B30** (1984) 4571.
- [35] L. ESAKI, L.L. CHANG, W.E. HOWARD and V.L. RIDEOUT in *Proceedings of the 11th International Conference on the Physics of Semiconductors*, Warsaw (PWN-Polish Scientific Warsaw) 1972, p. 431.
- [36] J.F. PALMIER, H. LE PERSON, C. MINOT, A. CHOMETTE, A. REGRENY, D. CALECKI, *Superlattices and Microstructures* **1** (1985) 67.
- [37] A. CHOMETTE, B. DEVEAUD, J.Y. EMERY, A. REGRENY and B. LAMBERT, in *Proceedings of the 17th International Conference on the Physics of Semiconductors*, San Francisco 1984, edited by J.D. Chadi and W.A. Harrison (Springer Verlag New York) 1985, p. 441.
- [38] B. RICCO and M.Y.A. AZBEL, *Phys. Rev.* **B29** (1984) 1970.
See also W.R. FRENSLEY, *J. Vac. Technol.* **B3** (1985) 1261.
- [39] L.L. CHANG, L. ESAKI and R. TSU, *Appl. Phys. Lett.* **24** (1974) 593.
- [40] E.E. MENDEZ, W.I. WANG, B. RICCO and L. ESAKI, *Appl. Phys. Lett.* **47** (1985) 415.
- [41] R. TSU and L. ESAKI, *Appl. Phys. Lett.* **22** (1973) 562.
- [42] J.A. BRUM and G. BASTARD, *Solid State Commun.* **53** (1985) 727.
- [43] see e.g. J.M. ZIMAN *Electrons and Phonons*, Clarendon Press, Oxford (1960) page 247.

CHAPTER VII

Optical properties of quasi bi-dimensional systems

In this chapter the discussion will be centred around a few of the basic optical properties of quasi bi-dimensional semiconductor heterostructures. Due to technological applications (e.g. quantum well lasers, improved photo-detectors) and characterization purposes, the optical properties of heterostructures have been the subject of a considerable body of literature. Many of the reported investigations have however dealt with a limited number of samples and this has often led to unresolved conflicting interpretations.

We shall restrict ourselves to absorption and photoluminescence experiments in quantum wells and superlattices. These are the two main tools used for characterizing the heterolayers, particularly their energy levels. Our aim is to point out the diversity of possible physical situations rather than to attempt a synthetic review of the optical properties of semiconductor heterolayers. In our opinion such a synthesis is impossible at the present time. We have to wait for improvements in the quality of the heterolayers, in the reproducibility of the growth conditions and in theoretical progress before drawing definite conclusions.

We shall first of all give a very simplified model of inter- and intra- subband absorption processes without including the excitonic effects. We shall then describe, again in a simplified fashion, the excitonic absorption in type I quantum wells. As regards photoluminescence, we shall insist on the fact that, despite its apparent simplicity, it is quite tricky to interpret. Photoluminescence may be the worst way of determining energy levels in heterostructures, as it essentially probes the lowest lying of these levels which are often defect-related.

To illustrate our considerations, a few experimental spectra will be presented. Whenever possible, the physical origin of the observed structures will be given.

I. Absorption (one-electron approximation).

An optical absorption experiment consists of measuring the attenuation of a light beam which passes through a sample of thickness t . If the incident electromagnetic wave is monochromatic and propagates perpendicularly to the sample surface (area S), which is supposed to be flat, it can be shown [1,2] that the transmitted intensity is given by :

$$I_t = \frac{I_0(1 - R)^2 \exp(-\alpha t)}{1 - R^2 \exp(-2\alpha t)} \quad (1)$$

where I_0 is the intensity of the incident beam, R the sample reflectance, and α the absorption coefficient at the angular frequency ω of the electromagnetic wave. For weakly absorbing media, R is given by

$$R = R_0 = \left(\frac{n-1}{n+1} \right)^2 \quad (2)$$

where n is the refractive index of the sample. For absorption measurements, which probe the vicinity of the bandgap of the heterostructures, the ω -dependence of R and R_0 in equations (1, 2) can often be neglected.

Generally, the investigated samples, which are epitaxial layers, are very small in thickness ($t \sim 1 - 10 \mu\text{m}$), whereas their area S is of the order of a few mm^2 . Thus, the geometry which is most currently used in characterization experiments corresponds to an electromagnetic wave which propagates parallel to the growth axis of the heterostructure (z axis). Absorption experiments on the edge of the samples, corresponding to an electromagnetic wave which propagates in the layer planes of the heterostructures, are much more difficult to conduct. Notice however that in heterostructure lasers, the wave propagates in the layer plane. This geometry is therefore very important practically.

To calculate the absorption coefficient $\alpha(\omega)$, the transition probability per unit time P for a photon to disappear is first of all evaluated and, as a result, $\alpha(\omega)$ can then be obtained from the knowledge of P .

Let us consider an electromagnetic wave characterized by an angular frequency ω and a propagation vector \mathbf{q} ($\omega = c/nq$). We assume that this monochromatic wave is linearly polarized. The electric field \mathbf{F} of this wave can thus be written

$$\mathbf{F}_{\text{e.m.}}(\mathbf{r}, t) = F \mathbf{\epsilon} \cos(\omega t - \mathbf{q} \cdot \mathbf{r}) \quad (3)$$

where $\mathbf{\epsilon}$ is the polarization vector which is perpendicular to \mathbf{q} . Instead of using the electric field and its associated electric potential ϕ , we shall use a gauge where ϕ is zero and where

$$\mathbf{F}_{\text{e.m.}} = -\frac{1}{c} \frac{\partial \mathbf{A}_{\text{e.m.}}}{\partial t} \quad (4)$$

where $\mathbf{A}_{\text{e.m.}}$ is the vector potential :

$$\mathbf{A}_{\text{e.m.}}(\mathbf{r}, t) = -\frac{ecF}{2i\omega} \{ \exp[i(\omega t - \mathbf{q} \cdot \mathbf{r})] - \exp[-i(\omega t - \mathbf{q} \cdot \mathbf{r})] \} \quad (5)$$

In the presence of the vector potential $\mathbf{A}_{\text{e.m.}}$, the electron momentum \mathbf{p} should be replaced by $\mathbf{p} + \frac{e\mathbf{A}_{\text{e.m.}}}{c}$ in the equations of motion, where e is the magnitude of the electron charge. Thus, to the first order in $\mathbf{A}_{\text{e.m.}}$ (linear absorption) and neglecting any contributions arising from the spin-orbit terms, the one-electron Hamiltonian of an heterostructure can be written :

$$\mathcal{H} = \mathcal{H}_0 + \frac{e}{2m_0c} [\mathbf{p} \cdot \mathbf{A}_{\text{e.m.}} + \mathbf{A}_{\text{e.m.}} \cdot \mathbf{p}] \quad (6)$$

where \mathcal{H}_0 is the heterostructure Hamiltonian in the absence of electromagnetic wave. In equation (6) the effects associated with the time varying magnetic field of the wave have been neglected.

The electromagnetic perturbation is time-dependent. It will thus induce transitions between the initial states $|i\rangle$ and the final states $|f\rangle$, where $|\nu\rangle$ denotes an eigenstate of \mathcal{H}_0 with energy ε_ν . In the following, we shall take $\varepsilon_f > \varepsilon_i$.

The Fermi Golden Rule [3] tells us how to calculate the transition probability per unit time \tilde{P}_{if} that an electron, under the action of $\mathcal{H} - \mathcal{H}_0$, makes a transition from the state $|i\rangle$ to the state $|f\rangle$. Since $\varepsilon_f > \varepsilon_i$, only the $+\omega$ component of $\mathbf{A}_{\text{e.m.}}$ induces transitions which yields :

$$\tilde{P}_{if} = \frac{2\pi}{\hbar} |\langle f | V | i \rangle|^2 \delta [\varepsilon_f - \varepsilon_i - \hbar\omega] \quad (7)$$

where :

$$V = \frac{ieF}{4m_0\omega} [\mathbf{e} \cdot \mathbf{p} e^{-i\mathbf{q} \cdot \mathbf{r}} + e^{-i\mathbf{q} \cdot \mathbf{r}} \mathbf{e} \cdot \mathbf{p}] \quad (8)$$

Under typical experimental conditions the photon wavevector \mathbf{q} can safely be neglected : the wave amplitude varies over distances (the wavelength $2\pi/q \sim 1 \mu\text{m}$) which are considerably larger than any characteristic dimension of electronic origin. Thus, under this electric dipole approximation V reduces to :

$$V = \frac{ieF}{2m_0\omega} \mathbf{e} \cdot \mathbf{p} \quad (9)$$

In the case where the levels $|i\rangle$ and $|f\rangle$ are partially or completely occupied, one has to account for the impossibility of allowing transitions from either an empty level or towards a filled one. The perturbation by the electromagnetic wave of the statistical distributions of the levels $|\nu\rangle$ of the heterostructure is assumed to be negligible. The probability per unit time that an electron makes a $|i\rangle \rightarrow |f\rangle$ transition or, equivalently, that a photon disappears is thus equal to :

$$P_{if} = \tilde{P}_{if} f(\varepsilon_i) [1 - f(\varepsilon_f)] \quad (10)$$

In equation (10), $f(\varepsilon_\nu)$ is the mean occupancy of the level $|\nu\rangle$:

$$f(\varepsilon_\nu) = \{1 + \exp[\beta(\varepsilon_\nu - \mu)]\}^{-1} \quad (11)$$

where $\beta = (k_B T)^{-1}$, T the temperature and μ the chemical potential of the electrons.

As a result of the electronic transitions $|i\rangle \rightarrow |f\rangle$, the electromagnetic wave loses an energy $\hbar\omega \tilde{P}_{if} dt$ during the time interval dt .

The transitions $|f\rangle \rightarrow |i\rangle$ can also be induced by the wave (stimulated emission). They occur at the rate of

$$\tilde{P}_{fi} = \frac{2\pi}{\hbar} |\langle i | V | f \rangle|^2 \delta [\varepsilon_f - \varepsilon_i - \hbar\omega] \quad (12)$$

transitions per unit time and contribute to the creation of photons by a quantity

$$P_{fi} = \tilde{P}_{fi} f(\varepsilon_f) [1 - f(\varepsilon_i)] \quad (13)$$

per unit time. The net energy loss of the electromagnetic wave per unit time associated with $|i\rangle \leftrightarrow |f\rangle$ transitions is thus

$$\hbar\omega [P_{if} - P_{fi}] \quad (14)$$

Finally, by summing over all states $|i\rangle$, $|f\rangle$, we obtain the decrease in the wave energy per unit time :

$$P(\omega) = \frac{2\pi}{\hbar} \frac{e^2 F^2}{4m_0^2 \omega^2} \hbar\omega \sum_{i,f} \delta[\varepsilon_f - \varepsilon_i - \hbar\omega] |\langle f | \mathbf{e} \cdot \mathbf{p} | i \rangle|^2 [f(\varepsilon_i) - f(\varepsilon_f)] \quad (15)$$

This rate can also be written as

$$P(\omega) = - \frac{d}{dt} \int_{\Omega} \mathcal{E}_{e.m.} d^3r \quad (16)$$

where Ω is the volume of the sample ($\Omega = \mathcal{L}S$) and $\mathcal{E}_{e.m.}$ is the energy density of the electromagnetic wave. The sample "volume" Ω is not a quantity which is as straightforward to define as in bulk material. The definition of the sample area S does not cause much trouble : like in bulk materials there exists a translational invariance in the layer planes of a heterolayer. On the other hand, the heterostructures have no translational invariance along their growth axis while bulk materials are, at the scale of envelope functions, homogeneous media along any direction. Since, to our knowledge, there is no first principle derivation of the absorption for quantum heterolayers we should define \mathcal{L} in an empirical way. In equation (16) we have equated the time derivative of the electromagnetic energy in Ω to the losses due to microscopic electronic transitions. Thus, one should identify for a *perfect* heterostructure \mathcal{L} to the length over which the heterostructure energy levels are well defined, i.e. to a length which, if changed slightly, does not appreciably alter the energy levels and their associated wavefunctions. For a single quantum well structure with thickness L one should take $\mathcal{L} \gg L$. For a superlattice comprising N periods (thickness d) one may take $\mathcal{L} \geq N d$ if $N \gg 1$. On the other hand, \mathcal{L} should not be too large to ensure that the assumption of a plane electromagnetic wave, which underlies equation (15), is still valid.

In equation (16) the time derivative of $\mathcal{E}_{e.m.}$ can be expressed in terms of the divergence of the Poynting vector [2]. Finally, by expressing F in terms of $\mathcal{E}_{e.m.}$ and by writing that time average $\langle \mathcal{E}_{e.m.} \rangle$ of $\mathcal{E}_{e.m.}$ over a period $\frac{2\pi}{\omega}$ decreases exponentially along the propagation direction we obtain :

$$\langle \mathcal{E}_{e.m.}(\mathcal{L}) \rangle = \langle \mathcal{E}_{e.m.}(0) \rangle \exp(-\alpha(\omega) \mathcal{L}) \quad (17a)$$

in the case of a propagation along the growth axis and

$$\langle \mathcal{E}_{e.m.}(x) \rangle = \langle \mathcal{E}_{e.m.}(0) \rangle \exp(-\alpha(\omega) x) \quad (17b)$$

when the wave propagates in the layer planes of the heterostructures. In equations (17 a,b) the absorption coefficient $\alpha(\omega)$ is given by

$$\alpha(\omega) = A \sum_{i,f} \frac{1}{m_0} |\mathbf{e} \cdot \mathbf{p}_{if}|^2 \delta(\varepsilon_f - \varepsilon_i - \hbar\omega) [f(\varepsilon_i) - f(\varepsilon_f)] \quad (18)$$

$$A = \frac{4\pi^2 e^2}{ncm_0\omega\Omega} \quad (19)$$

and

$$\mathbf{p}_{if} = \langle i | \mathbf{p} | f \rangle \quad (20)$$

In equation (18) we have neglected the spontaneous emission due to our classical treatment of the electromagnetic field. The difference between both expressions (17a) and (17b) stems for the lack (existence) of a translational invariance of the heterostructure along (perpendicular to) the growth direction. Since the x (or y) carrier envelope functions are plane waves a calculation of the absorption coefficient at the coordinate x which resembles that done for bulk materials is operative when the wave propagates in the layer plane (for a macroscopic sample area S) : one may ideally cut the sample into slices of the thickness Δx around any x along the propagation direction, define carrier plane waves in Δx (thus $\Delta x \gg$ host periodicities) and write that the drop in the light beam intensity is nearly equal to $\alpha(\omega) \Delta x$. Finally one goes to the limit $\Delta x \rightarrow 0$ to get the familiar equation (17b). Such a procedure is meaningless when the wave propagates along the growth axis. The thickness \mathcal{L} , of quantum origin, can not be cut into slices without altering the energy levels, wavefunctions and ultimately the whole light absorption process. Thus, it is impossible to know, at better than \mathcal{L} , where a photon has been absorbed and thus to define at point z what the beam intensity is. On the other hand equation (17a) is still valid because it only requires the validity of equation (16) and the fact that, inside \mathcal{L} , photons travel at the constant velocity $\frac{c}{n}$ where n is the index of refraction of the quantum heterostructure.

A has dimension of a [length] $^{-1}$ and \mathbf{p}_{if} , which is proportional to the dipolar matrix element $e\mathbf{r}_{if}$, will govern the selection rules of the allowed optical transitions. Note finally that the expression of the absorption coefficient involves the quantum thickness of the heterostructure \mathcal{L} . The latter does not coincide with the macroscopic thickness of a sample since usually the heterolayers are grown on a substrate with a buffer layer etc...

To exploit the general expression of $\alpha(\omega)$, we need to have a model for the energy levels. We choose the simplest one and write the wavefunctions of the initial state $|i\rangle$ in the form

$$F_i(\mathbf{r}) = u_{\nu_i}(\mathbf{r}) f_i(\mathbf{r}) \quad (21)$$

$$f_i(\mathbf{r}) = \frac{1}{\sqrt{S}} \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp) \chi_i(z) \quad (22)$$

where $u_{\nu_i}(\mathbf{r})$ is the periodic part of the Bloch functions at the zone centre (assumed to be the same in both type of layers) for the band ν_i . $\chi_i(z)$ is the envelope function

which describes the z motion of the electron in the subband i corresponding to the extremum ν_i and \mathbf{k}_\perp , \mathbf{r}_\perp are two-dimensional wave and position vectors respectively :

$$\mathbf{k}_\perp = (k_x, k_y); \quad \mathbf{r}_\perp = (x, y) \quad (23)$$

An expression similar to equations (21, 22) holds for the final state. These expressions of the wavefunctions are the lowest order of an expansion in ascending powers of $\mathbf{k} \cdot \mathbf{p}_{\nu\nu}/(\epsilon_\nu - \epsilon_{\nu'})$, where $\mathbf{k} = (\mathbf{k}_\perp, k_z^{(A)} \text{ or } k_z^{(B)})$ is the three-dimensional wavevector (real or imaginary) which characterizes the wavefunction in the A or B layers of the heterostructure.

As long as we are interested in optical transitions from one subband derived from a given host band (say a valence subband) to a subband derived from another host band (say a conduction subband) equations (22, 23) are sufficient. On the other hand, if we are dealing with transitions between the subbands derived from the same host band (say both initial and final states are conduction subbands) equations (22, 23) become insufficient even at the lowest order. To be consistent, we have to use either the eigenfunctions which include the $\mathbf{k} \cdot \mathbf{p}_{\nu\nu}/(\epsilon_\nu - \epsilon_{\nu'})$ corrections of the next order and keep a coupling Hamiltonian between carriers and light which is given by equation (9), or keep equations (21, 22) for the wavefunctions but use an *effective* coupling Hamiltonian between the carriers and the electromagnetic wave. This effective coupling can be calculated along the same lines as the one which was used in Appendices A and B of chapter II to build both the effective $\mathbf{k} \cdot \mathbf{p}$ matrix within the $\Gamma_6, \Gamma_7, \Gamma_8$ subspace and the effective slowly varying potential which acts on the conduction electrons of non-parabolic Γ_6 band. Here, in order to evaluate $\alpha(\omega)$ in the case of intraband transitions, we shall restrict ourselves to the parabolic dispersion relations of the host bands. The effective coupling procedure therefore amounts to replacing $1/m_0$ in equation (9) by $1/m^*(z)$, where $m^*(z)$ is the position-dependent effective mass in the heterolayer. This means that, under most circumstances, the selection rules for the intraband transitions will be the *same*, irrespective of the use of a bare or effective coupling between carrier and light. Thus, in the derivation of the intraband transition selection rules, we shall keep equations (21, 22) for the wavefunction, but restore m^* instead of m_0 where appropriate. This is relevant for the evaluation of the *magnitude* of the absorption coefficient.

Let us denote by Ω_0 the volume of the elementary cell of the materials involved in the considered structure. We have :

$$\epsilon \cdot \mathbf{p}_{if} = \epsilon \cdot \int_{\Omega} F_i^*(\mathbf{r}) \mathbf{p} F_f(\mathbf{r}) d^3r \quad (24)$$

$$\epsilon \cdot \mathbf{p}_{if} \simeq \epsilon \cdot \langle u_{\nu_i} | \mathbf{p} | u_{\nu_f} \rangle \int_{\Omega} f_i^* f_f d^3r + \delta_{if} \epsilon \cdot \int_{\Omega} f_i^* p f_f d^3r \quad (25)$$

where :

$$\langle u_{\nu_i} | \mathbf{p} | u_{\nu_f} \rangle = \int_{\Omega_0} u_{\nu_i}^*(\mathbf{r}) \mathbf{p} u_{\nu_f}(\mathbf{r}) d^3r \quad (26)$$

In equations (24-26) we have taken advantage of the rapid variations of u_{ν_i}, u_{ν_f} over k_\perp^{-1} or over the characteristic lengths of variation of $\chi_i(z), \chi_f(z)$.

The allowed optical transitions split into two categories. On the one hand, the *intraband* transitions ($\nu_i = \nu_f$) which involve the dipole matrix elements between envelope functions and on the other hand the *interband* transitions which occur between subbands originating from different extrema. The selection rules for the latter have two origins :

- i) the overlap integral between envelope functions selects the quantum numbers of the initial and final subbands
- ii) the atomic-like dipole matrix element $\langle u_{\nu_i} | \mathbf{p} | u_{\nu_f} \rangle$ gives rise to the selection rules on the polarization of the light wave.

Let us examine the intraband and interband transitions in closer details.

I.1 INTRABAND TRANSITIONS. — The dipole matrix element between envelope functions is :

$$\langle f_i | \mathbf{e} \cdot \mathbf{p} | f_f \rangle = \frac{1}{S} \int d^3r \chi_{n_i}^*(z) \exp(-i\mathbf{k}_\perp \cdot \mathbf{r}_\perp) [\varepsilon_x p_x + \varepsilon_y p_y + \varepsilon_z p_z] \times \chi_{n_f}(z) \exp(i\mathbf{k}'_\perp \cdot \mathbf{r}_\perp) \quad (27)$$

or :

$$\langle f_i | \mathbf{e} \cdot \mathbf{p} | f_f \rangle = (\varepsilon_x \hbar k_x + \varepsilon_y \hbar k_y) \delta_{n_i, n_f} \delta_{\mathbf{k}'_\perp, \mathbf{k}_\perp} + \varepsilon_z \delta_{\mathbf{k}'_\perp, \mathbf{k}_\perp} \times \int dz \chi_{n_i}^*(z) p_z \chi_{n_f}(z) \quad (28)$$

The polarizations ε_x , ε_y give rise to allowed transitions only if both the initial and final states coincide (i.e. if $\omega = 0$). The intrasubband absorption ($n_i = n_f$), which in the static limit cannot be reasonably treated without including scattering mechanisms, is the two-dimensional analogue of the free-carrier absorption. For perfect heterostructures, the free carrier absorption is forbidden in quasi bi-dimensional electron gases for the same reason as in bulk materials, i.e. the impossibility of conserving the energy and momentum simultaneously during the photon absorption by an electron. Free carrier absorption may be induced by defects (impurities, phonons) capable of providing the momentum necessary for the electron transition. The effect of reduced dimensionality on the free-carrier absorption has been calculated [4].

The polarization ε_z , which corresponds to an electromagnetic wave propagating in the layer plane with an electric field vector parallel to the growth axis of the structure, leads to optical transitions which are allowed between subbands provided that the in-plane wavevector of the carrier is conserved (vertical transitions in the \mathbf{k}_\perp space). In addition, if the heterostructure Hamiltonian has a definite parity the initial and final subbands should be of opposite parities. The occurrence of intraband transition without the participation of defects is specific to the quasi bi-dimensional materials and originates from the existence of a z -dependent potential in \mathcal{H}_0 . This potential makes up for the momentum necessary for the intraband absorption when the electric field vector of the light is parallel to z .

Table I. — Selection rules for intraband (inter-subband) transitions.

Polarization	ϵ_x	ϵ_y	ϵ_z
Propagation parallel to \hat{z}	forbidden if $\omega \neq 0$	forbidden if $\omega \neq 0$	impossible
Propagation parallel to \hat{x}	impossible	forbidden if $\omega \neq 0$	allowed
Propagation parallel to \hat{y}	forbidden if $\omega \neq 0$	impossible	allowed

Let us now evaluate the order of magnitude of such allowed intraband transitions in rectangular quantum well whose lowest subband (confinement energy E_1) is occupied by N_e electrons.

We first consider transitions which send the E_1 electrons deep into the quantum well continuum, a possible loss mechanism in quantum well lasers. For these highly excited continuum states we may neglect the reflection-transmission phenomena which take place at $z = \pm \frac{L}{2}$, where L is the quantum well thickness. Thus :

$$\chi_f(z) = \frac{1}{\sqrt{\mathcal{L}}} \exp(ikz) \quad (29)$$

and :

$$\chi_i(z) = \sqrt{\frac{2}{L}} \cos\left(\frac{\pi z}{L}\right); \quad |z| \leq \frac{L}{2} \quad (30)$$

where, for simplicity, we have neglected the exponential tails of the bound state wavefunction in the barrier. We evaluate $\langle \chi_i | p_z | \chi_f \rangle$ and obtain :

$$\langle \chi_i | p_z | \chi_f \rangle = \sqrt{\frac{2}{L\mathcal{L}}} \cos\left(\frac{kL}{2}\right) \left[\frac{1}{k + \pi/L} - \frac{1}{k - \pi/L} \right] \hbar k \quad (31)$$

The quantity

$$B = \frac{1}{m^*} \sum_{i,f} |\epsilon \cdot p_{if}|^2 \delta(\epsilon_f - \epsilon_i - \hbar\omega) [f(\epsilon_i) - f(\epsilon_f)] \quad (32)$$

is also easily evaluated if the continuum levels have a negligible population. We find :

$$B = \frac{2k_0 L N_e}{\pi} \cos^2\left(\frac{k_0 L}{2}\right) \left[\frac{1}{\pi + k_0 L} - \frac{1}{k_0 L - \pi} \right]^2 \quad (33)$$

where :

$$k_0^2 = \frac{2m^*}{\hbar^2} (\hbar\omega - V_b + E_1) \quad (34)$$

and V_b is the barrier height. The absorption coefficient is therefore :

$$\alpha(\omega) = \frac{8\pi e^2 n_e}{n c m^* \mathcal{L} \omega} x \cos^2 \left(\frac{x}{2} \right) \left[\frac{1}{x + \pi} - \frac{1}{x - \pi} \right]^2 \quad (35)$$

$$x = k_0 L \quad (36)$$

with $n_e = N_e/S$. At large ω , $\alpha(\omega)$ decreases like $\omega^{-5/2}$. We notice that the absorption coefficient is inversely proportional to \mathcal{L} , but does not explicitly depend on S . These results are not surprising. Both the initial and final states are delocalized in the layer plane. The carrier interacts with the electromagnetic wave all over the area S . On the other hand, since the initial state is confined within the quantum well, only a fraction L/\mathcal{L} of the electromagnetic energy is available for absorption.

As a second example, let us now consider the optical absorption which promotes the carrier from the ground (E_1) to the first excited (E_2) subbands of a quantum well.

To lighten the algebra, we again approximate the wavefunctions by those of an infinitely deep well, i.e. :

$$\chi_i(z) = \chi_1(z) = \sqrt{\frac{2}{L}} \cos \left(\frac{\pi z}{L} \right) \quad (37)$$

$$\chi_f(z) = \chi_2(z) = \sqrt{\frac{2}{L}} \sin \left(\frac{2\pi z}{L} \right) \quad (38)$$

We obtain :

$$\langle \chi_i | p_z | \chi_f \rangle = -\frac{8i\hbar}{3L} \quad (39)$$

and

$$B = \frac{64\hbar^2}{9m^*L^2} (N_1 - N_2) \delta(E_2 - E_1 - \hbar\omega) \quad (40)$$

where N_1 and N_2 are the number of electrons in the initial and final subbands whose confinement energies are E_1 and E_2 respectively. The absorption coefficient associated with $|1\rangle \rightarrow |2\rangle$ transitions is then equal to :

$$\alpha(\omega) = \frac{256\pi^2 e^2 \hbar^2}{9nc(m^*)^2 L^2 \omega \mathcal{L}} (n_1 - n_2) \delta[E_2 - E_1 - \hbar\omega] \quad (41)$$

where n_1 and n_2 are the areal density of carriers found in the E_1 and E_2 subbands respectively.

The delta singularity originates from the exact parallelism between the in-plane dispersion relations of the initial and final subbands. Introducing non-parabolicity or

any other effect which alters this parallelism, would change the delta peak into a line of finite energy extension. It may also be noticed that we have obtained the same \mathcal{L}^{-1} dependence for α as in equation (35). In both examples, the physical absorption mechanism takes place within the quantum well thickness, whereas the electromagnetic energy is delocalized over the whole heterostructure.

To get an order of magnitude of the absorption coefficient associated with $|1\rangle \rightarrow |2\rangle$ transitions, let us take $m^* = 0.07 m_0$, $n = 3.6$, $n_1 - n_2 = 10^{11} \text{ cm}^{-2}$, $\mathcal{L} = 1 \text{ } \mu\text{m}$, $\hbar\omega = E_2 - E_1 = 3 \frac{\hbar^2 \pi^2}{2m^* L^2}$ and replace the delta function by a Lorentzian

function with a full width at half maximum $2\Gamma = 4 \text{ meV}$. We obtain $\alpha \approx 67 \text{ cm}^{-1}$. The absorption coefficient associated with $|1\rangle \rightarrow$ continuum transitions is very small ($\alpha \sim 10^{-3} \text{ cm}^{-1}$) since x in equations (35, 36) is usually ~ 10 for material parameters adapted to GaAs wells. Note finally that the previous calculations have neglected electron-electron interactions and that collective excitations have not been taken into account.

We summarize in table I the selection rules for intraband transitions in perfect heterostructures.

I.2 INTERBAND TRANSITIONS. — We have to discuss here, on the one hand, the selection rules associated with the wave polarization and, on the other hand, the selection rules on the subband index which arise from the envelope functions overlap. In addition, interband transitions raise the question of the respective location of the initial and final states. It is clear that the optical absorption in a type II system (e.g. InAs-GaSb), where electrons and holes are spatially separated, should be weaker than in type I structures (e.g. GaAs-Ga(Al)As) where electrons and holes are essentially localized in the same layer.

A major complication arises due to the intricate valence subband dispersions in the layer plane. We have already discussed (chapter III) the decoupling of the $m_j = \pm \frac{3}{2}$ and $m_j = \pm \frac{1}{2}$ states at $\mathbf{k}_\perp = \mathbf{0}$. Strictly speaking, it is necessary to treat the problem of the in-plane dispersion relations at $\mathbf{k}_\perp \neq \mathbf{0}$ rigorously in order to be able to evaluate $\alpha(\omega)$. In fact, recent publications [5, 6] have dealt with this very complicated problem. Here, we shall bypass this difficulty by assigning an arbitrary in-plane effective mass to the valence subbands while retaining the $m_j = \pm 3/2$, $m_j = \pm 1/2$ decoupling. This will give us a reasonable description of the *onsets* of the absorptions. We notice parenthetically that if our model predicts that a given interband transition is allowed in the vicinity of $\mathbf{k}_\perp = \mathbf{0}$, it cannot in practice become forbidden at $\mathbf{k}_\perp \neq \mathbf{0}$. Symmetrically, any forbidden transition found in our simplified approach might become allowed due to the valence subband mixing between the $m_j = \pm 3/2$, $m_j = \pm \frac{1}{2}$, states at $\mathbf{k}_\perp \neq \mathbf{0}$. But it will present a smoother absorption edge and will be weaker than an allowed transition. The major weakness of our simplified model lies in its inability to account for peaks in the joint density of states which arise away from the zone centre when the initial and final subbands are nearly parallel. We believe this situation however to be rare.

Below, we thus adopt the following subband dispersion relations

$$\varepsilon_{HH_n}(\mathbf{k}_\perp) = -\varepsilon_g - HH_n - \frac{\hbar^2 k_\perp^2}{2 M_n} \quad (42)$$

$$\varepsilon_{LH_n}(\mathbf{k}_\perp) = -\varepsilon_g - LH_n - \frac{\hbar^2 k_\perp^2}{2 m_n} \quad (43)$$

$$\varepsilon_{c,n}(\mathbf{k}_\perp) = E_n + \frac{\hbar^2 k_\perp^2}{2 m_c} \quad (44)$$

where M_n and m_n can eventually become negative if the HH_n (or LH_n) subband has a positive curvature in the vicinity of $\mathbf{k}_\perp = \mathbf{0}$. We introduce the matrix element

$$\Pi = \frac{-i}{m_0} \langle S | p_x | X \rangle = \frac{-i}{m_0} \langle S | p_y | Y \rangle = \frac{-i}{m_0} \langle S | p_z | Z \rangle \quad (45)$$

which is related to the Kane matrix element E_p (~ 23 eV) by

$$E_p = 2 m_0 \Pi^2 \quad (46)$$

I.2.1 Polarization selection rules. — Examination of the periodic parts of the Bloch functions of the Γ_6 , Γ_7 , Γ_8 bands leads to the selection rules for the electromagnetic wave polarization. These are summarized in table II for the $HH_n \rightarrow E_m$, $LH_n \rightarrow E_m$ and $(\Gamma_7)_n \rightarrow E_m$ transitions respectively.

Table II shows that for interband transitions :

a) The heavy hole \rightarrow electron transitions are three times more intense than the light hole \rightarrow electron transitions. The $\Gamma_7 \rightarrow$ electron transitions have an intermediate strength (2/3 of those involving heavy holes).

b) For the propagation parallel to z , corresponding to the electric field of the wave in the layer planes, the three possible types of transitions are allowed. On the other hand, when the propagation occurs in the layer planes the polarization ε_z is forbidden for the $HH_n \rightarrow E_m$ transitions. This may have some influence on the lasing action in quantum well systems since, in this case, the light propagates perpendicularly to the growth axis. Anyhow, absorption (or spontaneous emission) studies in this geometry should show a strong dependence of the spectra on the exciting light polarization for the $HH_n \rightarrow E_m$ transitions. An example of such a dependence is presented in figure 1.

I.2.2 Selection rules on the envelope function quantum numbers : evaluation of $\langle f_i | f_f \rangle$. — We need to evaluate $\langle f_i | f_f \rangle$. This is equal to

$$\begin{aligned} \langle f_i | f_f \rangle = \langle \chi_n^{(h)} | \chi_m^{(c)} \rangle \langle \mathbf{k}_\perp | \mathbf{k}'_\perp \rangle = \int dz \chi_n^{*(h)}(z) \chi_m^{(c)}(z) \times \\ \frac{1}{S} \int d^2 r_\perp \exp[i(\mathbf{k}'_\perp - \mathbf{k}_\perp) \cdot \mathbf{r}_\perp] \quad (47) \end{aligned}$$

Table II. — Selection rules for interband (inter-subband) transitions obtained from the absolute value of the matrix element $\langle S^\dagger | \mathbf{e} \cdot \mathbf{p} | u_\alpha \rangle$, $\alpha = \Gamma_8$, $\Gamma_7 \cdot \mathbf{k}_\perp = \mathbf{0}$.

Polarization	ε_x	ε_y	ε_z	Type of transitions
Propagation parallel to $\hat{\mathbf{z}}$	$\frac{\Pi}{\sqrt{2}}$	$\frac{\Pi}{\sqrt{2}}$	impossible	$HH_n \rightarrow E_m$
Propagation parallel to $\hat{\mathbf{x}}$	impossible	$\frac{\Pi}{\sqrt{2}}$	forbidden	$HH_n \rightarrow E_m$
Propagation parallel to $\hat{\mathbf{y}}$	$\frac{\Pi}{\sqrt{2}}$	impossible	forbidden	$HH_n \rightarrow E_m$
Propagation parallel to $\hat{\mathbf{z}}$	$\frac{\Pi}{\sqrt{6}}$	$\frac{\Pi}{\sqrt{6}}$	impossible	$LH_n \rightarrow E_m$
Propagation parallel to $\hat{\mathbf{x}}$	impossible	$\frac{\Pi}{\sqrt{6}}$	$\frac{2\Pi}{\sqrt{6}}$	$LH_n \rightarrow E_m$
Propagation parallel to $\hat{\mathbf{y}}$	$\frac{\Pi}{\sqrt{6}}$	impossible	$\frac{2\Pi}{\sqrt{6}}$	$LH_n \rightarrow E_m$
Propagation parallel to $\hat{\mathbf{z}}$	$\frac{\Pi}{\sqrt{3}}$	$\frac{\Pi}{\sqrt{3}}$	impossible	$(\Gamma_7)_n \rightarrow E_m$
Propagation parallel to $\hat{\mathbf{x}}$	impossible	$\frac{\Pi}{\sqrt{3}}$	$\frac{\Pi}{\sqrt{3}}$	$(\Gamma_7)_n \rightarrow E_m$
Propagation parallel to $\hat{\mathbf{y}}$	$\frac{\Pi}{\sqrt{3}}$	impossible	$\frac{\Pi}{\sqrt{3}}$	$(\Gamma_7)_n \rightarrow E_m$

Thus, the first selection rule is the conservation of the in-plane wavevector in the optical transitions

$$\mathbf{k}'_\perp - \mathbf{k}'_\perp = \mathbf{0} \quad (48)$$

It arises from the translational invariance of the heterostructure Hamiltonian in the layer plane.

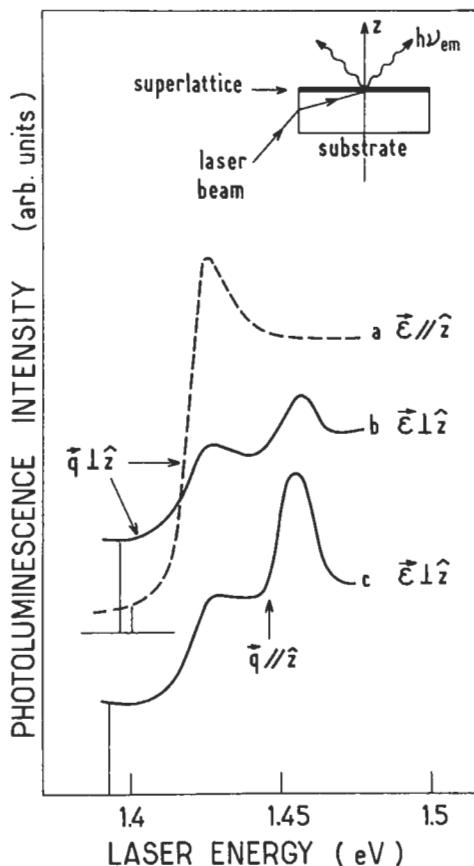


Fig. 1. — Polarization dependence of the photoluminescence excitation spectrum in a $\text{In}_{1-x}\text{Ga}_x\text{As-GaAs}$ superlattice at $T = 77 \text{ K}$. In this experiment the excitation spectrum mimics the absorption coefficient. For an electromagnetic wave propagating in the layer plane ($\mathbf{q} \perp \hat{z}$) two absorption peaks are seen in the $\epsilon \perp \hat{z}$ polarization while only one shows up in the $\epsilon \parallel \hat{z}$ polarization. This is in agreement with the predictions of table II. Moreover, when $\mathbf{q} \parallel \hat{z}$ and thus $\epsilon \perp \hat{z}$ two peaks are observed which coincide in energy with those seen in the $\epsilon \perp \hat{z}$, $\mathbf{q} \perp \hat{z}$ polarization. After reference [51].

The evaluation of the selection rules on the subband indices is obtained by calculating $\langle \chi_n^{(h)} | \chi_m^{(e)} \rangle$. At this point one should make a distinction between the type I and type II systems.

* Type I systems (GaAs-Ga(Al)As ; Ga(In)As-InP ; Ga(In)As-Al(In)As...).

These heterostructures are such that conduction and valence electrons are essentially confined in the same layers. For a type I symmetrical quantum well, the envelope functions have a definite parity with respect to the centre of the well. The overlap integral $\langle \chi_n | \chi_m \rangle$ is thus non zero only if $n + m$ is even.

In addition, if the valence and conduction quantum wells are rectangular and infinitely deep, only the transitions which fulfill $n = m$ are allowed. In the GaAs-Ga(Al)As rectangular wells the transitions $\Delta n = n - m = 0$ are much stronger than those which correspond to Δn even ($\neq 0$). To our knowledge, only the transition $HH_3 \rightarrow E_1$ has been identified in absorption [7] and photoluminescence excitation spectroscopy [8]. Other GaAs-Ga(Al)As wells which are not rectangular, e.g. the pseudo-parabolic wells [9] or Separate Confinement Heterostructures [10], have been optically studied. The former structures are obtained by imposing a quadratic variation of the conduction and valence edges by beam chopping during the epitaxial growth, whereas the latter structures consist of embedding a narrow GaAs-Ga_{1-x₁}Al_{x₁}As quantum well into Ga_{1-x₂}Al_{x₂}As barriers with $x_2 > x_1$. Both kinds of heterostructures have permitted the observation of more optical transitions than those found in the plain GaAs-Ga(Al)As rectangular quantum wells. This has led to a significant re-appraisal of the band discontinuities in the GaAs-Ga(Al)As system (for recent reviews see [11, 12]).

The selection rule $n + m$ even is in fact very strong and only very asymmetrical wells can lead to transitions with $n + m$ odd. Investigations of intentionally designed asymmetrical wells might yield a better determination of the band offsets, since more optical transitions are allowed.

* Type II systems (InAs-GaSb, InP-Al_{0.48}In_{0.52}As)

In these structures, the electrons and holes are spatially separated. Let us consider the case of a type II quantum well, whose band profiles are shown in figure 2. We denote the well-acting (barrier-acting) materials for electrons by A(B) and their

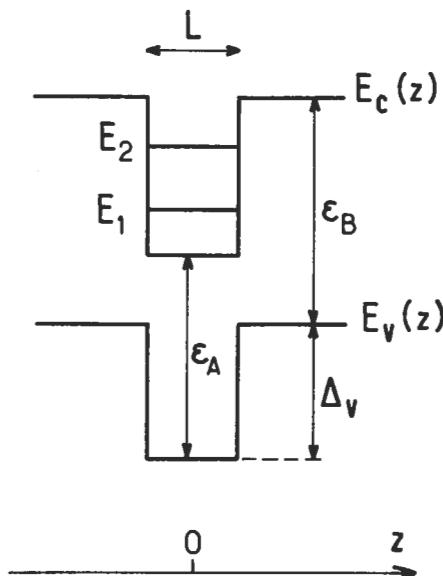


Fig. 2. — Conduction and valence band edge profiles in a type II quantum well.

respective bandgaps by ϵ_A , ϵ_B . The A material is a barrier for the holes and since we are considering a quantum well structure, the valence spectrum is continuous and twice degenerate. To simplify the discussion we shall later consider only photon energies such that $\hbar\omega < \epsilon_A$ and assume that $E_1 < \Delta_v$. For a given bound state E_n of the conduction band, a transition should involve only one of the two degenerate valence states : if $\chi_m^{(c)}$ is even in z , the initial state should also be even in z in order to participate to the optical transitions. It is also clear that the overlap of the conduction and valence envelope functions is only due to the exponential tails of the conduction and valence states in the layers B and A respectively : in the limit of large valence and conduction band discontinuities, the overlap integrals $\langle \chi_n^{(h)} | \chi_m^{(c)} \rangle$ tend to zero. In addition, $\langle \chi_n^{(h)} | \chi_m^{(c)} \rangle$ should increase with increasing n and m : the tails of the conduction and valence envelopes in their respective barriers become more and more important. Notice that the opposite is true in type I systems.

I.2.3 Order of magnitude of the absorption coefficient. Comparison between type I and type II systems. — As an example, let us consider the optical transitions between the ground heavy hole and electron subbands of a rectangular type I quantum well (ϵ_x polarization). In equation (18) the quantity $\frac{1}{m_0} |\epsilon \cdot p_{if}|^2$ is equal to :

$$\frac{1}{m_0} |\epsilon \cdot p_{if}|^2 = \frac{E_p}{4} |\langle \chi_1^{(h)} | \chi_1^{(c)} \rangle|^2 \quad (49)$$

and the absorption coefficient $\alpha_{HH_1 \rightarrow E_1}(\omega)$ to :

$$\begin{aligned} \alpha_{HH_1 \rightarrow E_1}(\omega) = & A \frac{E_p}{4} |\langle \chi_1^{(h)} | \chi_1^{(c)} \rangle|^2 \times \\ & 2 \sum_{k_\perp} \delta \left[\epsilon_A + E_1 + HH_1 + \frac{\hbar^2 k_\perp^2}{2} \left(\frac{1}{m_c} + \frac{1}{M_1} \right) - \hbar\omega \right] \end{aligned} \quad (50)$$

By performing the summation upon k_\perp we finally obtain :

$$\begin{aligned} \alpha_{HH_1 \rightarrow E_1}(\omega) = & \frac{\pi e^2 E_p}{ncm_0 \omega \mathcal{L} \hbar^2} \frac{m_c M_1}{m_c + M_1} |\langle \chi_1^{(h)} | \chi_1^{(c)} \rangle|^2 \times \\ & Y[\hbar\omega - \epsilon_A - E_1 - HH_1] \end{aligned} \quad (51)$$

In equations (50, 51) population effects for both initial and final states have been neglected (undoped quantum wells), ϵ_A is the bandgap of the bulk well-acting material and $Y(x)$ is the step function. For a $LH_1 \rightarrow E_1$ transition a similar expression is obtained with M_1 , HH_1 replaced by m_1 , LH_1 and E_p by $E_p/3$ (see Tab. II).

The absorption coefficient has a *staircase-like* shape at the onset of the absorption (see Fig. 3), as expected from the joint density of states of the two bi-dimensional subbands. For $HH_n \rightarrow E_m$ transitions, $n, m \neq 1$, the magnitude of the step will involve the square of $\langle \chi_n^{(h)} | \chi_m^{(c)} \rangle$ which is smaller than one and also the in-plane

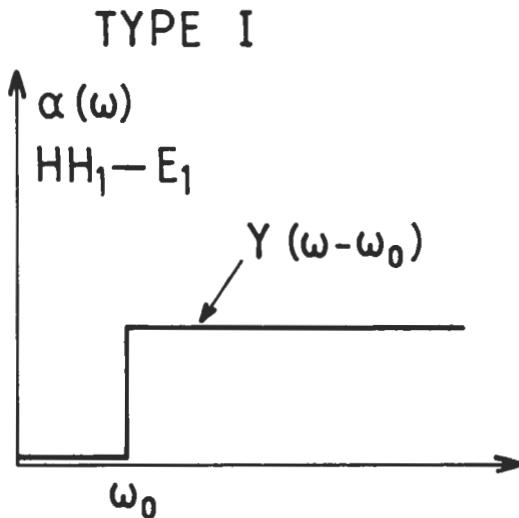


Fig. 3. — Staircase lineshape of the absorption coefficient of a type I quantum well.

effective mass of the n^{th} heavy hole subband in the vicinity of the zone centre, i.e. M_n will replace M_1 in equation (51). We have seen in chapter III that some of these M_n 's may be negative (i.e. may have electron-like behaviour near $\mathbf{k}_\perp = \mathbf{0}$). This, in turn, affects the magnitude of $\alpha_{HH_n \rightarrow E_m}$. Experimentally, it is often found that the edges of higher lying heavy hole-electron transitions are less marked than the $HH_1 \rightarrow E_1$ one, which is consistent with an increased broadening of the excited subband states.

However, the most significant effect in equation (51) is the *blue shift* of the quantum well fundamental absorption edge with respect to that of the bulk A material. Photons start to be absorbed at an energy $E_1 + HH_1$, above ϵ_A . This blue shift can be tuned by varying the quantum well thickness. For instance, bulk GaAs starts absorbing in the infra-red part of the spectrum ($\epsilon_A = 1.5192$ eV at low temperature). Narrow GaAs-Ga_{1-x}Al_xAs quantum wells ($L \sim 30$ Å) can be designed to start absorbing light only in the red part of the spectrum ($\epsilon_A + E_1 + HH_1 \approx 1.7275$ eV if $x = 0.5$). The blue shift of the fundamental absorption edge has been observed in a number of heterostructures ; for example in GaAs-Ga(Al)As [7] ; Ga(In)As-InP [13] ; Ga(In)As-Al(In)As [14] ; GaSb-AlSb [15] ; Ga(In)As-GaAs [16].

Finally, let us calculate the magnitude of $\alpha_{HH_1 \rightarrow E_1}$ in the case of a single quantum well embedded in a structure of total thickness $\mathcal{L} = 1\mu\text{m}$. For $\hbar\omega = 1.6$ eV, $m_c = 0.067 m_0$, $M_1 \gg m_c$, $n = 3.6$, $E_p = 23$ eV and taking $\langle \chi_1^{(\text{h})} | \chi_1^{(\text{e})} \rangle \sim 1$, we obtain $\alpha_{HH_1 \rightarrow E_1} = 60 \text{ cm}^{-1}$. Such a faint absorption can hardly be measured in practice, for it leads to an attenuation of the light beam intensity of only $1 - \exp(-\alpha_{HH_1 \rightarrow E_1} \mathcal{L}) \sim \alpha_{HH_1 \rightarrow E_1} \mathcal{L} = 0.6\%$. Therefore, one has to use *multiple quantum wells* to enhance

the absorption. These structures are obtained by separating N identical wells with $N - 1$ thick barriers, resulting in an absorption which is enhanced by a factor of N over that of a single well. The barriers have to be thick enough to prevent any tunnel coupling between the wells, a condition which is more difficult to achieve when higher lying transitions ($n, m > 1$) are involved. In practice N should be ≥ 10 in order to obtain a sizeable absorption.

We stress again that the low absorption coefficient is due to the different spatial locations of the electromagnetic energy (delocalized over \mathcal{L}) and the electrons which actually absorb light (essentially localized within the quantum well).

For a type II quantum well, the only difficulty consists in evaluating the overlap integral between valence states (which are extended in z and labelled by a wavevector k_v) and conduction states (which are localized in the well). For that purpose, we use a rather crude approach : the valence barrier Δ_v is assumed high enough to be considered as impenetrable by the valence electrons. In this case, two linearly independent valence envelope functions correspond to each k_v . They are :

$$\chi_r^{(h)}(z) = \frac{2}{\sqrt{\mathcal{L}}} \sin \left[k_v \left(z - \frac{L}{2} \right) \right] Y \left[z - \frac{L}{2} \right] \quad (52)$$

$$\chi_l^{(h)}(z) = \frac{-2}{\sqrt{\mathcal{L}}} \sin \left[k_v \left(z + \frac{L}{2} \right) \right] Y \left[-z - \frac{L}{2} \right] \quad (53)$$

These functions are normalized to 1 over $\frac{\mathcal{L}}{2}$ ($L \ll \mathcal{L}$) and for heavy holes we have :

$$\varepsilon_v(k_v, \mathbf{k}_\perp) = -\varepsilon_A + \Delta_v - \frac{\hbar^2 k_v^2}{2m_v} - \frac{\hbar^2 k_\perp^2}{2M_1} \quad (54)$$

where m_v is the heavy hole mass along the z direction. Note that in principle m_v should be set equal the M_1 in equation (54) since each of the B layers is a bulk material. We keep the m_v and M_1 notations in order to be able to make a comparison with the absorption coefficient in type I structures later on.

We can now introduce functions $\chi_e(z)$ and $\chi_0(z)$ which are even and odd with respect to z :

$$\chi_e^{(h)}(z) = \frac{1}{\sqrt{2}} [\chi_r^{(h)}(z) + \chi_l^{(h)}(z)] \quad (55)$$

$$\chi_0^{(h)}(z) = \frac{1}{\sqrt{2}} [\chi_r^{(h)}(z) - \chi_l^{(h)}(z)] \quad (56)$$

We are interested in calculating the absorption coefficient at the onset of the absorption, i.e. that which is associated with valence $\rightarrow E_1$ transitions. The envelope function $\chi_1^{(e)}(z)$ of the electrons in the E_1 subband can be written thus :

$$\chi_1^{(e)}(z) = A_e \cos(k_e z); \quad |z| \leq \frac{L}{2} \quad (57)$$

$$\chi_1^{(c)}(z) = B_c \exp\left[-\kappa_c\left(z - \frac{L}{2}\right)\right]; \quad z > \frac{L}{2} \quad (58)$$

$$\chi_1^{(c)}(-z) = \chi_1^{(c)}(z) \quad (59)$$

The overlap integral $\langle \chi_o^{(h)} | \chi_1^{(c)} \rangle$ vanishes while $\langle \chi_e^{(h)} | \chi_1^{(c)} \rangle$ is equal to

$$\langle \chi_e^{(h)} | \chi_1^{(c)} \rangle = 2 \sqrt{\frac{2}{\mathcal{L}}} B_c \frac{k_v}{k_v^2 + \kappa_c^2} \quad (60)$$

The coefficient B_c can easily be related to the probability $P_b(E_1)$ of finding the electron in the B layers, while in the E_1 states :

$$P_b(E_1) = \frac{B_c^2}{\kappa_c} \quad (61)$$

The quantity B defined in equation (32), with m^* replaced by m_0 (since we are dealing with interband transitions), can now be calculated. It is equal to

$$B = \frac{P_b(E_1) \kappa_c E_p}{\pi^2 \hbar^2} \times \frac{m_c M_1 S}{m_c + M_1} \int_0^{k_{\max}} \frac{dk_v k_v^2}{(\kappa_c^2 + k_v^2)^2} \quad (62)$$

where :

$$k_{\max}^2 = \frac{2 m_v}{\hbar^2} (\hbar\omega - \varepsilon_A + \Delta_v - E_1) \quad (63)$$

and finally the absorption coefficient is found to be equal to

$$\alpha(\omega)_{hh \rightarrow E_1} = \frac{2 e^2 E_p P_b(E_1)}{nc m_0 \omega \hbar^2 \mathcal{L}} \frac{m_c M_1}{m_c + M_1} \left[\frac{-x}{1 + x^2} + \text{Arctan } x \right];$$

$$x = \frac{k_{\max}}{\kappa_c} \quad (64)$$

The onset of the absorption is located at

$$\hbar\omega_0 = \varepsilon_A - \Delta_v + E_1 \quad (65)$$

In the vicinity of this onset, $\alpha_{hh \rightarrow E_1}$ behaves like $(\omega - \omega_0)^{3/2}$ (Fig. 4) instead of displaying the staircase-like behaviour characteristic of type I quantum wells. Thus, the indirect optical transitions in real space have the same smoothing effect on the absorption edge as the indirect transitions in reciprocal space in the case of bulk materials. The physical origin of the absorption (exponential tail of the E_1 wavefunction outside its confining layer) is confirmed by the presence of the

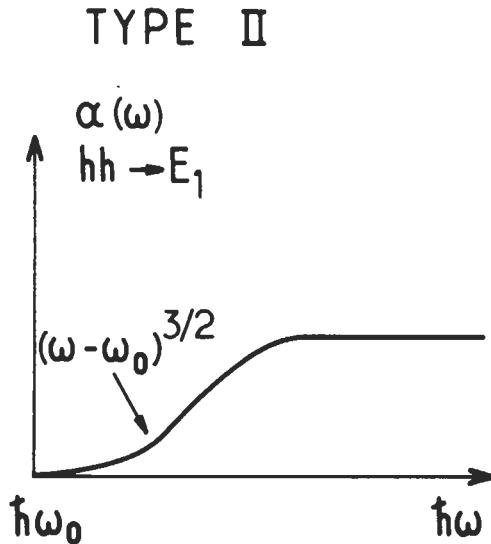


Fig. 4. — Absorption lineshape of a type II quantum well.

proportionality factor P_b in equation (64). At large ω , i.e. $k_{\max} \gg \kappa_c$, $\alpha(\omega)$ levels off at the value

$$\alpha_{hh \rightarrow E_1}^{\infty} = \frac{\pi e^2 E_p}{ncm_0 \omega \mathcal{L} \hbar^2} P_b(E_1) \frac{m_c M_1}{(m_c + M_1)} \quad (66)$$

which is just $P_b(E_1)$ times the magnitude of the plateau value for absorption in types I quantum wells (see Eq. (51)). This means that the use of multiple wells to enhance the single well absorption is even more imperative in type II than in type I structures. Since $P_b(E_1)$ is usually only a few per cent, it appears that hundreds of wells may be necessary to measure the absorption in type II quantum wells.

Finally, it may be added that light hole $\rightarrow E_1$ ($lh \rightarrow E_1$) transitions have the same threshold as heavy hole $\rightarrow E_1$ ($hh \rightarrow E_1$) transitions (Eq. (65)). Thus only the sum of the two contributions will actually be measured. In addition, transitions from the valence levels to the excited conduction subbands E_i , $i > 1$ will be more intense than the $hh \rightarrow E_1$ or $lh \rightarrow E_1$ transitions. This arises from the smaller κ_c and larger $P_b(E_i)$ for the excited subbands.

I.2.4 Interband optical transitions in superlattices. — When the barriers separating consecutive quantum wells become too thin, the wells are coupled due to the tunnel effect across the barriers. The localized levels of the wells hybridize, forming minibands along the z axis. The quantum number labelling the z motion, instead of simply being a subband index (E_i, HH_j, LH_k) , should also include a superlattice

wave vector q which can be restricted to the first superlattice Brillouin zone $[-\pi/d, +\pi/d]$ where d is the superlattice period. When analysing interband optical transitions in superlattices, it is essential to exploit the mirror symmetry (if the band edge profile is symmetric with respect to the centre of each kind of layer) and the translational invariance of the superlattice potential [17]. The fact that the translations T_d commute with the Hamiltonian leads to the Bloch form for the z -dependence of the envelope functions. This in turn results in a conservation of the superlattice wavevector q for allowed optical transitions. Photons are thus absorbed "vertically" in the superlattice Brillouin zone and in the layer plane. In addition, the relative spatial localization of electrons and holes play a major part in evaluating the selection rules. Under the same assumptions as those which were used (Eqs. (42-44)) to derive the absorption coefficients of type I and type II quantum wells, it has been shown [17] that for type II superlattices an optical transition allowed at $q = 0$ becomes forbidden at $q = \frac{\pi}{d}$ and *vice versa*. In other words, the quantity $|\langle \chi_i^{(h)} | \chi_f^{(e)} \rangle|^2$ varies like $1 + \cos qd$. This strongly affects the absorption lineshape, eliminating one of the two van Hove singularities displayed by the joint density of states for optical transitions in superlattices. This is exemplified in figures 5a,b where the absorption lineshapes for type I and type II superlattices are plotted against the dimensionless parameter ξ defined as

$$\hbar\omega = \hbar\omega_0 + (\delta_c + \delta_v)\xi$$

where $\hbar\omega_0$ is the minimum energy required to create an electron-hole pair for the subband-to-subband interband transition under considerations and δ_c and δ_v are the

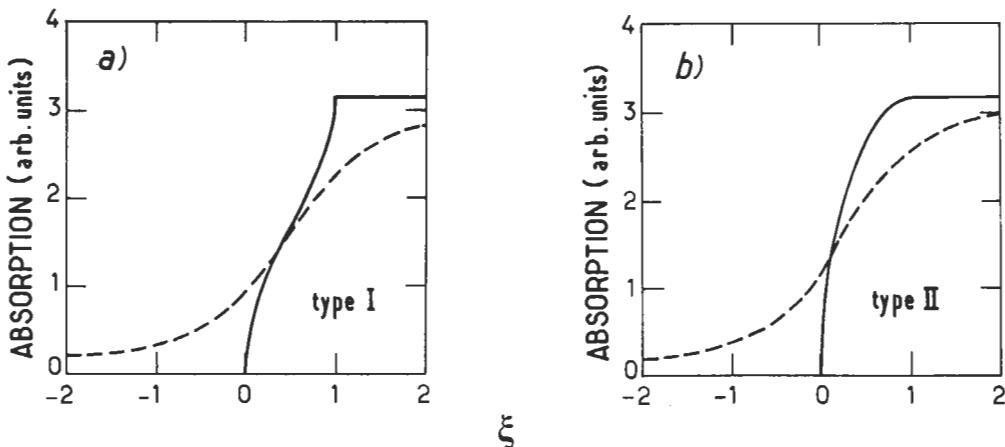


Fig. 5. — a) Absorption lineshape in a type I superlattice. solid line : no broadening included ; dashed line : phenomenological broadening ; the broadening parameter is equal to one half of the sum of the valence and conduction subband widths. b) same as in a) but for a type II superlattice. In both figures ξ is a dimensionless photon energy linked to ω by a linear relationship. $\xi = 0$ ($\xi = 1$) corresponds to a vertical photon absorption at $q = 0$ ($q = \pi/d$). After reference [17].

widths of the conduction and valence subband respectively. The solid lines correspond to materials which are supposed to be perfect. The van Hove anomaly at the band edge is missing in type II superlattices. The dashed lines give the absorption coefficient for imperfect materials, the damping being treated in a phenomenological way. It can be seen that the absorption lineshapes become very similar for both types of structures. Thus, rather than relying on a lineshape analysis, which is always difficult due to the rather unknown broadening effects, it is better to rely on the magnitude of the absorption coefficient. As in the case of quantum wells, the absorption in a type II superlattice is induced by the leakage of electron and hole wavefunctions outside the layers where they are principally localized. Therefore it is usually $P_b(E_i)$ or $P_b(HH_1)$ smaller than in the type I system, the latter being roughly N times that of a single quantum well where N is the number of superlattice periods.

From an experimental point of view, little information has been obtained on the dispersion relations along the growth axis and its influence on the optical absorption. The GaAs-Ga(Al)As superlattices are often analysed in terms of isolated quantum wells, even when the barriers are thin enough ($< 100 \text{ \AA}$) and the subbands sufficiently excited ($n > 1$) to ensure that these subbands have a non negligible width ($\delta_c, \delta_v \geq 10 \text{ meV}$). Figure 6 shows the energy dependence of the absorption coefficient in InAs-GaSb "semiconductor" superlattices [18] which presents smooth onsets of the subband to subband absorptions. It is difficult to know the respective parts played by the broadening and genuine "type II effects" in this smoothness.

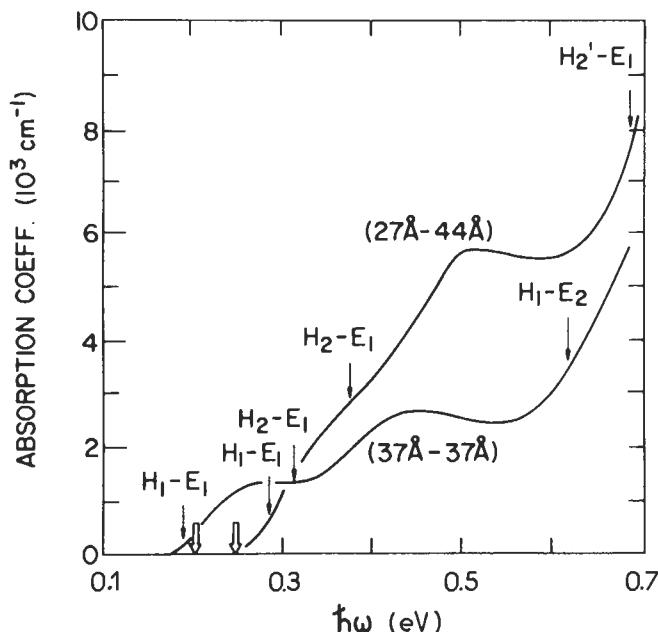


Fig. 6. — Measured energy dependence of the absorption coefficient in two InAs-GaSb superlattices. $T = 4.2 \text{ K}$. After reference [18].

II. Absorption : a simplified description of excitonic effects.

A convenient picture of an exciton is an electron and hole orbiting around each other. This picture is quite difficult to derive from first principles because an exciton is in fact a many-body effect.

The excitonic effects are seldom negligible in practice. They most clearly manifest themselves with the appearance of sharp peaks in the interband absorption spectra whereas, as discussed in the previous paragraph, non interacting electron models at most only predict staircase-like structures starting at $\hbar\omega_{\text{onset}}$. In the exciton (or electron-hole pairs) picture, the pair energies larger than $\hbar\omega_{\text{onset}}$ correspond to dissociated pairs. It is often forgotten that the electron-hole attraction also alters the optical response of the pair continuum [19]. This alteration is so deep that, in bulk materials for instance, the calculated one-electron absorption is *never* correct. Even deep in the continuum, where the large electron-hole relative velocity makes the non-interacting particle model at its best, it falls short of accounting for the correct absorption coefficient, i.e. the coefficient calculated by including the electron-hole interaction (although approaching it asymptotically from below [19]). We shall see later how this enhancement can be traced back algebraically. Physically, it arises from the long range nature of the Coulombic attraction between the electron and the hole, even when they do not form a bound state.

↙ The exciton picture emerges after lengthy many-body calculations [20]. All of these start by attempting to calculate an energy difference between the ground state of a crystal, which consists of N electrons occupying a filled valence band, and an excited state of the crystal obtained by optically promoting an electron from the aforesaid filled band to an empty conduction band. If there were no electron-electron interactions, this energy difference $\Delta\epsilon$ would simply be the sum of the bandgap ϵ_g (eventually including valence and conduction confinement energies in the case of heterostructures) and the difference between the kinetic energies in the final and excited states, i.e. for parabolic bands

$$\Delta\epsilon = \epsilon_g + \frac{\hbar^2 k_c^2}{2m_c} - \left(\frac{-\hbar^2 k_v^2}{2m_v} \right) \quad (67)$$

The onset of the interband absorption would occur at $\hbar\omega_{\text{onset}} = \epsilon_g$. The existence of lower excited states of the crystal are the origins of peaked structures below ϵ_g . They are obtained by rearranging the conduction and valence electronic densities (in a coupled way). To the excited state whose energy with respect to the ground state is $\Delta\epsilon$ (Eq. (67)), one associates valence and conduction charge distributions which are uniform in space (at the scale of the envelope functions) and which are spatially uncorrelated. By making wave packets of conduction states with various k_c and wave packets of valence states with various k_v , where k_c and k_v are univocally correlated ($k_v = k_c - \mathbf{K}$, \mathbf{K} fixed) the potential energy which is gained is greater (by a decrease of the repulsive electron-electron interaction) than the kinetic energy which is lost (by spatially correlating the electrons).

The following discussion is a survey of excitonic absorptions in semiconductor heterostructures and of their associated selection rules. This discussion does not

pretend to be complete : although relatively well documented on the experimental side, the optical properties of excitons in semiconductor heterostructures have not received enough attention on the theoretical side.

We shall adopt the same decoupling procedure for the excitons as for the free particle states (Eqs. (21, 22, 42-44)). Thus, in this crude model, there exists two kinds of decoupled excitons : those formed between heavy holes and electrons (heavy hole excitons) and those formed between light holes and electrons (light hole excitons). This decoupled exciton theory, first used by Miller *et al.* [21] and Greene and Bajaj [22] in the calculation of exciton binding energies (see chapter IV), neglects off-diagonal coupling terms in the valence band Hamiltonian. For our purpose, it proves to be very convenient, for it lightens the already heavy algebra involved in the derivation of the excitonic absorption.

Let us consider $HH_1 \rightarrow E_1$ excitonic transitions in a type I single quantum well. Other situations may be treated with minor modifications. The ground state of our quantum well consists of N valence electrons occupying the HH_1 subband and an empty E_1 conduction subband. The dispersion relations of HH_1 and E_1 are taken as parabolic upon the in-plane wavevector \mathbf{k}_\perp (see Eqs. (42, 44)), and the one electron wavefunctions are respectively

$$F_{\mathbf{k}_{\perp v}}^{(h)}(\mathbf{r}) = u_{hh}(\mathbf{r}) \chi_1^{(h)}(z) \frac{1}{\sqrt{S}} \exp(i\mathbf{k}_{\perp v} \cdot \mathbf{r}_\perp) \quad (68)$$

$$F_{\mathbf{k}_{\perp c}}^{(c)}(\mathbf{r}) = u_c(\mathbf{r}) \chi_1^{(c)}(z) \frac{1}{\sqrt{S}} \exp(i\mathbf{k}_{\perp c} \cdot \mathbf{r}_\perp) \quad (69)$$

where u_{hh} and u_c are the heavy hole (i.e. $u_{hh} \equiv u_{\frac{3}{2}, \pm \frac{3}{2}}$) and electron (i.e. $u_c \equiv iS^\dagger$ or iS^\downarrow) periodic parts of the Bloch function at the zone centre of the host layers.

It is clear that the quadratic dispersion law for HH_1 is unable to give rise to a filled band. It is understood that at large enough $\mathbf{k}_{\perp v}$ this dispersion flattens so that ultimately the HH_1 energy will be a periodic function upon $\mathbf{k}_{\perp v}$ with a period given by the in plane basis vectors of a suitable Brillouin zone. Actually, the knowledge of the whole dispersion of the HH_1 subband is not needed, for we are only interested in Wannier excitons which are weakly bound (with respect to the bandgaps). Their wavefunctions are built from HH_1 and E_1 states which involve $\mathbf{k}_{\perp c}$, $\mathbf{k}_{\perp v}$ values much smaller than the in-plane size of the Brillouin zone. We assume that for these states the quadratic laws are valid.

In the presence of the electromagnetic wave, specified in equations (3-5), the heterostructure Hamiltonian can be written

$$\mathcal{H} = \mathcal{H}_0 + \delta\mathcal{H}(t) \quad (70)$$

$$\mathcal{H}_0 = \sum_i \mathcal{H}_{0i} + \sum_{i>j} \frac{e^2}{r_{ij}} \quad (71)$$

$$\delta\mathcal{H}(t) = \frac{e}{2m_0c} \sum_i (\mathbf{p}_i \cdot \mathbf{A}_{\text{c.m.}}(\mathbf{r}_i, t) + \mathbf{A}_{\text{c.m.}}(\mathbf{r}_i, t) \cdot \mathbf{p}_i) \quad (72)$$

\mathcal{H}_0 is the N electrons Hamiltonian without the electromagnetic wave. It includes the sum of N one-electron terms and of the repulsive electron-electron interactions. The eigenstates of \mathcal{H}_0 are in the form given by equations (68, 69). If $|\psi_i\rangle$ and $|\psi_f\rangle$ denote two eigenstates of \mathcal{H}_0 with energies ε_i and ε_f they can be coupled by the time-dependent $\delta\mathcal{H}(t)$. Restricting our consideration once more to the dipolar approximation, the transition probability that any $|\psi_i\rangle \rightarrow |\psi_f\rangle$ transition occurs, i.e. that a photon disappears, is

$$P_{if} = \frac{2\pi}{\hbar} \left| \langle \psi_i | \sum_n V_n |\psi_f \rangle \right|^2 \delta(\varepsilon_f - \varepsilon_i - \hbar\omega) \quad (73)$$

where :

$$V_n = \frac{ieF}{2m_0\omega} \mathbf{e} \cdot \mathbf{p}_n \quad (74)$$

II.1 ABSORPTION IN THE ABSENCE OF COULOMBIC INTERACTIONS. EQUIVALENCE WITH THE ONE-ELECTRON MODEL. — In the absence of e^2/r_{ij} terms in \mathcal{H}_0 , the exact initial and final states of \mathcal{H}_0 are Slater determinants [3, 19, 20]. The initial state is unique and corresponds to the N electrons which occupy the HH_1 subband :

$$\psi_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{bmatrix} F_{\mathbf{k}_{11}}^{(h)}(\mathbf{r}_1) & F_{\mathbf{k}_{11}}^{(h)}(\mathbf{r}_2) & \dots & F_{\mathbf{k}_{11}}^{(h)}(\mathbf{r}_N) \\ F_{\mathbf{k}_{1v}}^{(h)}(\mathbf{r}_1) & F_{\mathbf{k}_{1v}}^{(h)}(\mathbf{r}_2) & \dots & F_{\mathbf{k}_{1v}}^{(h)}(\mathbf{r}_N) \\ \vdots & \vdots & & \vdots \\ F_{\mathbf{k}_{1N}}^{(h)}(\mathbf{r}_1) & F_{\mathbf{k}_{1N}}^{(h)}(\mathbf{r}_2) & \dots & F_{\mathbf{k}_{1N}}^{(h)}(\mathbf{r}_N) \end{bmatrix} \quad (75)$$

A final state ψ_f is obtained by removing a HH_1 electron with wavevector \mathbf{k}_{1v} and sending it into the E_1 subband where it acquires an in-plane wavevector \mathbf{k}_{1c} . Its wavefunction is obtained from equation (75) by changing the line $F_{\mathbf{k}_{1v}}^{(h)}(\mathbf{r}_i)$ into $F_{\mathbf{k}_{1c}}^{(e)}(\mathbf{r}_i)$, $i = 1, 2, \dots, N$.

The Slater determinants can be rewritten in the more compact form :

$$\langle \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N | \psi_i \rangle = \frac{1}{\sqrt{N!}} \sum_p (-1)^p P \prod_{\mathbf{k}_{1n}} F_{\mathbf{k}_{1n}}^{(h)}(\mathbf{r}_n) \quad (76)$$

$$\langle \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N | \psi_f \rangle = \frac{1}{\sqrt{N!}} \sum_p (-1)^p F_{\mathbf{k}_{1c}}^{(e)}(\mathbf{r}_c) \prod_{\mathbf{k}_{1n} \neq \mathbf{k}_{1v}} F_{\mathbf{k}_{1n}}^{(h)}(\mathbf{r}_n) \quad (77)$$

where P is the permutation operator which exchanges the particles between themselves. The case $N = 2$ leads to

$$\langle \mathbf{r}_1, \mathbf{r}_2 | \psi_i \rangle = \frac{1}{\sqrt{2}} [F_{\mathbf{k}_{11}}^{(h)}(\mathbf{r}_1) F_{\mathbf{k}_{12}}^{(h)}(\mathbf{r}_2) - F_{\mathbf{k}_{12}}^{(h)}(\mathbf{r}_1) F_{\mathbf{k}_{11}}^{(h)}(\mathbf{r}_2)] \quad (78)$$

$$\langle \mathbf{r}_1, \mathbf{r}_2 | \psi_f \rangle = \frac{1}{\sqrt{2}} [F_{\mathbf{k}_{11}}^{(h)}(\mathbf{r}_1) F_{\mathbf{k}_{1c}}^{(e)}(\mathbf{r}_2) - F_{\mathbf{k}_{11}}^{(h)}(\mathbf{r}_2) F_{\mathbf{k}_{1c}}^{(e)}(\mathbf{r}_1)] \quad (79)$$

The dipolar matrix element of the $\psi_i \rightarrow \psi_f$ transition can therefore be written :

$$\langle \psi_i | \sum_n V_n | \psi_f \rangle = \frac{1}{N!} \sum_n \sum_{PP'} (-1)^{P+P'} \int d^3r_1 \dots d^3r_N \left[P \prod_{k_{ln}} F_{k_{ln}}^{(h)}(\mathbf{r}_n)^* \right] \times \\ \frac{ieF}{2m_0\omega} \mathbf{\epsilon} \cdot \mathbf{p}_n \left[P F_{k_{lc}}^{(e)}(\mathbf{r}_c) \prod_{k_{ln} \neq k_{lc}} F_{k_{ln}}^{(h)}(\mathbf{r}_n) \right] \quad (80)$$

Fortunately this complicated expression simplifies into [3, 19, 20] :

$$\langle \psi_i | \sum_n V_n | \psi_f \rangle = \frac{ieF}{2m_0\omega} \mathbf{\epsilon} \cdot \int d^3r F_{k_{lc}}^{(h)*}(\mathbf{r}) \mathbf{p} F_{k_{lc}}^{(e)}(\mathbf{r}) \quad (81)$$

due to

$$\langle F_{k_{ln}}^{(h)} | F_{k_{lc}}^{(e)} \rangle = \langle F_{k_{ln}}^{(h)} | \mathbf{p} | F_{k_{lc}}^{(e)} \rangle = 0 \quad (82)$$

In the case $N = 2$, equation (80) can be checked by direct computations. Using equations (78, 79), one obtains 8 terms ; of these, 6 vanish owing to equation (82) and the last 2, which are identical, cancel the $1/2!$. Finally P_{if} is simply

$$P_{if} = \frac{2\pi}{\hbar} \frac{e^2 F^2}{4 m_0^2 \omega^2} \left| \langle F_{k_{lc}}^{(h)} | \mathbf{\epsilon} \cdot \mathbf{p} | F_{k_{lc}}^{(e)} \rangle \right|^2 \delta [\epsilon_A + E_1 + HH_1 + \frac{\hbar^2 k_{lc}^2}{2 m_c} + \frac{\hbar^2 k_{lc}^2}{2 M_1} - \hbar\omega] \quad (83)$$

This expression is exactly the same as the one obtained in the one electron picture, which after all is not very surprising. If the dipolar matrix element in equation (83) were written in terms of both the dipole matrix element between u_{hh} and u_c and the overlap integral of the envelope functions, we could re-derive the polarization selection rules on the one hand and the selection rules on the envelope functions quantum numbers on the other hand.

In particular we would find that only "vertical" transitions are optically allowed, i.e. $\mathbf{k}_{\perp v} = \mathbf{k}_{\perp c}$.

There is however another way to derive this selection rule. Let us consider the translation operator \mathcal{T}_\perp which is such that

$$\mathcal{T}_\perp f(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = f(\mathbf{r}_1 + \mathbf{a}_\perp, \mathbf{r}_2 + \mathbf{a}_\perp, \dots, \mathbf{r}_N + \mathbf{a}_\perp) \quad (84)$$

where \mathbf{a}_\perp is an in-plane lattice vector. The operator \mathcal{T}_\perp and \mathcal{H}_0 commute as well as \mathcal{T}_\perp and $\sum_n V_n$. The initial state $|\psi_i\rangle$ is an eigenstate of \mathcal{T}_\perp with an eigenvalue which can always be written $\exp(i \mathbf{K}_\perp^{\text{in}} \cdot \mathbf{a}_\perp)$.

On the other hand from equation (75) we can directly obtain

$$\mathcal{T}_\perp \psi_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \exp \left[i \left(\sum_n \mathbf{k}_{\perp n} \right) \cdot \mathbf{a}_\perp \right] \psi_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (85)$$

Thus

$$\mathbf{K}_\perp^{\text{in}} = \sum_n \mathbf{k}_{\perp n} \quad (86)$$

However, for a filled band

$$\sum_n \mathbf{k}_{\perp n} = \mathbf{0} \quad \text{i.e.} \quad \mathbf{K}_\perp^{\text{in}} = \mathbf{0} \quad (87)$$

The final state is also an eigenstate of \mathcal{C}_\perp with eigenvalue $\exp(i\mathbf{K}_\perp \cdot \mathbf{a}_\perp)$. By directly applying \mathcal{C}_\perp on $|\psi_f\rangle$ we obtain

$$\mathcal{C}_\perp \psi_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \exp[i(\mathbf{k}_{\perp c} - \mathbf{k}_{\perp v}) \cdot \mathbf{a}_\perp] \psi_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (88)$$

Thus

$$\mathbf{K}_\perp = \mathbf{k}_{\perp c} - \mathbf{k}_{\perp v} \quad (89)$$

From now on we shall denote the final state by $\psi_{\mathbf{k}_{\perp c}, \mathbf{k}_{\perp c} - \mathbf{k}_\perp}$. The selection rule on the envelope function quantum number means $\mathbf{K}_\perp = \mathbf{0}$. This is not surprising in view of the commutativity between $\sum_n V_n$ and \mathcal{C}_\perp . As the initial state was characterized by $\mathbf{K}_\perp^{\text{in}} = \mathbf{0}$ it could only be coupled by $\sum_n V_n$ to an excited state with the same \mathbf{K}_\perp value, resulting in the selection rule $\delta_{\mathbf{K}_\perp, \mathbf{0}}$. In short we have obtained

$$P_{\text{if}} = \frac{2\pi}{\hbar} \frac{e^2 F^2}{4m_0^2 \omega^2} |\langle u_{\text{hh}} | \mathbf{e} \cdot \mathbf{p} | u_c \rangle|^2 |\langle \chi_1^{(\text{h})} | \chi_1^{(\text{c})} \rangle|^2 \delta_{\mathbf{K}_\perp, \mathbf{0}} \times \\ \delta \left[\epsilon_A + E_1 + HH_1 + \frac{\hbar^2 k_{\perp c}^2}{2m_c} \left(\frac{1}{m_c} + \frac{1}{M_1} \right) - \hbar\omega \right] \quad (90)$$

From such an expression, one could easily derive the staircase-like absorption spectrum, equation (51) which is characteristic of subband \rightarrow subband interband transitions in type I quantum wells.

II.2 ABSORPTION IN THE PRESENCE OF ELECTRON-ELECTRON INTERACTIONS. — In the presence of Coulombic terms, single Slater determinants are no longer exact eigenstates of \mathcal{H}_0 . However they remain the best single determinant approximation of the ground state if the one-electron functions $F_{\mathbf{k}_{\perp v}}^{(\text{h})}(\mathbf{r})$ are solutions of the coupled integro-differential Hartree-Fock equations (3). We keep this approximation for $|\psi_i\rangle$. For the final state a single determinant is also not an exact solution. If we keep a single determinant of the form given by equation (77), we shall obtain a Hartree-Fock solution which describes a situation where a valence electron, uniformly spread over the area S , is replaced by a conduction electron whose charge density is also uniformly spread over S . The removal of a valence electron hardly affects the self-consistent Hartree-Fock potential experienced by the remaining $N-1$ valence

electrons (the charge density is of the order of $\frac{1}{S}$). Similarly the single conduction electron only slightly modifies the self-consistent potential. This results in an excited energy which differs from the ground state energy by almost (i.e. to the order of 0 ($\frac{1}{S}$)) the same energy $\Delta\varepsilon$ as calculated in equation (67). A lower lying excited state can be constructed by expanding $|\psi_f\rangle$ on the set of the single determinants $|\psi_{k_{\perp c}, k_{\perp c} - K_{\perp}}\rangle$. Because \mathcal{H}_0 commutes with \mathcal{G}_{\perp} we know that such an expansion can be restricted to a summation over $k_{\perp c}$, K_{\perp} being a good quantum number. Thus, even with the inclusion of electron-electron interactions (within the Hartree-Fock approximation) we can write

$$\langle \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N | \psi_f \rangle = \sum_{k_{\perp c}} A(k_{\perp c}) \psi_{k_{\perp c}, k_{\perp c} - K_{\perp}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (91)$$

If we now reason in terms of electron-hole states instead of $(N - 1) + 1$ electrons states, the wavefunction $\psi_{k_{\perp c}, k_{\perp c} - K_{\perp}}$ corresponds to an electron-hole pair with a total wavevector $k_{\perp c} - (k_{\perp c} - K_{\perp}) = K_{\perp}$ (since the hole wavevector is minus that of the missing electron). Thus, the physical meaning of equation (91) is that we are looking for an excited state of the heterostructure which consists of a wave packet of electron-hole pairs, which have the same total wavevector K_{\perp} . This wave packet is the exciton. The exciton can be seen either as a particular excitation of $(N - 1) + 1$ electrons or as a particular combination of electron-hole pairs. The exciton is a truly *delocalized* excitation. This can be checked by calculating the average in-plane electronic density $\frac{1}{N} \sum_m \delta(\mathbf{r}_{\perp} - \mathbf{r}_{\perp m})$ in the excited state equation (91) and by comparing it with that found in the ground state, equation (75). Both are found equal to $\frac{1}{S}$ (if only the slow spatial variations which occur at the scale of the envelope function are retained). What may give rise to localized states in the exciton are the internal degrees of freedom of the pairs. This is due to the attraction between the electron and the hole (or equivalently to the diminished electron-electron repulsion) in the excited state equation (91). The electron-hole attraction, together with the restricted class of electron-hole pairs that we retain in equation (91) (i.e. which all have the same total wavevector K_{\perp}), can be described in a better way if we use the exciton amplitude

$$\beta(\mathbf{r}_{\perp}) = \frac{1}{\sqrt{S}} \sum_{k_{\perp c}} A(k_{\perp c}) \exp(i k_{\perp c} \cdot \mathbf{r}_{\perp}) \quad (92)$$

The physical interpretation of $\beta(\mathbf{r}_{\perp})$ is that $|\beta(\mathbf{r}_{\perp})|^2$ is the areal probability density of finding the electron and the hole separated from each other by \mathbf{r}_{\perp} , while having a total wavevector K_{\perp} . Equations (91, 92) are the starting points of the exciton formalism, fully developed in, for instance, Knox's text book [20]. For our purposes, it is sufficient to know that $\beta(\mathbf{r}_{\perp})$ (after considerable manipulations and numerous assumptions) is the solution of

$$\left[\frac{p_{\perp}^2}{2\mu} - \frac{e^2}{\kappa} \iint \frac{dz_e dz_h |\chi_1^{(e)}(z_e)|^2 |\chi_1^{(h)}(z_h)|^2}{\sqrt{r_{\perp}^2 + (z_e - z_h)^2}} + \epsilon_A + E_1 + HH_1 \right] \times \beta(r_{\perp}) = \epsilon \beta(r_{\perp}) \quad (93)$$

where μ is the exciton reduced mass ($\mu^{-1} = m_c^{-1} + M_1^{-1}$) and κ the relative dielectric constant of the heterostructure. The energy difference between the excited state equation (91) and the ground state is equal to :

$$\Delta \epsilon = \epsilon_{\nu} + \frac{\hbar^2 K_{\perp}^2}{2(m_c + M_1)} \quad (94)$$

where ϵ_{ν} is one of the eigenvalues of equation (93). The ground state of equation (93) has been extensively discussed in chapter IV. We know that a simple trial wavefunction of this ground state is

$$\beta_{1s}(r_{\perp}) = \sqrt{\frac{2}{\pi \lambda^2}} \exp(-r_{\perp}/\lambda) \quad (95)$$

where λ is the effective Bohr radius of the quasi bi-dimensional exciton. The $1S$ exciton state lies at an energy lower than $\epsilon_g + E_1 + HH_1$ by a quantity R^* , which is the exciton binding energy. Thus, the minimum energy separation between the ground state and the excited state (characterized by the wavevector K_{\perp}) is :

$$\Delta \epsilon_{\min} = \epsilon_A + E_1 + HH_1 - R^* + \frac{\hbar^2 K_{\perp}^2}{2(m_c + M_1)} \quad (96)$$

which can be lower than $\epsilon_A + E_1 + HH_1$ if K_{\perp} is small enough.

There exists an infinite number of bound states to equation (93). They correspond to $\beta_{\nu}(r_{\perp})$ functions which decay to zero at large r_{\perp} . Unbound solutions also exist. They form a continuum and are characterized by $\beta_{\nu}(r_{\perp})$ exciton amplitudes, which extend over the whole area S , and by energies $\epsilon_{\nu} > \epsilon_A + E_1 + HH_1$. In the interacting electron-hole problem these unbound solutions are the counterparts of the plane wave solutions in the non-interacting electron-hole problem.

We now evaluate P_{if} and obtain the selection rules for the excitonic absorption from the ground state to the final state which is characterized on the one hand by K_{\perp} and on the other by a quantum number ν which labels the solutions of equation (93). This is indicated by affixing a subscript ν to the $A(k_{\perp c})$ of equation (91) :

$$\langle \psi_i | \sum_n V_n | \psi_f \rangle = \sum_{k_{\perp c}} A_{\nu}(k_{\perp c}) \langle \psi_i | \sum_n V_n | \psi_{k_{\perp c} - K_{\perp}} \rangle \quad (97)$$

The matrix element appearing on the right hand side of equation (97) has already been evaluated in paragraph (II.1). This leaves us with :

$$\langle \psi_i | \sum_n V_n | \psi_f \rangle = \frac{ieF}{2m_0 \omega} \langle u_{hh} | \epsilon \cdot p | u_c \rangle \langle \chi_1^{(h)} | \chi_1^{(e)} \rangle \times \delta_{K_{\perp}, 0} \sum_{k_{\perp c}} A_{\nu}(k_{\perp c}) \quad (98)$$

which, together with equation (92), leads to :

$$P_{\text{if}} = \frac{2\pi}{\hbar} \frac{e^2 F^2}{4m_0^2 \omega^2} |\langle u_{\text{hh}} | \mathbf{e} \cdot \mathbf{p} | u_{\text{c}} \rangle|^2 |\langle \chi_1^{(\text{h})} | \chi_1^{(\text{e})} \rangle|^2 \delta_{\mathbf{K}_\perp, \mathbf{0}} S |\beta_\nu(\mathbf{0})|^2 \delta(\varepsilon_\nu - \hbar\omega) \quad (99)$$

where ε_ν is the energy of the eigenstates $|\nu\rangle$ of equation (93). Again, if we neglect the electron-hole interaction, the result of the non-interacting electron model is recovered : in this case $\beta_\nu(\mathbf{r}_\perp)$ and ε_ν are given by

$$\beta_\nu(\mathbf{r}_\perp) = \frac{1}{\sqrt{S}} \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp); \quad \varepsilon_\nu = \frac{\hbar^2 k_\perp^2}{2\mu} + E_1 + HH_1 + \varepsilon_A \quad (100)$$

and the staircase-like shape of the absorption is readily derived.

In the case of excitonic absorption, the selection rule $\delta_{\mathbf{K}_\perp, \mathbf{0}}$ is very important since it considerably restricts the possibilities of creating (or annihilating) excitons. It expresses the delocalized nature of the exciton, an entity which corresponds to a *delocalized* Bloch state of the excited crystal, despite the fact that the exciton amplitude $\beta_\nu(\mathbf{r}_\perp)$ can be a spatially localized function of \mathbf{r}_\perp . The $\delta_{\mathbf{K}_\perp, \mathbf{0}}$ selection rule can also be associated with the conservation of the in-plane wavevector \mathbf{K}_\perp which, being zero in the ground state, must also be zero in the excited state if the dipolar approximation is used to describe the interaction between the electromagnetic wave and the carriers. Going beyond the dipolar approximation would give rise to the selection rule $\delta_{\mathbf{K}_\perp, \mathbf{Q}_\perp}$, where \mathbf{Q}_\perp is the in-plane projection of the photon wavevector.

As seen from equation (99) the excitonic absorption fulfills the *same* polarization selection rules as the band-to-band absorption model which neglects the electron-electron interaction. This is because the electron-electron interaction has more symmetry (spherical symmetry) than the heterostructure potential. Looking at table II we notice that the heavy hole excitonic transitions are forbidden (within our model) when the electromagnetic wave propagates in the layer plane and when its electric field vector is polarized along the growth (\hat{z}) axis.

An important new fact brought by the electron-hole interaction is the appearance of the term $|\beta_\nu(\mathbf{0})|^2$ in equation (99). It means that only excitons with a non-zero amplitude at $\mathbf{r}_\perp = \mathbf{0}$ can absorb light. Since the effective bi-dimensional potential which appears in equation (93) is radial, the $\beta_\nu(\mathbf{r})$ functions can be classified according to the eigenvalues of \mathbf{L} and L_z , where \mathbf{L} is the angular momentum. In addition, a radial quantum number n (discrete for the bound states, continuous for the extended states) labels the eigenstates. Thus, only the nS excitons (corresponding to $L = 0$) can be optically created.

The only situation where equation (93) admits exact solutions corresponds to a purely bi-dimensional case (i.e. $|\chi_1^{(\text{e})}(z)|^2 = |\chi_1^{(\text{h})}(z)|^2 = \delta(z)$). In this case R^* is equal to $4R_0$, where R_0 is the three-dimensional exciton Rydberg. The magnitude of $|\beta_{nS}(\mathbf{0})|^2$, and thus of the nS excitonic absorption, is proportional to $8/\pi a_0^2 (2n-1)^3$ where $n = 1, 2, \dots$ and a_0 is the bulk exciton Bohr radius [23]. This shows that the

magnitude of the nS exciton absorption decreases very rapidly with n . However the levels are more dense. Both effects combine to lead to a finite absorption when $\hbar\omega$ approaches $\epsilon_A + E_1 + HH_1$ from below. The absorption associated with the dissociated electron-hole hole pairs is also enhanced with respect to the non interacting electron model (see Fig. 7). It approaches the one-electron result when $\hbar\omega$ exceeds $\epsilon_A + E_1 + HH_1$ by many R^* .

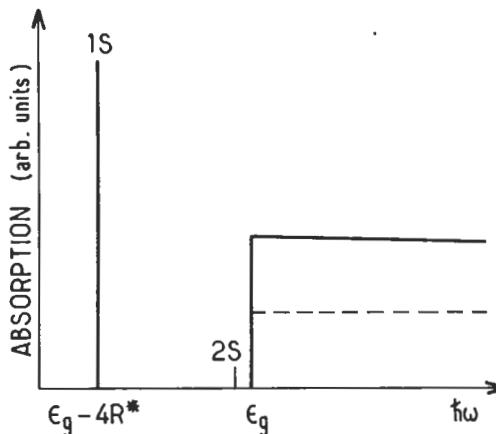


Fig. 7. — Theoretical (unbroadened) absorption spectrum of a purely bi-dimensional semiconductor (solid line). R^* is the three-dimensional exciton binding energy. Only the $1S$ and $2S$ exciton bound states contributing to the absorption have been represented. The dashed line plateau corresponds to the absorption coefficient of dissociated and uncorrelated electron-hole pairs. ϵ_g is the onset of the absorption due to the dissociated excitons.

Let us rewrite the $HH_1 \rightarrow E_1$ excitonic absorption coefficient in the quasi bi-dimensional case

$$\alpha_{HH_1 \rightarrow E_1}(\omega) = \frac{2\pi^2 e^2 E_p}{ncm_0 \omega \mathcal{L}} |\langle \chi_1^{(h)} | \chi_1^{(e)} \rangle|^2 \sum_{\nu} |\beta_{\nu}(\mathbf{0})|^2 \delta(\hbar\omega - \epsilon_{\nu}) \quad (101)$$

where $E_p, \mathcal{L}, n \dots$ have the same meaning as in paragraph I. We first of all focus our attention on the quantum well thickness (L) dependence of the magnitude of the $1S$ exciton peak. To account for the broadening mechanisms, we replace the delta functions in equation (101) by a Gaussian function (which sometimes fits the observed lineshapes better than a Lorentzian one) and we use equation (95) to evaluate $\beta_{1S}(\mathbf{0})$. For a single quantum well, we obtain the peak value

$$\mathcal{L} \alpha_{HH_1 \rightarrow E_1}^{\text{peak}} = \frac{4\pi e^2 E_p}{ncm_0 \omega \lambda^2} |\langle \chi_1^{(h)} | \chi_1^{(e)} \rangle|^2 \frac{1}{\Gamma \sqrt{2\pi}} \quad (102)$$

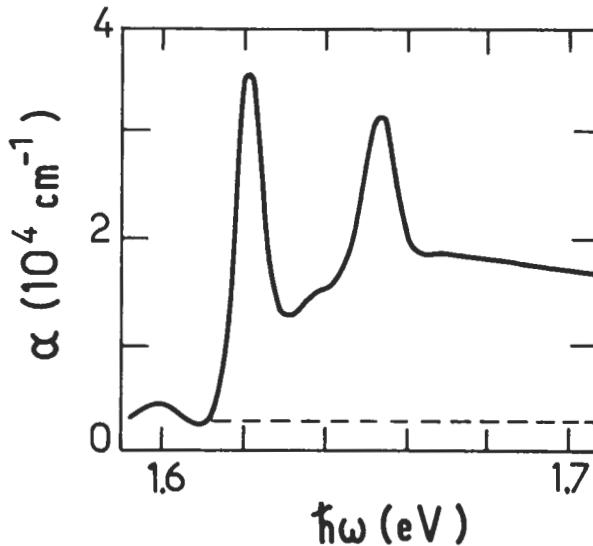


Fig. 8. — Measured absorption coefficient of a 76 Å-33 Å GaAs-AlAs multiple quantum well sample at $T = 2$ K. Between the two main absorption peaks corresponding to excitonic $E_1 - HH_1$, $E_1 - LH_1$ transitions respectively, one sees a weaker absorption due to either the $E_1 - HH_1$ 2S excitons or to the $E_1 - HH_1$ exciton continuum. After reference [25].

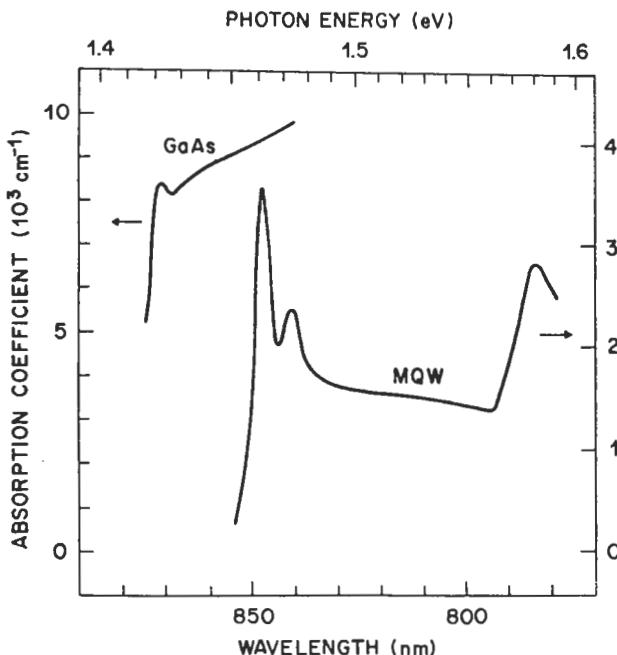


Fig. 9. — Room temperature absorption of a bulk GaAs (left scale) and GaAs- $Ga_{1-x}Al_xAs$ multiple quantum wells structure (right scale). Notice the strong excitonic resonance of the multiple quantum well. After reference [26].

where $2\Gamma\sqrt{2}$ is the full width at e^{-1} amplitude of the Gaussian peak. For a GaAs-Ga(Al)As quantum well a representative figure for $\mathcal{L}\alpha_{HH_1 \rightarrow E_1}^{\text{peak}}$ is obtained with $n = 3.6$, $\hbar\omega = 1.6 \text{ eV}$, $E_p = 23 \text{ eV}$, $\lambda = 70 \text{ \AA}$, $\Gamma\sqrt{2\pi} = 3 \text{ meV}$ and $\langle\chi_1^{(h)}|\chi_1^{(e)}\rangle \simeq 1$. We obtain $\mathcal{L}\alpha_{HH_1 \rightarrow E_1}^{\text{peak}} \sim 0.19$. This is ~ 32 times larger than the band to band plateau value obtained in the non-interacting electron model of the $HH_1 \rightarrow E_1$ absorption. Precise absorption measurements of the continuum edge are however difficult since it is often (energy) located in the low energy tail of the $1S E_1 - LH_1$ exciton.

When the quantum well thickness varies, two factors affect the peak value of the $1S$ excitonic absorption.

- * λ decreases with decreasing L . This strengthens the excitonic absorption and witnesses an increase in the bi-dimensionality of the exciton. Correlatively the energy distance between the $1S$ exciton peak and the continuum increases.

- * the broadening parameter Γ also increases. As shown by Weisbuch *et al.* [24] the layer thickness fluctuations make Γ to vary like L^{-3} .

Thus, depending on which mechanism prevails, the peak value of the excitonic absorption may increase or decrease. Masumoto *et al.* [25] have recently reported a complete study of the thickness dependence of the excitonic absorption in GaAs-Ga(Al)As multiple quantum wells.

The exciton continuum absorption ($\hbar\omega > \varepsilon_A + E_1 + HH_1$) is also enhanced by the electron-hole attraction compared with that calculated with a non-interacting electron model. Note however that this enhancement cannot be explicitly calculated because the solutions of the Schrödinger equation for the quasi bi-dimensional exciton equation (93) are not known analytically. Experimentally, a small hump is often visible in high quality GaAs quantum wells. It is located in energy between the $1S HH_1 - E_1$ and $1S LH_1 - E_1$ exciton peaks (Fig. 8) and can be attributed either to the $2S HH_1 - E_1$ broadened exciton absorption [21] or to the onset of the $HH_1 \rightarrow E_1$ exciton continuum [22]. These two energies are so close (1-2 meV) that the small hump probably arises from both of them.

Experimentally, the excitonic peak structures which appear on the low energy side of the continuum plateaus have been observed in all undoped quantum well systems of high quality (see Figs. 9-12), even in materials which hardly show exciton peaks in the bulk (e.g. $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$). This is due to the enhancement of the exciton binding energy in quasi bi-dimensional materials.

The same effect explains the persistence of pronounced excitonic structures in the room temperature absorption spectra of GaAs-Ga(Al)As multiple quantum wells. Chemla and co-workers [26] were able to describe the temperature dependence of the excitonic peak width by the formula

$$\Gamma(T) = \Gamma_0 + \frac{\Gamma_1}{-1 + \exp\left(\frac{\hbar\omega_{\text{LO}}}{k_b T}\right)} \quad (103)$$

where Γ_0 is an inhomogeneous broadening parameter associated with interface fluctuations and Γ_1 another broadening parameter which corresponds to the exciton-

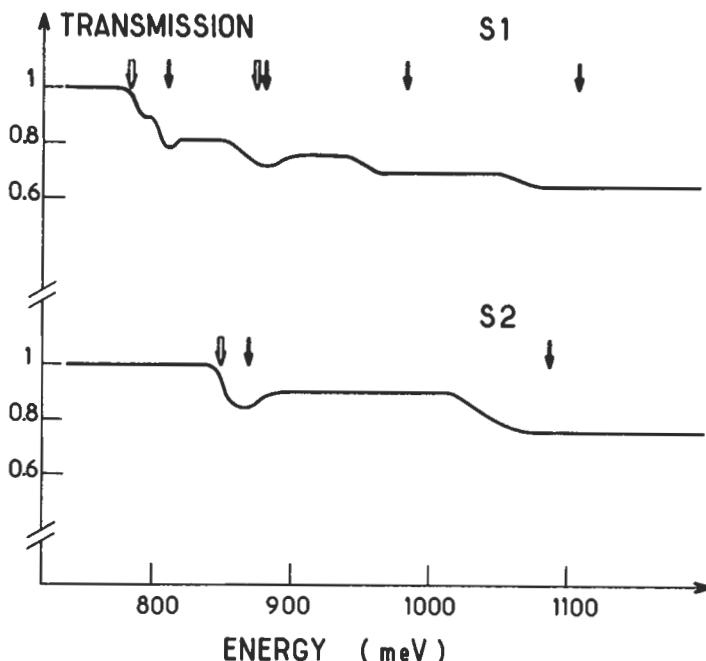


Fig. 10. — Energy dependence of the transmission coefficient in GaSb-AlSb superlattices at $T = 2$ K. The minima correspond to excitonic transitions between different valence and conduction subbands. After reference [15].

longitudinal optical phonon interaction. The latter mechanism is considered as being predominant as soon as the optical phonons are thermally activated. In the denominator of equation (103), $\hbar\omega_{LO} = 36$ meV is the GaAs longitudinal optical phonon energy. The temperature dependence of $I(T)$ arises entirely from the thermal population of phonons. For their 102 Å thick GaAs wells, Chemla *et al.* found that the best fit of their data to equation (103) was obtained for $\Gamma_0 = 2$ meV and $\Gamma_1 = 5$ meV.

Since excitons play an important role in the optical absorption, it is interesting to find out what happens to them once they have been optically created [26]. We consider the ideal case where finite lifetime effects can be neglected and where there is only a single "chemical reaction" between the excitons (all assumed to be in their bound 1S states) and the electrons and holes which arises from the thermal dissociation of the excitons.

The chemical reaction is



The areal densities of excitons, electrons and holes are denoted by n_x , n_e , n_h respectively and we assume that at the initial time a density N_0 of excitons was created in the 1S state by resonant excitation. After a short transient ($\sim 10^{-12}$ s) a

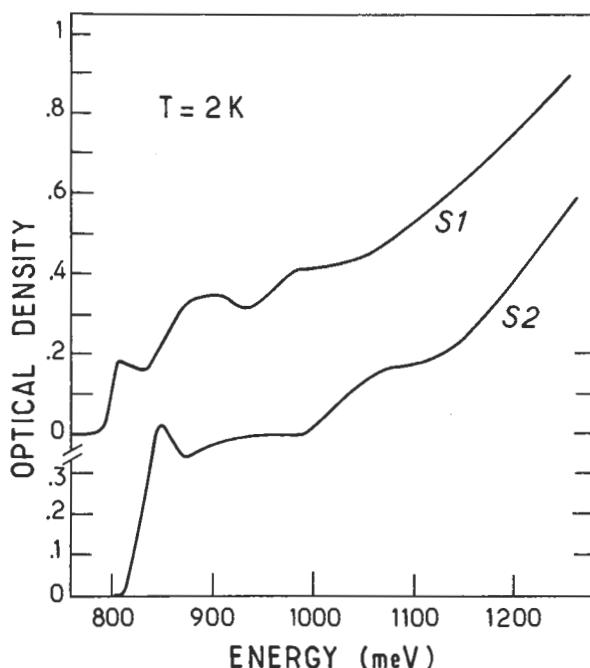


Fig. 11. — Excitonic absorption of $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ -InP superlattices at $T = 2$ K. After reference [13].

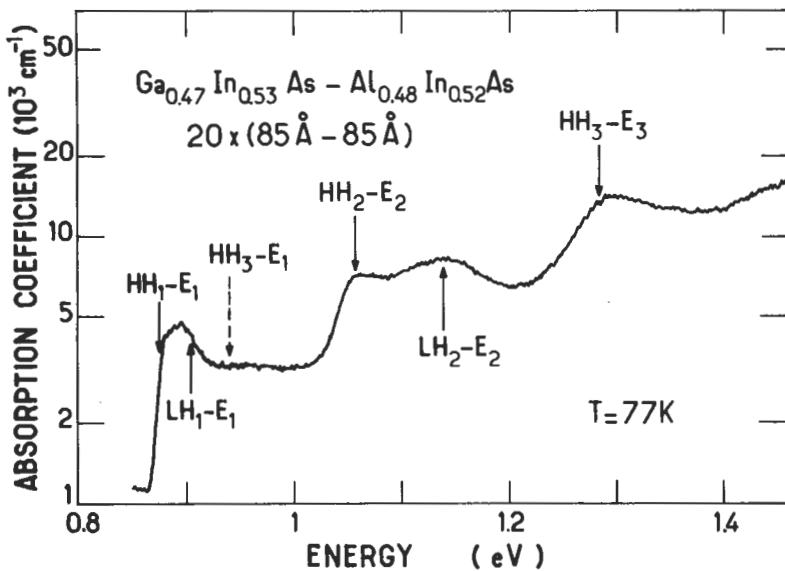


Fig. 12. — Excitonic absorption of $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ - $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$ superlattices at $T = 77$ K. Courtesy J.Y. Marzin.

thermal equilibrium between the different species is established. The assumption of infinite lifetime enables us to write the electrical neutrality equation

$$n_e = n_h \quad (105)$$

as well as the law of conservation of matter

$$N_0 = n_x + n_e \quad (106)$$

On the other hand, the assumption of thermal equilibrium leads to the law of mass action

$$\frac{n_e n_h}{n_x} = K(T) \quad (107)$$

The constant $K(T)$ can be expressed in terms of the one-particle partition functions ξ_e , ξ_h , ξ_x of the different species. If we assume that each species is diluted enough to be treated as a classical ideal gas, we obtain after some calculations :

$$K(T) = \frac{m_e M_1}{m_e + M_1} \frac{k_B T}{\pi \hbar^2} \exp \left(- \frac{R^*}{k_B T} \right) \quad (108)$$

The fraction of dissociated excitons is then

$$f = \frac{n_e}{N_0} = \frac{1}{2N_0} [-K(T) + \sqrt{K^2(T) + 4N_0 K(T)}] \quad (109)$$

When the number of photocreated excitons goes to zero $f \rightarrow 1$. On the other hand N_0 cannot be arbitrarily large : the maximum N_0 is $\sim (\pi \lambda^2)^{-1}$, where λ is the characteristic in-plane extension of the exciton amplitude. Beyond $(\pi \lambda^2)^{-1}$, the excitons would interact strongly and ultimately ionize to give rise to an electron-hole plasma. As an order of magnitude, let us take $\lambda = 70 \text{ \AA}$, $R^* = 10 \text{ meV}$, $k_B T = 25 \text{ meV}$, $M_1 = 0.4 m_0$ and $m_e = 0.07 m_0$. We obtain $K(T) \sim 4.2 \times 10^{11} \text{ cm}^{-2}$ and $f \sim 0.54$. Thus, the exciton dissociation is very important at room temperature according to this simple model, particularly at low excitation. It might even be asked if the presence of electrons and holes is unlikely, *via* screening effects, to accelerate the dissociation : the free carriers would reduce R^* by screening the excitons, which in turn would generate more free electrons and holes etc...

In their work on the optical properties of excitons in quantum wells at room temperature, Chemla *et al.* [26] have also investigated the non-linear absorption at the excitonic resonance. From the point of view of the absorption saturation, quantum wells present two interesting features :

- i) The exciton binding energy is larger than in the corresponding bulk materials. Consequently the excitonic transitions have an important oscillator strength in a narrow frequency range.
- ii) At room temperature, the excitons are easily dissociated into free electrons and holes which efficiently screen the attractive coulombic potential, and consequently participate in the quenching of the exciton peak.

As a result, the non-linear response of GaAs quantum wells is enhanced with respect to bulk GaAs, as shown in figure 13.

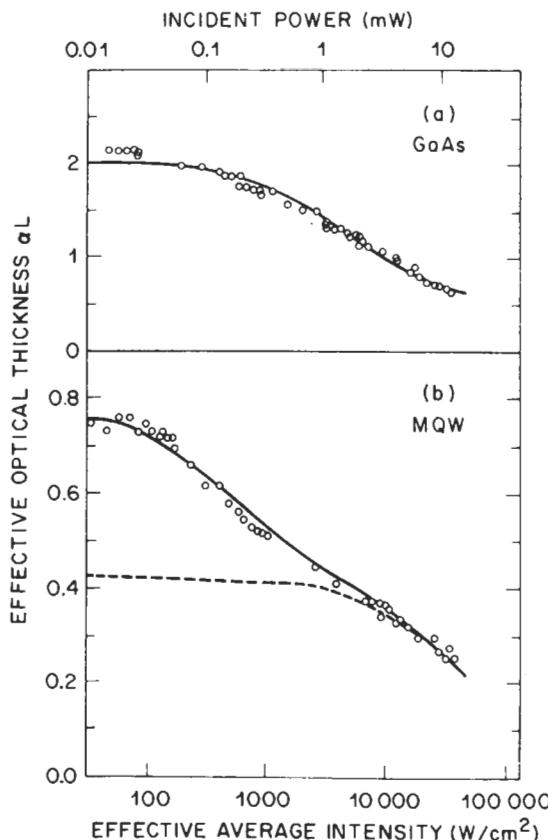


Fig. 13. — Non linear absorption coefficient of bulk GaAs and GaAs-Ga_{1-x}Al_xAs multiple quantum wells. After reference [26].

III. Photoluminescence of quasi bi-dimensional systems.

III.1 INTRODUCTION. — The optical spontaneous emission, for example photoluminescence, cannot be predicted when the electromagnetic field is described classically. However, it is possible to relate its intensity exactly to that of absorption by means of the Einstein coefficients [1, 2, 27]. Here, we are interested in presenting the theoretical considerations on the lineshape of some recombination processes in semiconductor heterostructures rather than calculating the photoluminescence intensities. When possible, these considerations will be illustrated by examples. In contrast with a widespread belief, the information one can extract from the

photoluminescence concerning the energy levels of a given heterostructure is often very scarce and depends to a large extent on an *a priori* knowledge of that heterostructure. It is true that photoluminescence experiments are much easier to perform than absorption experiments. It is also true that they are much more complicated to interpret.

Absorption and photoluminescence seem to be, at first sight, symmetrical processes. On the one hand a photon is absorbed promoting an electron from the level $|i\rangle$ to the level $|f\rangle$ and photons which have not been absorbed are detected. On the other hand, the emission involves the transition $|i\rangle \rightarrow |f\rangle$ with $\epsilon_i > \epsilon_f$ and the photons which are emitted at the energy $\epsilon_i - \epsilon_f$ can be observed. Such a process is possible only if the system has been initially excited in the state $|i\rangle$. Thus, it is not at equilibrium and one of the ways in which this can be reached is by optical emission. The radiative channel is in competition with the non radiative relaxation processes (phonon emission, capture by deep centres, Auger effect etc...) which send the excited carriers to lower states from which they can emit photons or relax non radiatively etc... It appears therefore that luminescence is different from absorption because, instead of having a 100 % efficiency (as in the case of absorption where one absorbed photon creates one electron-hole pair), spontaneous emission is only one of the mechanisms which might occur. This is a question of the lifetimes of the level $|i\rangle$ with respect to the radiative and non-radiative relaxations. To illustrate these remarks let us consider the four-level system shown in figure 14. The system is pumped by $|1\rangle \rightarrow |4\rangle$ transitions and we shall only consider the luminescence signal of levels $|3\rangle$ and $|2\rangle$ at energies $\epsilon_3 - \epsilon_1$ and $\epsilon_2 - \epsilon_1$ respectively. The parameters τ_{ij} and T_{ii} are time constants which characterize inter-excited levels non radiative channels and radiative channels from level $|i\rangle$ to level $|1\rangle$ respectively. The effect of the final state population will be neglected. As level $|3\rangle$ is assumed to be populated non-radiatively from $|4\rangle$, it suffices, since levels $|1\rangle$ and $|4\rangle$ are of no interest, to write the rate equations

$$\frac{dn_3}{dt} = \frac{n_0}{\tau_p} - \frac{n_3}{\tau_{32}} - \frac{n_3}{T_{31}} \quad (110)$$

$$\frac{dn_2}{dt} = + \frac{n_3}{\tau_{32}} - \frac{n_2}{\tau_{21}} - \frac{n_2}{T_{21}} \quad (111)$$

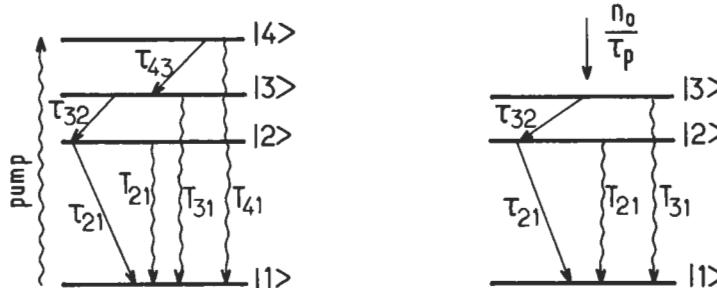


Fig. 14.

In the steady state, $dn_i/dt = 0$ and therefore

$$n_3 = \frac{n_0}{\tau_p} \frac{T_{31}}{1 + T_{31}/\tau_{32}} \quad (112)$$

$$n_2 = \frac{n_0}{\tau_p} \frac{T_{31}}{1 + T_{31}/\tau_{32}} \frac{T_{21}}{\tau_{32}} \frac{1}{1 + T_{21}/\tau_{21}} \quad (113)$$

The luminescence signals at energies $\hbar\omega_{31} = \epsilon_3 - \epsilon_1$ and $\hbar\omega_{21} = \epsilon_2 - \epsilon_1$ are respectively proportional to n_3 and n_2 . It is easy to see from equations (112, 113) that, if $\tau_{32} \ll T_{31}$, almost all the population of level $|3\rangle$ is used to populate $|2\rangle$ and does not participate in the emission which occurs at $\hbar\omega_{31}$. Even if the density of states of level $|3\rangle$ is more important (or, in our "atomic" scheme, if its degeneracy is larger) than that of $|2\rangle$, the spontaneous emission from $|3\rangle$ can be considerably smaller than that from $|2\rangle$. This is to be compared with an absorption experiment which would give a larger absorption at energy $\hbar\omega_{31}$ than at energy $\hbar\omega_{21}$ if the oscillator strengths of the two transitions are the same and the degeneracy of $|3\rangle$ is larger than that of $|2\rangle$. In fact, in bulk semiconductors at low temperature, the luminescence very often involves impurity levels (impurity-to-band, donor-acceptor transitions...) while the density of states of the impurity levels is much smaller (by several orders of magnitude) than the density of states of the free states in the bands. These impurity levels, which are so prevalent in the photoluminescence, are almost invisible in absorption spectra. While the absorption is sensitive to the density of states, the photoluminescence gives information on this quantity which is completely distorted by the relaxation effects. This occurs to such an extent that it favours the lowest lying excited states of the materials, even if their density of states is very small.

Let us now consider the transient phenomena (i.e. time-resolved photoluminescence) and assume that the pump rate of level $|3\rangle$ is described by a function $g(t)$ which is arbitrary, except that $g(t)$ vanishes if $t < 0$. If we denote the quantity $T_{21}^{-1} + \tau_{21}^{-1} - T_{31}^{-1} - \tau_{32}^{-1}$ by Ω_{23} we obtain the time dependences

$$n_2(t) = \frac{1}{\tau_{32} \Omega_{23}} \left[\exp(-t/T_3) \int_{-\infty}^t g(t') \exp(t'/T_3) dt' - \exp(-t/T_2) \int_{-\infty}^t g(t') e^{t'/T_2} dt' \right] \quad (114)$$

$$n_3(t) = \exp(-t/T_3) \int_{-\infty}^t g(t') \exp(t'/T_3) dt' \quad (115)$$

where : $T_3^{-1} = T_{31}^{-1} + \tau_{32}^{-1}$; $T_2^{-1} = T_{21}^{-1} + \tau_{21}^{-1}$ (116)

In the case where the pump state is the delta spike $n_0 \delta(t)$ we obtain

$$n_2(t) = \frac{n_0 Y(t)}{\tau_{32} \Omega_{23}} [\exp(-t/T_3) - \exp(-t/T_2)] \quad (117)$$

$$n_3(t) = n_0 Y(t) \exp(-t/T_3) \quad (118)$$

The population $n_2(t)$ first increases with time, (as long as the filling of $|2\rangle$ from state $|3\rangle$ is larger than the radiative and non-radiative depopulations), and then decreases (see Fig. 15). It is very important to notice that the luminescence decay time at energy $\epsilon_2 - \epsilon_1$ only yields the *total* lifetime T_2 of the excited level. This involves both the radiative and non-radiative lifetimes. The time-resolved luminescence experiments thus need to be cautiously interpreted if conclusions are to be drawn on the radiative lifetimes of the excited levels. In addition, level $|3\rangle$ is never pumped at a rate $n_0 \delta(t)$. The actual function $g(t)$ displays an apparatus-limited time width, as does the detection system which counts the photons emitted at the energy $\epsilon_2 - \epsilon_1$ or $\epsilon_3 - \epsilon_1$. This implies that deconvolutions of the experimental signal $n_2(t)$ are required in order to distinguish between the "true" $n_2(t)$ and the time-dependences caused by instrument limitations.

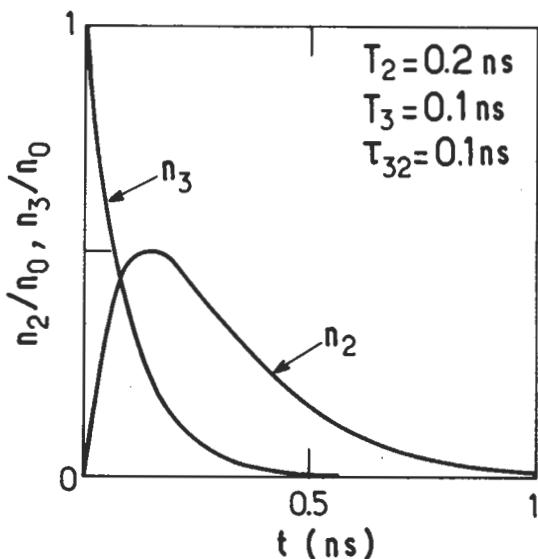


Fig. 15. — Calculated time evolutions of the n_2 and n_3 populations in a three levels system. The level $|3\rangle$ has been assumed to be populated by a $\delta(t)$ pulse.

Excitation spectroscopy

A technique closely associated with photoluminescence is the photoluminescence excitation spectroscopy. The detection spectrometer is set at some energy inside the emitted photoluminescence band ($\epsilon_2 - \epsilon_1$ in the "atomic" model of Fig. 16) and the energy of the exciting light is scanned. As a result, the various excited levels of the solids $|2\rangle, |3\rangle\dots$ are populated at rates g_2, g_3, \dots respectively. These rates are proportional to the absorption coefficients $\alpha(\epsilon_2), \alpha(\epsilon_3)\dots$ Once populated, the excited levels relax either radiatively or non radiatively towards lower energies. In

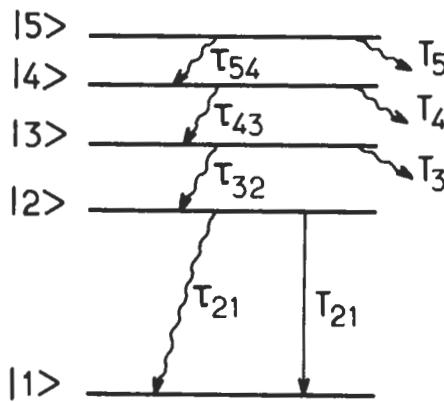


Fig. 16.

particular, a fraction of their populations ends up in the lowest lying excited state $|2\rangle$. In the steady state and for the situation depicted in figure 16 one obtains

$$n_2 = \frac{g_2}{T_{21}^{-1} + \tau_{21}^{-1}} \quad (119)$$

for an excitation energy equal to $\varepsilon_2 - \varepsilon_1$,

$$n_2 = \frac{g_3}{(T_{21}^{-1} + \tau_{21}^{-1})(1 + \tau_{32} T_3^{-1})} \quad (120)$$

for an excitation energy equal to $\varepsilon_3 - \varepsilon_1$, and

$$n_2 = \frac{g_4}{(T_{21}^{-1} + \tau_{21}^{-1})(1 + \tau_{32} T_3^{-1})(1 + \tau_{43} T_4^{-1})} \quad (121)$$

$$n_2 = \frac{g_5}{(T_{21}^{-1} + \tau_{21}^{-1})(1 + \tau_{32} T_3^{-1})(1 + \tau_{43} T_4^{-1})(1 + \tau_{54} T_5^{-1})} \quad (122)$$

for excitation energies equal to $\varepsilon_4 - \varepsilon_1$ and $\varepsilon_5 - \varepsilon_1$ respectively. In equations (119-122) τ_{ij} is now the relaxation time associated with the non-radiative path $i \rightarrow j$ and T_i is the decay time due to all the mechanisms which empty $|i\rangle$, apart from those which contribute to the population of $|i-1\rangle$.

The previous model is too crude to be applied as such to semiconductor heterostructures but it helps us to understand the different kinds of information given by the photoluminescence excitation spectra. It should be noticed that the photoluminescence signal at energy $\varepsilon_2 - \varepsilon_1$, which is proportional to n_2 has an amplitude which is governed by two competing factors :

- the absorption coefficients of the exciting light of energy $\varepsilon_2 - \varepsilon_1$ via the rates g_i .

ii) the relative orders of magnitude of the τ_{ij} 's and the T_i 's

In the limiting case where all the τ_{ij} 's are much shorter than the T_i 's, the excitation spectrum gives information equivalent to that provided by absorption (see Fig. 17a). The advantage of the former technique over the latter is its sensitivity. Like photoluminescence, excitation spectroscopy does not require thick samples, e.g. it can easily be performed on a single GaAs quantum well, whereas the absorption experiments, as shown previously, require multiple quantum wells.

The excitation spectroscopy is moreover selective, allowing the physical origin of the photoluminescence signal in the heterostructure to be traced back. Let us suppose that the photoluminescence band contains several peaks. By setting the detection wavelength at each of these peaks, one obtains excitation spectra whose shape is characteristic of the various excited levels which, after relaxation, ultimately

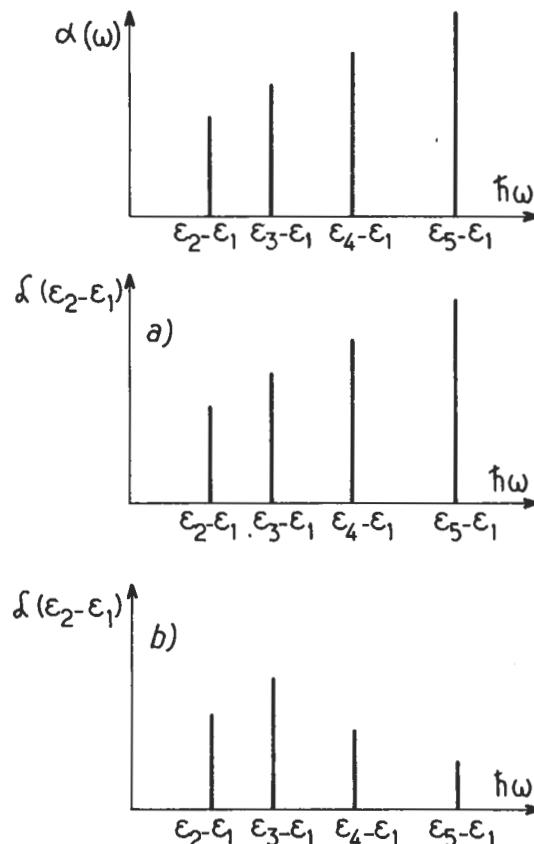


Fig. 17. — Comparison between absorption and luminescence excitation spectra in a five levels system. Upper figure : absorption coefficient ; middle and lower figures : luminescence excitation spectra at $\epsilon_2 - \epsilon_1$. case a) $\tau_{ij} \ll T_i$: the excitation spectrum mimics the absorption spectrum ; case b) resonant relaxation $\tau_{32} \ll T_3$ and $\tau_{ij} = T_i$, $i \neq 3$.

give rise to these different photoluminescence peaks. If the excitation spectra are independent of the detection wavelength, one may safely assume that the whole luminescence band has a single physical origin. If, on the other hand, the excitation spectra are markedly different, one can conclude that the various photoluminescence lines have different physical origins. As is often the case in heterostructures, these differences arise from the different locations of the recombining levels. This case is exemplified in figure 18. The maximum of the photoluminescence signal, as well as its high energy tail, originates from a modulation-doped GaAs well, whereas the two low energy peaks in the photoluminescence band arise from the photoluminescence signal of the GaAs buffer layer grown prior to the doped GaAs well.

In another limiting case, the excitation spectrum is very different from the absorption spectrum. This happens when, for instance, the τ_{ij} 's are all of the same order of magnitude or larger than the T_i 's but one, say τ_{32} , which is much shorter than T_3 (resonant relaxation from $|3\rangle$ to $|2\rangle$). In this case, the excitation spectrum shows a peak at $\epsilon_3 - \epsilon_1$, which is not found in the absorption spectrum (see Fig. 17b). This peak arises because the levels $|4\rangle$ and $|5\rangle$, once populated, contribute little to the population of $|3\rangle$ and $|2\rangle$, leading to a reduced photoluminescence signal (at

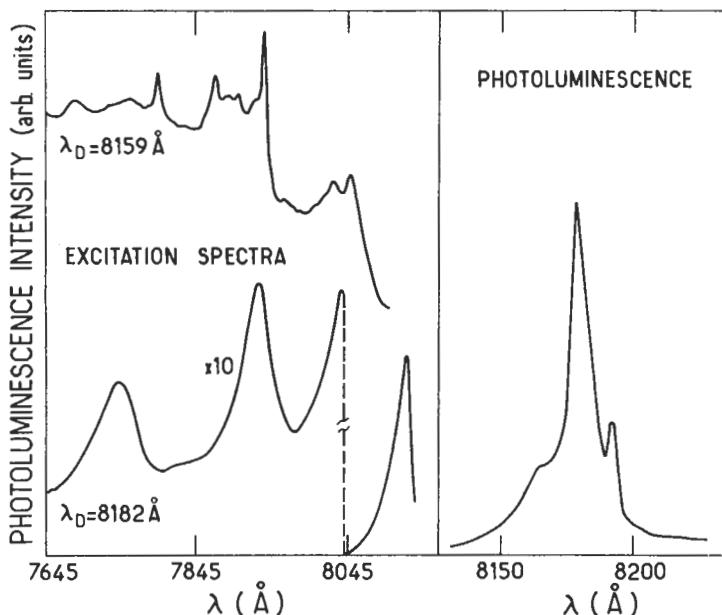


Fig. 18. — Photoluminescence (right panel) and luminescence excitation spectra (left panel) of a modulation doped GaAs-Ga_{0.61}Al_{0.39}As quantum well at $T = 2$ K. Setting the detecting spectrometer on the low energy side of the photoluminescence line reveals a bulk GaAs behaviour, i.e. oscillations due to resonant relaxation of the conduction electrons. The excitation spectrum taken on the high energy side of the photoluminescence line is characteristic of the GaAs quantum well. Courtesy M.H. Meynadier, J. Orgonasi and C. Delalande.

energy $\epsilon_2 - \epsilon_1$) over what could have been inferred uniquely from the rates g_4, g_5 .

Thus excitation spectra *may* or *may not* be equivalent to absorption spectra. Figure 19 exemplifies the resonant relaxation regime. This figure shows the excitation spectrum of a *single* GaAs-Ga(Al)As modulation-doped heterojunction at low temperature [28]. There is no significant luminescence band involving the two-dimensional electron gas confined near the GaAs-Ga(Al)As hetero-interface. This is due to the very quick relaxation of holes far away in the acceptor depletion length. Strong luminescence peaks are associated with the radiative recombination of the bulk GaAs buffer layer. Setting the detection spectrometer on one of these peaks produces the excitation spectrum shown in the upper panel of figure 19. It consists of a narrow peak associated with the bulk GaAs free exciton and several structures which are more or less equidistant. These structures arise whenever the absorbed light creates an electron whose energy differs by $n\hbar\omega_{LO}$ from the conduction band edge, where $n = 1, 2, \dots$ and $\hbar\omega_{LO}$ is the energy of the longitudinal optical phonon

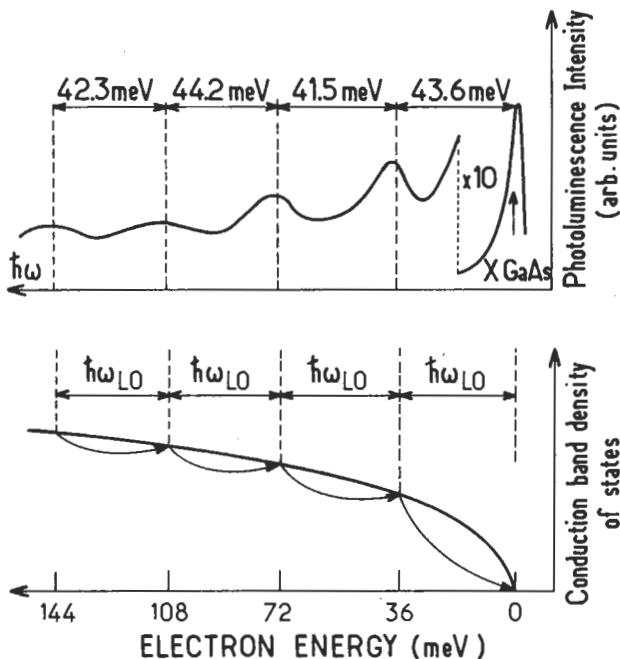


Fig. 19. — Luminescence excitation spectrum of a high quality GaAs-Ga_{0.7}Al_{0.3}As modulation-doped single heterojunction at $T = 2$ K. The detection spectrometer is set on one of the luminescence lines of the GaAs epilayer (or substrate). The oscillations seen in the excitation spectrum are not associated with the singularities of the GaAs absorption coefficient but reflects the resonant electron relaxations which take place any time a photon creates an electron hole pair such that the electron energy is $n\hbar\omega_0$, $n = 1, 2, \dots$ above the GaAs conduction band edge. Courtesy M.H. Meynadier and C. Delalande.

(see Fig. 19 lower panel)... The resonant photon energies which produce structures in the excitation spectrum are :

$$\hbar\omega = \varepsilon_g + n\hbar\omega_{LO}(1 + m_c/m_{lh}) \quad (123)$$

$$\hbar\omega = \varepsilon_g + n\hbar\omega_{LO}(1 + m_c/m_{hh}) \quad (124)$$

where m_c , m_{lh} , m_{hh} are the electron, light hole, heavy hole effective masses respectively and ε_g the GaAs bandgap.

The other limiting regime, where the excitation spectrum roughly mimics the absorption spectrum, is exemplified is the GaAs-Ga(Al)As quantum wells and superlattices (see Fig. 20). In the vast majority of these heterostructures any trace of peaks or structures in the excitation spectra can be fairly associated with an absorption process. The lack of structures associated with optical phonons, which are

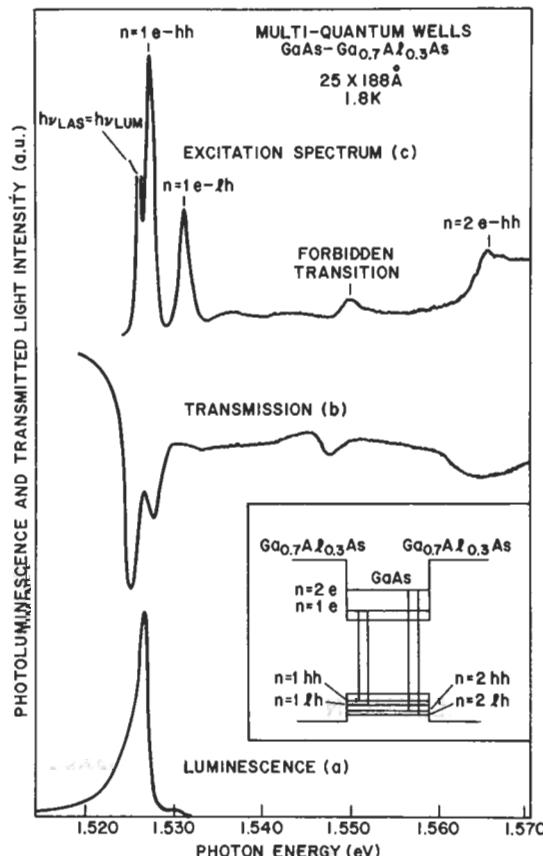


Fig. 20. — Comparison between the luminescence (a), transmission (b) and luminescence excitation spectra (c) of a multiple GaAs-Ga_{0.7}Al_{0.3}As quantum well. $T = 1.8$ K. After reference [24].

so prevalent in bulk GaAs (see above), has not been clearly explained to our knowledge. It calls either for a strongly diminished electron-phonon coupling in quantum wells or for a very fast carrier-carrier interaction which would overshadow the resonant emission.

III.2 QUANTUM WELL LUMINESCENCE (STEADY STATE). — Through the Einstein relationships the spontaneous emission can be related to the absorption. However since we are only interested in the lineshape of the luminescence spectrum and not in its absolute magnitude, we shall discard all the proportionality constants and treat the spontaneous emission and absorption on the same footing. We shall denote by $f(\varepsilon_i)$ and $f(\varepsilon_f)$ the stationary distribution functions of the initial ($|i\rangle$) and final ($|f\rangle$) states. Thus the luminescence signal $\mathcal{L}(\omega)$ associated with that particular $|i\rangle \rightarrow |f\rangle$ transition will be proportional to :

$$\mathcal{L}_{i \rightarrow f}(\omega) \propto P_{if} f(\varepsilon_i) [1 - f(\varepsilon_f)] \quad (125)$$

where P_{if} is the transition probability per unit time that the carrier undergoes a transition $|i\rangle \rightarrow |f\rangle$ due to the effect of the (dipolar) coupling Hamiltonian between the light and the carrier

$$P_{if} = \frac{2\pi}{\hbar} \frac{e^2 F^2}{4m_0^2 \omega^2} |\langle i | \mathbf{e} \cdot \mathbf{p} | f \rangle|^2 \delta(\varepsilon_i - \varepsilon_f - \hbar\omega) \quad (126)$$

The optical transition $|i\rangle \rightarrow |f\rangle$ is symmetrical to the absorption one ($|i\rangle \rightarrow |f\rangle$) and therefore follows the *same selection rule*. As for the absorption, the dipole matrix element can be split into two parts (see Eqs. (24-26)) ¹²

$$\langle i | \mathbf{e} \cdot \mathbf{p} | f \rangle = \langle u_i | \mathbf{e} \cdot \mathbf{p} | u_f \rangle \int d^3r f_i^*(\mathbf{r}) f_f(\mathbf{r}) \quad (127)$$

The first matrix element on the right hand side of equation (127) gives the selection rule on the polarization of the emitted light (in the layer plane of the layers or along the growth axis, see Tab. II) and the overlap integral selects the subband indices which govern interband recombination.

In a quantum well with band edge profiles which are symmetric in z we find that, as with absorption, the only optical allowed transitions between subbands preserve the parity of the z dependent envelope functions. Thus, using the same labelling as in paragraphs I and II, $n + m$ should be even. In addition, if the broadening is neglected, the radiative recombination occurs with conservation of the in-plane wavevector of the carrier ($\mathbf{k}_\perp^{(i)} = \mathbf{k}_\perp^{(f)}$). Similarly, in superlattices the optical transitions take place vertically in the superlattice Brillouin zone ($\mathbf{k}_\perp^{(i)} = \mathbf{k}_\perp^{(f)}$); $q_z^{(i)} = q_z^{(f)}$.

The explicit form of the stationary distribution functions $f(\varepsilon_i)$ and $f(\varepsilon_f)$ is often not known. Thus, they are (usually) *taken* as being Fermi Dirac distribution functions characterized by T_e, μ_e for conduction electrons and T_h, μ_h for the valence electrons :

$$f(\varepsilon_i) = \{1 + \exp[\beta_e(\varepsilon_i - \mu_e)]\}^{-1}; \quad \beta_e = (k_B T_e)^{-1} \quad (128)$$

$$1 - f(\varepsilon_f) = \{1 + \exp[-\beta_h(\varepsilon_f - \mu_h)]\}^{-1}; \quad \beta_h = (k_B T_h)^{-1} \quad (129)$$

This formulation assumes that the photo-created conduction (valence) electrons thermalize amongst themselves much quicker than they recombine and that they acquire a temperature $T_c(T_h)$. This temperature is often different from the lattice temperature T and depends on the density of injected carriers, external perturbations (e.g. heating by an in-plane electric field [29])... The reader should be aware that a great deal of the photoluminescence spectra interpretations are based on the assumptions of thermalized carriers. This assumption is very convenient but is seldom justified by detailed calculations. With this reservation in mind, let us discuss some emission mechanisms in more details.

III.2.1 Band to band emission. — For an emitted light which propagates along the z axis, we know (see Tab. II) that the $HH_n \leftrightarrow E_m$ or $LH_n \leftrightarrow E_m$ transitions are allowed provided that $n + m$ is even and that the heterostructure potential is even in z . For light emitted in the layer plane, the ϵ_z polarization is forbidden for the $E_m \leftrightarrow HH_n$ transitions and allowed for the $LH_n \leftrightarrow E_m$ transitions. We retain the quadratic dispersions, (Eqs. (42-44)) for the various subbands and compute the luminescence lineshape of the subband-to-subband transitions. For the $E_m \leftrightarrow HH_n$ recombination we obtain

$$\begin{aligned} \mathcal{L}(\omega) \propto \langle u_{hh} | \epsilon \cdot p | u_c \rangle^2 | \langle \chi_n^{(h)} | \chi_m^{(e)} \rangle |^2 \int d^2 k_\perp \delta(\epsilon_g + E_m + HH_n + \\ \frac{\hbar^2 k_\perp^2}{2\mu_{nm}} - \hbar\omega) \times \frac{1}{1 + \exp \left[\beta_e \left(E_m + \frac{\hbar^2 k_\perp^2}{2m_e} - \mu_e \right) \right]} \times \\ \times \frac{1}{1 + \exp \left[-\beta_h \left(-\epsilon_g - HH_n - \frac{\hbar^2 k_\perp^2}{2M_n} - \mu_h \right) \right]} \quad (130) \end{aligned}$$

where μ_{nm} is the reduced electron-hole mass ($\mu_{nm}^{-1} = m_e^{-1} + M_n^{-1}$)

Let us introduce the chemical potentials η_e , η_h of the electrons and holes measured from E_m and HH_n respectively (see Fig. 21) :

$$\eta_m^{(e)} = \mu_e - E_m; \quad \eta_n^{(h)} = -\mu_h - HH_n - \epsilon_g \quad (131)$$

They are related to the steady state areal concentrations of electrons (n_m) and holes (p_n) in the E_m and HH_n subbands by :

$$\eta_m^{(e)} = k_B T_e \ln \left[-1 + \exp \left(\frac{\pi \hbar^2 n_m}{k_B T_e m_e} \right) \right] \quad (132)$$

$$\eta_n^{(h)} = k_B T_h \ln \left[-1 + \exp \left(\frac{\pi \hbar^2 p_n}{k_B T_h M_n} \right) \right] \quad (133)$$

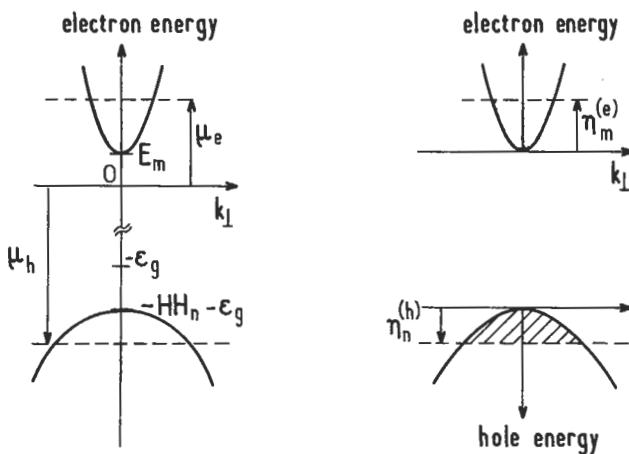


Fig. 21. — Definition of the reduced electron ($\eta_m^{(e)}$) or hole ($\eta_n^{(h)}$) Fermi levels in terms of the conduction (μ_e) and valence (μ_h) Fermi levels.

The total steady state concentrations n , p of the electrons and holes are respectively :

$$n = \sum_m n_m ; \quad p = \sum_n p_n \quad (134)$$

Let us denote by y the excess emitted photon energy over $\varepsilon_g + E_m + HH_n$:

$$\hbar\omega = \varepsilon_g + E_m + HH_n + y \quad (135)$$

We then obtain

$$\mathcal{L}_{m \rightarrow n}(y) \propto \frac{1}{1 + \exp \left[\beta_e \left(y \frac{\mu_{nm}}{m_c} - \eta_n^{(e)} \right) \right]} \times \frac{1}{1 + \exp \left[\beta_h \left(y \frac{\mu_{nm}}{M_n} - \eta_n^{(h)} \right) \right]} Y(y) \quad (136)$$

It is interesting to notice that this lineshape is obtained analytically whatever the degeneracies of the electron and holes gases (namely, whatever β_e , β_h , μ_e , μ_h). This is a particular feature of bi-dimensional systems which display a constant density of states if the dispersion relations are quadratic in k_\perp . Several limiting cases can be considered and two are discussed below.

i) Non degenerate electrons and holes ($\eta_n^{(e)} < 0, \eta_n^{(h)} \leq 0$).

Such distributions are obtained at low concentrations of injected carriers and/or high temperatures

$$\mathcal{L}(y) \propto Y(y) \exp(-y/k_B T^*) \quad (137)$$

$$\text{with : } \frac{1}{k_B T^*} = \frac{1}{k_B T_e} \frac{M_n}{M_n + m_c} + \frac{1}{k_B T_h} \frac{m_c}{m_c + M_n} \quad (138)$$

The step-like onset reflects the quasi bi-dimensional nature of the carrier motions. In practice it is rounded off by band-tailing, damping etc... However the characteristic feature is the exponential decay at large y (i.e. ω) which may allow T^* to be deduced from the luminescence spectrum. In addition, electrons and holes are often assumed to be in thermal equilibrium with each other ($T_e = T_h = T^*$), which is likely for delocalized carriers. When $k_B T^*$ becomes comparable to the energy separation between two consecutive subbands of a given band, the photoluminescence spectrum displays several lines. For instance in GaAs-Ga(Al)As quantum wells with a GaAs layer thickness ≥ 100 Å, one observes at low carrier injection and room temperature two luminescence lines associated with the $E_1 - HH_1$ and $E_1 - LH_1$ recombinations respectively. By decreasing T^* , the $E_1 - LH_1$ feature disappears.

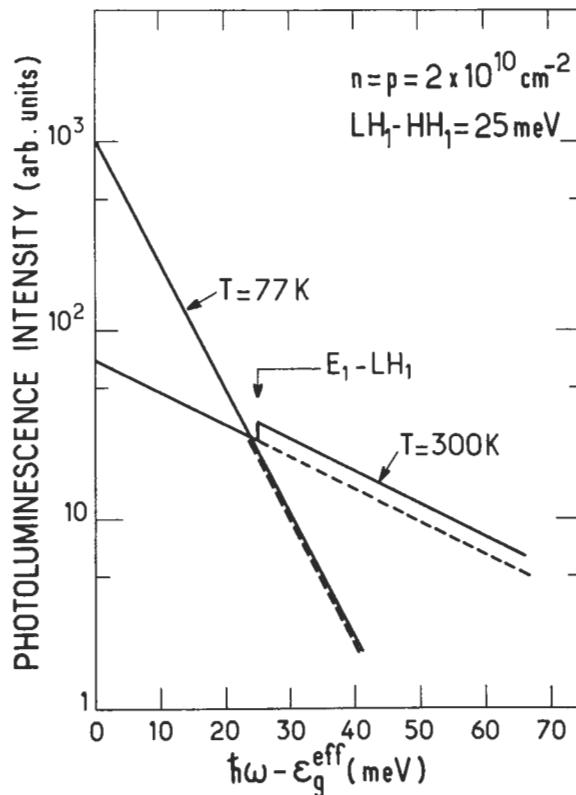


Fig. 22a. — Calculated band-to-band recombination lineshape in a GaAs-Ga_{0.67}Al_{0.33}As quantum well for non degenerate populations of electrons and holes. At $T = 77$ K only the $E_1 - HH_1$ recombination is significant while at $T = 300$ K the thermal population of the LH_1 subband is sufficient enough to allow the existence of a sizeable $E_1 - LH_1$ recombination. ϵ_g^{eff} denotes the effective bandgap of the quantum well viz. $\epsilon_g(\text{GaAs}) + E_1 + HH_1$. For $\hbar\omega < \epsilon_g^{\text{eff}}$ there is no band-to-band emission.

ii) Degenerate electrons and holes ($\beta_e \eta_m^{(c)} > 3$, $\beta_h \eta_n^{(h)} > 3$)

This situation corresponds to high injected carrier concentrations and/or low temperatures. The lineshape is again very simple :

$$\mathcal{L}(y) \propto Y(y) Y\left(\eta_m^{(c)} - \frac{\mu_{nm}}{m_c} y\right) Y\left(\eta_n^{(h)} - \frac{\mu_{nm}}{M_n} y\right) \quad (139)$$

The photoluminescence signal *versus* ω is rectangular but the edges are actually smoothed by the damping (low energy side) and finite temperature effects (high energy side). The high energy cut-off corresponds to y_{\max} where

$$y_{\max} = \text{Inf} \left[\frac{m_c}{\mu_{nm}} \eta_m^{(c)}, \frac{M_n}{\mu_{nm}} \eta_n^{(h)} \right] \quad (140)$$

The previous limiting cases are illustrated in figures (22a, b) where plots of the photoluminescence lineshape *versus* energy are calculated for band-to-band emission

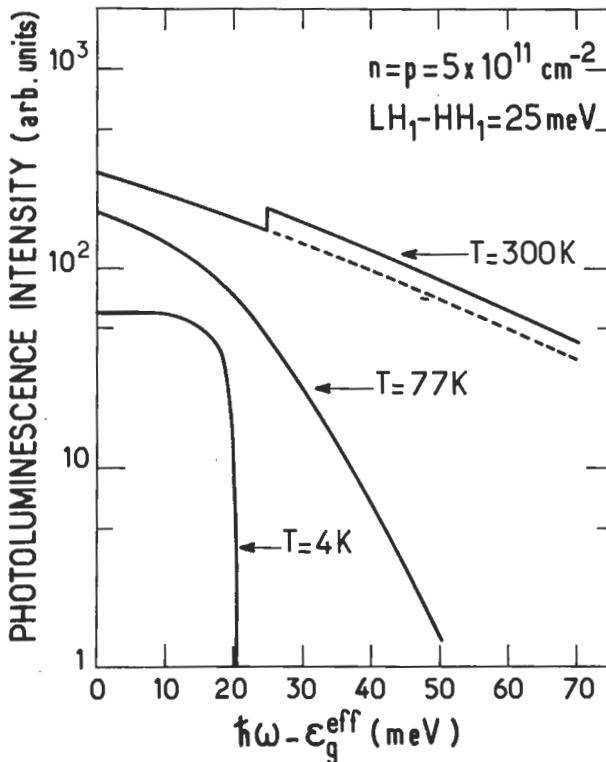


Fig. 22b. — Same as in figure 22a except that the recombining populations have much larger densities. At high temperature the electrons and holes approximately follow a Boltzmann distribution while at low temperature the electron and hole gases are degenerate, which affects the luminescence lineshape. For $\hbar\omega < \epsilon_g^{\text{eff}}$ there is no band-to-band emission.

in a GaAs-Ga_{0.67}Al_{0.33}As quantum well (thickness 55 Å) at different temperatures and for different concentrations of injected carriers. Only the $E_1 - HH_1$ and $E_1 - LH_1$ recombinations have been considered and equal carrier concentrations ($n = p$) and temperatures ($T_e = T_h$) have been assumed in the calculations.

Band to band recombination takes place when excitonic effects can be discarded. This may be the case in undoped GaAs quantum wells at room temperature (see the discussion in paragraph III.2.2) but the situation is controversial [30, 31]. We show in figures 23, 24 two examples where the radiative recombination does not involve excitons because their binding energies are insignificant. Firstly, in InAs-GaSb type II superlattices at room temperature [32] the spatial separation between the electrons and holes has weakened the excitonic binding so much that the excitons, if any, are ionized into free electron-hole pairs. The theoretical lineshape in these "true" superlattices is modified with respect to the formula given in equation (136) due to the subband dispersions along the growth axis and the peculiar selection rules prevailing in type II superlattices (see [17]): one of the two van Hove singularities of the joint density of states is suppressed. From figure 23 it can be seen that the high energy part of the spectrum is well reproduced by the calculations whereas the low energy side cannot be interpreted by a model which assumes an ideal superlattice. Recombination involving impurities or defects have to be invoked to account for the low energy side of the recombination spectrum.

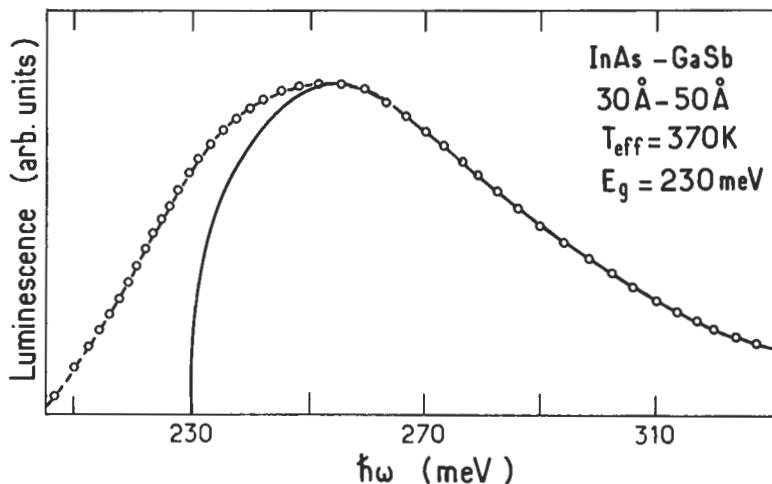


Fig. 23. — Calculated and measured band-to-band recombination in an InAs-GaSb superlattice. Solid line: theory. Open circles: experiments. After reference [32].

Notice that the effective carrier temperature is significantly higher than the lattice temperature. This hot carrier luminescence may have arisen due to the large kinetic energies supplied to the injected carriers by the exciting laser light which is well in excess of the effective bandgap of the superlattice.

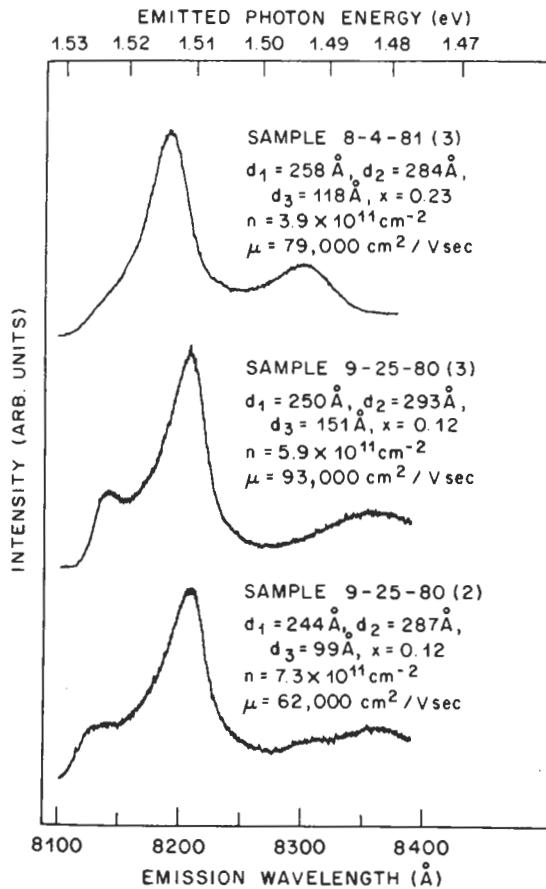


Fig. 24. — Photoluminescence lines of several n-type modulation-doped $\text{GaAs-Ga}_{1-x}\text{Al}_x\text{As}$ multiple quantum wells at $T = 2$ K. d_1 is the thickness of the GaAs wells, d_2 that of the $\text{Ga}_{1-x}\text{Al}_x\text{As}$ layers, and d_3 that of the undoped $\text{Ga}_{1-x}\text{Al}_x\text{As}$ spacer layers. The quoted mobilities refer to $T = 2$ K. After reference [33].

As a second example (Fig. 24) we present the radiative recombination of n-type modulation-doped GaAs quantum wells [33]. The large electron concentration effectively decreases the excitonic binding in these structures and consequently band to band emission predominates. The luminescence band contains three lines. The two high energy lines have been interpreted in terms of $E_1 \rightarrow H_1$ and $E_2 \rightarrow H_1$ recombinations. To account for the spectral positions of the lines a "renormalized" GaAs bandgap had to be invoked, the shrinkage of this gap being equal to 20 meV.

It is worth noticing that it is possible to observe the luminescence spectra calculated in this section (and later on), only if the absorption by the sample of the emitted photons can be neglected. If $\mathcal{L}_0(\omega)$ is the luminescence spectrum originating from a point M at a distance z from the sample surface, the spectrum effectively

observed corresponds to photons which have travelled through the sample to reach the surface. It is given by

$$\mathcal{L}(\omega) = \mathcal{L}_0(\omega) [1 - R] \exp(-z \alpha(\omega)) \quad (141)$$

where $\alpha(\omega)$ and R are the absorption coefficient and the reflectivity of the considered sample respectively. It is sensible to assume that the spontaneous emission is homogeneous, which means that $\mathcal{L}_0(\omega)$ is independent of z , so that the apparent luminescence spectrum is given by

$$\mathcal{L}(\omega) = (1 - R) \mathcal{L}_0(\omega) \frac{[1 - \exp(-h\alpha(\omega))]}{h\alpha(\omega)} \quad (142)$$

where h is the distance that a photon should travel in the sample. The luminescence

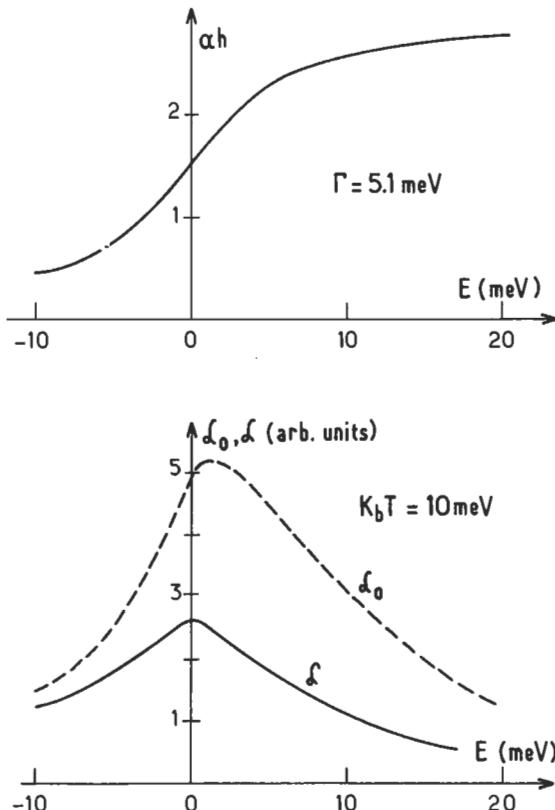


Fig. 25. — Calculated reabsorption effects on a photoluminescence line. Upper figure : absorption coefficient α times sample thickness versus photon energy (measured from the effective bandgap). Lower figure band-to-band luminescence intensities without reabsorption effect (\mathcal{L}_0) and including reabsorption effects (\mathcal{L}). In both figures a phenomenological damping coefficient of $\Gamma = 5.1$ meV has been assumed.

spectrum is thus modified by photon re-absorption, the low-energy side being favoured with respect to the high-energy one (see Fig. 25). For photons which propagate along the growth axis of the heterostructure this effect is weak and for single quantum wells negligible since $h \leq 1 \mu\text{m}$ and $\alpha(\omega) \sim 60 \text{ cm}^{-1}$. However, the situation may be different for photons which propagate in the layer planes because, in this case, h is of the order of several millimeters.

III.2.2 Excitonic recombination. — The excitonic recombination fulfils the same selection rules as the excitonic absorption, i.e. $\mathbf{K}_\perp = \mathbf{0}$ and $n + m$ even if the heterostructure potential is symmetric in z (see section II.2). In addition, only the nS excitons can radiatively recombine. This means that the excitonic luminescence emitted below the $E_m - HH_n$ band edge should consist of monochromatic lines. Most likely a single line caused by the $1S$ exciton attached to the $E_1 - HH_1$ bandgap will be observed, owing to the fast relaxation of exciton states towards the ground one. We have already had difficulties to interpret the low energy side of the band-to-band recombination line of non-interacting electrons, all the luminescence signal below the $E_1 - HH_1$ edge arising from defects.

However, the high energy side at least could be successfully described by models which neglect defects. In the case of excitonic recombination, a proper theory should take into account the broadening effects from the very beginning. The broadening effects are two-fold. Firstly they provide bound exciton states below $\epsilon_g + E_1 + HH_1 - R^*$. Secondly they relax the $\mathbf{K}_\perp = \mathbf{0}$ selection rules, allowing a finite luminescence above $\epsilon_g + E_1 + HH_1 - R^*$ to be predicted. At low temperatures the broadening mechanisms are either the exciton-defect interactions or the exciton-acoustical phonon interactions. Careful studies should be able to discriminate between both kinds of effects.

The second difficulty with excitons is the hybrid statistics which they obey. Diluted excitons can be treated as bosons to a first approximation, essentially because they arise from the pairing of two fermions, the electron and the hole. Excitons retain an increasing fermion nature when their density increases to the point where the concept of bound electron-hole pairs fades away to be replaced by that of an electron hole fluid, i.e. interacting fermions systems.

These two difficulties (relaxation of the \mathbf{K}_\perp selection rule and the nature of the exciton statistics) have been the subjects of a considerable body of literature in bulk materials [see, e.g., 27, 34]. They have not been studied much in heterostructures, both from the experimental and theoretical points of view. Therefore, we shall limit our considerations to a presentation of some experimental results.

One of the dominant features of the optical properties of GaAs-Ga(Al)As quantum wells is the strong intensity of the "free" exciton luminescence at low temperatures with respect to the intensity of the lines involving impurities. This is the opposite of what is usually observed in bulk GaAs. This is remarkable but has not been satisfactorily explained, at least to our knowledge. Possibly, this may be related to the very small effective volume in which the carriers and the electromagnetic wave interact in the case of quantum wells. Once created the photon has fewer chances of

being re-absorbed to create another exciton than in bulk materials, where the effective volume of interaction is the whole crystal.

The main arguments used to assign the photoluminescence line to excitonic recombination in GaAs quantum wells is its energy position and the nature of the recombining species, as supported by the spin orientation measurements [35]. In "good" samples, the maximum of the photoluminescence line is very close to (and sometimes coincides with) the maximum of the $HH_1 \rightarrow E_1$ absorption line (or photoluminescence excitation line) : see figures 20, 26. In such a case, the Stokes shift s , i.e. the energy separation between these two lines (luminescence and absorption) is equal to zero but in other samples s can be as large as several meV's. The existence of such a shift seems to be independent of the residual doping level of the GaAs (usually p- and n-type in MBE and MOCVD grown materials respectively). Thus it seems reasonable to conclude that the observed excitons are not bound to extrinsic defects (acceptors or donors). A model of exciton trapping on intrinsic interface defects [36] has been proposed, which correlates s to the sizes of the defects (extension along the z axis and in the layer plane). This model predicts that beyond $L \sim 100 \text{ \AA}$ in GaAs-Ga_{0.47}Al_{0.53}As quantum wells the trapped exciton binding energy becomes smaller than 1 meV. Thus, in this model, the Stokes shift could only be observed in thin quantum wells. The luminescence line would be associated with trapped excitons with a low density of states ($\leq 10^{10} \text{ cm}^{-2}$), with respect to that of the delocalized excitons (\sim a few times 10^{11} cm^{-2}), while the absorption, which is essentially sensitive to the large density of states regions of the energy spectrum, would exhibit features due to the delocalized excitons. The thermally activated

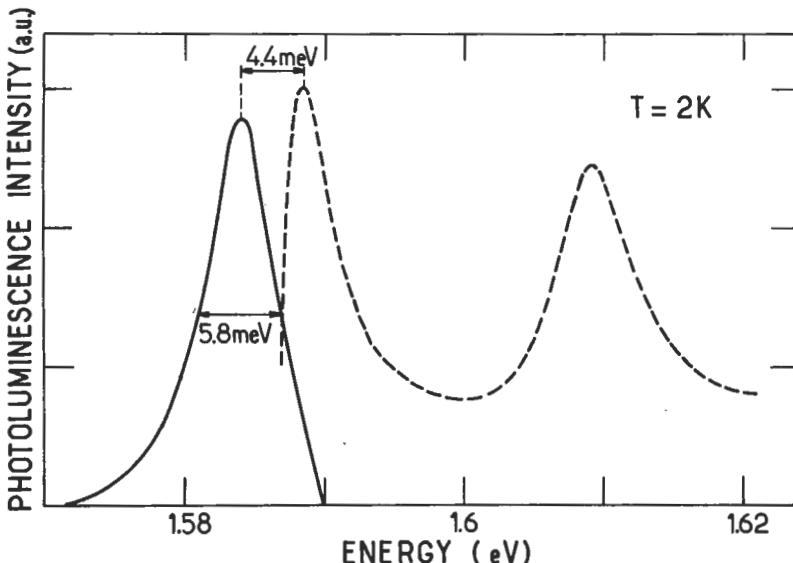


Fig. 26. — Photoluminescence (solid line) and photoluminescence excitation spectra of a 70 Å thick GaAs-Ga_{0.48}Al_{0.52}As single quantum well at $T = 2 \text{ K}$. After reference [36].

detrapping of excitons has in fact been observed [37]. It should be noted that the previous model was designed for defects whose in-plane dimensions are not too large (≤ 500 Å in GaAs-Ga(Al)As quantum well). It is clear that the excitonic photoluminescence in quantum wells with defects, which extend over 1000 Å or more in the layer plane, is better described by models which consider *free* excitons moving in "micro quantum wells" whose areas are equal to the defect areas and whose thickness is equal to the local quantum well thickness in the defect. In other words, the criterion for deciding which model is more appropriate is to count the number N of bound states (for the exciton centre of mass) that the defect supports. For $N \leq 5$ the motion of the exciton centre of mass is size-quantized and one recovers a photoluminescence due to excitons bound to these defects. For $N \geq 5$, micro-

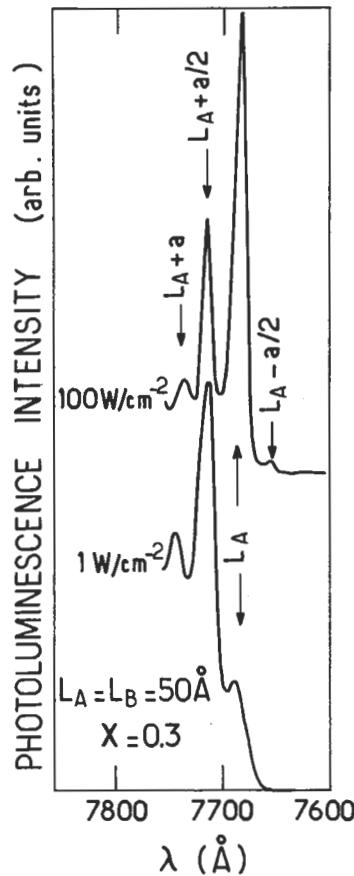


Fig. 27.— Photoluminescence lines of a 50 Å-50 Å GaAs-Ga_{0.7}Al_{0.3}As superlattice at $T = 2$ K for two different laser intensities. The peaks labelled L_A , $L_A \pm \frac{a}{2}$, $L_A + a$ refer to the excitonic recombination lines associated with islands of the GaAs wells having a thickness $L_A = 50$ Å, or differing from L_A by one or two monolayers. $a = 2.83$ Å. Courtesy B. Deveaud.

quantum well models are preferable and if the diffusion of excitons from one large defect to another one is difficult, structures in the excitonic photoluminescence should be observed. These structures appear at the *local* exciton energies in these various islands. Typically, with one monolayer fluctuations of the quantum thickness L , the photoluminescence shows three peaks due to the excitons which move in micro-quantum wells of thickness L , $L \pm a$, where $a = 2.83 \text{ \AA}$ is the thickness of a GaAs monolayer.

Such structures have been evidenced [38] in high quality MBE grown GaAs-Ga(Al)As multiple quantum wells and superlattices (Fig. 27). It is now admitted that the depth of interface fluctuations in GaAs-Ga(Al)As quantum wells grown by MBE can be reduced to one monolayer. On the other hand, the in-plane extensions of the interface defects depend significantly on the growth conditions.

Thus, a variety of experimental results concerning the Stokes shift and the shape of the excitonic photoluminescence line may and do occur. A series of sophisticated optical measurements (transient gratings, Rayleigh scattering, hole burning etc...) have been undertaken by Sturge *et al.* [39, 40] to ascertain the nature of exciton states in quantum wells. The aim is to discover whether quasi bidimensional excitons can become localized by a weak disorder, the localization being the result of constructive interferences of the exciton wavefunctions by randomly located defects. These experiments tend to show that, at low temperature, excitons are localized rather high in energy and that even a large part of the absorption spectrum arises from these localized excitons (Fig. 28).

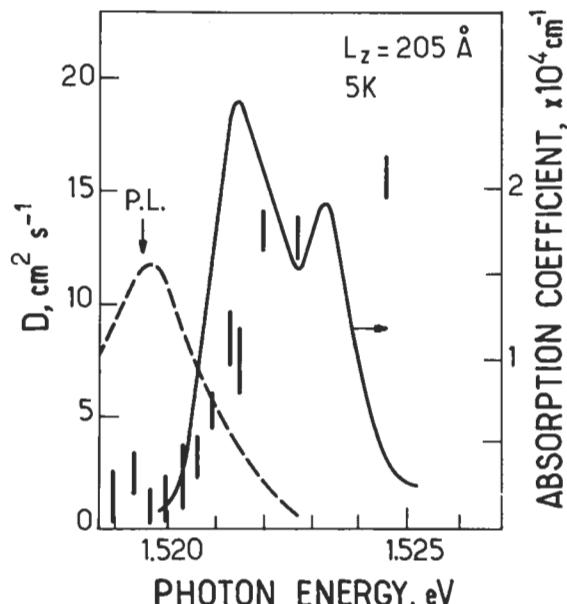


Fig. 28. — Photoluminescence (dashed line), absorption coefficient (solid line) and diffusion coefficient (vertical bars) measured in a GaAs-Ga_{1-x}Al_xAs multiple quantum well at $T = 5 \text{ K}$. After reference [39].

To our knowledge, the free exciton luminescence lineshape in a quantum well has never been calculated. One currently characterizes the quality of quantum well structures by the width δ at half-maximum of this line. When δ is smaller than some meV's the sample is claimed to be good. However, it is worth noticing that quantum wells which exhibit broad lines display a very intense luminescence, which shows the arbitrary nature of the criterion. Empirically the width δ and the Stokes shift are correlated: the wider the luminescence, the larger the shift.

A model of the width γ of the luminescence excitation spectrum has been proposed by Weisbuch *et al.* [24]. If the local quantum well thickness is $L + \delta L$ inside an interface defect, the exciton absorption occurs at an energy given by:

$$\hbar\omega(L + \delta L) = E_1(L + \delta L) + HH_1(L + \delta L) - R^*(L + \delta L) + \varepsilon_g \quad (143)$$

thus the linewidth γ is given by

$$\gamma \sim \delta L \left[\frac{dE_1}{dL} + \frac{dHH_1}{dL} \right] \quad (144)$$

since the variation of the exciton binding is negligible if $\delta L \ll L$. As the confinement energies vary like L^{-2} , γ should vary like L^{-3} . This behaviour has been observed in GaAs-Ga(Al)As multiple quantum wells (see Fig. 29). Recently, Singh *et al.* [41] have proposed a model of the photoluminescence linewidth δ . These authors have

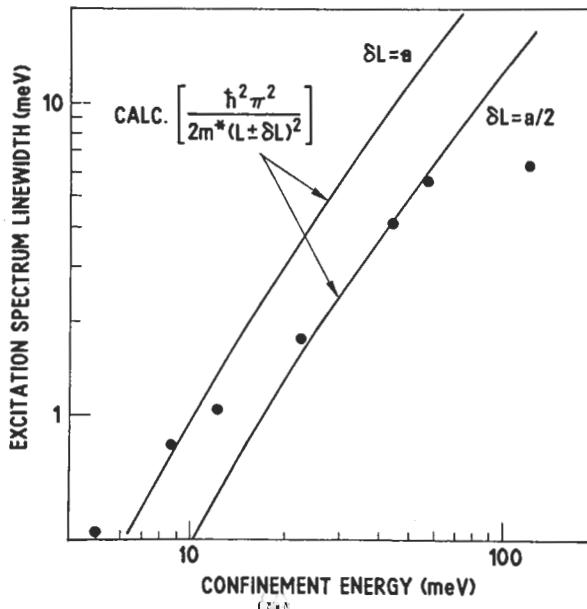


Fig. 29. — Calculated and measured excitation spectrum linewidth as a function of the confinement energy (viz. $E_1 + HH_1$) in GaAs-Ga_{1-x}Al_xAs multiple quantum wells. After reference [24].

correlated δ to the distribution of fluctuations in the well thickness, neglecting however the effect of carrier relaxation towards lower energy states.

The nature of the radiative recombination at room temperature in GaAs quantum wells is rather controversial. Taking into account the dissociation of the excitons (see Eqs. (104-109)), it seems unlikely that this recombination is entirely due to excitons. However Dawson *et al.* [31] and Bimberg *et al.* [30] have interpreted their steady-state photoluminescence and time-resolved cathodoluminescence lines emitted by the GaAs well and the GaAs substrate in terms of excitonic recombination (see Fig. 30). For the three lines shown in this figure, the recombination energy decreases by 0.4 meV/K when $T > 100$ K. If it were due to excitons at low temperature and to band-to-band processes at room temperature, the quantum well radiative recombination line would have a temperature dependence which differs from that of the substrate luminescence, contrary to what is shown in the figure. On the other hand, it may be argued that most of the temperature dependences of the various lines arise from the well known GaAs bandgap shrinkage with increasing T (the bandgap varies by ~ 106 meV between $T = 0$ K and $T = 300$ K) which is considerably larger than the exciton binding energy (10-15 meV in most quantum wells).

Investigations of the time dependence of the photoluminescence of GaAs-Ga(Al)As quantum wells give results which differ significantly from one group to the other. The simple model described in equations (114-118) would lead us to interpret these differences in terms of sample-dependent non-radiative lifetimes. Nonetheless,

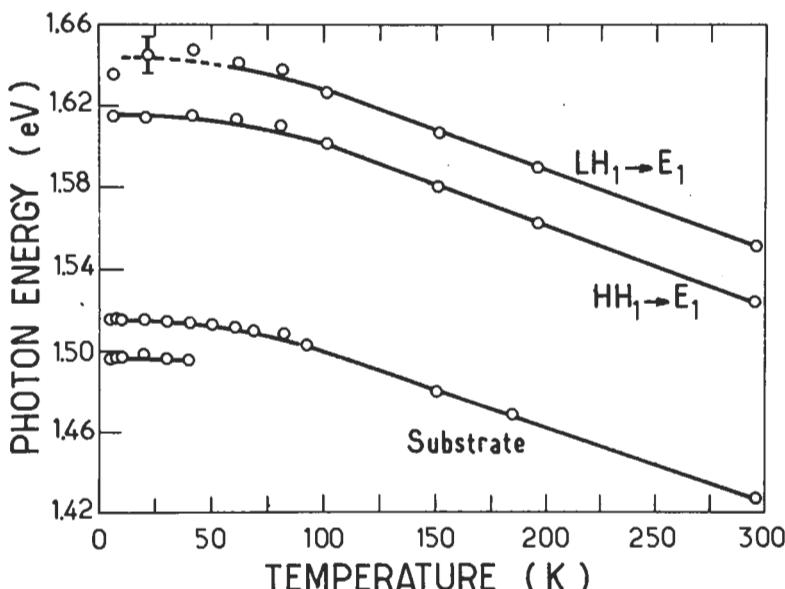
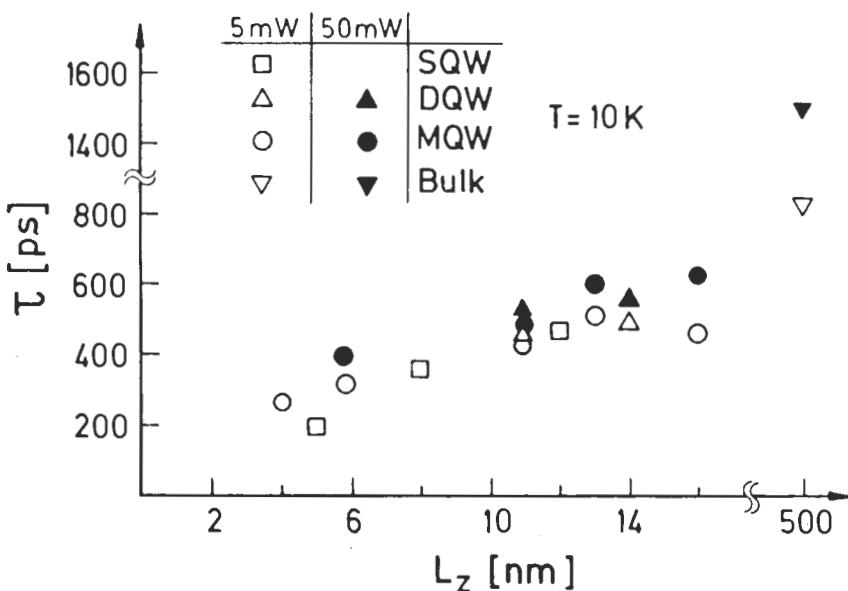


Fig. 30. — Temperature dependence of the energy positions of GaAs-Ga_{1-x}Al_xAs multiple quantum wells and GaAs substrate luminescence lines. For $T > 100$ K the energy shift is -0.4 meV/K for all three lines. After reference [30].

it seems agreed that the characteristic time of the exciton luminescence decay at low temperature shortens when the quantum well thickness L is decreased (see e.g. [42, 43]). If one identifies the decay time with the radiative lifetime τ_r and assumes the recombination to be excitonic, the lifetime reduction can be understood from the squeezing of the in-plane exciton Bohr radius λ with decreasing L . Indeed, the inverse of the radiative lifetime τ_r^{-1} , which is proportional to the transition probability per unit time that a photon is emitted, contains (as in the case of excitonic absorption) an enhancement factor $|\beta_{1s}(0)|^2$ which varies (see Eq. (95) like $\lambda^{-2}(L)$). In the ideal situation where valence and conduction barriers are infinitely high and where all the valence band mixings are neglected, the in-plane extension λ decreases from a_0 to $a_0/2$ when L decreases from infinity to zero. Thus, the excitonic luminescence lifetime should decrease by a factor of four between these two limits. It so happens that the photoluminescence decay time decreases approximately by a factor of four from bulk GaAs to GaAs quantum well with $L = 40$ Å (see Fig. 31). This excellent agreement between theory and experiment is probably fortuitous.



intrinsic interface defects. One should incriminate the participation of extrinsic defects (donors, acceptors, ...) in the radiative recombination. In fact, semiconductor quantum wells other than GaAs-Ga(Al)As display strong photoluminescence lines of extrinsic origin (e.g. GaSb-AlSb [15], InP-Ga_{0.47}In_{0.53}As [54]). Even GaAs-Ga(Al)As quantum wells often exhibit a weak extrinsic photoluminescence line which is situated on the low energy side of the excitonic recombination (the excitons being eventually trapped on intrinsic interface defects) (see e.g. [45-48]). The extrinsic photoluminescence line in GaAs-Ga(Al)As heterostructures is usually attributed to $E_1 \rightarrow$ acceptor recombination processes (see Fig. 32). The extrinsic origin of this line is also witnessed by its saturation or tendency to saturation behaviour at high pumping levels. If the acceptors participating in the $E_1 \rightarrow$ acceptor emission are not too numerous, an exciting pump intensity I_1 should exist where the optically injected holes occupy all the available acceptor sites. Increasing the pump intensity beyond I_1 does not lead to any further increase in the extrinsic photoluminescence signal. On the other hand the carriers involved in intrinsic recombinations (eventually paired into excitons) dispose of a much larger density of states : the subbands one. The intrinsic line is thus not easily saturable. Therefore by increasing the pump intensity one should observe the progressive disappearance of

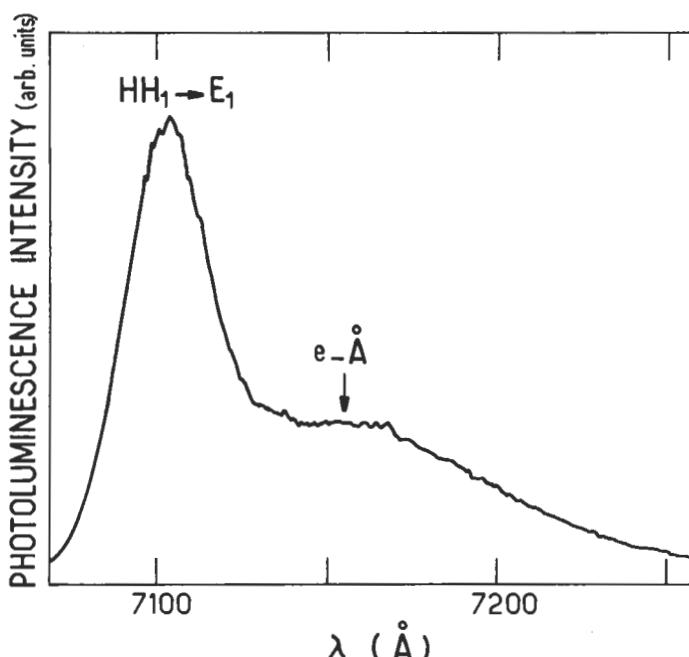


Fig. 32. — An example of $T = 2$ K photoluminescence of a GaAs-Ga_{1-x}Al_xAs quantum well at low exciting light power. The high energy peak corresponds to $E_1 - HH_1$ excitonic recombination while the low energy structure is associated with $E_1 \rightarrow$ Acceptor recombination processes. Courtesy M.H. Meynadier and C. Delalande.

the low energy recombination peak to the benefit of the intrinsic line. This is illustrated in figure 33 in MBE grown GaAs-Ga(Al)As quantum well [47].

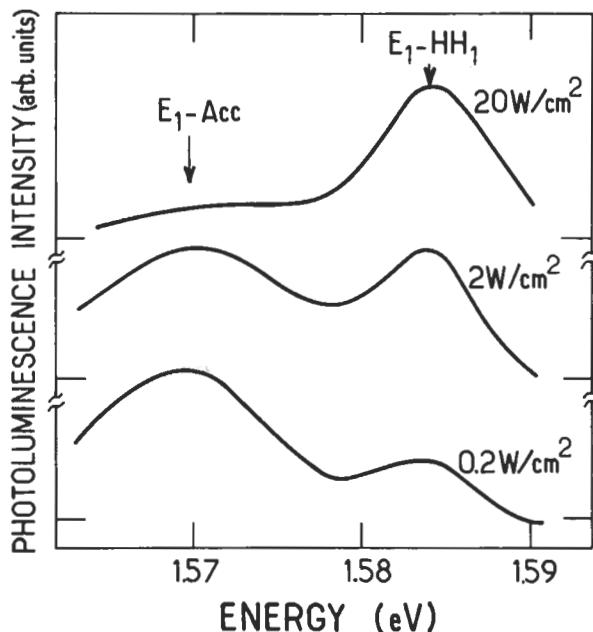


Fig. 33. — Saturation of the $E_1 \rightarrow$ Acceptor line with increasing power of the exciting light in a GaAs-Ga_{1-x}Al_xAs quantum well at $T = 2$ K. This heterostructure contained more acceptors than the one shown in figure 32. After reference [47].

We have previously seen (chapter IV) that, in quantum wells, the degeneracy of the impurity levels with respect to their position along the growth axis is lifted. For uniformly doped quantum wells, impurities near the centre of the well are statistically dominant and are therefore the easiest to observe. For $L < a_0$ (L : quantum well thickness, a_0 : bulk effective impurity Bohr radius), the impurities located near the interfaces also play a non-negligible role.

In fact, it has been shown in GaAs heterostructures that the residual doping is not spatially uniform and that impurity segregation (presumably carbon) occurs near the inverted interface (binary material grown on top of the ternary material) [48]. This phenomenon appreciably enhances the luminescence from E_1 to the acceptors on one edge of the sample ($E_1 \rightarrow A_c$). The residual impurity profile is usually very difficult to discover experimentally. It might be interesting to try to extract from the photoluminescence lineshape some information concerning the spatial position of the maximum of the impurity profile ; its spatial extension etc... However, this is not a straightforward task : the knowledge of the maximum of the luminescence peak and

of the curves $\varepsilon_A(z_i)$, where ε_A is the acceptor binding energy, is not sufficient because the radiative recombination matrix element depends heavily on z_i . It involves the square of the overlap integral I between the envelope functions $\chi_1^{(c)}(z)$ and $\chi_{\text{Acc}}(z, z_i)$ which varies appreciably with z_i for $|z_i| > \frac{L}{2}$. In the $E_1 \rightarrow \text{A}(z_i)$ radiative recombination, acceptors close to the centre of the well are much more efficient than the remote ones, and their efficiency may compensate for their small number. A second factor which may falsify the conclusions drawn from the sole knowledge of the maximum of the luminescence peak is the eventual thermalization of the holes. A typical energy extension of $\varepsilon_A(z_i)$ is 10 meV. This is considerably larger than $k_B T$ at low temperature. If the hole statistical distribution function is characterized by a low temperature T_h , as it is usually at low pumping level, only the deepest bound holes (within $k_B T_h$) will participate in the luminescence process, even though they are small in number. To circumvent this difficulty [47], the study of the extrinsic photoluminescence lineshape in the saturation regime (or close to it) has been proposed. All the acceptor sites are therefore occupied by a hole and the distortion of the lineshape by statistical occupancy is minimized.

The luminescence lineshape associated with the $E_1 \rightarrow \text{A}(z_i)$ recombination is not very difficult to calculate if one adopts a hydrogenic description of the acceptor :

$$\chi_{\text{Acc}}(z, z_i) = N(\lambda, z_i) \chi_1^{(h)}(z) \exp\left(-\frac{1}{\lambda} \sqrt{\rho^2 + (z - z_i)^2}\right) \quad (145)$$

where λ , which is a function of z_i , is a variational parameter. The electron distribution function $f^{(c)}$ is supposed to be a Fermi Dirac one (temperature T_c , chemical potential μ_c). The hole distribution is 1. The luminescence spectrum $\mathcal{L}(\omega, z_i)$ is therefore

$$\mathcal{L}(\omega, z_i) \propto \int d^2 k_\perp f^{(c)}\left(E_1 + \frac{\hbar^2 k_\perp^2}{2m_c}\right) \tau^{-1}[|E_1, \mathbf{k}_\perp\rangle \rightarrow |\text{A}(z_i)\rangle] \quad (146)$$

where :

$$\begin{aligned} \tau^{-1}[|E_1, \mathbf{k}_\perp\rangle \rightarrow |\text{A}(z_i)\rangle] &\propto \frac{2\pi}{\hbar} \delta[\varepsilon_g + E_1 + HH_1 - \varepsilon_A(z_i) - \hbar\omega] \times \\ &\times \left| \int_{-\infty}^{+\infty} \chi_1^{(c)}(z) \chi_1^{(h)}(z) dz \int_0^\infty \rho d\rho \exp\left(-\frac{1}{\lambda} \sqrt{\rho^2 + (z - z_i)^2}\right) \times \right. \\ &\left. \times \int_0^{2\pi} d\varphi \exp[i k_\perp \rho \cos \varphi] \right|^2 \end{aligned} \quad (147)$$

and ε_g is the bandgap of the well acting material. We shall denote by $I_{k_\perp}^2(\lambda, z_i)$ the square of the overlap integral between the envelope functions of the initial and final states. By introducing the wavevector k_i defined as

$$k_i = \left\{ \frac{2m_c}{\hbar^2} [\hbar\omega - \varepsilon_g - E_1 - HH_1 + \varepsilon_A(z_i)] \right\}^{1/2} \quad (148)$$

we obtain :

$$\mathcal{L}(\omega, z_i) \propto I_{k_i}^2(\lambda, z_i) \times \frac{Y[\hbar\omega - \varepsilon_g - E_1 - HH_1 + \varepsilon_A(z_i)]}{1 + \exp\left[\beta_c\left(E_1 + \frac{\hbar^2 k_i^2}{2m_c} - \mu_c\right)\right]} \quad (149)$$

The luminescence lineshape $\mathcal{L}(\omega)$ is finally given by weighing each $\mathcal{L}(\omega, z_i)$ by $N_A(z_i)$, the acceptor profile in the heterostructure.

$$\mathcal{L}(\omega) = \int dz_i N_A(z_i) \mathcal{L}(\omega, z_i) \quad (150)$$

It is easy to see from equations (151, 152) that the maximum of the luminescence line is not necessarily obtained for the value $\hbar\omega_0$ of $\hbar\omega$ which is associated with the maximum of the distribution $N_A(z_i)$. Attempts to extract some parameters of pre-supposed $N_A(z_i)$ in GaAs-Ga(Al)As heterostructures have shown a reasonable agreement between the experimental spectra and equations (149, 150) [47]. An example of such a lineshape fit is shown in figure 34.

Photoluminescence lines associated with donor $\rightarrow HH_1$ recombination [49] or with the emission due to excitons bound to neutral acceptors [50] have been reported in

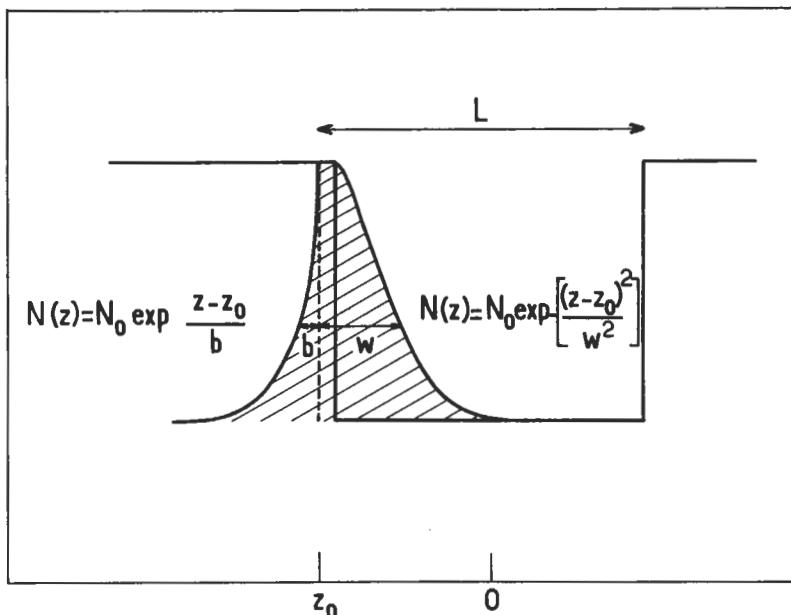


Fig. 34a. — Assumed impurity profile in GaAs-Ga_{1-x}Al_xAs quantum wells. z_0 is the position of the maximum of the impurity distribution measured from the centre of the quantum well, b is the characteristic decay length of the impurity distribution function in the Ga_{1-x}Al_xAs barrier and w the one in the GaAs well. After reference [47].

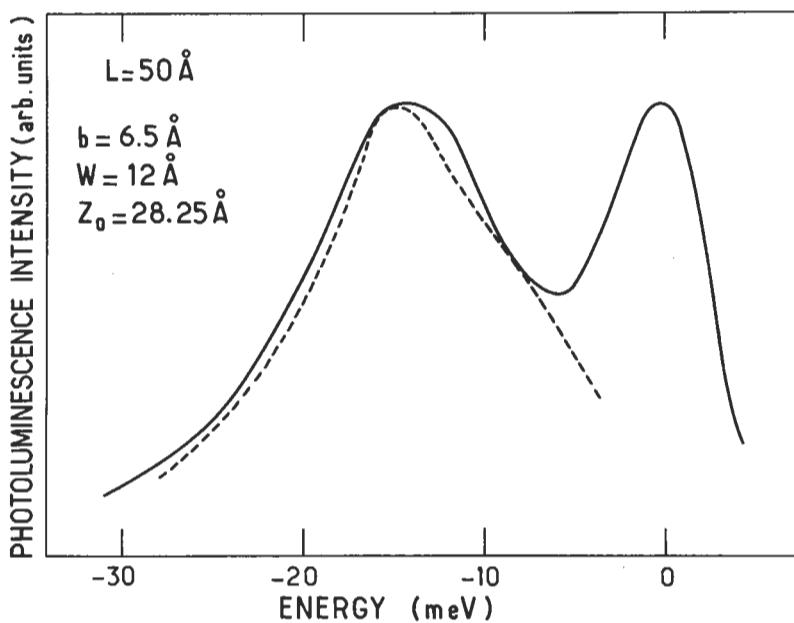


Fig. 34b. — Fit of the luminescence lineshape of a 50 Å thick GaAs-Ga_{1-x}Al_xAs quantum well at $T = 2$ K using the impurity distribution function shown in figure 34a. After reference [47].

GaAs-Ga(Al)As quantum wells which were intentionally doped. It seems that the binding energy $\varepsilon_{X\text{A}}$ of the exciton bound to neutral acceptors is, as in bulk materials, proportional to that of the acceptor ε_{A} ($\varepsilon_{X\text{A}} = 0.133 \varepsilon_{\text{A}}$), the proportionality constant being 30 % larger than in bulk GaAs [50].

General References

Several review papers have recently been devoted to the optical properties of semiconductor heterostructures. The reader may consult the following :

- [1] P. VOISIN, "Optical and Magneto-Optical Absorption in Quantum Wells and Superlattices", in the Proceedings of the Les Houches Winterschool *Heterojunctions and Semiconductor superlattices* edited by G. Allan, G. Bastard, N. Boccara, M. Lannoo and M. Voos (Springer Verlag) 1986.
- [2] C. DELALANDE, Photoluminescence in Semiconductor Quantum Wells, *ibidem* as reference [1].
- [3] C. WEISBUCH, Fundamental Properties of III-V Semiconductor 2 D Quantum Structures : The Basis for Optical and Electronic Devices in *Device and Circuit, Applications of III-V Superlattices and Modulation-Doping* edited by R. Dingle in *Semiconductors and Semimetals* edited by R.K. Willardson and A.C. Beer (Academic Press, New York) 1987, in press.

References

- [1] M. BORN and E. WOLF, *Principles of Optics* (Pergamon Press Oxford) 1964.
- [2] J. PANKOVE, *Optical Processes in Semiconductors* (Dover New York) 1975.
- [3] A MESSIAH, *Mécanique Quantique* (Dunod Paris) 1959 ; see also G. Baym, *Lectures on Quantum Mechanics* (Benjamin Readings) 1969.
- [4] H.N. SPECTOR, *Phys. Rev.* **B28** (1983) 971.
- [5] G.D. SANDERS and YIA-CHUNG CHANG, *J. Vac. Sci. Technol.* **B 3(4)** (1985) 1285.
- [6] J.N. SCHULMAN and YIA-CHUNG CHANG, *Phys. Rev.* **B31** (1985) 2056.
- [7] R. DINGLE, in *Festkörperprobleme* edited by H.J. Queisser, *Adv. Solid State Phys.* vol. **15** (Pergamon/Vieweg, Braunschweig) 1975.
- [8] R.C. MILLER, D.A. KLEINMAN and A.C. GOSSARD, *Phys. Rev.* **B29** (1984) 7085.
- [9] R.C. MILLER, A.C. GOSSARD, D.A. KLEINMAN and O. MUNTEANU, *Phys. Rev.* **B29** (1984) 3740.
- [10] M. H. MEYNADIER, C. DELALANDE, G. BASTARD, M. VOOS, F. ALEXANDRE and J.L. LIEVIN, *Phys. Rev.* **B31** (1985) 5539.
- [11] G. DUGGAN, *J. Vac. Sci. Technol.* **B3(4)** (1985) 1224.
- [12] H. KROEMER, in *Proceedings of the Second International Conference on Modulated Semiconductor Structures* (Kyoto 1985) *Surf. Sci.* **174** (1986) 299.
- [13] P. VOISIN, M. VOOS, M. RAZEGHI (1984) unpublished.
- [14] A.F.S. PENNA, J. SHAH, A. PINCZUK, D. SIVCO and A.Y. CHO, *Appl. Phys. Lett.* **46** (1985) 184.
- [15] P. VOISIN, C. DELALANDE, M. VOOS, L.L. CHANG, A. SEGMULLER, C.A. CHANG and L. ESAKI, *Phys. Rev.* **B30** (1984) 2276.
- [16] J.Y. MARZIN, M. QUILLEC, E.V.K. RAO, G. LEROUX and L. GOLDSTEIN, *Surf. Sci.* **142** (1984) 509.
- [17] P. VOISIN, G. BASTARD and M. VOOS, *Phys. Rev.* **B29** (1984) 935.
- [18] L.L. CHANG, G.A. SAI-HALASZ, L. ESAKI and R.L. AGGARWAL, *J. Vac. Sci. Technol.* **19** (1981) 589.
- [19] J.D. DOW, in *Optical Properties of Solids, New Developments*, edited by B.O. Seraphin (North Holland, Amsterdam) 1976.
- [20] R.S. KNOX, *Theory of Excitons, Solid State Phys. Suppl.* 5 (Academic Press, New York) 1963.
- [21] R.C. MILLER, D.A. KLEINMAN, W.T. TSANG and A.C. GOSSARD, *Phys. Rev.* **B24** (1981) 1134.
- [22] R.L. GREENE and K.K. BAJAJ, *Solid State Commun.* **45** (1983) 831.
- [23] F.L. LEDERMAN and J.D. DOW, *Phys. Rev.* **B13** (1976) 1633.
- [24] C. WEISBUCH, R. DINGLE, A.C. GOSSARD and W. WIEGMANN, *Solid State Commun.* **38** (1981) 709.
- [25] Y. MASUMOTO, M. MATSUURA, S. TARUCHA and H. OKAMOTO, *Phys. Rev.* **B32** (1985) 4275.
- [26] D.S. CHEMLA, *J. Lumin.* **30** (1985) 502.
- [27] H.B. BEBB and E.W. WILLIAMS, in *Semiconductors and Semimetals*, vol. 8, edited by R.K. Willardson and A.C. Beer (Academic Press, New York) 1972.
- [28] M.H. MEYNADIER and C. DELALANDE (1983) unpublished.
- [29] J. SHAH, A. PINCZUK, H.L. STORMER, A.C. GOSSARD and W. WIEGMANN, *Appl. Phys. Lett.* **42** (1983) 55.
- [30] D. BIMBERG, J. CHRISTEN, A. STECKENBORN, G. WEIMANN and W. SCHLAPP, *J. Lumin.* **30** (1985) 562.

- [31] P. DAWSON, G. DUGGAN, H.I. RALPH and K. WOODBRIDGE, *Phys. Rev.* **B28** (1983) 7381.
- [32] P. VOISIN, Thèse de Doctorat d'Etat, Paris (1983) unpublished.
- [33] A. PINCZUK, J. SHAH, H.L. STORMER, R.C. MILLER, A.C. GOSSARD and W. WIEGMANN, *Surf. Sci.* **142** (1984) 492.
- [34] T.M. RICE, *Solid State Physics*, edited by H. Ehrenreich, F. Seitz and D. Turnbull, vol. 32 (Academic Press, New York) 1977 ; see also P. Nozières, *Physica* **B117** and **B118** (1983) 16.
- [35] C. WEISBUCH, R.C. MILLER, R. DINGLE, A.C. GOSSARD and W. WIEGMANN, *Solid State Commun.* **37** (1981) 219.
- [36] G. BASTARD, C. DELALANDE, M.H. MEYNADIER, P.M. FRIJLINK and M. VOOS, *Phys. Rev.* **B29** (1984) 7042.
- [37] C. DELALANDE, M.H. MEYNADIER and M. VOOS, *Phys. Rev.* **B31** (1985) 2497.
- [38] B. DEVEAUD, J.Y. EMERY, A. CHOMETTE, B. LAMBERT and M. BAUDET, *Superlattices Microstruct.* **1** (1985) 205.
- [39] M.D. STURGE, J. HEGARTY and L. GOLDNER, in the *Proceedings of the 17th International Conference on the Physics of Semiconductors*, San Francisco (1984) edited by J.D. Chadi and W.A. Harrison (Springer Verlag, New York) 1985.
- [40] J. HEGARTY, M.D. STURGE, C. WEISBUCH, A.C. GOSSARD and W. WIEGMANN, *Phys. Rev. Lett.* **49** (1982) 930.
- [41] J. SINGH, K.K. BAJAJ and S. CHAUDHURI, *Appl. Phys. Lett.* **44** (1984) 805.
- [42] E.O. GOBEL, J. KUHL and R. HOGER, *J. Lumin.* **30** (1985) 541.
- [43] D. BIMBERG, J. CHRISTEN, A. STECKENBORN, G. WEIMANN and W. SCHLAPP, *Appl. Phys. Lett.* **44** (1984) 84.
- [44] M. VOOS, *J. Vac. Sci. Technol.* **B1(2)** (1983) 404.
- [45] R.C. MILLER, A.C. GOSSARD, W.T. TSANG and O. MUNTEANU, *Phys. Rev.* **B25** (1982) 3871.
- [46] W.T. MASSELINK, YIA-CHUNG CHANG and H. MORKOC, *Phys. Rev.* **B28** (1983) 7373, and *Phys. Rev.* **B32** (1985) 5190.
- [47] M.H. MEYNADIER, J.A. BRUM, C. DELALANDE, M. VOOS, F. ALEXANDRE and J.L. LIEVIN, *J. Appl. Phys.* **58** (1985) 4307.
- [48] R.C. MILLER, W. T. TSANG and O. MUNTEANU, *Appl. Phys. Lett.* **41** (1982) 374.
- [49] B.V. SHANABROOK and J. COMAS, *Surf. Sci.* **142** (1984) 504.
- [50] R.C. MILLER, A.C. GOSSARD, W. T. TSANG and O. MUNTEANU, *Solid State Commun.* **43** (1982) 519.
- [51] J.Y. MARZIN, M.N. CHARASSE and B. SERMAGE, *Phys. Rev.* **B31** (1985) 8298.

CHAPTER VIII

Effect of static external electric and magnetic fields on the energy levels of quasi bi-dimensional electron gases.

In this chapter we deal with the effects of static external fields (electric or magnetic) on the energy levels of idealized quantum wells or single heterojunctions.

The electric field effects (hereafter termed Stark effects) on insulating quantum well structures have recently received attention [1, 2] due to the possibility of making fast electro-optical devices [2] (e.g. fast electro-optical modulators).

The magnetic field effects (Landau quantization [3]) give access to fine characterization methods of conducting quasi bi-dimensional structures through cyclotron resonance or Shubnikov-de Haas experiments [4]. At the same time, they prove to be difficult to understand. As we shall see below, a magnetic field \mathbf{B} completely reorganizes the energy spectrum of quasi bi-dimensional electrons by splitting the continuum of the quasi bi-dimensional density of states ($B = 0$) into highly degenerate levels which are separated by finite bandgaps ($B \neq 0$). Thus, what was a metallic system at $B = 0$ is now either metallic or insulating at finite B depending on whether the Fermi level sits in the gaps or in the levels. The existence of finite bandgaps is responsible for the quantization of the Hall conductivity into multiples of the fundamental constant e^2/h . The quantized Hall effect, which may turn out to be the most spectacular experimental effect discovered in solid states physics this last decade, is at the moment only roughly understood. Much work remains to be done before all its intricacies are mastered.

In the following discussions band structure effects are ignored. The carriers are characterized by an isotropic effective mass (m_e for electrons, m_h for holes) which is assumed to be position-independent.

I. The Stark effects.

Two kinds of Stark effects can be found in quasi bi-dimensional structures, depending on whether the electric field \mathbf{F} is applied parallel to the growth (z) axis (longitudinal Stark effect) or perpendicular to it (transverse Stark effect). Usually, the microstructures designed for Stark effect experiments are insulating. The field is applied either *via* a Schottky barrier [1, 7] or by inserting the structure into the intrinsic part of a p.i.n junction [8] (longitudinal Stark effect). In order to study the transverse Stark effect, contacts have to be diffused throughout the structure. As a result of these techniques, the magnitude of the electric field is difficult to assess precisely and may, moreover, show inhomogeneities. In some experiments, the

electric field was estimated with $\sim 30\%$ accuracy [7]. In others, the uncertainty was ± 10 kV/cm over the whole investigated field range [2, 8]. Longitudinal field in excess of 10^5 V/cm could be applied on multiple quantum well structures either by biasing a p.i.n diode [2] or by using a Schottky contact [7]. These fields are considerable : over 100 Å a 10^5 V/cm electric field bends the band by 0.1 eV.

I.1 TRANSVERSE STARK EFFECT IN A QUANTUM WELL ($F \parallel \hat{x}$). — Classically, due to the presence of the confining barriers (we do not consider the unbound quantum well states) the z motion is bound ($|z| < \frac{L}{2}$ where L is the quantum well thickness) and periodic upon time. Along the electric field (\hat{x} axis) the carrier is uniformly accelerated whereas along the \hat{y} direction the particle is free. The total energy of the carrier is a constant of motion and can be split into contributions from the \hat{x} , \hat{y} , \hat{z} motions respectively :

$$\varepsilon = \varepsilon_x + \varepsilon_y + \varepsilon_z = \frac{p_x^2}{2m^*} + eF x + \frac{p_y^2}{2m^*} + \frac{p_z^2}{2m^*} + V_b Y \left[-\frac{L^2}{4} + z^2 \right] \quad (1)$$

where V_b is the quantum well barrier height. The allowed classical motions should be such that $p_x^2 \geq 0$. Thus the x motion is limited to :

$$x < \frac{1}{eF} \varepsilon_x \quad (2)$$

The stationary states of the quantum motion can be factorized into

$$\phi(\mathbf{r}) = \varphi(x) \chi_n(z) \exp(ik_y y) \quad (3)$$

where $\chi_n(z)$ is the n^{th} quantum well bound state wavefunction (energy ε_n) and the plane wave takes care of the free motion along the \hat{y} axis. Consequently

$$\varepsilon = \varepsilon_x + \frac{\hbar^2 k_y^2}{2m^*} + \varepsilon_n \quad (4)$$

and $\varphi(x)$ is the solution of the one-dimensional Schrödinger equation

$$\left(\frac{p_x^2}{2m^*} + eFx \right) \varphi(x) = \varepsilon_x \varphi(x) \quad (5)$$

with the boundary condition $|\varphi(x = \pm \infty)|$ finite. The change of variable

$$x = \frac{1}{eF} \varepsilon_x - \left(\frac{\hbar^2}{2m^* eF} \right)^{1/3} \eta \quad (6)$$

results in transforming equation (5) into the Airy equation [9]

$$\frac{d^2 \varphi}{d\eta^2} + \eta \varphi = 0 ; \quad |\varphi(\pm \infty)| \text{ bound} \quad (7)$$

whose solutions are

$$\varphi(\eta) = \begin{cases} \sqrt{\frac{|\eta|}{3\pi}} K_{1/3} \left[\frac{2}{3} |\eta|^{3/2} \right]; & \eta \leq 0 \\ \frac{1}{3} \sqrt{\pi\eta} \left[J_{-1/3} \left(\frac{2}{3} \eta^{3/2} \right) + J_{1/3} \left(\frac{2}{3} \eta^{3/2} \right) \right]; & \eta \geq 0 \end{cases} \quad (8)$$

where $J_\nu(x)$ and $K_\nu(x)$ are the Bessel function of real and imaginary arguments of order ν respectively. The wavefunction decays exponentially in the classically forbidden region ($\eta \leq 0$):

$$\varphi(\eta) \approx |\eta|^{-1/4} \exp \left[-\frac{2}{3} |\eta|^{3/2} \right]; \quad \eta \rightarrow -\infty \quad (9)$$

In the classically accessible region the wavefunction oscillates

$$\varphi(\eta) \approx |\eta|^{-1/4} \sin \left[\frac{2}{3} \eta^{3/2} - \frac{\pi}{4} \right]; \quad \eta \rightarrow \infty \quad (10)$$

Since the potential energy eFx is negative and arbitrarily large at negative x , the whole energy spectrum of equation (7) is continuous and an allowed motion may be associated with any energy ϵ_x . This means that below the $F = 0$ edge (i.e. $\epsilon_x = 0$), allowed energy states can be found at finite F . Symmetrically, above the

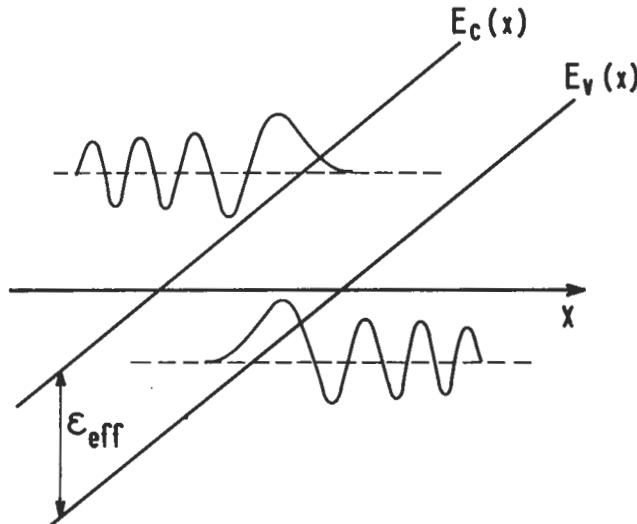


Fig. 1. — Conduction and valence band profiles along the \hat{x} direction for the transverse Stark effect geometry. ϵ_{eff} is an effective bandgap separating two conduction and valence subbands which arise from the size quantization along the \hat{z} direction. Two Airy functions are drawn for conduction and valence states respectively.

$F = 0$ edge of valence electrons one may find allowed valence levels at finite F . In other words, the electric field suppresses the bandgap.

At zero field, band to band optical absorption is only possible for photon energy $\hbar\omega$ such that $\hbar\omega > \epsilon_g + E_1 + H_1$, where ϵ_g is the bandgap of the bulk well-acting material and E_1, H_1 the electron and hole ground states confinement energies. At finite F , absorption should be found below the $F = 0$ edge, and the absorption coefficient should oscillate with ω . These effects have been known for a long time in bulk materials (Franz-Keldysh effect) [10] and should not be very different in quantum well structures. They have not yet been observed as such in quantum wells. Instead, a blurring of the excitonic structures was observed. We have already seen in chapters IV and VII that pronounced exciton effects take place in type I quantum wells due to the spatial confinement of the electrons and holes within the quantum well slab. What happens to these excitons when an electric field is applied in the layer plane ?

In a purely bi-dimensional system, the exciton Hamiltonian in the presence of an electric field ($F \parallel \hat{x}$) reads

$$\mathcal{H}_{\text{exc}} = \frac{P_{x_e}^2}{2m_e} + \frac{P_{y_e}^2}{2m_e} + \frac{P_{x_h}^2}{2m_h} + \frac{P_{y_h}^2}{2m_h} - \frac{e^2}{\kappa \sqrt{(x_e - x_h)^2 + (y_e - y_h)^2}} + eF(x_e - x_h) \quad (11)$$

As usual, \mathcal{H}_{exc} can be separated into contributions from the centre of mass and relative motions :

$$\mathcal{H}_{\text{exc}} = \frac{P_x^2 + P_y^2}{2(m_e + m_h)} + \frac{p_x^2 + p_y^2}{2\mu} - \frac{e^2}{\kappa\rho} + eFx \quad (12)$$

$$\boldsymbol{\rho} = \boldsymbol{\rho}_e - \boldsymbol{\rho}_h \quad (13)$$

$$\mathbf{R} = (m_e \boldsymbol{\rho}_e + m_h \boldsymbol{\rho}_h) / (m_e + m_h) \quad (14)$$

$$\mu = \frac{m_e m_h}{(m_e + m_h)} \quad (15)$$

If we discard the centre of mass degrees of freedom, the ground state of the relative motion at $F = 0$ is bound by $4R^*$, where R^* is the three-dimensional effective Rydberg, the two-dimensional Bohr radius being equal to one half of the three-dimensional one a^* . Figure 2 shows the potential energy profile along the \hat{x} axis for the relative motion at both $F = 0$ and $F \neq 0$. It can be seen that if the potential energy difference induced by the field over one Bohr radius $\left(\frac{1}{2}eF a^*\right)$ is comparable to the zero field binding energy ($4R^*$), the relative motion becomes unbound on the negative x side. This result is qualitatively the same as that found in bulk materials, where it is known that a critical field $F_c \sim \frac{R^*}{ea^*}$ ionizes the exciton.

Since quasi-2D excitons are more tightly bound than the corresponding bulk ones, they will be field-ionized at larger fields. Typically, a factor 2 - 4 on F_c is gained by decreasing the quantum well thickness : the excitonic binding is increased by a factor

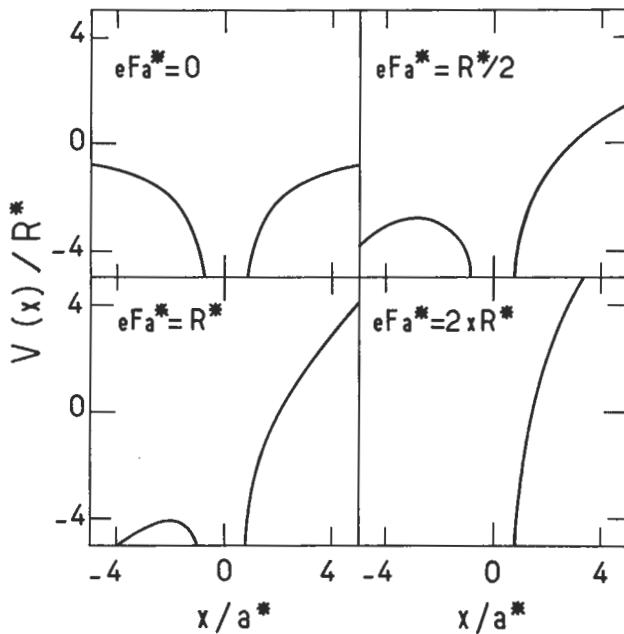


Fig. 2. — The potential energy profiles along the \hat{x} direction for the relative motion of a two-dimensional exciton subjected to a static electric field $\mathbf{F} \parallel \hat{x}$ are shown for four different magnitudes of the applied field F : $eFa^* = 0, 0.5 R^*, R^*$ and $2R^*$ respectively. R^* and a^* denoting the three-dimensional effective Rydberg and Bohr radius respectively.

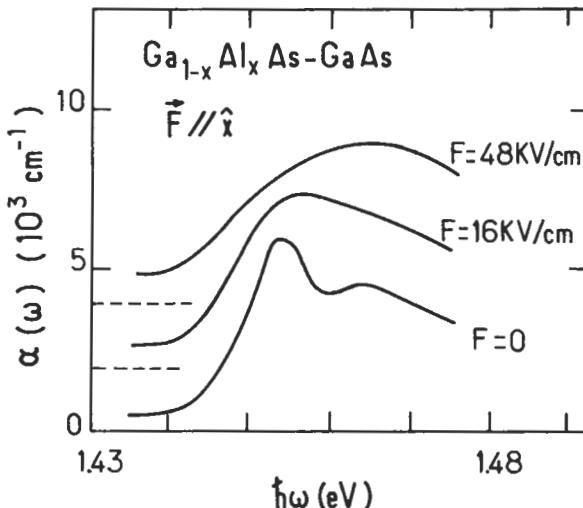


Fig. 3. — Room temperature electro-absorption in $\text{GaAs-Ga}_{0.68}\text{Al}_{0.32}\text{As}$ multiple quantum wells. $\mathbf{F} \parallel \hat{x}$. After reference [2].

2-3 and the effective exciton Bohr radius shrinks from a^* to $\sim 0.7 a^*$. Using numbers appropriate for GaAs, we find $F_c \sim 1.9 \times 10^4$ V/cm in GaAs quantum wells. Chemla *et al.* [2] have performed a detailed study of electro-absorption effects in multiple GaAs-Ga_{1-x}Al_xAs quantum well structures at room temperature. Their results for the transverse configuration is shown in figure 3. Clearly at $F = 1.6 \times 10^4$ V/cm the light hole excitonic resonance has been washed out and the heavy hole excitonic structure is already significantly blurred. At $F = 4.8 \times 10^4$ V/cm the absorption spectrum is featureless.

I.2 LONGITUDINAL STARK EFFECT IN A QUANTUM WELL ($F \parallel \hat{z}$). — The potential energy profile for such a configuration is shown in figure 4. The genuine feature of the longitudinal configuration is the appearance of a finite dipole \mathbf{D} between the

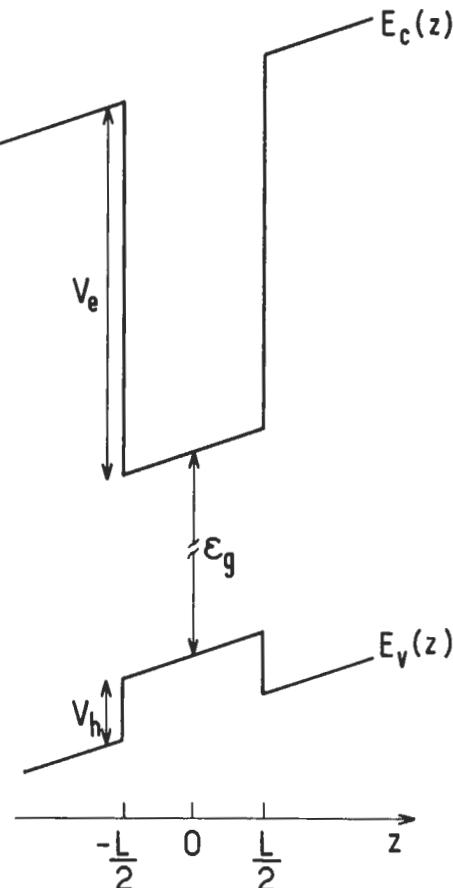


Fig. 4. — Position dependences of the conduction and valence band edges of a type I quantum well subjected to a longitudinal electric field \mathbf{F} ($\mathbf{F} \parallel \hat{z}$). ϵ_g denotes the bandgap of the well-acting material. V_c and V_h denote the conduction and valence barrier heights respectively.

electron and the hole of a photocreated electron-hole pair. Classically, the lowest lying state of the conduction electron is to remain immobile at the left-hand corner of the tilted quantum well. Since valence electrons have a negative mass, the topmost valence state corresponds to a carrier which remains immobile on the right-hand corner. Compared with the classical results, the wave-mechanical description brings to light several new features.

- i) The confinement over a distance L leads to size-quantization. At zero field the lowest lying conduction and valence states are E_1 and H_1 respectively.
- ii) The expectation value of the dipole operator $D = |e|(z_h - z_e)$ can be smaller or larger than L . This is due to the fact that the wavefunction leaks into the barrier and spreads over the whole well.
- iii) In principle this problem has no stationary bound states since the potential energy eFz becomes arbitrarily large and negative (positive) for conduction (valence) electrons when $z \rightarrow -\infty$ ($z \rightarrow +\infty$).

Over a large field range, however, the field-induced ionization of the quantum well remains of secondary importance. Let us respectively denote the confinement energy, conduction barrier height and characteristic penetration length of the ground conduction state in the barrier at zero field by $E_1^{(0)}$, V_e , κ_c^{-1} respectively. The conduction electron escape towards $z = -\infty$ will remain unimportant if the barrier drop over the penetration length $eF\kappa_c^{-1}$ is much smaller than the effective barrier height for the $E_1^{(0)}$ state $V_e - E_1$ [11] :

$$eF\kappa_c^{-1} \ll V_e - E_1^{(0)} \quad (16)$$

In this way, the effective barrier height does not change in the barrier region where the conduction electron is mostly localized (the integrated probability of finding the carrier between $-\kappa_c^{-1} - \frac{L}{2}$ and $-\frac{L}{2}$ is $\approx 86\%$ of the integrated probability of finding it over the whole left hand side barrier).

Figure 5 sketches the qualitative evolution of the ground state wavefunction of conduction electrons in a quantum well which is subjected to a longitudinal field. The energy zero is taken at the centre of the well, which also is the origin of the coordinates. At zero field (Fig. 5a), the average electron position in the ground state $\langle z_e \rangle$ is zero. By increasing F , $\langle z_e \rangle$ becomes negative. At low field (Fig. 5b), $\langle z_e \rangle$ is proportional to F . The ground state energy E_1 experiences a downward shift ΔE_1 which is equal to half of the product of the induced dipole $-e\langle z_e \rangle$ by the field. This is the domain of the quadratic Stark shift. At larger fields, the carriers tend to accumulate near the left hand side barrier (Fig. 5c). The induced dipole tends to saturate, as does ΔE_1 . If the barrier is infinitely high, this situation prevails for arbitrarily large F . Instead of spreading over the whole well, the carrier becomes essentially localized by a triangular well. Its energy varies like :

$$E_1(F) = -eF \frac{L}{2} + eF L \left[\frac{\hbar^2}{2m_e eF L^3} \right]^{1/3} \eta_0 \quad (17)$$

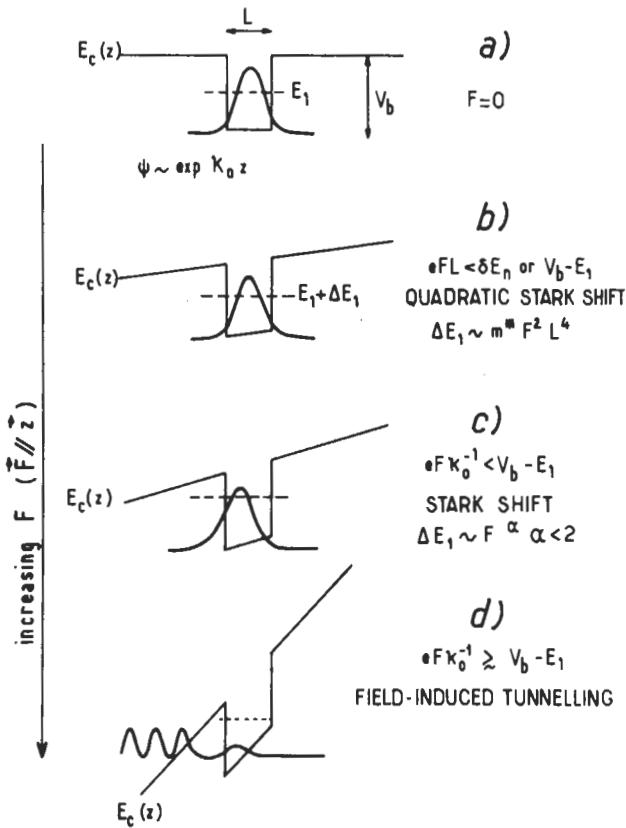


Fig. 5. — Qualitative evolution with increasing field strength ($F \parallel \hat{z}$) of the ground bound state of a quantum well subjected to a longitudinal electric field. a) $F = 0$, b) quadratic Stark shift, c) carrier accumulation regime, d) field-ionization of the quantum well.

where η_0 is the first positive zero of the Airy function ($\eta_0 \sim 2.33$) [9]. It is worth pointing out that E_1 becomes negative when F is large enough.

If, on the other hand, the barrier height is finite, the carrier accumulation is never complete but gradually transforms into the carrier escape phenomenon when the inequality (Eq. (16)) is no longer fulfilled. When F is sufficiently large, the carrier is swept out from the quantum well (Fig. 5d).

The conduction electron Hamiltonian which describes the longitudinal Stark effect in a quantum well reads :

$$\mathcal{H} = \mathcal{H}_\perp + \mathcal{H}_\parallel = \frac{p_x^2 + p_y^2}{2m_e} + \frac{p_z^2}{2m_e} + eF z_e + V_c Y \left[-\frac{L^2}{4} + z_c^2 \right] \quad (18)$$

The in-plane motion (\mathcal{H}_\perp) is free and decouples from the z motion (\mathcal{H}_\parallel). Let $\chi_j^0(z)$ and E_j^0 denote the zero field eigenfunctions and eigenenergies of \mathcal{H}_\parallel . At weak

field the quadratic Stark shift of the ground state may be sought by perturbation calculus :

$$\Delta E_1 = -e^2 F^2 \sum_{n \neq 1} \frac{|z_{1n}|^2}{E_n^0 - E_1^0} \quad (19)$$

where :

$$z_{1n} = \int_{-\infty}^{+\infty} \chi_1^0(z) \chi_n^0(z) dz \quad (20)$$

The summation in equation (19) runs over all the quantum well states (bound or unbound). Thus ΔE_1 is quite tedious to evaluate. However, it appears that the dominant contribution is due to the $n = 1 \rightarrow n = 2$ virtual transitions when the $n = 2$ state is bound [11]. Since $E_2^0 - E_1^0 \sim \frac{\hbar^2}{2m_e L^2}$, $\Delta E_1 \sim F^2 m_e L^4$; thus it increases rapidly with L . However the field range where equation (19) is justified, is such that $\Delta E_1 \ll E_2^0 - E_1^0$. Thus it narrows quite rapidly with L , and is such that $F^2 m_e^2 L^6$ is constant.

In order to describe the field range as a whole, variational approaches have been developed. A convenient and rather precise trial wavefunction is obtained by multiplying the ground quantum well function $\chi_1^0(z)$ by a decreasing exponential term [4].

$$\chi_1(z) = \chi_1^0(z) \exp(-\beta_e z) \quad (21)$$

Near the $z = -\frac{L}{2}$ interface, $\chi_1(z)$ has the same shape as the Fang-Howard wavefunction (corrected to account for the wavefunction penetration in the barrier). At weak field, we notice that $\chi_1(z)$ accounts for the essential effect of the field, which is to admix the even ground state with the odd excited states. For valence electrons, $-\beta_e$ should be replaced by $+\beta_h$ in order to account for the valence polarization which is opposite to that of the conduction polarization. Figures 6, 7, which were obtained by using equation (21), illustrate the results of the variational calculations. In agreement with previous discussions, one notices that the electric field effects depend markedly on L and are more pronounced for valence than for conduction electrons.

Finally, it should be noticed that the exact eigenstates of \mathcal{H}_1 can be obtained. Within each layer, $\chi_1(z)$ is a linear combination of the two independent solutions of the Airy equation. Thus, the eigenstates are obtained via matching conditions at the $z = \pm \frac{L}{2}$ interfaces. However, the Airy functions are inordinarily complicated for convenient use.

When near bandgap light is shone on the quantum well structure, excitons are formed between the quantum well valence and conduction subbands. We shall be mainly concerned with the lowest lying heavy hole ($E_1 - HH_1$) and light hole ($E_1 - LH_1$) excitons. The electric field polarizes E_1 and HH_1 (LH_1) along opposite directions and thus weakens the excitonic binding. However, and this is the major

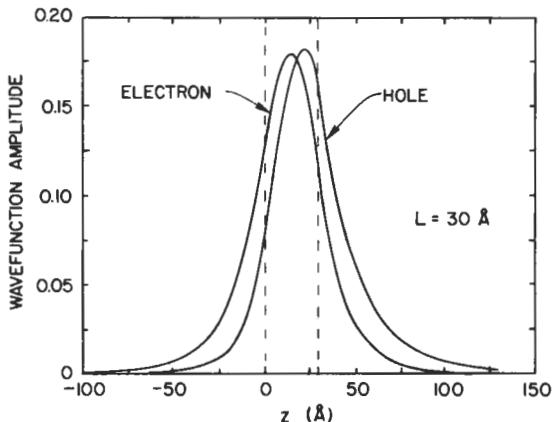


Fig. 6. — Wavefunction amplitudes *versus* distance for electrons and heavy holes in a GaAs quantum well of thickness 30 Å, subjected to a longitudinal electric field of 10^5 V/cm. The origin is taken at the edge of the well. The dashed lines represent the well boundaries. After reference [11].

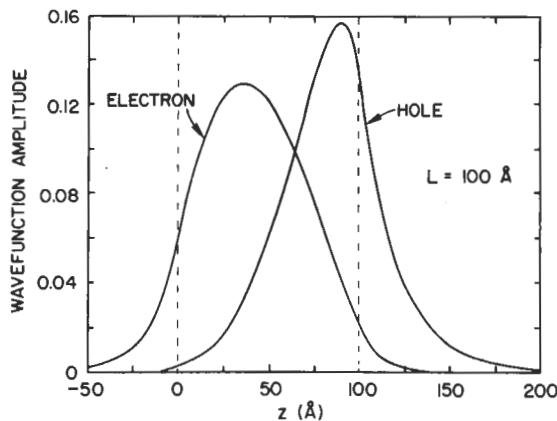


Fig. 7. — Wavefunction amplitudes *versus* distance for electrons and heavy holes in a GaAs quantum well of thickness 100 Å, subjected to a longitudinal electric field of 10^5 V/cm. The origin is taken at the edge of the well. The dashed lines represent the well boundaries. After reference [11].

advantage of the longitudinal configuration, the exciton dissociation is considerably hindered by the conduction and valence potential barriers. We expect the exciton binding energy $R(F)$ to show little variation with F and to reflect the carrier accumulation at the interfaces. However, if the field becomes so large that the carriers are swept out of the quantum well figure 5d, the exciton will finally be field-ionized. We saw in figure 3 that roughly 20 kV/cm are enough to wash out the

excitonic resonances in the transverse configuration. Chemla *et al.* [2] have demonstrated that fields as large as 100 kV/cm can be applied in the longitudinal configuration without destroying the excitonic binding. This improved excitonic stability, which leads to peaked structures in the absorption coefficient, is accompanied by a tunability of the excitonic resonance energy. For example, the lowest lying heavy hole exciton peak is found at

$$\hbar\omega_{hh} = \varepsilon_g + E_1(F) + HH_1(F) - R_{hh}(F) \quad (22)$$

where ε_g is the bandgap of the well-acting material, $R_{hh}(F)$ the heavy hole exciton binding energy and $E_1(F)$, $HH_1(F)$ the ground state electron and heavy hole confinement energies. The tunability of $\hbar\omega_{hh}$ is due, not so much to the variations of $R_{hh}(F)$, but rather to the field dependent band edge energy $\varepsilon_g + E_1(F) + HH_1(F)$ as long as the exciton is reasonably bound. The resonance energy can be easily changed by ~ 30 meV, which is significantly larger than the excitonic linewidth in high quality materials.

Let us study in more detail [12] the quantity $R_{hh}(F)$. At vanishing F , the electron and the hole move, on average, in the same plane ($z = 0$). At finite F , they move on two different planes which are separated by $D/|e|$, where D is the magnitude of the field-induced dipole between the electron and hole charge distributions. To the lowest order in F , the energy $R_{hh}(F)$ decreases by a quantity which is proportional to $D^2/e^2\lambda_0^2$ where λ_0 is the zero field, in-plane, Bohr radius of the exciton. As we saw previously, the dipole D is proportional at low field to :

$$D \propto e^2 FL (m_e + m_h) \frac{L^3}{\hbar^2} \quad (23)$$

Thus, $R_{hh}(F)$ also shows a quadratic Stark shift which varies like L^8 . However, this weak field analysis is valid for a field range where FL^3 is constant. Therefore, as L increases, the excitonic binding decreases more and more quickly with respect to F , but in a field range which narrows quite rapidly.

The quadratic excitonic Stark shift is followed by a smoother field dependence, which is reminiscent of the carrier accumulation regime. The electron and the hole tend to accumulate on each side of the quantum well. Thus, D tends to saturate. If the carrier accumulation were complete (impenetrable barriers), the excitonic binding energy would extrapolate to that of an exciton whose electron and hole move on two different planes separated by L .

However, we know that the carrier wavefunctions leak more and more heavily into the finite barriers when F increases. Ultimately, the carriers are swept out of the quantum well and the exciton finally field-ionizes.

The exciton Hamiltonian in the presence of a longitudinal electric field is written

$$\begin{aligned} \mathcal{H}_{exc} = & \varepsilon_g + \frac{P_\perp^2}{2\mu} + \frac{P_{z_e}^2}{2m_e} + \frac{P_{z_h}^2}{2m_h} + V_e Y \left[z_e^2 - \frac{L^2}{4} \right] + V_h Y \left[z_h^2 - \frac{L^2}{4} \right] + \\ & + eF(z_e - z_h) \cdot \frac{e^2}{\kappa} [\rho^2 + (z_e - z_h)^2]^{-1/2} \end{aligned} \quad (24)$$

where the centre of mass kinetic energy term has been dropped. We already know approximate solutions of equation (24) when $\frac{e^2}{\kappa}$ is set equal to zero:

$\epsilon_g + E_1(F) + HH_1(F)$ for the lowest lying state of an electron-hole pair. Although the z and (x, y) motions are not exactly separable, we have already seen in previous chapters that the exciton z motion is forced by the quantum well barriers, whereas the coulombic term binds the electron-hole pair in the layer plane. Thus, a reasonable trial wavefunction for the ground heavy hole exciton state will be [12]:

$$\psi_{\text{exc}}(\mathbf{r}_e, \mathbf{r}_h) = \mathcal{N} \chi_{E_1}(z_e) \chi_{HH_1}(z_h) \exp[-\rho/\lambda] \quad (25a)$$

where λ is the variational parameter, ρ the in-plane electron-hole distance, \mathcal{N} a normalization coefficient and $\chi_{E_1}(z_e) \chi_{HH_1}(z_h)$ are the field-dependent wavefunctions for the electron and the hole which move in the quantum well tilted by the electric field (see Eq. (21)). Since the exciton wavefunction is taken as separable in ρ and

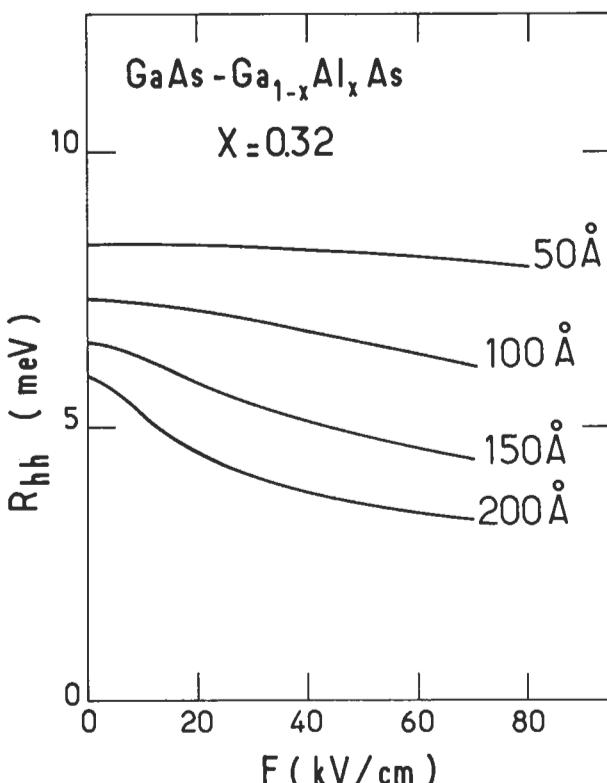


Fig. 8. — Calculated variations of the heavy hole exciton binding energy R_{hh} versus the electric field $F(F/\sqrt{2})$ in GaAs-Ga_{0.68}Al_{0.32}As quantum wells of different thicknesses. After reference [12].

z_e, z_h , the required minimization amounts to find the minimum over λ of :

$$\hbar\omega_{hh}(\lambda) = \varepsilon_g + E_1(F) + HH_1(F) + \frac{\hbar^2}{2\mu\lambda^2} - \frac{2e^2}{\kappa\lambda} \int_0^\infty x \, dx \, e^{-x} \times$$

$$\iint_{-\infty}^{+\infty} \frac{dz_e \, dz_h \, \chi_{E_1}^2(z_e) \chi_{HH_1}^2(z_h)}{\sqrt{x^2 + \frac{4}{\lambda^2} (z_e - z_h)^2}} \quad (25b)$$

Note however that as $E_1(F), HH_1(F)$ are not known exactly (except at $F = 0$), there is no warrant that $\{\varepsilon_g + E_1(F) + HH_1(F) - \text{Min}_\lambda(\hbar\omega_{hh}(\lambda))\}$ will be a lower bound to $R_{hh}(F)$. Figures 8, 9 show the L dependence at fixed L or the F dependence at fixed L of the heavy hole exciton binding energy R_{hh} in GaAs-Ga_{0.68}Al_{0.32}As quantum wells. In the $R_{hh}(F)$ curves, one clearly sees the quadratic Stark shift decrease at low field, becoming steeper when L increases and observable on a narrower field range. This regime is followed by a weaker F dependence, reminiscent of the carrier accumulation regime.

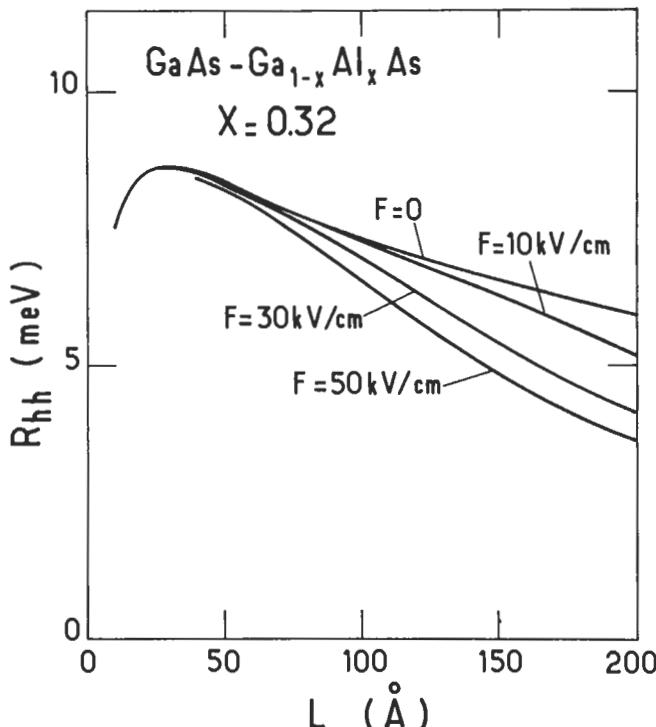


Fig. 9. — Calculated variations of the heavy hole exciton binding energy R_{hh} versus the quantum well thickness L for different field strengths in GaAs-Ga_{0.68}Al_{0.32}As (F/\hat{z}). After reference [12].

Detailed electro-absorption experiments were performed by Wood *et al.* [8, 2] on multiple GaAs-Ga_{0.68}Al_{0.32}As quantum wells at room temperature. The field was applied by biasing a p.i.n. junction. Figure 10 shows the electro-absorption data. It demonstrates unambiguously that much larger fields can be applied in the longitudinal than in the transverse configuration (and, *a fortiori*, in bulk GaAs) without blurring the excitonic resonance too severely. The second important point noticeable in figure 10 is the magnitude of the energy shifts experienced by $\hbar\omega_{hh}(F)$. Suppose that light is shone on the multiple quantum wells with a photon energy smaller (by ~ 10 meV) than the zero field exciton resonance energy $\hbar\omega_{hh}(F=0)$. It is not absorbed. By applying a field of ~ 50 kV/cm, the photon energy becomes coincident with the $E_1 - HH_1$ exciton resonance and the light beam is significantly absorbed. Thus, by switching the field on and off, the beam intensity can be controlled [13]. One advantage of this (ideally) on-off control is that it can be very fast.

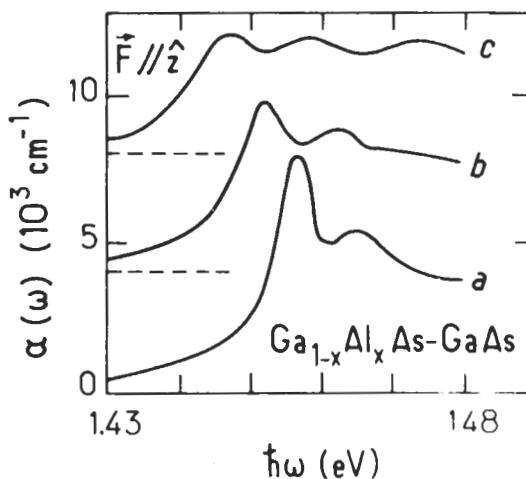


Fig. 10. — Room temperature electro-absorption data in GaAs-Ga_{0.68}Al_{0.32}As multiple quantum wells ($L = 95 \text{ \AA}$) $F \parallel \hat{z}$. Curve a : $F = 10 \text{ kV/cm}$. Curve b : $F = 47 \text{ kV/cm}$. Curve c : $F = 73 \text{ kV/cm}$. After reference [2].

Figure 11 presents a comparison between the calculated and observed field dependences of the heavy hole exciton resonance in GaAs-Ga_{0.68}Al_{0.32}As ($L = 95 \text{ \AA}$). Since the calculations involve no adjustable parameters, the model and the data agree reasonably. Let us stress again that the main source of the red-shift of the exciton resonance is the field dependence of $E_1(F) + HH_1(F)$ whereas $R_{hh}(F)$ contributes little to this shift.

Earlier electro-optical experiments were performed by Mendez *et al.* [1], again on GaAs-Ga_{1-x}Al_xAs multiple quantum wells. These authors observed a spectacular quenching of the photoluminescence intensity of their structures when a longitudinal field was applied to the multiple quantum wells *via* a Schottky contact (Fig. 12). A

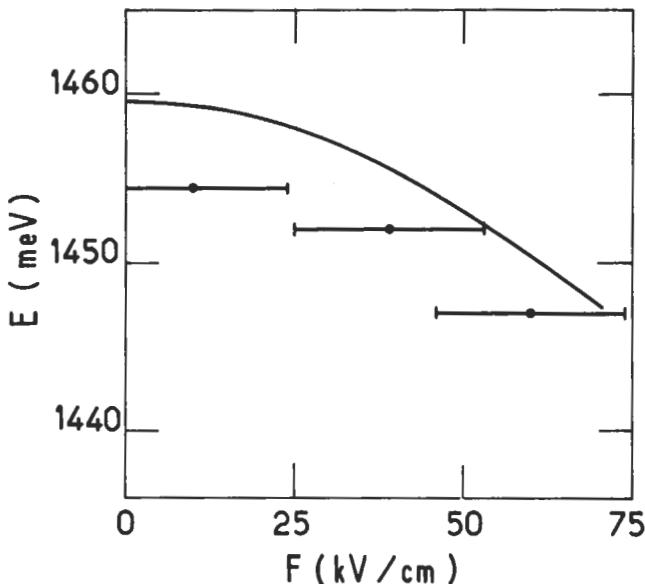


Fig. 11. — Field dependence of the heavy hole exciton resonance energy in GaAs-Ga_{0.68}Al_{0.32}As multiple quantum wells ($L = 95 \text{ \AA}$) at room temperature. Solid line: theory. Bars: experiments. After reference [12].

red shift of the peak position was also reported. The luminescence drop could only be qualitatively explained by a field-induced electron-hole separation in a single well.

Recent electro-reflectance experiments, performed by Alibert *et al.* [7] on single GaAs-Ga_{0.4}Al_{0.54}As quantum wells, have shown red shifts of the excitonic structures which correlate well with the longitudinal Stark effect model previously discussed. However, the magnitude of the electric field, as deduced from the Franz-Keldysh oscillations of the thick Ga_{0.46}Al_{0.54}As barriers, had to be corrected by $\sim 30\%$ to quantitatively explain the data.

II. Landau quantization of a quasi bi-dimensional electron gas.

II.1 ENERGY LEVELS. — In this paragraph we consider the motion of a quasi bi-dimensional electron subjected to a static and uniform magnetic field \mathbf{B} , tilted by an angle θ with respect to the growth ($\hat{\mathbf{z}}$) axis of the structure (see Fig. 13). Thus we can choose :

$$\mathbf{B} = [0, B \sin \theta, B \cos \theta] \quad (26)$$

It can be shown that the Lorentz force $q \frac{\mathbf{v}}{c} \wedge \mathbf{B}$ acting on a particle of charge q can be derived by using the Hamiltonian equations $\left[\frac{d\mathbf{r}}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \frac{d\mathbf{p}}{dt} = -\frac{\partial \mathcal{H}}{\partial \mathbf{r}} \right]$ provided

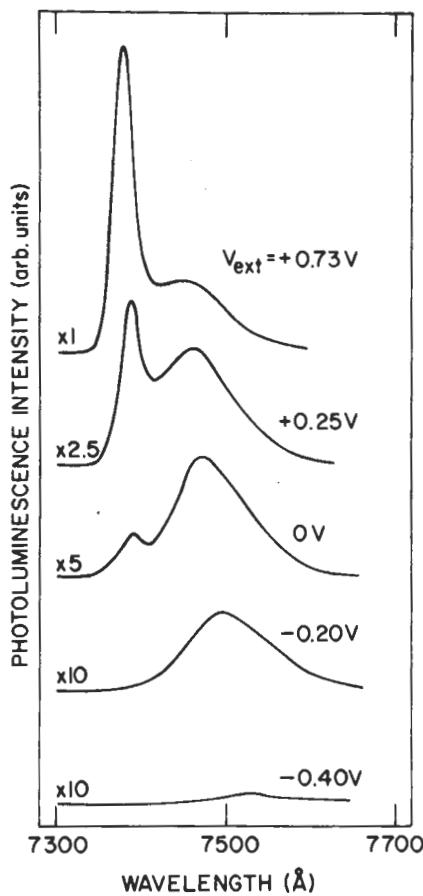


Fig. 12. — Photoluminescence intensity for various applied voltages *versus* emission wavelength in a multiple GaAs-Ga_{0.65}Al_{0.35}As quantum well. The field is applied *via* a Schottky contact ($F//\hat{z}$). V_{ext} denotes the voltage drop between the metallic electrode and the GaAs substrate. $T = 6$ K. After reference [1].

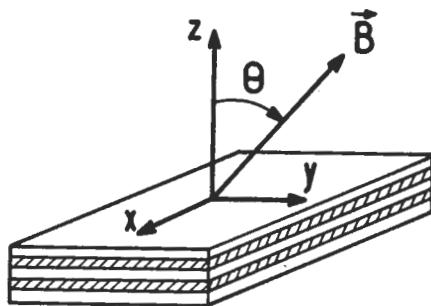


Fig.13.

the particle momentum \mathbf{p} is replaced by $\mathbf{\Pi}$ in the classical Hamiltonian [14]. The vector $\mathbf{\Pi}$ is equal to

$$\mathbf{\Pi} = \mathbf{p} - \frac{q}{c} \mathbf{A} \quad (27)$$

where \mathbf{A} is the vector potential of the magnetic field \mathbf{B}

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (28)$$

$$\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right] \quad (29)$$

and q is the algebraic charge of the carrier ($q = -e$ for electrons, $e > 0$). Unfortunately \mathbf{A} is only a mathematical tool which is in fact not very convenient since, if we replace \mathbf{A} by $\mathbf{A}' = \mathbf{A} + \nabla f$ where f is an arbitrary differentiable function, the magnetic field $\mathbf{B}' = \nabla \times \mathbf{A}'$ coincides with \mathbf{B} . Thus, there is large ambiguity in the choice of the vector potential which corresponds to a given field \mathbf{B} . Since only the field \mathbf{B} retains a physical significance, the results (classical and quantum-mechanical) should only depend on \mathbf{B} and not on the precise choice of \mathbf{A} , i.e. should be gauge-invariant. In the following we will use

$$\mathbf{A} = [Bz \sin \theta, Bx \cos \theta, 0] \quad (30)$$

as well as the notations :

$$B \sin \theta = B_{\perp}; \quad B \cos \theta = B_{\parallel} \quad (31)$$

$$\omega_{c\perp} = \frac{eB_{\perp}}{m^*c}; \quad \omega_{c\parallel} = \frac{eB_{\parallel}}{m^*c} \quad (32)$$

Where m^* is the carrier effective mass and e the absolute value of the electron charge. The electron spin σ interacts with the field adding to the particle energy the quantity

$$g^* \mu_B \sigma \cdot \mathbf{B} \quad (33)$$

where $\mu_B = \frac{e\hbar}{2m_0c}$ is the Bohr magneton and g^* the effectice Landé g factor of the carrier ($g^* = 2$ for electrons in the vacuum, $g^* \leq 0$ in semiconductors). The operator σ has the eigenvalues $\pm \frac{1}{2}$.

The Hamiltonian we have to investigate is (for electrons)

$$\begin{aligned} \mathcal{H} = & \frac{1}{2m_e} \left[-i\hbar \frac{\partial}{\partial x} + \frac{e}{c} B_{\perp} z \right]^2 + \frac{1}{2m_e} \left[-i\hbar \frac{\partial}{\partial y} + \frac{e}{c} B_{\parallel} x \right]^2 + \\ & + \frac{P_z^2}{2m_e} + V_{\text{conf}}(z) + g^* \mu_B \sigma \cdot \mathbf{B} \end{aligned} \quad (34)$$

where $V_{\text{conf}}(z)$ is the potential energy due to the presence of the confining barriers,

and/or due to the averaged electron-electron interaction (self consistent electrostatic potential)

The orbital and spin variables separate. Thus the eigenfunctions of \mathcal{H} can be written as

$$\Psi(\mathbf{r}, \sigma) = \Psi(\mathbf{r}) \alpha(\sigma) \quad (35)$$

Suppose we quantize σ along \mathbf{B} . Let us denote by $|\uparrow\rangle$ and $|\downarrow\rangle$ the two eigenvectors of σ_z , and by $|\nu\rangle$ the eigenstates of the orbital part of \mathcal{H} associated with the eigenvalue ε_ν . We will obtain the eigenstates and eigenenergies of equation (34) in the form :

$$\psi(\mathbf{r}, \sigma) = \langle \mathbf{r} | \nu \rangle \otimes |\uparrow\rangle; \quad \varepsilon_\nu + \frac{1}{2} g^* \mu_B B \quad (36)$$

$$\psi(\mathbf{r}, \sigma) = \langle \mathbf{r} | \nu \rangle \otimes |\downarrow\rangle; \quad \varepsilon_\nu - \frac{1}{2} g^* \mu_B B \quad (37)$$

The spin term is rotationally invariant. Therefore, if we quantize σ along a fixed direction (say z) which is not necessarily colinear to \mathbf{B} , we will obtain (after diagonalizing $g^* \mu_B \sigma \cdot \mathbf{B}$) the same spin eigenvalues $\pm \frac{1}{2} g^* \mu_B B$. Thus the spin-splitting $g^* \mu_B B$ depends on the *modulus* of \mathbf{B} , not on its direction.

Since we now know how to treat the spin effects, we will leave them for a while, and focus our attention on the search for the eigenvalues ε_ν . The spin effects, however, will have to be taken into account when evaluating the carrier density of states.

As far as the orbital part of \mathcal{H} is concerned, we notice that it is y -independent. Thus

$$\psi_\nu(\mathbf{r}) = \frac{1}{\sqrt{L_y}} \exp(i k_y y) \varphi_\mu(x, z) \quad (38)$$

Generally ($\theta \neq 0$, $V_{\text{conf}}(z)$ arbitrary), the x and z variables cannot be separated (for a solvable case : the ni-pi structures, where $V_{\text{conf}}(z) = V_0 \times \frac{z^2}{a^2}$, see Ref. [15]).

However, except when θ is close to $\pi/2$, the separability between the x and z motions appears to be a good zeroth order approximation. In fact, let us rewrite \mathcal{H} in the form :

$$\mathcal{H} = \mathcal{H}_{\parallel} + \mathcal{H}_{\perp} + \delta \mathcal{H} \quad (39a)$$

where

$$\mathcal{H}_{\parallel} = \frac{P_z^2}{2m_e} + V_{\text{conf}}(z) \quad (39b)$$

$$\mathcal{H}_{\perp} = \frac{P_x^2}{2m_e} + \frac{1}{2m_e} \left[\hbar k_y + \frac{e}{c} B_{\parallel} x \right]^2 \quad (39c)$$

$$\delta \mathcal{H} = \delta \mathcal{H}_1 + \delta \mathcal{H}_2 = \frac{1}{2m_e c^2} e^2 B_{\perp}^2 z^2 + P_x z \frac{e B_{\perp}}{m_e c} \quad (39d)$$

where we have made use of equation (38). If the splitting of our original Hamiltonian proves to be correct, the zeroth order energies (i.e. those of $\mathcal{H}_{\parallel} + \mathcal{H}_{\perp}$) will only depend on B_{\parallel} and not on B_{\perp} . This means that, in contrast to the spin term, the eigenvalues ε_{ν} will depend on the component of \mathbf{B} parallel to the growth axis.

An immediate experimental check of this property for a given heterostructure is to rotate the sample with respect to \mathbf{B} and to see whether or not the physical property (of orbital origin) which is under study depends on B_{\parallel} i.e. on the angle θ . For a three-dimensional isotropic system with vanishing $V_{\text{conf}}(z)$, it is very easy to show that the eigenvalues of \mathcal{H} only depend on the modulus of \mathbf{B} and not on its direction (see Appendix A). Thus, provided the decoupling of \mathcal{H} into a main separable term and a small non separable one is valid, a clear cut proof of the existence of quasi bi-dimensional electron gas in a given heterostructure is obtained by studying the angular dependence of magnetic field dependent properties.

Let us then examine the validity of the decoupling procedure. First we investigate the eigenstates of $\mathcal{H}_{\parallel} + \mathcal{H}_{\perp}$. The wavefunctions $\varphi_{\mu}(x, z)$ factorize into

$$\varphi_{\mu}(x, z) = \chi_m(z) \varphi_n(x) \quad (40)$$

where :

$$\mathcal{H}_{\parallel} \chi_n(z) = \xi_m \chi_m(z) \quad (41a)$$

and

$$\mathcal{H}_{\perp} \varphi_n(x) = \left\{ \frac{P_x^2}{2m_e} + \frac{1}{2m_e} \left[\hbar k_y + \frac{eB_{\parallel}}{c} x \right]^2 \right\} \varphi_n(x) = \varepsilon_n \varphi_n(x) \quad (41b)$$

$\chi_m(z)$ are the wavefunctions for the z motion which corresponds to the m^{th} state (bound or unbound) of energy ξ_m for the heterostructure Hamiltonian at zero magnetic field. The functions $\varphi_n(x)$ are the eigensolutions of an harmonic oscillator problem (frequency ω_q) centred at :

$$x_0 = - \frac{\lambda^2 k_y}{\cos \theta} \quad (42)$$

where λ is the magnetic length

$$\lambda = \sqrt{\frac{\hbar c}{eB}} \quad (43)$$

The eigenvalues of equation (41b) are k_y -independent and evenly spaced in $\hbar\omega_q$. They are the celebrated Landau levels [5] :

$$\varepsilon_n(k_y) = \left(n + \frac{1}{2} \right) \hbar\omega_q ; \quad n = 0, 1, 2, \dots \quad (44)$$

The eigenfunctions of equation (39b) which correspond to the eigenvalue $\left(n + \frac{1}{2} \right) \hbar\omega_q$, can be expressed in terms of the Hermite polynomials [9]

$$\varphi_n(x) = \frac{\cos \theta^{1/4}}{\sqrt{\lambda}} \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} \exp \left[-\frac{\cos \theta}{2\lambda^2} (x - x_0)^2 \right] \mathcal{H}_n \left[\left(\frac{x - x_0}{\lambda} \right) \sqrt{\cos \theta} \right] \quad (45)$$

where

$$\mathcal{H}_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} [e^{-x^2}] \quad (46)$$

It can be verified that

$$\langle n-1 | x | n \rangle = \int_{-\infty}^{+\infty} \varphi_{n-1}(x) x \varphi_n(x) dx = \frac{\lambda}{\sqrt{\cos \theta}} \sqrt{\frac{n}{2}} \quad (47)$$

$$\langle n | x | n \rangle = x_0 \quad (48)$$

$$\left\langle n-1 \left| \frac{\partial}{\partial x} \right| n \right\rangle = \frac{\sqrt{\cos \theta}}{\lambda} \sqrt{\frac{n}{2}} \quad (49)$$

We have thus succeeded in classifying the eigenstates of $\mathcal{H}_{\parallel} + \mathcal{H}_{\perp}$. These eigenstates consist of Landau level ladders which are attached to each of the zero field heterostructures bound (or unbound) states. Each of the levels is degenerate with respect to k_y or to x_0 , the centre of the n^{th} harmonic oscillator function. Thus, apart from the spin quantum number we may write :

$$| \nu \rangle = | m, n, k_y \rangle \quad (50)$$

which corresponds to the energy

$$\varepsilon_{\nu} = \xi_m + \left(n + \frac{1}{2} \right) \hbar \omega_{\parallel} \quad (51)$$

The remarkable point is that the carrier motion is entirely quantized $V_{\text{conf}}(z)$ has quantized the z component of the carrier motion, whereas the \mathbf{B}_{\parallel} component, which is parallel to the growth axis, has quantized the in-plane motion.

Each of the eigenvalues ε_{ν} is however enormously degenerate. This orbital degeneracy g_{orb} can be calculated as follows : the centre of the harmonic oscillator function has to be in the crystal. Thus

$$-\frac{L_x}{2} < x_0 < +\frac{L_x}{2} \quad (52)$$

where the dimension L_x is so large that the crystal boundaries were justifiably ignored in equation (39).

The allowed k_y values are uniformly distributed and separated by the interval $2\pi/L_y$ (periodic boundary conditions applied to the y axis). Thus, the degeneracy of the eigenvalue ε_{ν} is equal to

$$g_{\text{orb}} = \frac{L_x L_y}{2\pi\lambda^2} \cos \theta. \quad (53)$$

Notice that g_{orb} is exact at $\theta = 0$ ($\delta\mathcal{H}_1 = \delta\mathcal{H}_2 = 0$ in this case) but only approximate at $\theta \neq 0$ since the effects of $\delta\mathcal{H}_1$, $\delta\mathcal{H}_2$ have not yet been evaluated.

We now have to investigate the effects of $\delta\mathcal{H}$ on the zeroth order eigenstates $|m, n, k_y\rangle$. The first term $\delta\mathcal{H}_1$ does not affect the separability between x and z (as a matter of fact $\delta\mathcal{H}_1$ could have been included into \mathcal{H}_{\parallel}). The only effect is that the different quantum states of the z motion are admixed. For a state $|m\rangle$ which is sufficiently energy-separated from the others (e.g. the ground state), $\delta\mathcal{H}_1$ can be treated by first-order perturbation calculus. Thus ξ_m is changed by $\delta\xi_m$ where :

$$\delta\xi_m = \frac{e^2}{2m_e c^2} B_{\perp}^2 \langle \chi_m | z^2 | \chi_m \rangle. \quad (54)$$

To ensure the convergence of our perturbative scheme, $\langle \chi_n | \delta\mathcal{H}_1 | \chi_m \rangle$ should remain much smaller than the level spacings $\xi_n - \xi_m$. For the states of a particle confined by a potential well of range $\sim L$, $\xi_n \sim \hbar^2 \pi^2 n^2 / 2m^* L^2$ whereas $\langle \chi_n | z^2 | \chi_m \rangle \sim L^2$. Thus, $\langle \chi_1 | \delta\mathcal{H}_1 | \chi_2 \rangle / (\xi_2 - \xi_1) \sim \frac{L^4 \sin^2 \theta}{3\pi^2 \lambda^4}$ and the perturbative approach will be justified for the ground state if the Landau quantization along the z direction remains much smaller than the size quantization arising from $V_{\text{conf}}(z)$.

Suppose, for instance, that the heterostructure consists of a quantum well of thickness L , clad between impenetrable barriers. For the ground state we obtain :

$$\delta\xi_1 = \frac{e^2 B^2}{4m_e c^2} \sin^2 \theta L^2 \left(\frac{1}{6} - \frac{1}{\pi^2} \right) \quad (55)$$

If $B = 10$ T, $\theta = \frac{\pi}{4}$, $L = 100$ Å, $m_e = 0.067 m_0$, we obtain $\delta\xi_1 \sim 0.2$ meV which is much smaller than $\xi_2 - \xi_1 \sim 168$ meV. Note however that $\delta\xi_1 \sim \xi_2 - \xi_1$ if $L \geq 500$ Å. In the latter situation, the first order perturbation becomes insufficient. A better approach would have been to include $\delta\mathcal{H}_1$ into \mathcal{H}_{\parallel} from the very beginning. However, the failure of the perturbative approach indicates that the whole decoupling procedure underlying equations (39a-39b) becomes dubious. This is what happens when the system is weakly bi-dimensional, either because the bound states supported by $V_{\text{conf}}(z)$ are too dense or (when $V_{\text{conf}}(z)$ is finite) too close to the continuum. A better zero-order Hamiltonian is the whole magnetic field dependent term (bulk-like). The perturbation becomes $V_{\text{conf}}(z)$ and, if B is large enough, the effects associated with $V_{\text{conf}}(z)$ can be treated within the subspace which is spanned by all the states of a given n . For more complete treatments of tilted magnetic effects in superlattices, see references [15, 16].

The second kind of perturbing terms, $\delta\mathcal{H}_2$, can be rewritten as $z p_x \omega_{c\perp}$. $\delta\mathcal{H}_2$ has no non-vanishing diagonal elements in the $|m, n, k_y\rangle$ basis since $\langle n | p_x | n \rangle = 0$. To the second order of perturbation theory we obtain :

$$\delta[\xi_m + \varepsilon_n] = \sum_{\substack{(m', n') \\ \neq (m, n)}} \frac{\omega_{c\perp}^2 | \langle \chi_m | z | \chi_{m'} \rangle |^2 | \langle \varphi_n | p_x | \varphi_{n'} \rangle |^2}{\xi_m - \xi_{m'} + (n - n') \hbar \omega_q} \quad (56)$$

Thus, $\delta\mathcal{H}_2$ couples Landau levels of different subbands whose indexes n' , n differ by one. In addition, if $V_{\text{conf}}(z)$ is even in z , only subbands of opposite parities are coupled by $\delta\mathcal{H}_2$. Note that $\delta[\xi_m + \varepsilon_n]$ can become very large (but cannot be described by means of perturbative treatments) when the energy denominators in equation (56) vanish. Practically, this means that two consecutive Landau levels which belong to two adjacent subbands intersect. For the sake of definiteness, let us assume that the involved levels and subbands are $n = 0$, $n = 1$ and $m = 1$, $m = 2$ respectively. Notice that at $\theta = 0$ the two levels $|1, 1, k_y\rangle$ and $|2, 0, k_y\rangle$ can cross since $\delta\mathcal{H}_2$ vanishes. Let us denote by B_c the crossing fields at $\theta = 0$: $B_c = (\xi_2 - \xi_1) \frac{m_e c}{\hbar e}$. At finite θ , the crossing field would take place at $B_c/\cos \theta$. However, the $\delta\mathcal{H}_2$ matrix element which connects the two levels is now non-vanishing and increases with θ . Thus, the crossing is suppressed and replaced by an anticrossing (see Fig. 14). Retaining the dominant $\delta\mathcal{H}_2$ contributions and neglecting $\delta\mathcal{H}_1$, we may write

$$|\psi\rangle \sim \alpha |1, 1, k_y\rangle + \beta |2, 0, k_y\rangle \quad B \sim B_c/\cos \theta \quad (57)$$

The eigenvalues are the solutions of

$$\begin{bmatrix} \xi_1 + \frac{3}{2} \hbar \omega_q - \varepsilon & \langle 2, 0, k_y | \delta\mathcal{H}_2 | 1, 1, k_y \rangle \\ \langle 1, 1, k_y | \delta\mathcal{H}_2 | 2, 0, k_y \rangle & \xi_2 + \frac{1}{2} \hbar \omega_q - \varepsilon \end{bmatrix} = 0 \quad (58)$$

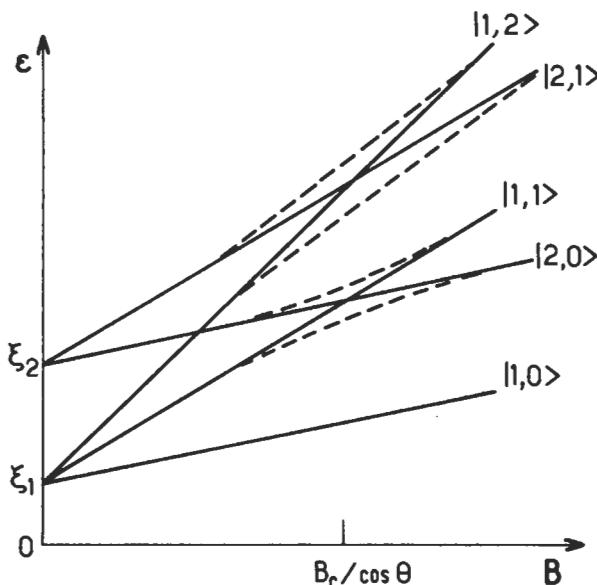


Fig. 14. — Schematic fan charts of Landau levels attached to two subbands 1 and 2 with confinement energies ξ_1 and ξ_2 respectively. The dashed lines represent the effect of the anticrossing induced by $\delta\mathcal{H}_2$.

or

$$\varepsilon_{\pm} = \frac{\xi_1 + \xi_2}{2} + \hbar\omega_q \pm \sqrt{\left(\frac{\xi_2 - \xi_1 - \hbar\omega_q}{2}\right)^2 + |\langle 2, 0, k_y | \delta\mathcal{H}_2 | 1, 1, k_y \rangle|^2} \quad (59)$$

The magnitude of the anticrossing (which is the energy difference $\delta\varepsilon = \varepsilon_+ - \varepsilon_-$ evaluated at $B = B_c/\cos\theta$) is

$$|\delta\varepsilon| = \frac{1}{\lambda(B_c)} \langle \chi_1 | z | \chi_2 \rangle \frac{\hbar e B_c}{m_e c} \tan\theta \sqrt{2} \quad (60)$$

which varies like $B_c^{3/2}$. At the anticrossing the mixing between the two interacting levels is complete: $|\alpha| = |\beta| = \frac{1}{\sqrt{2}}$. Note finally the close formal analogy

between this type of anticrossing and those which arise when the cyclotron energy $\hbar\omega_q$ becomes equal to the energy of the longitudinal optical phonon energy $\hbar\omega_{LO}$ (magneto-polaron [17]).

II.2 MAGNETIC FIELD DEPENDENT DENSITY OF STATES. — Since the effects of $\delta\mathcal{H}_1$ are only approximately known, our attention in this paragraphe will be restricted to the case of $\theta = 0$, taking into account the spin-splitting effect. As the z and (x, y) motions are separable, we can decompose the total density of states into contributions which arise from different subbands m :

$$\rho(\varepsilon) = \sum_m \rho_m(\varepsilon) \quad (61)$$

where

$$\rho_m(\varepsilon) = \sum_{\sigma_z, k_y, n} \delta \left[\varepsilon - \xi_m - \left(n + \frac{1}{2} \right) \hbar\omega_q - g^* \mu_B \sigma_z \right]; \sigma_z = \pm \frac{1}{2} \quad (62)$$

Since the eigenenergies are k_y -independent, the summation over k_y can immediately be performed. This results in the appearance of the orbital degeneracy factor g_{orb} . Thus

$$\rho_m(\varepsilon) = \frac{S}{2\pi\lambda^2} \sum_{n, \sigma_z} \delta \left[\varepsilon - \xi_m - \left(n + \frac{1}{2} \right) \hbar\omega_q - g^* \mu_B B \sigma_z \right] \quad (63)$$

As the carrier motion is entirely quantized, the density of states is zero except at the discrete energies

$$\varepsilon_{mn\sigma_z} = \xi_m + \left(n + \frac{1}{2} \right) \hbar\omega_q + g^* \mu_B B \sigma_z \quad (64)$$

where it is infinite. The results of equations (62-64) are quite remarkable when compared with the zero field situation:

$$\rho_{m\sigma_z}^{B=0}(\varepsilon) = \frac{m_e S}{2\pi \hbar^2} Y(\varepsilon - \xi_m) \quad (65)$$

where $Y(x)$ is the step function. At zero field, the quasi bi-dimensional electron gas is metallic, with a gapless density of states for $\varepsilon > \xi_1$. A finite magnetic field dramatically alters this situation, replacing the continuum equation (65) by point-like singularities equation (63) separated by gaps. Thus, depending on the respective locations of the characteristic electron energy η (the Fermi energy for a degenerate electron gas at $T = 0$ K) and $\varepsilon_{mn\sigma_z}$, the electron gas will behave either like an insulator ($\eta \neq \varepsilon_{mn\sigma_z}$) or a metal ($\eta = \varepsilon_{mn\sigma_z}$). Notice that similar effects cannot be found in bulk materials : the carrier free motion along the field prevents the density of states from vanishing at any energy larger than $\frac{1}{2}\hbar\omega_c - \frac{1}{2}|g_c^*|\mu_B B$ (see Appendix A).

Of course, we know that imperfections (impurities, interface defects, etc...) will alter the delta-like singularities of $\rho_m(\varepsilon)$, rounding off the peaks, adding states into the gaps etc... . However, it remains generally accepted that, if B is large enough (i.e. $B \geq$ few Teslas in good GaAs-Ga_{1-x}Al_xAs heterostructures at low temperature), the previous conclusions retain their validity, at least for the following meaning : for energies belonging to a certain bandwidth of finite (yet undefined) extension, the density of states is large and corresponds to conducting states, whereas for other energy segments, the states are localized, i.e. at $T = 0$ K are unable to contribute to the electrical conduction at vanishingly small electric fields.

Some words have to be added concerning the broadening effects. The present status of the art is far from being satisfactory. Firstly, there is, to the authors' knowledge, no general consensus on the microscopic origin of these effects. Most likely is it sample-dependent : in some samples, charged impurities play a dominant part, in others, interface roughness or alloy scattering (if it exists) may dominate. Secondly, even if we assume that the disorder potential has a simple algebraical form (e.g. uncorrelated delta-like scatterers), it proves difficult, if not impossible, to obtain precise information on $\rho_m(\varepsilon)$ by any methods other than by numerical simulations. The complications are twofold :

i) the unperturbed density of states is highly singular. This precludes the use of non-self-consistent perturbative treatments of the disorder potential.

ii) when the required self-consistency is included, calculations are performed up to a finite order in the perturbation approach. By doing so, the nature of the states which occur in the tails of the broadened delta functions are ill-treated : a localized level requires an inclusion of the electron-disorder potential to an infinite order.

A simple model of a broadened density of states has been derived by Ando and Uemura [18]. It consists of a self-consistent treatment performed at the Born approximation (second order) of the electron-disorder potential interaction. The disorder is assumed to originate from randomly distributed impurities. The scattering potentials, due to a given impurity, is taken as being Gaussian-shaped in space. Thus, the effects associated with the short-range or long-range nature of the impurity

potential can be studied in this model. Finally, the magnetic field was assumed to be strong enough to neglect the disorder-induced mixing of the Landau levels. The calculations result in replacing the delta functions $\delta(\varepsilon - \varepsilon_{mn\sigma_z})$ in equation (63) by semi-ellipses :

$$\delta(\varepsilon - \varepsilon_{mn\sigma_z}) \rightarrow \frac{2}{\pi \Gamma_n} \sqrt{1 - \left(\frac{\varepsilon - \varepsilon_{mn\sigma_z}}{\Gamma_n} \right)^2} \quad (66)$$

where, if the scatterers are of short range, the Landau level broadening $2\Gamma_n$ is independent of n and is proportional to $(B/\mu_0)^{1/2}$ where μ_0 is the zero-field electron mobility of the quasi bi-dimensional electron gas.

The drawback of the semi-ellipses is their abrupt cut-off at $\varepsilon_{mn\sigma_z} \pm \Gamma_n$. We should expect $\rho_m(\varepsilon)$ to fall off more gently near the band extrema. A popular, empirical formula for $\rho_m(\varepsilon)$ is provided by replacing the delta function in equation (63) by Gaussian functions :

$$\delta(\varepsilon - \varepsilon_{mn\sigma_z}) \rightarrow \frac{1}{\Gamma_n \sqrt{2\pi}} \exp \left[-\frac{(\varepsilon - \varepsilon_{mn\sigma_z})^2}{2\Gamma_n^2} \right] \quad (67)$$

However one does not know in equation (67) how to relate Γ_n to B , n etc... In spite of this weakness, equation (67) has proved to be much closer to the exact shape of the density of states than equation (66).

The latter has been calculated by Ando who performed a computer diagonalization of the one-electron Hamiltonian which includes disorder [19] (see Fig. 15).

A more severe drawback than not reproducing the exact shape of the density of states is that equations (66, 67) overlook a crucial point, i.e. the spatial nature (extended or localized) of the disorder-allowed energy levels in the density of states. For instance, in equation (66) all the states with energies such that $|\varepsilon - \varepsilon_{mn\sigma_z}| < \Gamma_n$ are capable of conducting the electrical current, even at $T = 0$ K. In reality, one would expect that by going deeper into the bandgaps, the allowed electron states should be characterized by wavefunctions which become more and more strongly localized in space. The nature of the localization process is still disputed. Does it correspond to a simple type of localization, i.e. is it associated with a single local extremum of the disorder potential ? If so, it would resemble the familiar situation of the non-interacting impurity states which exist in the bandgap of semiconductors. Is it on the other hand the result of constructive interference effects which are experienced by the carrier wavefunction when it interacts with many randomly distributed local extrema ? In the latter case would it be another example of the celebrated Anderson localization ?

In addition to providing the shape of the density of states, Ando's computer simulations [19] have revealed that the localization length (i.e. the characteristic distance over which the wavefunction of a localized state decays) markedly decreases for states whose energies are found deeper and deeper in the forbidden gaps (see Fig. 16). Thus, if we assimilate to a delocalized electron state, a state whose calculated localization length is comparable to or exceeds the sample size, there should exist

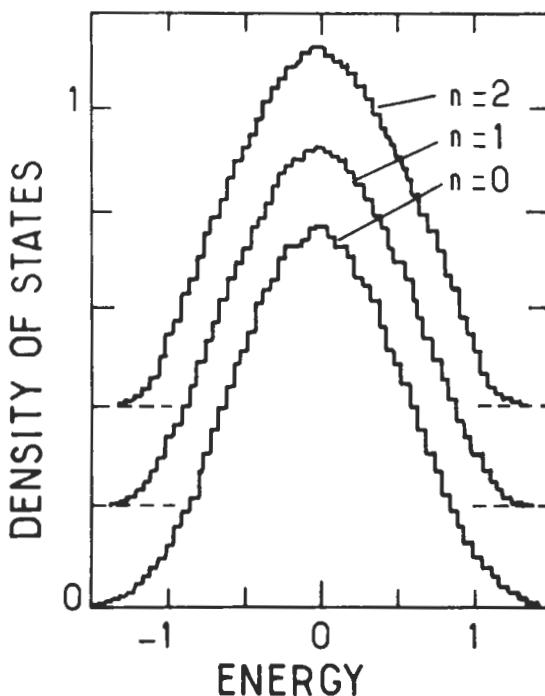


Fig. 15. — The histograms of the densities of states of the $n = 0, 1$ and 2 Landau levels broadened by short-range scatterers are plotted *versus* energy. For each curve, the energy origin is taken at the centre of the unperturbed Landau level. The energy unit is Γ , the level broadening calculated in the self consistent Born approximation and the density of states unit is $\frac{1}{2\pi\lambda^2\Gamma}$, where λ is the magnetic length $\left(\lambda = \frac{\hbar c}{eB}\right)$. After reference [19].

some kind of critical energy in the density of states of a broadened Landau level which marks the border between localized and delocalized levels. Granted that such a transition exists, the nature of the localized \rightarrow delocalized transition is for the present, rather mysterious. Nevertheless, it has become customary to consider that in Gaussian density of states (Eq. (67)), only a central fringe of states is capable of conducting the electrical current at $T = 0$ K. All the others, that is to say, the tail states may however participate in the conduction at $T \neq 0$ K, but only through phonon-assisted hopping between the localized states.

There have been few attempts to directly measure the magnetic field-dependent density of states of quasi bi-dimensional electron gases in heterostructures. This is due to the smallness of the signals in the required experiments. One of the most direct ways to measure this density of states is by using the de Haas-van Alphen effect, where the oscillations of the magnetic moment are measured *versus* the field. The magnetic moment is an extensive quantity, i.e. it is directly proportional to the number of conduction electrons in the sample. To illustrate the difficulty of de Haas-van Alphen experiments in heterostructures, let us compare a typical heterostructure

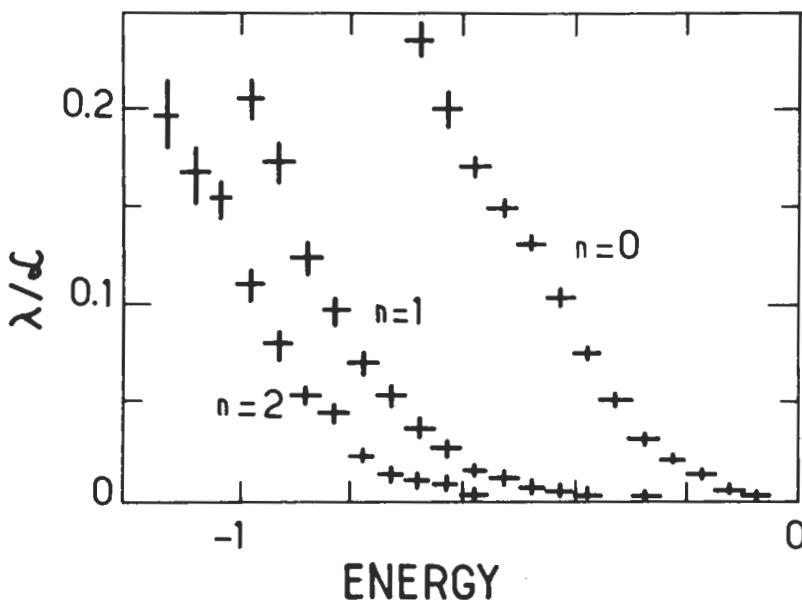


Fig. 16. — The inverses of the localization lengths \mathcal{L} (in units of the magnetic length λ) of the $n = 0, 1$ and 2 Landau levels broadened by short-range scatterers are plotted *versus* the energy. For each set of numerical data, the energy origin is taken at the center of the corresponding unperturbed Landau level. The energy unit is Γ , the level broadening calculated in the self consistent Born approximation. After reference [19].

situation with that of a bulk metal. Such a heterostructure consists of 100 active layers, each of which contains 10^{12} cm^{-2} electrons. Since sample areas are (for homogeneity purposes) only a few mm^2 we find that $10^{12} - 10^{13}$ electrons contribute to the de Haas-van Alphen signal. For a bulk metallic sample, whose volume is $\sim 1 \text{ mm}^3$, there are $\sim 10^{18}$ electrons which contribute to the signal. Thus, the de Haas-van Alphen effect in heterostructures is many orders of magnitude too small to allow for any clear cut experiments. The only unambiguous information extracted from de Haas-van Alphen experiments [20] is that the position of their oscillations compare favorably with those of the Shubnikov-de Haas experiments (Fig. 17). Thus, from both theoretical and experimental points of view, the shape of the density of states and the nature of the tail states remain open to question.

II.3 MAGNETOCONDUCTIVITY OF A QUASI BI-DIMENSIONAL ELECTRON GAS.

II.3.1 Macroscopic derivation. — Let us consider a quasi bi-dimensional electron gas subjected to a strong magnetic field. We are interested in its linear response to a weak static electric field $\mathbf{F} \perp \mathbf{z}$, i.e. applied in the layer plane of the heterostructure. The electron gas responds to this external disturbance which an electrical current

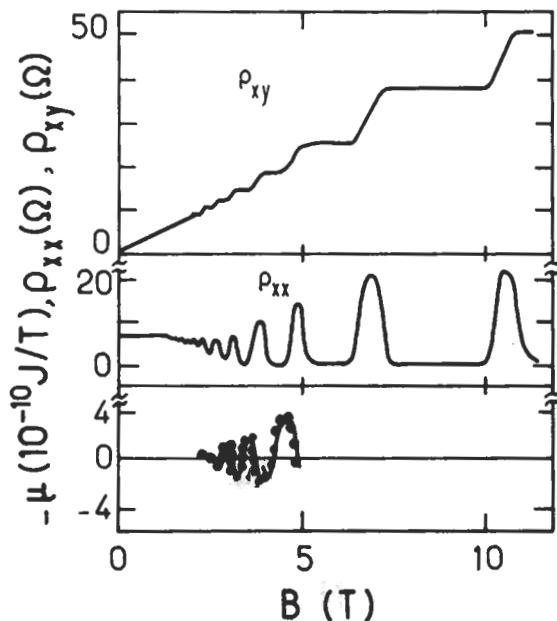


Fig. 17. — The oscillatory behaviour, with the magnetic field B , of the magnetic moment μ , the Hall resistivity ρ_{xy} and the transverse resistivity ρ_{xx} in a $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ modulation-doped superlattice. $T = 1.5$ K. After reference [20].

which, since F is assumed to be small, will be proportional to F . Thus, we may write that the x, y components of the areal current density \mathbf{J} are related to F by

$$\begin{aligned} J_x &= \sigma_{xx} F_x + \sigma_{xy} F_y \\ J_y &= \sigma_{yx} F_x + \sigma_{yy} F_y \end{aligned} \quad (68)$$

As we are dealing with cubic host materials $\sigma_{xx} = \sigma_{yy}$ and the Onsager relationships [21] show that

$$\sigma_{xy} = -\sigma_{yx} \quad (69)$$

The conductivity tensor $\vec{\sigma}$ is the quantity which is most naturally calculated. Practically however, this is the resistivity tensor :

$$\vec{\rho} = \vec{\sigma}^{-1} \quad (70)$$

which is the most accessible to the experiments. If we invert equation (68), we obtain :

$$\rho_{xx} = \rho_{yy} = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{xy}^2} \quad (71)$$

$$\rho_{yx} = -\rho_{xy} = \frac{\sigma_{xy}}{\sigma_{xy}^2 + \sigma_{xx}^2} \quad (72)$$

Let us consider a sample with a "spider" geometry. It consists of a long bar ($L_x \gg L_y$) with side contacts (see Fig. 18). The voltage V_x is imposed via a battery and the voltage V_H is measured across the small dimension of the bar. Since J_y vanishes we have

$$\begin{aligned} F_x &= \rho_{xx} J_x \\ F_y &= \rho_{yx} J_x \end{aligned} \quad (73)$$

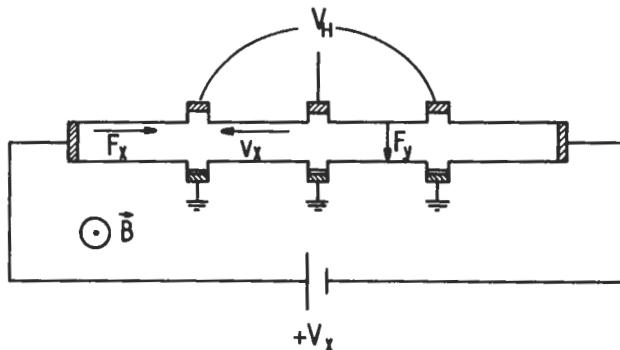


Fig. 18. — Schematic representation of the sample geometry used to measure the resistivity tensor $\rho(B)$. V_x and V_H denote the potential with respect to the ground.

If one assumes that the current densities J_x, J_y are uniform in the sample (the spatial variation of J is still a matter of controversy) we may write

$$J_x = \frac{I_x}{L_y} ; \quad V_H = F_y L_y ; \quad V_x = F_x L_x \quad (74)$$

Thus

$$V_x = \frac{\rho_{xx}}{L_y} L_x J_x = R_{xx} I_x \quad (75)$$

$$V_H = \rho_{yx} I_x = R_{yx} I_x$$

where V_x, V_H are expressed in volts, I_x in amperes and R_{xx}, R_{yx} in ohms. Finally

$$R_{xx} = \frac{L_x}{L_y} \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{xy}^2} \quad (76)$$

$$R_{yx} = \frac{\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2} \quad (77)$$

The resistance R_{xx} involves the geometrical factor L_x/L_y and thus is sample-dependent. In literature, people often get rid of this inconvenience by quoting the resistance which would be measured in samples with a square shape but which are otherwise identical to their investigated samples. Thus, to recall this feature, R_{xx} measurements are given in units of ohm square (Ω_{\square}).

On the other hand, the Hall resistance R_{yx} is independent of the sample shape. This explains why several Bureaux of Standards, in the hope of defining a new resistance standard, are interested in accurate R_{yx} data [22]. The former will be independent of sample size and thus free of dilatation effects which often plague the usual ohm standards which are, presently, based on the resistance of platinum rods. Finally, let us stress once more that both R_{xx} , R_{yx} as well as ρ_{xx} , ρ_{xy} are expressed in ohms. Nothing is wrong with the latter property. Only the resistivities ρ_{xx} , ρ_{yx} are two-dimensional resistivities. We know that, in the dimension equation

$$[\rho] = \frac{m_e}{n_e e^2 \tau} \quad (78)$$

as in bulk materials (τ has the dimension of a relaxation time). However in equation (78), n_e is an areal electron concentration instead of a volumic concentration as usually found in bulk materials.

III.3.2 Microscopic discussion. — In Appendix B the linear response theory, originally derived by Kubo, has been used to evaluate the components σ_{xx} and σ_{yx} of the static magnetoconductivity tensor in the independent electron approximation [23]. We found that at $T = 0$ K

$$\sigma_{xx} = \frac{\pi \hbar e^2}{(m_e)^2 S} \text{Trace} \{ \delta(\varepsilon_F - \mathcal{H}) \Pi_x \delta(\varepsilon_F - \mathcal{H}) \Pi_x \} \quad (79)$$

$$\sigma_{yx} = \frac{-ie^2 \hbar}{(m_e)^2 S} \int_{-\infty}^{+\infty} f(\varepsilon) d\varepsilon \text{Trace} \left\{ \delta(\varepsilon - \mathcal{H}) \left[\Pi_y \frac{1}{(\varepsilon - \mathcal{H})^2} \Pi_x - \Pi_x \frac{1}{(\varepsilon - \mathcal{H})^2} \Pi_y \right] \right\} \quad (80)$$

where $f(\varepsilon)$ is the Fermi Dirac distribution function which, at zero temperature, is a step function of the argument $\varepsilon_F - \varepsilon$ where ε_F is the Fermi level. In equations (79, 80) Π is $\mathbf{p} + \frac{e}{c} \mathbf{A}$ (Eq. (27)) and \mathcal{H} is the quasi bi-dimensional Hamiltonian which includes disorder :

$$\mathcal{H} = \frac{p_x^2}{2m_e} + \frac{1}{2m_e} \left(p_y + \frac{e}{c} Bx \right)^2 + E_1 + g^* \mu_B B \sigma_z + V_{\text{disorder}}(\mathbf{p})$$

$$\mathcal{H} = \mathcal{H}_0 + V_{\text{disorder}}(\mathbf{p}) \quad (81)$$

where $V_{\text{disorder}}(\mathbf{p}) = \int dz \chi_1^2(z) V_{\text{disorder}}(\mathbf{r}) \quad (82)$

and E_1 is the confinement energy of the ground subband (wavefunction χ_1) which arises from the size quantization along the growth axis of the heterostructures. In equations (81, 82) $\mathbf{p} = (x, y)$, $\mathbf{r} = (x, y, z)$. To make sense, equation (81) requires a

pronounced bi-dimensionality of the electron gas, i.e. that the excited subbands $E_2\dots$ must be far enough away from E_1 to allow a factorization of the carrier wavefunction $\psi(\mathbf{r})$ in $\chi_1(z) \varphi(\boldsymbol{\rho})$. The assumed strong bi-dimensionality also allows us to treat the case of a tilted magnetic field with respect to the growth axis. Only the magnetic field \mathbf{B} which appears in equation (81) has to be understood as the projection of \mathbf{B} along the $\hat{\mathbf{z}}$ axis. The matrix elements of Π_x , Π_y on the Landau levels basis (eigenstates of \mathcal{H}_0) are :

$$\langle n, k_y, \sigma_z | \Pi_x | n', k'_y, \sigma'_z \rangle = \frac{i\hbar}{\lambda \sqrt{2}} \times \\ [\sqrt{n} \delta_{n', n-1} - \sqrt{n+1} \delta_{n', n+1}] \delta_{k'_y, k_y} \delta_{\sigma'_z, \sigma_z} \quad (83)$$

$$\langle n, k_y, \sigma_z | \Pi_y | n', k'_y, \sigma'_z \rangle = \frac{\hbar}{\lambda \sqrt{2}} \times \\ [\sqrt{n} \delta_{n', n-1} + \sqrt{n+1} \delta_{n', n+1}] \delta_{k'_y, k_y} \delta_{\sigma'_z, \sigma_z} \quad (84)$$

In a perfect heterostructure $V_{\text{disorder}}(\boldsymbol{\rho})$ vanishes. Thus the states $|n, k_y, \sigma_z\rangle$ are eigenstates of \mathcal{H} . Consequently the matrix elements

$$\langle n, k_y, \sigma_z | \delta(\varepsilon - \mathcal{H}) | n', k'_y, \sigma'_z \rangle$$

are equal to

$$\langle n, k_y, \sigma_z | \delta(\varepsilon - \mathcal{H}) | n', k'_y, \sigma'_z \rangle = \delta(\varepsilon - \varepsilon_{n\sigma_z}) \delta_{nn'} \delta_{k'_y k_y} \delta_{\sigma'_z \sigma_z} \quad (85)$$

Equations (83-85) immediately lead to

$$\sigma_{xx} = 0 ; \quad \sigma_{yx} = \frac{n_e e^2}{m_e \omega_c} \quad (86)$$

Thus, for a perfect system the microscopic evaluation of the conductivity tensor leads to the same result as the classical Drude-type of calculations (see Appendix A for a three-dimensional version of the Drude model).

Equations (86) are not surprising but merely serve as a reminder that, in the absence of scattering, a carrier subjected to crossed electric and magnetic fields actually drifts along a direction perpendicular to these two fields. The electrical transport along the electric field is entirely due to collisions, i.e. to a non-vanishing $V_{\text{disorder}}(\boldsymbol{\rho})$.

We have already stated that very little is known about the exact shape of $V_{\text{disorder}}(\boldsymbol{\rho})$. Its spatial average over a macroscopic area should, however vanish. Thus, in the (x, y) plane, $V_{\text{disorder}}(\boldsymbol{\rho})$ shows uncorrelated fluctuations which are either repulsive or attractive for electrons. (Notice that for very large magnetic fields, *both* kinds of fluctuations, when considered as isolated, are capable of supporting bound states). Taking the randomness of the disorder potential into account implies that equations (79, 80) should be completed by a spatial average over V_{disorder} and, by

taking the limit $S \rightarrow \infty$. If V_{disorder} can be written as the sum of one-defect potentials centred at ρ_i

$$V_{\text{disorder}}(\rho) = \sum_{\rho_i} W(\rho - \rho_i) \quad (87)$$

the spatial average signifies an average over the positions ρ_i of the defects which are assumed to be uncorrelated

i) Magnetic field dependence of σ_{xx} (Shubnikov-de Haas effect)

Let us evaluate σ_{xx} by taking the eigenstates $|\nu\rangle$ of \mathcal{H} as the basis vectors

$$\sigma_{xx} = \frac{\pi e^2 \hbar}{(m_e)^2 S} \sum_{\mu, \nu} \delta(\varepsilon_F - \varepsilon_\mu) \delta(\varepsilon_F - \varepsilon_\nu) |\langle \nu | \Pi_x | \mu \rangle|^2 \quad (88)$$

Equation (88) shows that the zero temperature conductivity σ_{xx} is proportional to the number of pairs of states which have exactly the Fermi energy ($\varepsilon_F = \varepsilon_\mu = \varepsilon_\nu$), weighted by the square of the Π_x matrix element between them.

Let us now suppose that we adopt the scheme of broadened Landau levels whose tail states are localized. Then, if the Fermi energy stays within the tail states, the matrix element $\langle \nu | \Pi_x | \mu \rangle$ will decay (say exponentially) with $|\rho_\mu - \rho_\nu|$ which is the distance separating the two localized states $|\mu\rangle, |\nu\rangle$. In addition, due to the randomness of $V_{\text{disorder}}(\rho)$ the number of states with exactly the same energy $\varepsilon_\mu = \varepsilon_\nu$ is extremely small. All these facts lead us to the conclusion that σ_{xx} is very small at $T = 0$ K. More sophisticated analysis [24] show that σ_{xx} does indeed vanish when the limit of infinite S is taken. The previous argument does not hold for the extended states of the broadened Landau levels at least because $\langle \nu | \Pi_x | \mu \rangle$ is $\sim \frac{1}{\lambda}$ for those states. Thus, when the magnetic field is swept and the Fermi energy passes through the consecutive Landau levels, σ_{xx} undergoes oscillations, and is zero when ε_F is in the localized levels (or in a region where the density of states vanishes) and positive when ε_F coincides with the extended states of the spectrum.

Such behaviour is analytically derived within the framework of the self-consistent Born approximation. In such a model, one obtains [18] :

$$\sigma_{xx} = 0 \quad \text{if} \quad \left| \varepsilon_F - \left(n + \frac{1}{2} \right) \hbar \omega_c - E_1 - g^* \mu_B B \sigma_z \right| > \Gamma_n$$

$$\sigma_{xx} = \frac{e^2}{\pi^2 \hbar} \left(n + \frac{1}{2} \right) \left\{ 1 - \left[\frac{\varepsilon_F - \left(n + \frac{1}{2} \right) \hbar \omega_c - E_1 - g^* \mu_B B \sigma_z}{\Gamma_n} \right]^2 \right\} \quad (89)$$

elsewhere. Here $2\Gamma_n$ is the width of the n^{th} broadened Landau level whose density of states is given by equations (63, 66). Despite the incorrect treatment of the tail states, the formula (Eq. (89)) is very helpful in explicitly showing the oscillatory behaviour of σ_{xx} upon B .

These oscillations are well known (Shubnikov-de Haas effect) [4, 23] and have been observed in a variety of bulk materials and heterostructures. They are of quantum origin (being absent in a Drude-type analysis of the magnetoconductivity tensor) and evidence the magnetic field-induced singularities of the density of states.

The σ_{xx} maxima (or equivalently the ρ_{xx} maxima since $\sigma_{xx} \ll \sigma_{xy}$ in Eq. (71)) are evenly spaced in $\frac{1}{B}$. The fields $B_{n\sigma_i}$ where σ_{xx} is maximum are such that :

$$\varepsilon_F - E_1 = \left[n + \frac{1}{2} + \frac{g^* \mu_B \sigma_z m_e}{\hbar e} \right] \frac{\hbar e}{m_e} B_{n\sigma_i} \quad (90)$$

This means that a plot of $B_{n\sigma_i}^{-1}$ versus integer n should be a straight line (Fig. 19) whose slope S_B yields direct information on $\varepsilon_F - E_1$ or, equivalently, on the areal carrier concentration n_e which is present in the heterostructure :

$$n_e = \frac{m_e}{\pi \hbar^2} (\varepsilon_F - E_1) \quad (91)$$

$$S_B = \frac{\hbar e}{m_e} \frac{1}{(\varepsilon_F - E_1)} = \frac{e}{\pi \hbar n_e} \quad (92)$$

The determinations of n_e by means of the Shubnikov de Haas effect are very accurate. In addition they do not require any information concerning the sample geometrical details (shape, distance between contacts etc...)

ii) Magnetic field dependence of σ_{xy} : the quantized Hall effect.

In the self consistent Born approximation, there should exist corrections to the perfect Hall effect ($\sigma_{xy} = -n_e e^2/m_e \omega_c$). However, it is not very interesting to discuss these corrections insofar as the most spectacular effects are completely missed by any theory which neglects the localized states in the tails of the Landau levels.

What was experimentally observed is the appearance of *plateaus* in the ρ_{xy} curves versus B (Fig. 20) at low temperatures ($T \leq 4.2$ K). The ρ_{xy} plateaus are exactly (within at least one part per million) quantized, being subharmonics of h/e^2 . The so-called quantized Hall effect, discovered by von Klitzing *et al.* [6] in Si-MOSFET's, can be termed a universal effect. It has also been observed in other two-dimensional systems $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ [25] ; $\text{InP-Ga}_{0.47}\text{In}_{0.53}\text{As}$ [26] ; GaSb-InAs [28] ; AlSb-InAs [27]. Its occurrence is independent of the band structure of the host materials : the quantized Hall effect can be observed when the carriers are electrons belonging to a single conduction band ($\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$) [25] or to a multi-valleys conduction band (Si-MOSFET) [6]. It is also observed when the carriers are holes belonging to a degenerate valence band ($\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs}$) and [29] even when there is a mixed conduction (GaSb-InAs) [28].

The width ΔB of a given ρ_{xy} plateau increases as the temperature decreases, reaching an amplitude which is close to the maximum one at very low T (milli-Kelvin range) [30]. It appears established that ρ_{xy} develops plateaus when the Fermi level coincides with the localized states of the broadened Landau levels. In fact, in the

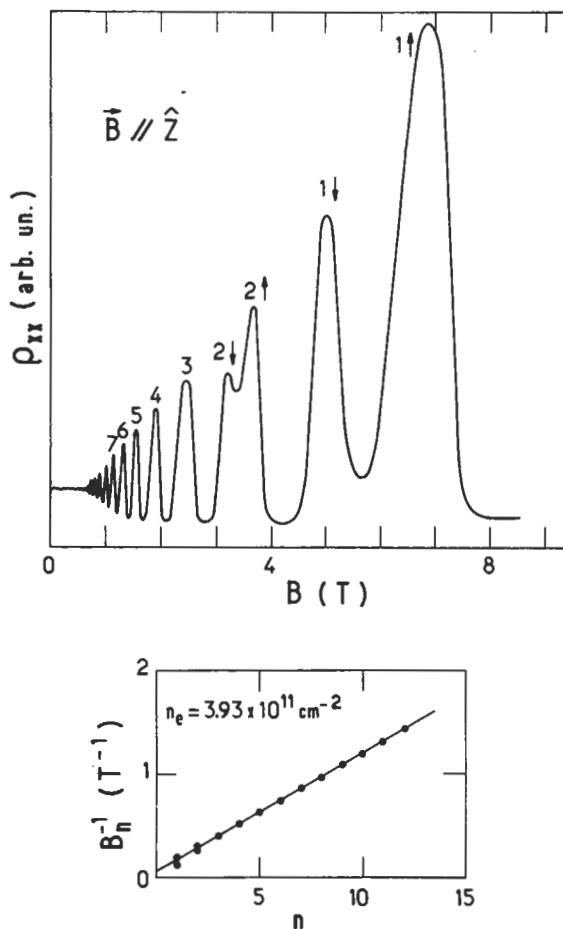


Fig. 19. — A representative example of the oscillatory behaviour displayed by $\rho_{xx}(B)$ upon B (Shubnikov de Haas effect) in a GaAs-Ga_{0.7}Al_{0.3}As heterojunction at $T = 2$ K (upper panel). The inverses of the magnetic fields $1/B_n$ corresponding to maxima of $\rho_{xx}(B)$ are plotted against consecutive integers n (lower panel). The slope of the straight line B_n^{-1} versus n directly yields the areal concentration of quasi bi-dimensional electrons n_e (see Eq. (92)). Courtesy Y. Guldner.

plateau regime for ρ_{xy} , ρ_{xx} (and σ_{xx}) are almost zero. The deviation from the zero resistance state exponentially decreases as temperature decreases. Such a feature evidences a hopping type of transport along the electric field. The details of the hopping processes remain uncertain : the hopping laws

$$\sigma_{xx}(T) \approx \exp \left[- \left(\frac{T_0}{T} \right)^\alpha \right] \quad (93)$$

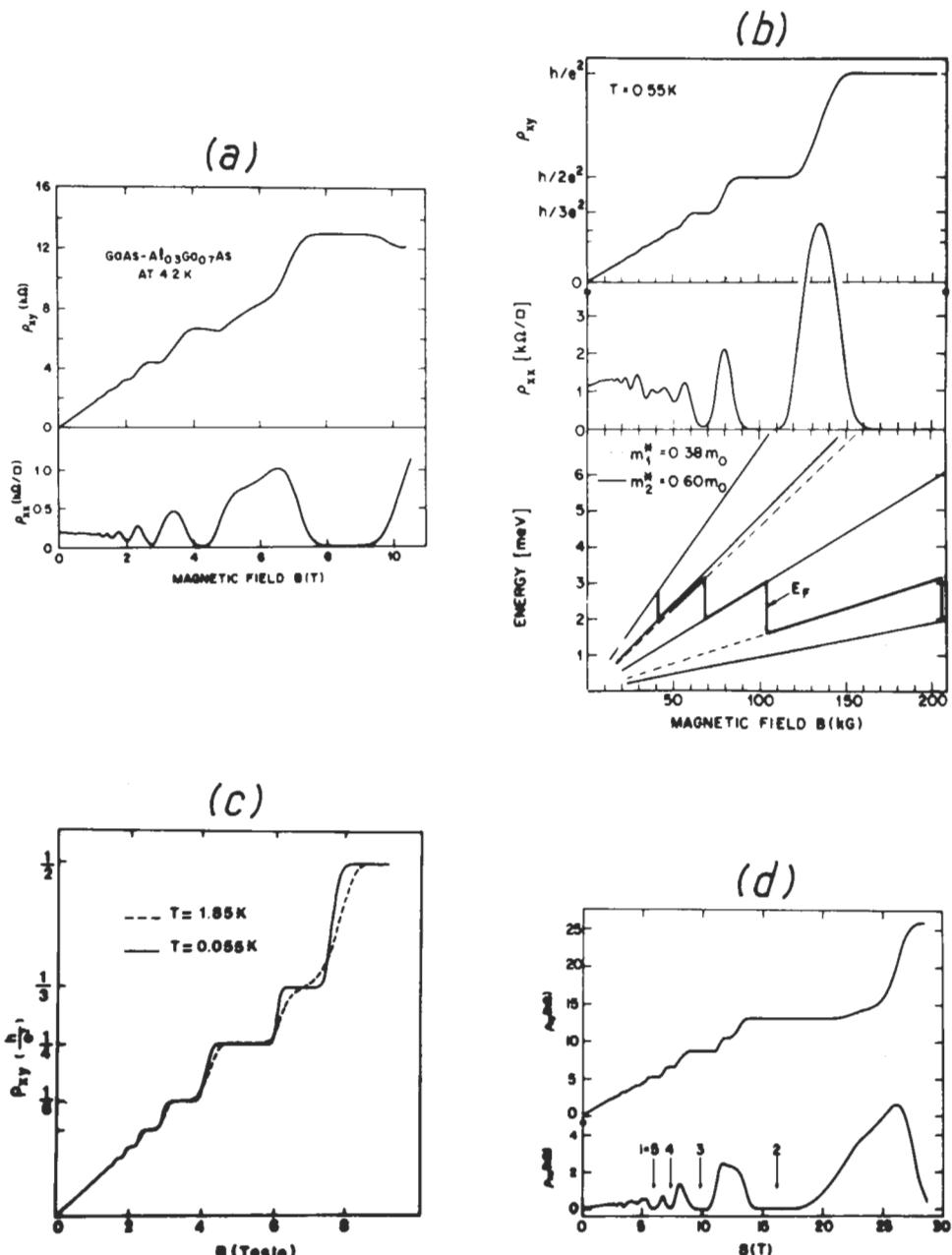


Fig. 20. — Ubiquity of the quantized Hall effect in III-V heterostructures. The ρ_{xy} curves plotted versus B are shown for n-Ga_{1-x}Al_xAs-GaAs [25] (Fig. a), n-InP-Ga_{0.47}In_{0.53}As [26] (Fig. c), p-Ga_{1-x}Al_xAs-GaAs [29] (Fig. b) and n-GaSb-InAs-GaSb [28] (Fig. d) heterostructures respectively.

fit the data reasonably well (Fig. 21) but the exponent α is sample-dependent [28, 31, 33]. Perhaps the disorder potential responsible for the existence of localized states is markedly different in different types of heterostructures.

That σ_{xy} could be exactly quantized in units of a fundamental constant $\left(\frac{e^2}{h}\right)$ irrespective of the band structure, defects etc... clearly called for a theoretical explanation which would go beyond the usual perturbative approaches of the magnetoconductivity tensor and which would exploit the genuine characteristic of

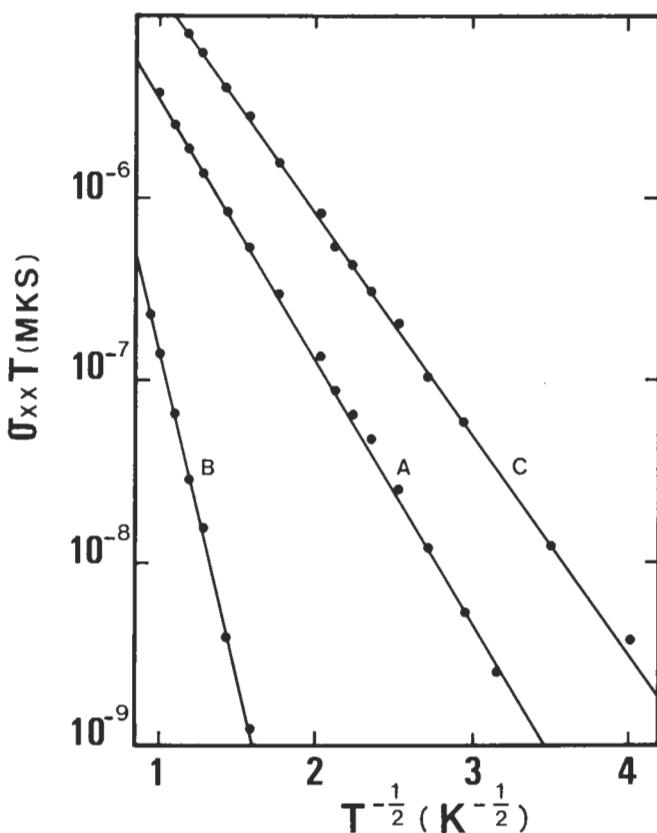


Fig. 21. — An experimental evidence of a hopping conductivity which occurs when the Fermi level stays in the located states of the tails of a Landau level : the product $T\sigma_{xx}(T)$ is plotted versus $T^{-1/2}$ (semi logarithmic scale). For this particular $\text{InP-Ga}_{0.47}\text{In}_{0.53}\text{As}$ heterostructure, the activated conductivity is well described by a formula $\sigma_{xx}(T) \sim T^{-1} \exp - \left[\frac{T_0}{T} \right]^{1/2}$ between $5 \times 10^{-2} \text{ K}$ and 1 K . Notice, however, that the T_0 's depend on the Landau level index : the three lines labelled A, B, C correspond to the Fermi level sitting between the $n = 2$ and $n = 3$ (curve A) ; $n = 1^-$ and $n = 2$ (curve B) and $n = 1^+$ and 1^- (curve C) Landau levels respectively. After reference [32].

two-dimensional magnetotransport, which is the existence of non-conducting states separating conducting ones.

It was Aoki and Ando [34] and Prange [35] who first pointed out that localized levels do not contribute to σ_{xy} and that, moreover, delocalized states, in the presence of localized ones, conduct more efficiently, thus ensuring the pinning of σ_{xy} to a multiple of $\frac{e^2}{h}$. The theoretical breakthrough was made by Laughlin [5] who showed that the gauge invariance of the electromagnetism implies the quantization of σ_{xy} , at $T = 0$ K, provided that the Fermi level sits in the localized levels ($\sigma_{xx} = 0$). For a detailed discussion of the theoretical aspects of the quantized Hall effects, see reference [36].

There is no available theory of the quantized Hall effect at finite temperature. Such a theory is however very desirable in order to ascertain the magnitude and temperature dependence of the corrections to the $T = 0$ K value, and thus the accuracy of new resistance standards based on the quantized Hall effect.

iii) The fractionally quantized Hall effect.

When the Fermi level has passed the 0^- Landau level, nothing is expected to occur. A wealth of extra features was in fact detected for fractional occupancies of the remaining 0^+ sublevel (Fig. 22) [37]. These features are only seen in samples with very high zero-field mobilities. Like the "normal" quantized Hall effect, the fractionally quantized Hall effect does not depend on the band structure of the host materials. It has been observed in both electrons [37, 38] and holes [39] quasi bi-dimensional systems. Of the fraction $\frac{p}{q}$ defined by

$$\rho_{xy} = \frac{q}{p} \frac{h}{e^2} \quad (94)$$

only those with odd denominators have been unambiguously observed. The fractionally quantized Hall effect has been explained by Laughlin [40, 41] in terms of a condensation of the electron (or hole) system into a microscopic collective ground state due to electron-electron (or hole-hole) interaction. This ground state is separated from excited one by $\sim 0.03 e^2/\lambda$, where λ is the magnetic length defined in equation (43). That a repulsive interaction between carriers may give rise to a condensation is intimately linked to the quasi bi-dimensionality of the electron motion. In a bulk material two such electrons can never form a bound state, whereas the two dimensional motion prevents the escape of the two electrons at infinite distance from each other. Laughlin's wavefunction of the condensed ground state has been generalized by Halperin [42]. The exact microscopic nature of the ground state is still disputed [43].

II.4 CYCLOTRON RESONANCE. — As shown in Appendix A the frequency-dependent magnetoconductivity tensor $\sigma_{\alpha\beta}(\omega)$ (which describes the linear response of the electron gas to a periodically varying electric field) exhibits singularities at $\omega = \omega_c$. Let us work in the Drude model (classical description) and assume that

$$\mathbf{F}(\mathbf{r}, t) = F \hat{\mathbf{x}} \cos \omega t \quad (95)$$

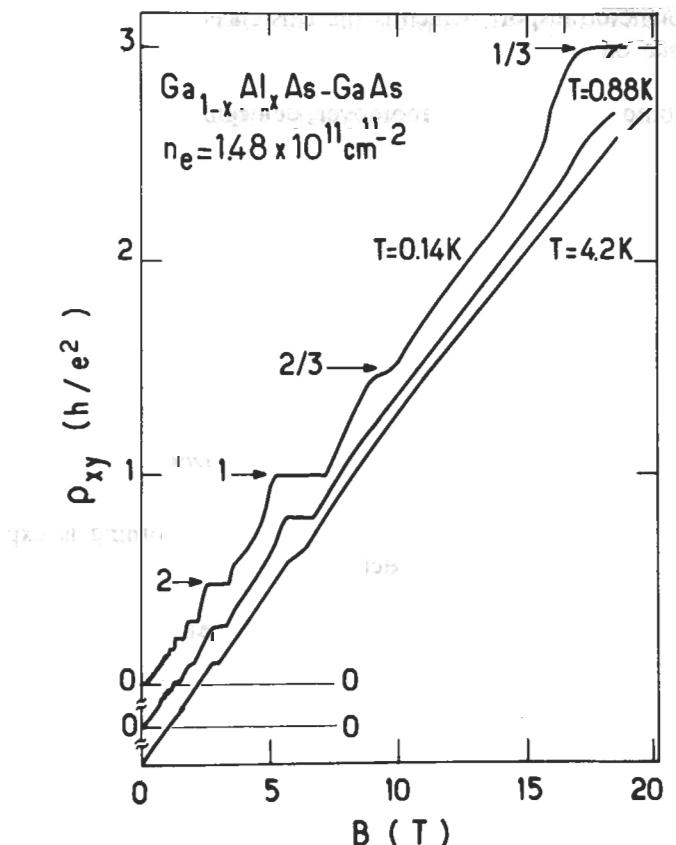


Fig. 22. — Fractionally Quantized Hall effect in a n-Ga_{1-x}Al_xAs-GaAs single heterojunction. The plateaus labelled 2/3 and 1/3 take place when the fractional occupation of the last (0⁺) Landau level is equal to 2/3 and 1/3 respectively. After reference [37].

The equation of motion of the complex velocity

$$\xi = v_x + i v_y \quad (96)$$

is

$$\frac{d\xi}{dt} = -\frac{\xi}{\tau} + i \omega_c \xi - \frac{eF}{m_e} \cos \omega t \quad (97)$$

where a viscous friction force $-m_e v/\tau$ accounts for the scattering mechanisms. In the permanent regime, ξ oscillates at the frequency ω and the \hat{x} component of the areal current density $j = -n_e e v$ is found to behave like

$$j_x = \frac{n_e e^2 \tau}{2 m_e} F \left\{ \frac{\cos \omega t + \tau(\omega - \omega_c) \sin \omega t}{1 + (\omega - \omega_c)^2 \tau^2} + \frac{\cos \omega t + \tau(\omega + \omega_c) \sin \omega t}{1 + (\omega + \omega_c)^2 \tau^2} \right\} \quad (98)$$

Over a period $2\pi/\omega$, the average power taken to the electric field is

$$\langle P \rangle = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} j_x F \cos \omega t \, dt \quad (99)$$

or

$$\langle P \rangle = \frac{n_e e^2 \tau}{4 m_e} F^2 \left\{ \frac{1}{1 + (\omega - \omega_c)^2 \tau^2} + \frac{1}{1 + (\omega + \omega_c)^2 \tau^2} \right\} \quad (100)$$

There are two contributions to $\langle P \rangle$. The first one presents a resonance at $\omega = \omega_c$, while the second one is a decreasing function of ω_c or ω . The resonant term arises from the $\frac{1}{2} e^{+i\omega t}$ part of the electric field. Indeed, the linear vibration (Eq. (95)) can be rewritten as the sum of two clockwise and counter-clockwise circular vibrations :

$$F(t) = \begin{cases} F \cos \omega t & = \begin{cases} \frac{1}{2} F \cos \omega t & + \begin{cases} \frac{1}{2} F \cos(-\omega) t \\ \frac{1}{2} F \sin(-\omega) t \end{cases} \\ \frac{1}{2} F \sin \omega t & + \begin{cases} \frac{1}{2} F \cos(-\omega) t \\ \frac{1}{2} F \sin(-\omega) t \end{cases} \end{cases} \\ 0 & \end{cases} \quad (101)$$

The first kind of circular vibration is, for $\omega = \omega_c$, constantly in phase with the electron motion. As a result, it transfers a finite amount of energy to the electron during a cyclotron period. On the other hand, the second type of circular vibration partly accelerates and partly decelerates the electron and, on average, does not transfer energy to the electron. Note that, ultimately, the carrier does not store the electromagnetic energy. Instead, this energy is given up to the lattice via the viscous friction force. In the limit $\tau \rightarrow \infty$, $\langle P \rangle$ would be zero unless $\omega = \omega_c$. Figure 23 shows a plot of $\langle P \rangle$ versus $\frac{\omega_c}{\omega}$ for three values of $\omega\tau$. Two features are remarkable :

i) $\langle P \rangle$ displays an asymmetric lineshape at low $\omega\tau$ and a rather symmetric one at large $\omega\tau$.

ii) $\langle P \rangle$ gets narrower for larger $\omega\tau$.

Suppose we consider a heterostructure with a d.c. mobility of $10^4 \text{ cm}^2/\text{Vs}$. The condition $\omega_c \tau \geq 1$, which ensures that the electron has completed a revolution around B without being scattered, is fulfilled as soon as $B \geq 1 \text{ Tesla}$. The cyclotron resonance absorption $\omega = \omega_c$ is observed in the far infrared or millimeter parts of the electromagnetic spectrum for the actual masses of III-V heterostructures. Two experimental techniques are used : either to sweep the magnetic field while keeping ω fixed, or to work at fixed B by sweeping the frequency ω . The first technique requires the use of far-infrared lasers or carcinotrons as monochromatic sources and a magnet to produce the field. It is more practical than the second one which requires

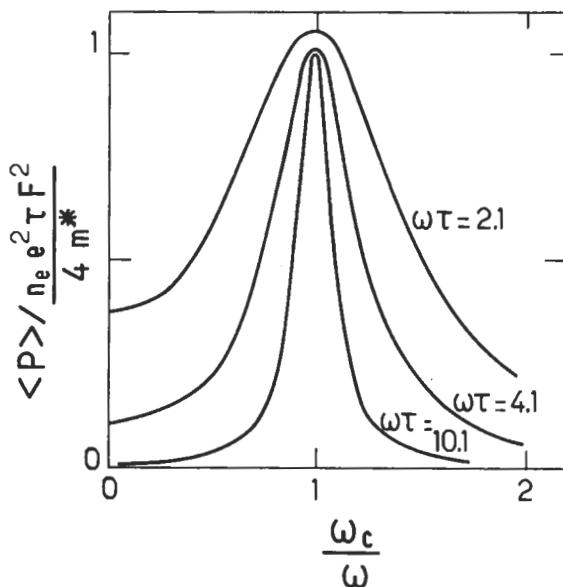


Fig. 23. — The power absorbed by a classical electron gas from an electromagnetic wave (angular frequency ω), averaged over a period of the cyclotron motion, is plotted *versus* the dimensionless ratio $\frac{\omega_c}{\omega}$ for three different values of the product $\omega\tau$: $\omega\tau = 2.1, 4.1, 10.1$ respectively.

the use of Fourier interferometers combined with a magnet. By reporting the photon energies $\hbar\omega_n$ *versus* the observed resonance fields B_n on a $(B, \hbar\omega)$ diagram one gets a straight line (for parabolic band structures) whose slope $\hbar e/m_e$ yields a very precise determination of the carrier effective mass m_e .

It is clear that the Drude model cannot describe very accurately all the experimental observations in quasi bi-dimensional electron gases. In particular, the quantity τ which accounts for the scattering effects has, in practice, no reason to be ω and B -independent as postulated in the Drude model.

A better model of the cyclotron resonance absorption has been derived by Ando and is based on the microscopic calculation of $\sigma_{\alpha\beta}(\omega)$ within the framework of the self-consistent Born approximation [44]. This model predicts resonance absorptions of the electromagnetic wave, not only at $\omega = \omega_c$ but also at the harmonics $\omega = n\omega_c$. In addition, when the width ΔB at half maximum of the main cyclotron resonance line ($\langle P \rangle$ *versus* B at fixed ω) is large enough (i.e. when the d.c. mobility is not too high) to allow the Fermi level to be crossed by several Landau levels (i.e. when n_e is large enough), the cyclotron resonance lineshape is calculated to exhibit extra oscillations which are ω -independent. These Ando oscillations are of the Shubnikov-

de Haas type. They have been observed in Si-MOS structures [45], as well as in modulation-doped $\text{Ga}_{1-x}\text{Al}_x\text{As}$ -GaAs heterostructures [46] (Fig. 24).

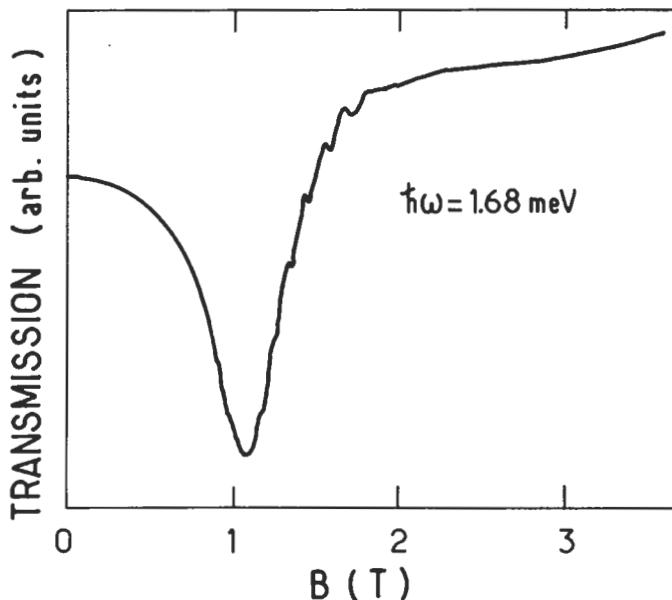


Fig. 24. — Lineshape of a cyclotron resonance absorption observed in $\text{Ga}_{1-x}\text{Al}_x\text{As}$ -GaAs modulation-doped heterojunction with $n_e = 7.5 \times 10^{11} \text{ cm}^{-2}$ and a low temperature mobility of $7.1 \times 10^4 \text{ cm}^2/\text{Vs}$. The wiggles taking place within the main cyclotron resonance curve correspond to Ando oscillations. $T = 2 \text{ K}$. After reference [46].

Apart from the Si-MOSFET's, the most complete study of the cyclotron absorption has been performed in $\text{Ga}_{1-x}\text{Al}_x\text{As}$ -GaAs heterostructures [46-49]. The first output of these studies is the effective mass of the carrier in the GaAs channel. It has been found that this mass is slightly enhanced over the band edge conduction mass of bulk GaAs ($0.067 m_0$). The mass enhancement is however modest (few percents) and is in general attributed to the weak non-parabolicity of the GaAs conduction band. It has not yet been proved that part of the enhancement does not originate from the matching conditions for the wavefunction and its derivative at the interface between GaAs and $\text{Ga}_{1-x}\text{Al}_x\text{As}$.

The most striking results of cyclotron resonance experiments pertain to the cyclotron resonance linewidth and lineshape. If interference effects are carefully excluded from the experiments, the lineshape of the cyclotron absorption is rather symmetric in $\text{Ga}_{1-x}\text{Al}_x\text{As}$ -GaAs heterostructures. For heterostructures with modest mobilities ($\leq 10^5 \text{ cm}^2/\text{Vs}$) the width at half maximum of the absorption curves follows the predictions of Ando's model [44]. In high mobility heterostructures it has been

shown [49] that the cyclotron resonance linewidth ΔB plotted *versus* the resonance field, displays oscillations (see Fig. 25), whereas Ando's model (at least with point-like scatterers) would rather predict a smooth increase of ΔB . Such oscillations have been tentatively correlated with the oscillations of the screening of the scattering potentials by the carriers.

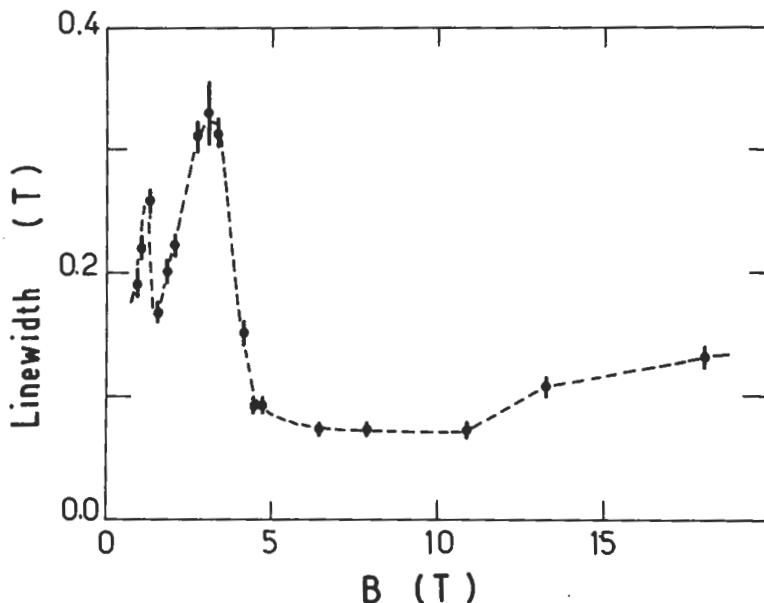


Fig. 25. — Oscillations of the cyclotron resonance linewidth versus the resonance fields in a high mobility $n\text{-Ga}_{1-x}\text{Al}_x\text{As-GaAs}$ heterojunction ($n_e = 1.2 \times 10^{11} \text{ cm}^{-2}$, low temperature mobility $\mu = 1 \times 10^5 \text{ cm}^2/\text{Vs}$). $T = 5 \text{ K}$. After reference [49].

In fact, if the resonance field corresponds to a situation where the Fermi level stays in the localized levels of a broadened Landau level or in the gap separating two such levels, the screening is inefficient, since the heterostructure is, for that particular field, an insulator. If, on the other hand, the resonance field is such that the Fermi level coincides with the extended states of a Landau level, the screening effects are very pronounced due to the large polarizability of the electron gas for that particular configuration. When the screening effects are weak, both the damping effects and the cyclotron resonance linewidth are large and vice versa.

When the host materials of the heterostructure are markedly non parabolic (e.g. GaSb-InAs [50], InSb MOS structures [51]) many subbands are often populated (both effects arise from the smallness of the effective mass). These subbands have different cyclotron resonance effective masses and the ground subband displays the

heavier mass. Thus, the cyclotron resonance absorption shows structures which can be associated with the cyclotron resonance absorption of each individual subband (see Fig. 26).

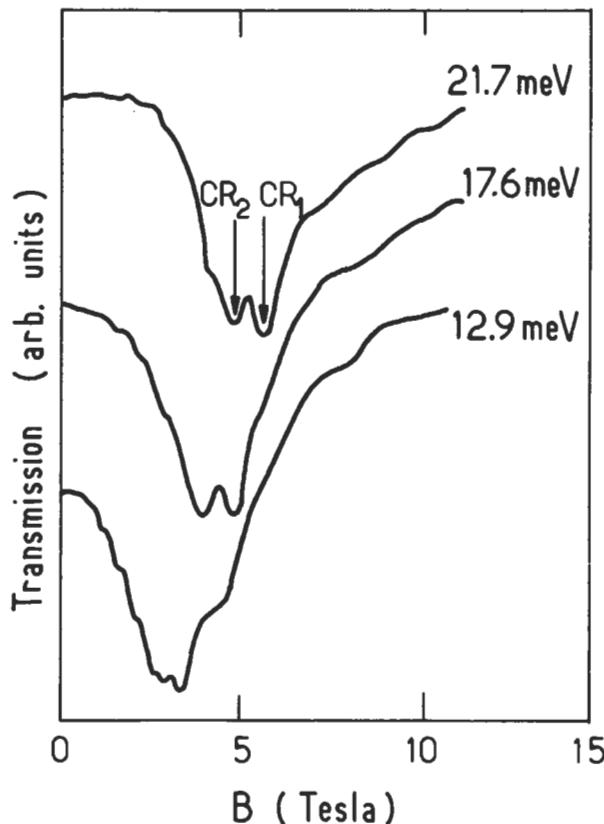


Fig. 26. — Double cyclotron resonance absorption observed in InAs-GaSb multi-heterojunctions. The two minima, labelled CR_1 and CR_2 , correspond to the cyclotron resonances of the E_1 and E_2 subbands respectively. $T = 2$ K. After reference [50].

Appendix A.

Motion of three-dimensional electrons in crossed electric and magnetic fields.

In this Appendix we remind the reader of some results concerning the three-dimensional motion of electrons in crossed electric and magnetic fields. In the first part we describe the Drude model which is essentially a classical calculation of the frequency dependence of the magnetoconductivity tensor. The scattering mechan-

isms are phenomenologically taken into account by means of a viscous friction force. In the second part, we will discuss the Landau level model, which accounts for the quantum motion of independent electrons in a magnetic field. The electric field effects on the Landau levels will be briefly discussed.

A.I THE DRUDE MODEL. — The electron gas is approximated by a collection of independent carriers whose equations of motion are classical. The volume density of electrical current is thus given by

$$\mathbf{J} = -\frac{N_e}{\Omega} e \mathbf{v} \quad (\text{A1})$$

where Ω is the sample volume and \mathbf{v} the velocity of any of the N_e electrons. If m_e denotes the effective mass, the Newton law for a carrier which moves in the presence of static magnetic field \mathbf{B} and oscillating electric field \mathbf{F} is

$$m_e \frac{d\mathbf{v}}{dt} = -\frac{e}{c} \mathbf{v} \times \mathbf{B} - m_e \frac{\mathbf{v}}{\tau} - e\mathbf{F} e^{i\omega t} \quad (\text{A2})$$

In equation (A2) the scattering by defects are embodied into the viscous friction force $-m_e \frac{\mathbf{v}}{\tau}$. In the permanent regime, the time dependence of \mathbf{v} follows that of the electric field. Thus

$$\mathbf{v} = \mathbf{v}_0 e^{i\omega t} \quad (\text{A3})$$

with

$$\left(i\omega + \frac{1}{\tau} \right) \mathbf{v}_0 + \frac{e}{m_e c} \mathbf{v}_0 \times \mathbf{B} = -\frac{e\mathbf{F}}{m_e} \quad (\text{A4})$$

With $\mathbf{F} \perp \mathbf{B}$ the solutions of equation (A4) are of the form

$$\mathbf{v} = a\mathbf{F} + b\mathbf{F} \times \mathbf{B} \quad (\text{A5})$$

Finally, after inserting equation (A5) into equation (A4) and writing

$$\mathbf{J} = \mathbf{J}_0 e^{i\omega t} \quad (\text{A6})$$

we obtain :

$$\mathbf{J}_0 = \frac{\sigma_0 (1 + i\omega\tau) \mathbf{F}}{\omega_c^2 \tau^2 + (1 + i\omega\tau)^2} - \frac{e\tau}{m_e c} \frac{\sigma_0 \mathbf{F} \times \mathbf{B}}{\omega_c^2 \tau^2 + (1 + i\omega\tau)^2} \quad (\text{A7})$$

where σ_0 is the conductivity at zero magnetic field :

$$\sigma_0 = \frac{N_e e^2 \tau}{\Omega m_e} \quad (\text{A8})$$

By definition of the magnetoconductivity tensor $\vec{\sigma}$ we obtain :

$$\sigma_{xx}(\omega) = \frac{\sigma_0 (1 + i \omega \tau)}{\omega_c^2 \tau^2 + (1 + i \omega \tau)^2} \quad (\text{A9})$$

$$\sigma_{yx}(\omega) = \frac{\omega_c \tau \sigma_0}{\omega_c^2 \tau^2 + (1 + i \omega \tau)^2} \quad (\text{A10})$$

with

$$\omega_c = \frac{e B}{m^* c} \quad (\text{A11})$$

At zero frequency $\sigma_{xx}(\omega)$ and $\sigma_{yx}(\omega)$ reduce to :

$$\sigma_{xx}(0) = \frac{\sigma_0}{1 + \omega_c^2 \tau^2} \quad (\text{A12})$$

$$\sigma_{yx}(0) = \frac{N_e e^2}{\Omega m_e \omega_c} \times \frac{\omega_c^2 \tau^2}{1 + \omega_c^2 \tau^2} \quad (\text{A13})$$

By inversion of the conductivity tensor we obtain the resistivity tensor $\vec{\rho}$ which is such that

$$\mathbf{F} = \vec{\rho} \mathbf{J} \quad (\text{A14})$$

$$\vec{\rho} = \begin{bmatrix} \rho_{xx} & \rho_{xy} \\ \rho_{yx} & \rho_{yy} \end{bmatrix} \quad (\text{A15})$$

with

$$\rho_{xx}(0) = \rho_{yy}(0) = \frac{1}{\sigma_0} \quad (\text{A16})$$

$$\rho_{yx}(0) = -\rho_{xy}(0) = \frac{-\omega_c \tau}{\sigma_0} = \frac{-m_e \omega_c \Omega}{N_e e^2} \quad (\text{A17})$$

Thus in the Drude model

- i) the magnetoresistivity $\rho_{xx}(B) - \rho_{xx}(0)$ identically vanishes.
- ii) the Hall resistivity ρ_{yx} increases linearly with the field.

The slope $\frac{d\rho_{yx}}{dB}$ is related to the carrier concentration by :

$$\frac{d\rho_{yx}}{dB} = \frac{-\Omega}{N_e e c} \quad (\text{A18})$$

Thus, a measurement of the Hall voltage V_H in a sample displaying a rectangular geometry yields direct information on the carrier concentration $\frac{N_e}{\Omega} (V_H = \rho_{yx} \frac{I_x}{L_x})$,

where I_x is the current flowing in the x direction and L_x the sample size along the \hat{x} direction).

Although the Drude model fails to reproduce the oscillatory behaviour of the magnetoresistivity ρ_{xx} it is nevertheless helpful in pointing out the important inequality

$$|\sigma_{xy}| \gg \sigma_{xx} \quad (\text{A19})$$

valid when $\omega_c \tau \gg 1$. This inequality, also valid for quasi bi-dimensional electron gases, arises from the fact that the current along the x direction is entirely due to scatterings ($\sigma_{xx}(0) = 0$ if $\tau \rightarrow \infty$) whereas the Hall current (or equivalently the Hall voltage) persists even in the absence of collisions. These two features result from the action of the Lorentz force which bends the carrier trajectories from a direction parallel to $\mathbf{F}(B = 0)$ to a direction perpendicular to $\mathbf{F}(B \neq 0)$.

It should be noted that the leading term in $\sigma_{xx}(0)$ is proportional to $\frac{1}{\tau} (\omega_c \tau \gg 1)$ whereas the deviation of σ_{xy} from the limiting value $-N_e e^2 / \Omega m^* \omega_c$ is proportional to $1/\tau^2$. Thus, thinking in terms of a quantum description of the tensor $\tilde{\sigma}$, we may anticipate that σ_{xx} will have a non-zero contribution to the lowest order of the electron-impurity interaction (Born approximation) whereas at this order of perturbation, σ_{xy} will remain unchanged.

A.II QUANTUM MOTION OF BULK ELECTRONS IN A MAGNETIC FIELD. — The Hamiltonian to be solved is

$$\mathcal{H} = \frac{p_x^2}{2m_e} + \frac{p_z^2}{2m_e} + \frac{1}{2m_e} \left(p_y + \frac{e}{c} Bx \right)^2 + \sigma_z g^* \mu_B B \quad (\text{A20})$$

The eigenfunctions of \mathcal{H} can be factorized into

$$\psi(\mathbf{r}, \sigma_z) = \frac{1}{\sqrt{L_y L_z}} \exp[i(k_y y + k_z z)] \varphi(x) \alpha(\sigma_z) \quad (\text{A21})$$

where $\alpha(\sigma_z)$ is the spin eigenstate ($|\uparrow\rangle$ or $|\downarrow\rangle$) and $\varphi(x)$ is the solution of

$$\left(\frac{p_x^2}{2m_e} + \frac{1}{2} m_e \omega_c^2 (x + x_0)^2 + \frac{\hbar^2 k_z^2}{2m_e} \right) \varphi(x) = \varepsilon \varphi(x) \quad (\text{A22})$$

with

$$x_0 = \lambda^2 k_y; \quad \lambda^2 = \frac{\hbar c}{eB} \quad (\text{A23})$$

As shown in equation (46) the function $\varphi(x)$ can be expressed in terms of the Hermite polynomials centred at $-x_0$. The eigenenergies are k_y -independent and equal to

$$\varepsilon(n, k_y, k_z, \sigma_z) = \left(n + \frac{1}{2} \right) \hbar \omega_c + \frac{\hbar^2 k_z^2}{2m_e} + \sigma_z g^* \mu_B B \quad (\text{A24})$$

i.e. consist of a set of one-dimensional parabolas evenly spaced by $\hbar\omega_c$ (see Fig. 27). The density of states

$$\rho(\varepsilon) = \sum_{n k_y k_z \sigma_z} \delta[\varepsilon - \varepsilon(n, k_y, k_z, \sigma_z)] \quad (\text{A25})$$

can be decomposed into the sum of the contributions which arise from each subband (n, σ_z) :

$$\rho(\varepsilon) = \sum_{n \sigma_z} \rho_{n\sigma_z}(\varepsilon) \quad (\text{A26})$$

The degeneracy, with respect to k_y , contributes to $\rho_{n\sigma_z}$ by a multiplicative factor $\frac{L_x L_y}{2\pi\lambda^2}$, whereas the summation over k_z finally leads to :

$$\rho_{n\sigma_z}(\varepsilon) = \frac{\Omega}{4\pi^2\lambda^2} \sqrt{\frac{2m_e}{\hbar^2}} \left[\varepsilon - \left(n + \frac{1}{2} \right) \hbar\omega_c - \sigma_z g^* \mu_B B \right]^{-\frac{1}{2}} \quad (\text{A27})$$

The singularity at the onset of the n^{th} subband is characteristic of the one-dimensional free motion of the carrier along the field. Notice, that $\rho_{n\sigma_z}(\varepsilon)$ never vanishes for $\varepsilon > \left(n + \frac{1}{2} \right) \hbar\omega_c + \sigma_z g^* \mu_B B$ since it is always possible to find an allowed electron state with the energy ε , the kinetic energy associated with the z motion matching the energy difference between ε and $\left(n + \frac{1}{2} \right) \hbar\omega_c + \sigma_z g^* \mu_B B$.

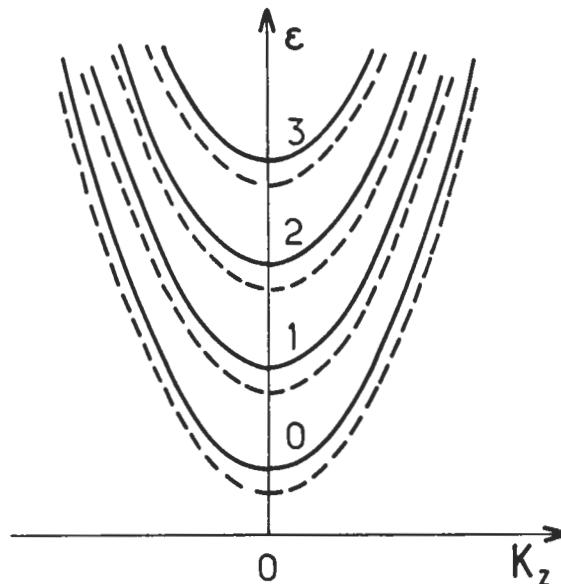


Fig. 27. — Dispersions of the Landau subbands $n = 0, 1, \dots$ in a bulk material. k_z denotes the carrier wavevector along the magnetic field. The solid and dashed lines correspond to the spin sub levels.

This feature is in marked contrast to the quasi bi-dimensional situation where we saw that the size-quantization in the confining potential $V_{\text{disorder}}(z)$ leads to finite bandgaps between the Landau levels.

As the density of states (Eq. (A27)) is so singular near the Landau subband extrema (see Fig. 28), it leads to important modifications of e.g., the optical and transport properties. For example, instead of displaying a smooth $\sqrt{\hbar\omega - \epsilon_g}$ behaviour near the bandgap of a semiconductor ($B = 0$ situation), the absorption coefficient becomes infinite (in practice very large) at certain fields B_m . This happens whenever the photon energy $\hbar\omega$ coincides with the energy difference between the onset of two conduction and valence Landau subbands. The study of interband magneto-optics has proved to be one the most powerful experimental techniques for accurately measuring the band parameters of bulk semiconductors.

The static conductivity also becomes radically altered by the application of a strong magnetic field. In a degenerate material, σ_{xx} oscillates with B (Shubnikov-de Haas oscillations) and becomes small whenever the Fermi energy ϵ_F coincides with the onset of a Landau level subband. For the fields B_n such that

$$\epsilon_F = \left(n + \frac{1}{2} \right) \frac{\hbar e B_n}{m^*} + g^* \mu_B B_n \sigma_z \quad (\text{A28})$$

the elastic scatterings at the Fermi energy become very efficient since both the initial and final states are numerous. This leads to pronounced σ_{xx} minima in B^{-1} and the slope of the curve B^{-1} versus the integers directly yields the Fermi energy.

The Schrödinger equation for an electron which moves in crossed electric and magnetic fields can also be solved exactly. The potential energy eFx associated with

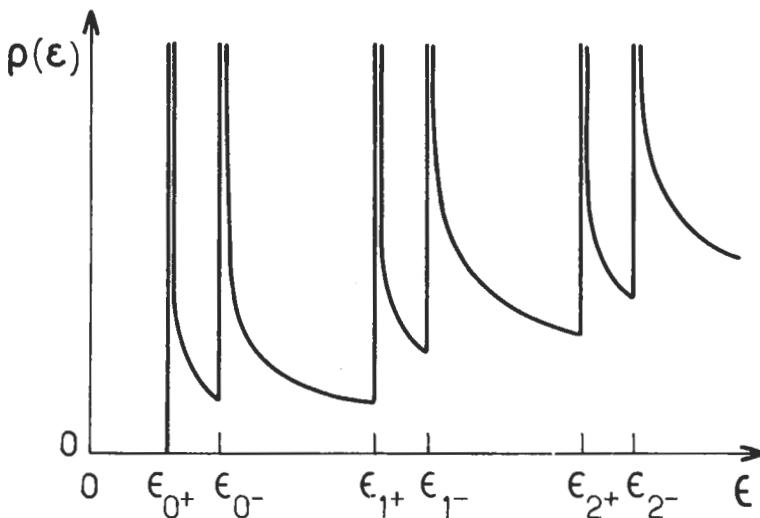


Fig. 28. — Density of states (arbitrary units) associated with the Landau subbands in a bulk material. $\rho(\epsilon)$ vanishes only if $\epsilon < \epsilon_0^+$.

the electric field merely displaces the centre of the wavefunctions $\varphi(x)$. In fact $\varphi(x)$ is the solution of

$$\left[\frac{p_x^2}{2m_e} + \frac{1}{2m_e} \left(\hbar k_y + \frac{e}{c} Bx \right)^2 + eFx + \frac{\hbar^2 k_z^2}{2m_e} + g^* \mu_B B \sigma_z \right] \varphi(x) = \varepsilon \varphi(x) \quad (\text{A29})$$

where we have made use of the factorization of the wavefunction (see equation (A21)). Equation (A29) can be rewritten

$$\left[\frac{p_x^2}{2m_e} + \frac{1}{2} m_e \omega_c^2 \left(x + \lambda^2 k_y + \frac{eF}{m_e \omega_c^2} \right)^2 - \lambda^2 k_y eF - \frac{1}{2} \frac{e^2 F^2}{m_e \omega_c^2} + \frac{\hbar^2 k_z^2}{2m_e} + g^* \mu_B B \sigma_z \right] \varphi(x) = \varepsilon \varphi(x) \quad (\text{A30})$$

Since the eigenenergies of a harmonic oscillator do not depend on the center of the harmonic potential the eigenvalue of (A30) are

$$\begin{aligned} \varepsilon(n, k_y, k_z, \sigma_z) = & \left(n + \frac{1}{2} \right) \hbar \omega_c + \frac{\hbar^2 k_z^2}{2m_e} + \\ & \sigma_z g^* \mu_B B - \lambda^2 k_y eF - \frac{1}{2} \frac{e^2 F^2}{m_e \omega_c^2} \end{aligned} \quad (\text{A31})$$

and the eigenfunctions $\varphi(x)$ can be expressed in terms of Hermite polynomials centred at $-x_F$ where

$$x_F = \lambda^2 k_y + \frac{eF}{m_e \omega_c^2} = x_0 + \frac{eF}{m_e \omega_c^2} \quad (\text{A32})$$

instead of $-x_0$ ($F = 0$)

Thus, the electric field has two effects. Firstly, it displaces the centre of the cyclotron orbit by $\frac{eF}{m_e \omega_c^2}$ and secondly it lifts the degeneracy with respect to k_y .

Appendix B.

Microscopic evaluation of the magnetoconductivity tensor.

Let us consider a heterostructure whose Hamiltonian is given by :

$$\begin{aligned} \mathcal{H} = & \frac{p_x^2}{2m_e} + \frac{1}{2m_e} \left(p_y + \frac{e}{c} Bx \right)^2 + \frac{p_z^2}{2m_e} + \\ & + V_{\text{conf}}(z) + V_{\text{disorder}}(\mathbf{r}) + g^* \mu_B B \sigma_z \end{aligned} \quad (\text{B1})$$

The disorder potential is such that its spatial average over a volume, which is large compared to the characteristic distances of the electron wavefunction, vanishes. In equation (B1) \mathbf{B} is assumed to be parallel to the z axis. $V_{\text{conf}}(z)$ the Hartree contribution of the electron-electron interaction. We assume that the electron gas is markedly bi-dimensional, i.e. the ground level E_1 of (B1) for the z motion is well energy-separated from the excited levels $E_2\dots$. In this way the three-dimensional Hamiltonian can be reduced to a quasi-bidimensional one by choosing

$$\psi(\mathbf{r}) = \chi_1(z) \nu(\boldsymbol{\rho}) \quad (\text{B2})$$

where

$$\left[\frac{p_z^2}{2m_e} + V_{\text{conf}}(z) \right] \chi_1 = E_1 \chi_1 \quad (\text{B3})$$

and the $|\nu\rangle$ are the eigenstates of

$$\mathcal{H} = \frac{p_x^2}{2m_e} + \frac{1}{2m_e} \left(p_y + \frac{e}{c} Bx \right)^2 + V_{\text{disorder}}(\boldsymbol{\rho}) + g^* \mu_B B \sigma_z \quad (\text{B3}')$$

with $\boldsymbol{\rho} = (x, y)$ and where

$$V_{\text{disorder}}(\boldsymbol{\rho}) = \int dz \chi_1^2(z) V_{\text{disorder}}(z) \quad (\text{B4})$$

It may be parenthetically remarked that, owing to the weighting factor $\chi_1^2(z)$, the dominant contributions to $V_{\text{disorder}}(\boldsymbol{\rho})$ are those which correspond to defects sitting close to the quasi bi-dimensional electron gas even if these defects are much less numerous than the remote ones.

Actually the electron gas is not insulated from the surroundings but permanently interacts with the lattice (*via* the electron-phonon interaction). There is a continuous exchange of energy between these two systems. As a result of these energy exchanges, and due to the much bigger size of the lattice system, the occupancy of an electron state $|\nu\rangle$ at thermal equilibrium is given by the Fermi Dirac distribution function $f(\varepsilon_\nu)$. The latter is the diagonal element of the density matrix operator ρ_0 :

$$f(\varepsilon_\nu) = \langle \nu | \rho_0 | \nu \rangle \quad (\text{B5})$$

Suppose now that an electric field oscillating at frequency ω is adiabatically switched on. This adds to \mathcal{H}_0 the time-dependent operator

$$h(t) = eFx e^{\eta t} \cos \omega t = \frac{1}{2} eFx [e^{i(\eta + i\omega)} + e^{i(\eta - i\omega)}] \quad (\text{B6})$$

The small term η ensures that the field has been adiabatically turned on and that consequently the electron gas has not been heated. The perturbing term $h(t)$ drives the electron gas out of thermal equilibrium. As a result, an electrical current will occur. The statistical distribution of the energy levels $|\nu\rangle$ is no longer described by

ρ_0 but by a time dependent density matrix $\rho(t)$ which evolves with time according to the equation

$$i\hbar \frac{\partial \rho}{\partial t} = [\mathcal{H} + h(t), \rho] \quad (B7)$$

At any time t the statistical average of any operator A is the mean value of $A\rho(t)$ or

$$\langle A \rangle = \sum_{\nu} \langle \nu | A \rho(t) | \nu \rangle = \text{Trace} [A \rho(t)] \quad (B8)$$

Notice that the Trace can be evaluated over any complete basis $|\nu\rangle$. We are interested in the linear response to the electron gas to the perturbation $h(t)$. Thus, we may linearize equation (B7) :

$$\rho(t) = \rho_0 + \rho_1(t) \quad (B9)$$

where

$$i\hbar \frac{\partial \rho_1}{\partial t} = [\mathcal{H}, \rho_1(t)] + [h(t), \rho_0] \quad (B10)$$

The areal density of the electrical current is given by the operator

$$\mathbf{J} = \frac{-e}{m^* S} \mathbf{I} \quad (B11)$$

where

$$\mathbf{I} = \mathbf{p} + \frac{e}{c} \mathbf{A}_0 \quad (B12)$$

\mathbf{A}_0 being the vector potential of the static magnetic field. Since there is no current on average at thermal equilibrium, we obtain :

$$\langle \mathbf{J}(t) \rangle = \frac{-e}{m^* S} \text{Trace} [\mathbf{I} \rho_1(t)] \quad (B13)$$

At time $t = 0$ we easily find from equation (B6, B10) that

$$\langle \mu | \rho_1(t=0) | \nu \rangle = \alpha_{\mu\nu} \left[\frac{1}{\varepsilon_{\nu} - \varepsilon_{\mu} - \hbar\omega + i\hbar\eta} + \frac{1}{\varepsilon_{\nu} - \varepsilon_{\mu} + \hbar\omega + i\hbar\eta} \right] \quad (B14)$$

with

$$\alpha_{\mu\nu} = \frac{1}{2} eF \langle \mu | x | \nu \rangle [f(\varepsilon_{\nu}) - f(\varepsilon_{\mu})] \quad (B15)$$

The matrix element $\langle \mu | x | \nu \rangle$ can be conveniently reexpressed in terms of $\langle \mu | \Pi_x | \nu \rangle$ by means of the identity

$$[x, \mathcal{H} + h(t)] = \frac{i\hbar \Pi_x}{m_e} \quad (B16)$$

After some manipulations $\langle J_x(t = 0) \rangle$ can be rewritten

$$\begin{aligned} \langle J_x(t = 0) \rangle = & \frac{\pi e^2 F}{(m_e)^2 S \omega} \sum_{\mu, \nu} |\langle \mu | \Pi_x | \nu \rangle|^2 f(\varepsilon_\nu) \times \\ & [\delta(\varepsilon_\nu - \varepsilon_\mu + \hbar\omega) - \delta(\varepsilon_\nu - \varepsilon_\mu - \hbar\omega)] \end{aligned} \quad (\text{B17})$$

which at vanishing frequency ω leads to

$$\langle J_x(t = 0) \rangle = \frac{2\pi \hbar e^2 F}{(m_e)^2 S} \sum_{\mu, \nu} |\langle \mu | \Pi_x | \nu \rangle|^2 f(\varepsilon_\nu) \delta'(\varepsilon_\nu - \varepsilon_\mu) \quad (\text{B18})$$

By definition σ_{xx} is the proportionality coefficient between the x components of the average electrical current density to that of the electric field. Thus

$$\sigma_{xx}(\omega = 0) = \frac{2\pi e^2 \hbar}{(m_e)^2 S} \sum_{\mu, \nu} |\langle \nu | \Pi_x | \mu \rangle|^2 f(\varepsilon_\nu) \delta'(\varepsilon_\nu - \varepsilon_\mu) \quad (\text{B19})$$

Equation (B19) is not convenient since it involves the exact eigenstates of \mathcal{H}_0 which, due to the presence of $V_{\text{disorder}}(\rho)$, are not known exactly. However, if we write

$$f(\varepsilon_\nu) = \int_{-\infty}^{+\infty} \delta(\varepsilon - \varepsilon_\nu) f(\varepsilon) d\varepsilon \quad (\text{B20})$$

we easily obtain :

$$\sigma_{xx}(\omega = 0) = \frac{2\pi e^2 \hbar}{(m_e)^2 S} \int_{-\infty}^{+\infty} f(\varepsilon) d\varepsilon \text{Trace} \{ \delta(\varepsilon - \mathcal{H}) \Pi_x \delta'(\varepsilon - \mathcal{H}) \Pi_x \} \quad (\text{B21})$$

Equation (B20) has the advantage of being independent of the basis on which the Trace is evaluated. However, the disorder potential, included in \mathcal{H} , precludes the possibility of evaluating σ_{xx} in closed forms. Equation (B21) can be further transformed by integrating by part. At $T = 0$ K, $\frac{\partial f}{\partial \varepsilon}$ is minus $\delta(\varepsilon - \varepsilon_F)$ where ε_F is the Fermi level. Thus

$$\sigma_{xx}^{T=0}(\omega = 0) = \frac{\pi e^2 \hbar}{(m_e)^2 S} \text{Trace} \{ \delta(\varepsilon_F - \mathcal{H}) \Pi_x \delta(\varepsilon_F - \mathcal{H}) \Pi_x \} \quad (\text{B22})$$

Proceeding along the same line as before we can easily calculate J_y and thus σ_{yx} . We obtain

$$\sigma_{yx}(\omega = 0) = \frac{-ie^2 \hbar}{2(m_e)^2 S} \sum_{\mu, \nu} \frac{f(\varepsilon_\nu) - f(\varepsilon_\mu)}{(\varepsilon_\nu - \varepsilon_\mu)^2} [\Pi_y^{\nu \mu} \Pi_x^{\mu \nu} - \Pi_y^{\mu \nu} \Pi_x^{\nu \mu}] \quad (\text{B23})$$

After some manipulations equation (B23) can be rewritten

$$\sigma_{yx}(\omega = 0) = \frac{-ie^2\hbar}{(m_e)^2 S} \int_{-\infty}^{+\infty} f(\varepsilon) d\varepsilon \text{ Trace} \left\{ \delta(\varepsilon - \mathcal{H}) \left[\Pi_y \frac{1}{(\varepsilon - \mathcal{H})^2} \Pi_x - \Pi_x \frac{1}{(\varepsilon - \mathcal{H})^2} \Pi_y \right] \right\} \quad (\text{B24})$$

Equations (B22, B24) together with

$$\sigma_{xx}(\omega = 0) = \sigma_{yy}(\omega = 0) \quad (\text{B25})$$

and the Onsager relationships

$$\sigma_{xy}(\omega = 0) = -\sigma_{yx}(\omega = 0) \quad (\text{B26})$$

complete our task of evaluating on a microscopic basis the magnetoconductivity tensor $\vec{\sigma}$ of a quasi bi-dimensional electron gas.

References

- [1] E.E. MENDEZ, G. BASTARD, L.L. CHANG, L. ESAKI, H. MORKOC, and R. FISHER, *Physica* **B 117** and **B 118** (1983) 711.
- [2] D.S. CHEMLA and D.A.B. MILLER, in *Heterojunction B and Discontinuities : Physics and Device Applications*, Eds. F. Capasso and G. Margaritondo (North Holland, Amsterdam) 1987, see also D.S. Chemla, *J. Lumin.* **30** (1985) 502.
- [3] L.D. LANDAU *Zeit. Phys.* **64** (1930) 629.
- [4] P.N. ARGYRES and L.M. ROTH, in "Physics of III-V compounds" edited by R.K. Willardson and A.C. Beer Volume 1 (Academic Press, New York).
- [5] R. LAUGHLIN, *Phys. Rev.* **B 23** (1981) 5632.
- [6] K. VON KLITZING, G. DORDA and M. PEPPER, *Phys Rev. Lett.* **45** (1980) 494.
- [7] C. ALIBERT, S. GAILLARD, J.A. BRUM, G. BASTARD, P. FRIJLINK and R. ERMAN, *Solid State Commun.* **53** (1985) 457.
- [8] T.H. WOOD, C.A. BURRUS, D.A.B. MILLER, D.S. CHEMLA, T.C. DAMEN, A.C. GOSSARD and W. WIEGMANN, *Appl. Phys. Lett.* **44** (1984) 16.
- [9] See e.g. M. ABRAMOWITZ and I.A. STEGUN "Handbook of Mathematical Functions" (National Bureau of Standards Applied Mathematics Series N° 55) (U.S. Government Printing Office, Washington) 1964.
- [10] L.V. KELDYSH, *Soviet Phys. J.E.T.P.* **6** (1958) 763 ;
W. FRANZ, *Z. NATURFORSCH* **13a** (1958) 484.
- [11] G. BASTARD, E.E. MENDEZ, L.L. CHANG and L. ESAKI, *Phys. Rev. B* **28** (1983) 3241.
- [12] J.A. BRUM and G. BASTARD, *Phys. Rev. B* **31** (1985) 3893.
- [13] D.A.B. MILLER, D.S. CHEMLA, T.G. DAMEN, A.C. GOSSARD, W. WIEGMANN, T.H. WOOD and C.A. BURRUS, *Appl. Phys. Lett.* **45** (1984) 13.
- [14] J.D. JACKSON, *Classical Electrodynamics* (Wiley, New York) 1962.

- [15] J.C. MAAN, in *Two-Dimensional Systems, Heterostructures and Superlattices*, edited by G. Bauer, F. Kuchar and H. Heinrich, Springer Series in Solid States Sciences 53 (1984) 183.
- [16] T. ANDO, *J. Phys. Soc. Japan* **50** (1981) 2978.
- [17] S. DAS SARMA, *Surf. Sci.* **142** (1984) 341.
- [18] T. ANDO and Y. UEMURA, *J. Phys. Soc. Japan* **36** (1974) 959.
- [19] T. ANDO, *J. Phys. Soc. Japan* **52** (1983) 1740 ; **53** (1984) 3101 ; **53** (1984) 3126.
- [20] T. HAAVASOJA, H.L. STORMER, D.J. BISHOP, V. NARYANAMURTI, A.C. GOSSARD and W. WIEGMANN, *Surf. Sci.* **142** (1984) 294.
- [21] L. ONSAGER, *Phys. Rev.* **37** (1931) 405 and *Phys. Rev.* **38** (1931) 2265.
- [22] K. VON KLITZING, in "Festkörperprobleme/Advances in Solid State Physics", edited by J. Treusch (Vieweg, Braunschweig) **XXI** (1981) 1.
- [23] R. KUBO, S.B. MIYAKE and N. HASHITSUME, *Solid State Phys.* **17** (1965) 269.
- [24] See e.g. H. AOKI and T. ANDO, *Phys. Rev. Lett.* **54** (1985) 831.
- [25] D.C. TSUI and A.C. GOSSARD, *Appl. Phys. Lett.* **38** (1981) 550.
- [26] Y. GULDNER, J.P. HIRTZ, J.P. VIEREN, P. VOISIN, M. VOOS and M. RAZEGHI, *J. Phys. Lett. France* **43** (1982) L-613 ;
See also Y. GULDNER, J.P. VIEREN, M. VOOS, F. DELAHAYE, D. DOMINGUEZ, J.P. HIRTZ and M. RAZEGHI, *Phys. Rev. B* **33** (1985) 3990.
- [27] C.A. CHANG, E.E. MENDEZ, L.L. CHANG and L. ESAKI, *Surf. Sci.* **142** (1984) 598.
- [28] E.E. MENDEZ, L.L. CHANG, C.A. CHANG, L.F. ALEXANDER and L. ESAKI, *Surf. Sci.* **142** (1984) 215.
- [29] H.L. STORMER, A.M. CHANG, Z. SCHLESINGER, D.C. TSUI, A.C. GOSSARD and W. WIEGMANN, *Phys. Rev. Lett.* **51** (1983) 126.
- [30] M.A. PAALANEN, D.C. TSUI, A.C. GOSSARD and J.C.M. HWANG, *Solid State Commun.* **50** (1984) 841.
- [31] G. EBERT, K. VON KLITZING, C. PROBST, E. SCHUBERTH, K. PLOOG and G. WEIMANN, *Solid State Commun.* **45** (1983) 625.
- [32] A. BRIGGS, Y. GULDNER, J.P. VIEREN, M. VOOS, J.P. HIRTZ and M. RAZEGHI, *Phys. Rev. B* **27** (1983) 6549.
- [33] M.A. PAALANEN, D.C. TSUI and A.C. GOSSARD, *Phys. Rev. B* **25** (1982) 5566.
- [34] H. AOKI and T. ANDO, *Solid State Commun.* **38** (1981) 1079.
- [35] R.E. PRANGE, *Phys. Rev. B* **23** (1981) 4802.
- [36] R.B. LAUGHLIN, in *Two-Dimensional Systems, Heterostructures and Superlattices*, edited by G. Bauer, F. Kuchar and H. Heinrich. Springer Series in Solid State Sciences 53 (1984) 272.
- [37] D.C. TSUI, H.L. STORMER and A.C. GOSSARD, *Phys. Rev. Lett.* **48** (1982) 1559 ;
D.C. TSUI, H.L. STORMER, J.C.M. HWANG, J.S. BROOKS and M.J. MAUGHTON, *Phys. Rev. B* **28** (1983) 2274.
- [38] E.E. MENDEZ, M. HEIBLUM, L.L. CHANG and L. ESAKI, *Phys. Rev. B* **28** (1983) 4886.
- [39] H.L. STORMER, A.M. CHANG, D.C. TSUI, J.C.M. HWANG, A.C. GOSSARD and W. WIEGMANN, *Phys. Rev. Lett.* **50** (1983) 1953.
- [40] R.B. LAUGHLIN, *Phys. Rev. B* **27** (1983) 3383.
- [41] R.B. LAUGHLIN, *Phys. Rev. Lett.* **50** (1983) 1395.
- [42] B.I. HALPERIN, *Helv. Phys. Acta* **56** (1983) 75.
- [43] R. TAO and D.J. THOULESS, *Phys. Rev. B* **28** (1983) 1142.
- [44] T. ANDO, *J. Phys. Soc. Japan* **38** (1975) 989.
- [45] G. ABSTREITER, J.P. KOTHAUS, J.F. KOCH and G. DORDA, *Phys. Rev. B* **14** (1976) 2480.

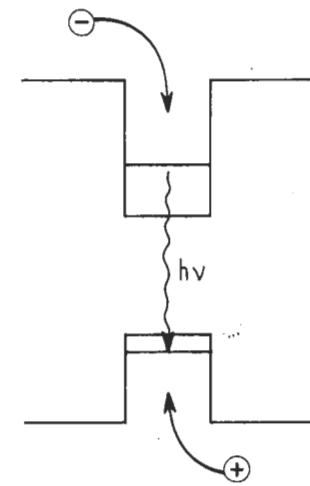
- [46] P. VOISIN, Y. GULDNER, J.P. VIEREN, M. VOOS, J.C. MAAN, P. DELESCLOUSE and NUYEN T. LINH, *Physica* **B 117** and **B 118** (1983) 634.
 - [47] H.L. STORMER, R. DINGLE, A.C. GOSSARD, W. WIEGMANN and M.D. STURGE, *Solid State Commun.* **29** (1979) 705.
 - [48] E. GORNICK, W. SEIDENBUSH, R. LASSNIG, H.L. STORMER, A.C. GOSSARD and W. WIEGMANN, same as reference [15] p. 60.
 - [49] Th. ENGLERT, J.C. MAAN, Ch. UIHLEIN, D.C. TSUI and A.C. GOSSARD, *Physica* **B 117** and **B 118** (1983) 631.
 - [50] Y. GULDNER, J.P. VIEREN, P. VOISIN, M. VOOS, J.C. MAAN, L.L. CHANG and L. ESAKI, *Solid State Commun.* **41** (1982) 755.
 - [51] A. DARR, A. HUBER and J.P. KOTTHAUS, *Physics of Narrow Gap Semiconductors*, edited by J. Rauluszkiewicz, M. Gorska and E. Kaczmarek (P.W.N. - Polish Scientific, Warsaw) 1978, p. 417.
-

Imprimé en France. — JOUVE, 18, rue Saint-Denis, 75001 PARIS
N° 13430. Dépôt légal : Juillet 1990

MONOGRAPHIES
DE PHYSIQUE

**wave mechanics
applied to
semiconductor
heterostructures**

GERALD BASTARD



ISBN 2-86883-092-7
Prix : **400,00 F.F.**