1 Spintronics: Transport and Materials

Phivos Mavropoulos, Kazunori Sato, and Stefan Blügel
Institut für Festkörperforschung, Forschungszentrum Jülich,
D-52425 Jülich, Germany

Contents

1.1 Introduction

The origin of *magnetoelectronics* or *spintronics* can be traced back to two independent experiments carried out in 1988 by Albert Fert [1] in Paris and Peter Grünberg in Jülich [2]. In these experiments it was found that an electrical current passing through ferromagnetic films separated by non-magnetic metallic spacer layers is subject to a resistivity which changes unexpectedly largely (gigantically at those days) with the change of the relative alignment of the magnetization in these films from ferromagnetic to antiferromagnetic alignment. This observation was coined the giant-magneto-resistance effect or GMR. Obviously GMR made it possible to turn the information of a two-state magnetic configuration (parallel or anti-parallel associated with bit 0 or 1) into an electrical one, or in a more abstract sense, turn spin information into charge current information. Already 8 years after the discovery, this effect is used for example in sensors embodied in read heads in hard disks of common PCs.

Soon after this discovery experiments have been carried out in which the non-magnetic metallic spacer layer was replaced by a non-magnetic insulator. In this set-up spin-polarized electrons tunnel from one ferromagnetic layer through an insulating barrier film into the second ferromagnetic layer, and again a strong dependence of the resistance upon the relative orientation of the magnetization is found. The effect is called the tunnel magneto-resistance (TMR) and the set-up is called a magnetic tunnel junction (MTJ). As opposed to the GMR systems, TMR systems exhibit a large voltage drop across the MTJ and operate with small electrical currents. Recently nonvolatile magnetic random access memories (MRAM) are made from an array of MTJs currently for special applications.

Inspired by the many new possibilities of spin-dependent transport properties, a concept of spin-transistors – semiconductor based transistors in which spin-polarized electrons are injected from a magnetic source, manipulated and controlled before they are collected at the magnetic drain – has been put forward by Datta and Das [3]. In these types of devices not only the charge property of the electron is used but also the fact that the electron has a spin degree of freedom. Connecting spin-dependent transport with the semiconductor world opened a completely new vista which is summarized under the term *spintronics* (often spintronics is also used in a more restrictive sense of transport, manipulation and control of only coherent spin transport). This new vista is based on the great functionality, engineerability of the semiconductor material and the scalability of the devices. Particular to semiconductors is the variability of the electron density by the dopant concentration, the manipulation and control of the electron density by external gate control and easy access to light. Therefore, spintronics offers the possibility of transferring information stored in the electron spin into the charge or light state. Programmable logics or application, e.g. qubit manipulation or read-out of the qubit state embodied in the electron spin maybe envisaged in the future.

In order to make this happen we deal with a variety of problems. One of these is the problem of *spin-injection*: getting the spin-polarized electrons into the semiconductor valence or conduction band. Schmidt *et al.* [4], Rashba, and Fert and Jaffres [5] noticed that the conductivity mismatch and the mismatch of the mean free spin-flip length between a ferromagnetic metal and a semiconductor reduces the spin-injection of the spin-polarized electrons from a ferromagnetic metal into the semiconductor to a nearly undetectable level. There are several ways out of this dilemma: (1) including a tunneling barrier between the semiconductor and the metal, (2) replacing the ferromagnetic metal by a ferromagnetic half-metallic system (majority states are occupied at the Fermi energy, E_F, minority states show a gap around E_F, which adds up to a spin-injected current with 100% spin polarization), or (3) replace the ferromagnetic metal by a ferromagnetic semiconductor (if possible with a Curie temperature above the room temperature). Finding appropriate ferromagnetic half-metallic systems and ferromagnetic semiconductors with a Curie temperature above the room temperature are currently key issues from the material side. In particular magnetic semiconductors offer completely new functionalities, as the collective magnetic state and Curie temperature depend on carrier concentration which is altered easily by the gate voltage or light. Without a detailed basic physical understanding of the magnetic interactions in these semiconductors a further development of this field is inconceivable.

For completeness we would like to add that for the manipulation of the injected spin in the semiconductor the use of the Rashba-effect with an external gate voltage was suggested. This idea, also conceptionally very nice, may be difficult to realize or may not be applicable, as in real semiconductors the potential landscape is unknown due to the unknown dopant distribution. The scientific community welcomes new suggestions to manipulate the spin direction. The second point we would like to mention is that under certain conditions along with the injection and accumulation of electron spins at the interface, goes a switching of the collective spin state [6]. This may in the future lead to a down-scalable fast switch of the magnetization bypassing the law of induction which requires large volumes difficult to down-scale.

A further interesting aspect of using the spin-degree of freedom for transport (instead of just the charge degree) is that an electron has spin 1/2 and obeys the symmetry group $SU(2)$. This

bears interesting consequences for the spin-transport of coherent electrons as it can be realized in nano-scale devices at low temperatures. The key to the interesting transport properties is the non-commutativity of the $SU(2)$ spin algebra, which breaks the time inversion symmetry. When a conducting electron in a conductor is scattered by some magnetic object, the electron wavefunction is multplied by a $U(1)$ phase factor $A(\mathbf{n}) = \alpha \exp(\beta\, \mathbf{n} \cdot \boldsymbol{\sigma})$, which is generally spin-dependent and is represented by a 2×2 matrix in spin space. Here α and β are complex numbers, $\mathbf{n}$ is a three-component unit vector characterizing the magnetization direction of the scattering objects, and $\boldsymbol{\sigma}$ are the Pauli matrices. Due to the non-commutativity of σ_i, after two scattering events $A(\mathbf{n}_1)$ and $A(\mathbf{n}_2)$, the amplitude depends in general on the sequence of the scattering events: $A(\mathbf{n}_1)A(\mathbf{n}_2) \neq A(\mathbf{n}_2)A(\mathbf{n}_1)$. Various features of coherent spin transport arize from this non-commutativity. Gen Tatara [7] has recently shown that an anomaly in charge transport arizes after three coherent scattering events. Such events may then open a perspective for new logical gates. We will not continue further on this point as it will be beyond the subject of this lecture.

1.2 Basic Considerations

In a perfect, infinite periodic crystal, electrons can travel forever. This is because the electronic eigenfunctions are Bloch states, $\Psi_{\mathbf{k}}(\mathbf{r})$, with a definite crystal momentum $\mathbf{k}$ and a definite group velocity

$$\mathbf{v}_{\mathbf{k}} = \frac{1}{\hbar} \nabla_{\mathbf{k}} E_{\mathbf{k}}. \tag{1.1}$$

In equilibrium, of course, there is no net current, since the electrons traveling in one arbitrary direction are as many as the ones traveling in the opposite one. But if one would create a current by placing an electron in a previously unoccupied state, this current would keep on forever.

However, in reality there are always deviations from the perfect periodicity. On the one hand, the crystal ends at some surface or interface; on the other hand, defects or even the thermal motion of the nuclei result in non-periodicity. Then the Bloch wavefunctions are no longer the eigenfunctions of the system, but rather evolve into each other as time passes. Under these circumstances any current cannot be kept constant without a voltage to sustain it, and electrical resistance appears.

The quantum mechanical interpretation of the resistance (and generally of the electronic transport) is a quite complicated task. It is easier when small currents and voltages are involved, in the so-called linear response regime, when the current I is proportional to the voltage ΔV via the conductance g:

$$I = g \cdot \Delta V \tag{1.2}$$

This is the case we will be concerned with here. Basically, one can distinguish two regimes of electronic transport: the ballistic and the diffusive. The former is relevant when there is no dissipation of energy within the sample of our study, *i.e.*, when the electron scattering is elastic and the electron energy is conserved. Then, as the electrons flow through our system, we can follow the Schrödinger wavefunctions through time with 100% determinism, and calculate

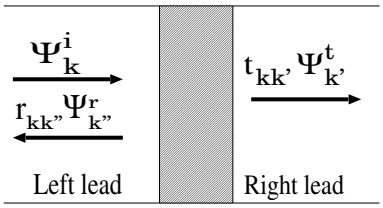

Fig. 1.1: Setup for the ballistic transport.

the transport coefficients. Low temperatures and well-ordered samples bring us close to this approach. The latter regime, the one of diffusive transport, is relevant when there is inelastic scattering (due to thermal events usually), or when our sample is randomly constructed, as in the case of random alloys where we know the structure only on the average but not in detail, atom for atom. Then we cannot follow the wavepackets with determinism, but must account for the uncertainty of the scattering processes by some kind of averaging.

It can be understood that the distinction between the two approaches is also one of length scale. Thus, even in a fairly well ordered sample at low temperatures, if an electron would travel long enough it would be scattered in a thermal or other random way. This notion of "long enough" is quantified by the *mean free path* and the *relaxation time*, which is understood as the average length (or average time) that the electron has to travel before it suffers a random collision. So, we can work in the ballistic approach when our sample dimension is less or comparable to the mean free path, and in the diffusive approach in the opposite case. The scale of modern nanoelectronics devices is so small, that in many cases their basic transport properties can be understood in the ballistic regime.

The ballistic regime is basically approached in the following way. We assume that our sample is sandwiched between two perfect infinite leads, as shown in Fig. 1.1. Within the leads, the eigenfunctions have the form of Bloch waves, but not at the interfaces and in the sample. Assume that a Bloch state $\Psi^i_{\mathbf{k}}$ is incident from the left. In general, this will finally evolve into a set of transmitted Bloch functions in the right lead, $\Psi^t_{\mathbf{k}'}$, plus a set of reflected Bloch functions travelling back into the left lead, $\Psi^r_{\mathbf{k}''}$. Important are the transmission and reflection amplitudes, $t_{\mathbf{k}\mathbf{k}'}$ and $r_{\mathbf{k}\mathbf{k}''}$ characterising this event. The *total* wavefunction within the leads is then written as

$$\Psi^{\mathrm{tot}}_{\mathbf{k}} = \begin{cases} \Psi^i_{\mathbf{k}} + \sum_{\mathbf{k}''} r_{\mathbf{k}\mathbf{k}''}\Psi^r_{\mathbf{k}''} & \text{in the left lead} \\ \\ \sum_{\mathbf{k}'} t_{\mathbf{k}\mathbf{k}'}\Psi^t_{\mathbf{k}'} & \text{in the right lead.} \end{cases} \tag{1.3}$$

In this equation we said nothing about the wavefunction within the sample. But of course, in order to find the transmission and reflection amplitudes, one has the difficult task of solving the Schrödinger equation within the sample and matching the wavefunctions and their derivatives with the Bloch states left and right. The techniques used for this purpose are beyond the scope of this text. Now once we have the amplitudes $t_{\mathbf{k}\mathbf{k}'}$, it is straightforward to calculate the transmission probability through the sample, and use it for the calculation of the conductance. The transmission probability from the incoming state $\Psi^i_{\mathbf{k}}$ to the transmitted $\Psi^t_{\mathbf{k}}$ is just

$$T_{\mathbf{k}\mathbf{k}'} = |t_{\mathbf{k}\mathbf{k}'}|^2 \frac{v_{\mathbf{k}'z}}{v_{\mathbf{k}z}}. \tag{1.4}$$

We see that a normalisation term, $v_{\mathbf{k}'z}/v_{\mathbf{k}z}$, is present, involving the z components of the group velocities for both wavefunctions (z is the direction in which the junction has been grown and in which we want to measure the current). To understand this necessary term, we must think that the transmission probability must depend not only on the amplitude of the transmitted wave in comparison to the incident, but also on *how fast* this wave is traveling. Thus, this term is a normalisation of the outgoing flux to the incoming flux. A rigorous proof of this formula involves the consideration of wavepackets which are taken to the limit of itinerant Bloch states at the end. We note here that some authors prefer to normalise the Bloch functions not to unit probability in space, but to unit flux, by the substitution $\Psi_{\mathbf{k}} \to \Psi_{\mathbf{k}}/\sqrt{v_{\mathbf{k}z}}$; then the normalisation term is contained in the Bloch functions and does not appear expicitly in the transmission probability.

As we know, for low temperatures and low voltage only the electrons at the Fermi level can contribute to transport, since they are the only ones that can be excited from occupied states (just below E_F) to unoccupied ones (just above E_F) by the weak perturbing external field. Then the conductance g can be directly related to the transmission probability $T_{\mathbf{k}\mathbf{k}'}$ for states at E_F via the Landauer formula:

$$g = \frac{e^2}{2\pi\hbar} \sum_{\mathbf{k}\mathbf{k}'} T_{\mathbf{k}\mathbf{k}'}, \text{ with } v_{\mathbf{k}z} > 0,\ v_{\mathbf{k}'z} > 0,\ E(\mathbf{k}) = E(\mathbf{k}') = E_F. \tag{1.5}$$

We take the states for which $v_{\mathbf{k}z} > 0$ and the same for $v_{\mathbf{k}'z}$ to distinguish the relevant incoming and outgoing states.

To understand the Landauer formula, think that the application of a very low voltage ΔV raises the Fermi level of the left lead by $e\Delta V$. At an energy E within this range, the current flowing via a particular outgoing state $\mathbf{k}'$ in the right lead is proportional to the group velocity of this state times the transmission probability of all incoming states into this state: $I_{\mathbf{k}'}(E) = e v_{\mathbf{k}'z} \sum_{\mathbf{k}} T_{\mathbf{k}\mathbf{k}'}$. This must be summed up for all outgoing states $\mathbf{k}'$ and integrated in the energy range from E_F to $E_F + \Delta V$, also taking into account the energetic density of outgoing states, $n_{\mathbf{k}'_\parallel}(E)$:[1]

$$I = e \int_{E_F}^{E_F + e\Delta V} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} T_{\mathbf{k}\mathbf{k}'}(E) v_{\mathbf{k}'z}(E)\, n_{\mathbf{k}'_\parallel}(E)\, dE \tag{1.6}$$

Remembering that the density of states is just $n_{\mathbf{k}'_\parallel}(E) = 1/(2\pi\hbar v_{\mathbf{k}'z}(E))$, we see that the group velocity cancels out. If we take then the limit of small ΔV, we arrive at the current-voltage relation 1.2 with the conductance g given by eq. 1.5.

After the considerations above we can make the link to spin-dependent transport. In many materials where spin magnetism is present, one can examine the two spin directions seperately, having spin-up and spin-down Bloch functions as eigenfunctions of the system. Then one has two different conductance coefficients g, one for each spin direction, and the current becomes spin dependent. We can imagine the situation as if we had parallel resistances. The aim of spin electronics is to exploit such materials in order to manipulate the current (or the resistance) by switching the direction of the magnetic moment in parts of a junction. In this way one can, for example, create a magnetic switch, using a contact between two materials which has

[1] Here we have decomposed $\mathbf{k}$ into the components $(\mathbf{k}_\parallel, k_z)$, and chosen k_z as the variable to change continuously with E during the integration, while $\mathbf{k}_\parallel := (k_x, k_y)$ is kept constant. We can imagine that our sample is large but finite in the (x, y) directions, so that the $\mathbf{k}_\parallel$'s form a very dense but discrete set, while k_z is continuous.

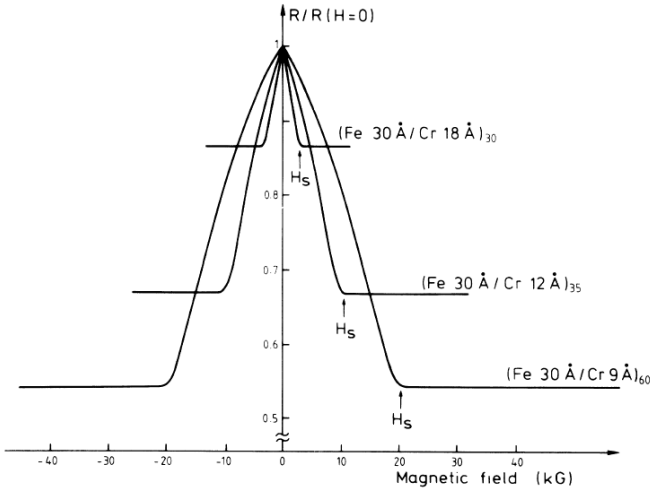

Fig. 1.2: Magnetoresistance as a function of the applied field in Fe/Cr multilayers, after ref. [1].

low resistance when their magnetic moments are aligned in parallel, but high resistance when they are alligned antiparallel. This is a form of magnetoresistance, though very different in nature from the one that is connected to the usual Hall effect. We shall examine such systems in the next sections; in particular, we shall focus on the Giant Magnetoresistance (GMR), the Tunneling Magnetoresistance (TMR), and also the spin injection.

1.3 Giant Magnetoresistance

As said in the introduction, in 1988, Fert and co-workers [1] and independently Grünberg and co-workers [2] announced the discovery of the Giant Magnetoresistance (GMR) effect in magnetic multilayers. This boosted the technology of magnetic storage media and brought new products in the market, within less than a decade. In this section we shall briefly describe the GMR effect and its interpretation through quantum-mechanical transport theory.

A *magnetic multilayer* is a structure of alternating ferromagnetic and nonmagnetic layers, for example Co/Cu/Co/Cu... Although each ferromagnetic layer has a single magnetisation direction, alternating magnetic layers can be aligned with their moments parallel (P) or antiparallel (AP). Suppose now that electrical current passes through the multilayer. It is natural to assume that the transmission probability of electrons and the resistance will be different in the two cases (P and AP), since the potential landscape encountered in each case is different. This assumption was verified by the experiments, which showed a very strong decrease of the resistance when one swiched from the AP to the P configuration by applying an external magnetic field. This is the GMR effect. The characteristic quantity is the so-called *GMR ratio*, meaning the relative change in conductance (sometimes also defined as its inverse, the relative change in resistance). For example, Fig. 1.2 shows that very high GMR ratio is obtained in Fe/Cr multilayers.[2]

Two types of geometry are used in GMR experiments: the Current In Plane (CIP) and the Current Perpendicular to the Plane (CPP) geometry. In the former case, the current flows parallel

[2]In reality, Cr is not nonmagnetic but antiferromagnetic; nevertheless the explanation for the GMR effect is similar.

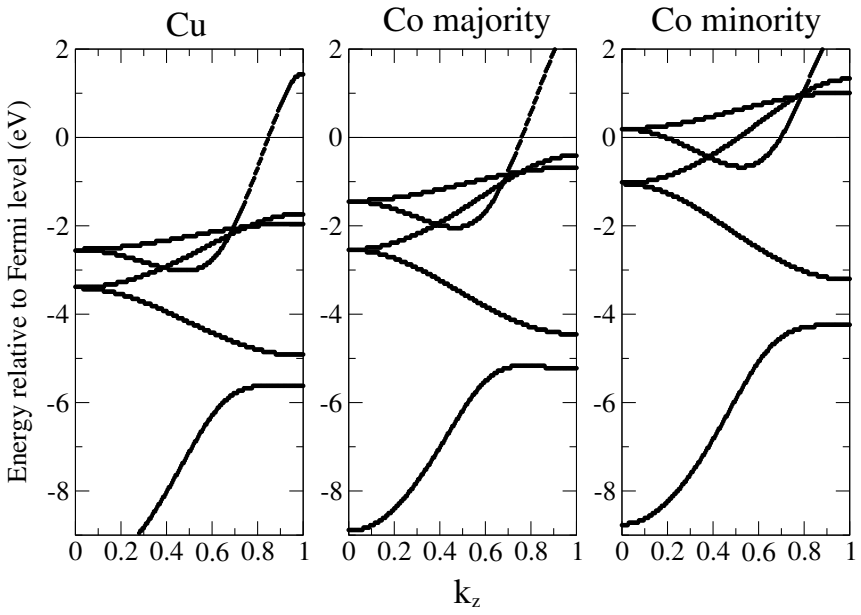

Fig. 1.3: Energy bands of Cu and Co in the $\Gamma - X$ direction.

to the layers, while in the latter it crosses the layers flowing in the perpendicular direction. This case (CPP) is easier to understand, and in what follows we shall concentrate on it.

Consider an electron traveling in a Co/Cu multilayer, in the direction perpendicular to the planes. As it encounters the Co/Cu interface, it will have a certain transmission amplitude t and reflection amplitude r. The values of t and r will depend on the electronic structure on the two sides of the interface. Since the system is spin-polarised, the electronic structure and the potential of the two spin directions is different and the spin-up electrons will have different t and r from the spin-down electrons.

We focus now to the electronic states at E_F, which actually carry the current. For the majority spin, these are very similar for Co and Cu; this is because the magnetic exchange splitting in Co shifts the majority-spin bands to lower energies, so that finally the majority d band becomes occupied and the s-band is left alone at E_F, just as it is in Cu. On the other hand, the minority-spin states at E_F are very different for Co and Cu, because of the presence of the minority d states of Co in that region. This can be seen by inspection of the band structure of Co and Cu, presented in Fig. 1.3. As a result of this, one finds that the majority-spin electons at E_F are transmitted easily through the interface, while the minority-spin ones are rather reflected.

Consider now the Co/Cu multilayer, first in the P configuration. The spin-up electrons have easy transmission through each subsequent interface, since they are majority electrons in all Co layers; on the other hand the spin-down electrons are always strongly reflected. But the net result is a pretty high transmission probability and strong current in total, in the same way as two parallel resistances allow for a strong current if one of them is low. Now consider the AP configuration. While spin-up electrons have high transmission through one Co layer, they suffer strong reflection at the next one, since its moment is reversed and they belong to the minority-spin there. The analogous happens to spin-down electrons: they can transmit through

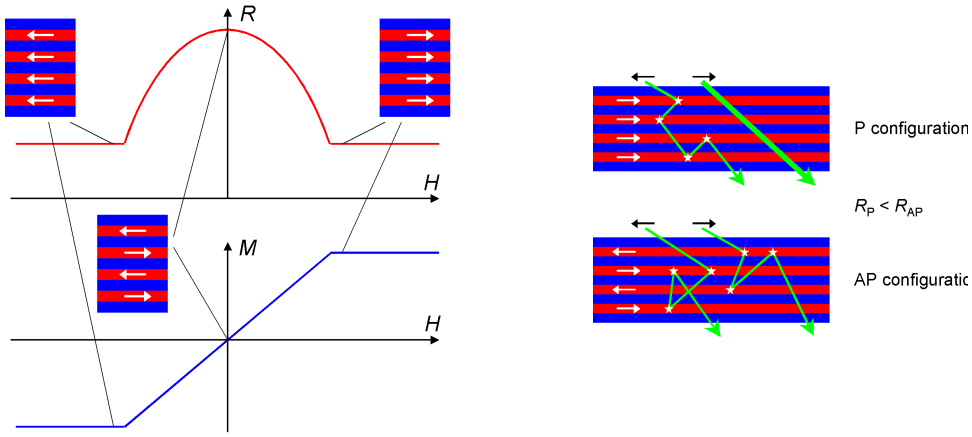

Fig. 1.4: Left: The application of an external magnetic field H aligns the moments of the magnetic layers. In this way the resistance decreases (upper plot), and the total magnetic moment increases (lower plot). Right: Electron transmission in the Parallel and Antiparallel configurations.

the Co layers where the spin-up electrons are reflected, but suffer reflection at the rest of the layers. As a result, the resistance for both spin-up and spin-down electrons is high, so that the combined resistance is also high. The situation is presented schematically in Fig. 1.4 (right).

As a conclusion, if in the absence of an external magnetic field the multilayer is initially in the AP configuration, with a high resistance, we can lower the resistance by applying a field and aligning the moment of all magnetic layers in the same direction, as shown in Fig. 1.4 (left). This is the essence of the GMR effect.

A prerequisite for all this to come true is that the ground state of the multilayer corresponds to the AP configuration, which then can be brought to the P configuration with an external field. On the contrary, if the ground state is the P one, it is quite impossible to swich the moment only in every other magnetic layer. It has been shown that one can construct systems that have an AP ground state by manipulating the thickness of the nonmagnetic layers. The effect is called *interlayer exchange coupling*, and the relevant theory was first given by Bruno [8].

The basics of GMR presented here help in the interpretation of the effect, but a deep understanding includes many other aspects. For one, the role of the defects and chemical interdiffusion at the layer interfaces has been studied extensively [9] by the use of the Boltzmann, rather than the Landauer, formula. Also, for the case that the mean free path is smaller than the sample size, resistor models have been developed [10]. In these models the multilayer is viewed as a series of resistors, seperately for spin up and spin down. Each resistor represents the reflection at an interface or the resistance coming from the diffuse scattering in the (imperfect) layer. Finally, there are cases where the spin-orbit scattering can be strong, which means that the two spin channels cannot be viewed as separate resistances but rather communicate.

1.4 Tunneling Magnetoresistance

The tunnel effect is a well-known example where the quantum-mechanical nature of electrons (or other particles) is demonstrated. In short, the wavefunction can penetrate regions where

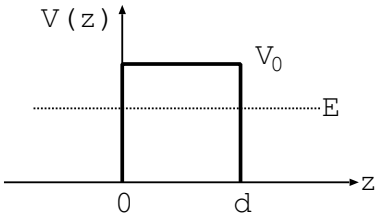

Fig. 1.5: The tunnel effect: electrons of energy E can transmit through a potential barrier of height $V_0 > E$.

the potential is higher than the particle total energy. As a result, a high potential barrier can be overcome by an incident particle of low energy. A typical example is that of a square one-dimensional potential step of height V_0 and of width d, as shown schematically in Fig. 1.5. The wavefunction can be written as

$$\Psi(z) = \begin{cases} e^{ikz} + re^{-ikz} & z < 0 \\ ae^{-\kappa z} + be^{\kappa z} & 0 < z < d \\ te^{ikz} & d < z \end{cases} \tag{1.7}$$

where $k = \sqrt{2mE/\hbar^2}$ and $\kappa = \sqrt{2m(V_0 - E)/\hbar^2}$. The coefficients are easily calculated by matching the wavefunction and its derivative at $z = 0, d$, and finally the transmission probability as a function of the energy E is found to be [3]

$$T = |t|^2 = \left(1 + \frac{1}{4}\left(\frac{\kappa}{k} + \frac{k}{\kappa}\right)^2 \sinh^2(\kappa d)\right)^{-1}. \tag{1.8}$$

In the limit of an extremely thick or extremely high barrier, i.e., when $\kappa d >> 1$, the above expression becomes

$$T = \left(\frac{4k\kappa}{k^2 + \kappa^2}\right) e^{-2\kappa d} + O(e^{-4\kappa d}). \tag{1.9}$$

We see that the transmission probability decays exponentially with increasing thickness. The parameter κ characterising this decay is known as *decay parameter* (its inverse $1/\kappa$ is the *decay length*).

The exponential decay of the transmission probability is typical for tunneling. In three dimensions the result is similar. Thus, if we have such a barrier in the z direction, while the motion is free in x and y directions, we can decouple the motion in z from the motion in x and y. The wavefunction has then the form

$$\Phi(x, y, z) = e^{i(k_x x + k_y y)} \Psi(z) \tag{1.10}$$

and we make the abbreviation $\mathbf{k}_\parallel = (k_x, k_y)$. Here, $\Psi(z)$ has again the form of eq. (1.7) but with $k = k_z$ and κ and k_z dependent on $k_\parallel$:

$$\kappa(\mathbf{k}_\parallel) = \kappa = \sqrt{\frac{2m}{\hbar^2}(V_0 - E) + k_\parallel^2}\ ,\ k_z = \sqrt{\frac{2m}{\hbar^2}E - k_\parallel^2}. \tag{1.11}$$

[3] The left and right group velocities are equal to $\hbar k/m$

Then we get a $\mathbf{k}_\parallel$-dependent transmission probability as

$$T(\mathbf{k}_\parallel) = \left(1 + \frac{1}{4}\left(\frac{\kappa}{k_z} + \frac{k_z}{\kappa}\right)^2 \sinh^2(\kappa d)\right)^{-1} \tag{1.12}$$

Since the system is translationally invariant in the x and y directions, $\mathbf{k}_\parallel$ is a constant of motion. The total transmission probability per unit-cell area is given by an integral over all $\mathbf{k}_\parallel$ vectors that correspond to a given E_F:

$$T_{\text{tot}} = \frac{\Omega}{(2\pi)^2} \int d\mathbf{k}_\parallel \, T(\mathbf{k}_\parallel) \tag{1.13}$$

with Ω the unit cell area in the x-y directions.

For large thicknesses or high barriers we should look for the minimum in $\kappa(\mathbf{k}_\parallel)$, since the exponential decay will cause the other states to vanish much faster compared to this one. We can see easily that the minimum is at $\mathbf{k}_\parallel = 0$.

The notion of tunneling just presented has to do with electrons passing through a region where their energy would be classically insufficient to take them. But this notion is generalised to cases when the electrons pass through regions where quantum-mechanical electronic states should normally not exist. For example, a semiconductor or an insulator possess no electronic states at E_F. But when such materials are brought in contact with a metal, the metallic states at E_F can penetrate into the insulator gap for a short distance, decaying exponentially and vanishing after a few monolayers. These are the so-called *metal-induced gap states* (MIGS). If one adds a second metal at the other side of the insulator, one can even have a low transmission, with a probability which decays exponentially with the insulator thickness. This effect is also called tunneling, because of the passing of electrons through the "forbidden" insulator region. As in the free-electron case, the MIGS and the tunneling are characterised by a decay parameter κ. Mathematically the function $\kappa(\mathbf{k}_\parallel)$ is derived by the analytical continuation of the band structure for complex $\mathbf{k}$ vectors, therefore it is named *complex band structure*. As we shall see, in many cases the tunneling properties can be understood by using these ideas.

In 1975 Julliere reported the first results on tunneling magnetoresistance (TMR) [11]. The experiment was done on a junction made of a Ge semiconducting slab, sandwiched between two Co ferromagnetic leads. The experiment showed that the resistance depended on whether the two Co leads had their magnetic moments alligned in a parallel or antiparallel fashion. Although in both cases the electrons had to tunnel through the Ge slab, giving a high resistance, in the case of antiparallel alignment the resistance was higher. This was the TMR effect.

In order to interpret these results, Julliere proposed a simple model: he suggested that, for each spin direction, the tunneling probability and current is proportional to the density of states (DOS) at E_F in the region of the interfaces. For Co the spin-down DOS at E_F is higher than its spin-up counterpart: $n_\downarrow > n_\uparrow$ (by convension, spin-up means majority spin and spin-down minority spin). If we denote with a prime the DOS at the second interface, then the current in the parallelly-alligned case is, according to Julliere:

$$I_P \sim n_\uparrow n'_\uparrow + n_\downarrow n'_\downarrow. \tag{1.14}$$

But if we reverse the moment of the second electrode, the spin-up is interchanged there with spin-down, and the current becomes

$$I_{AP} \sim n_\uparrow n'_\downarrow + n_\downarrow n'_\uparrow. \tag{1.15}$$

Since $n_\downarrow > n_\uparrow$, the two currents are unequal ($I_P > I_{AP}$) and magnetoresistance occurs. We see that the basic argument of Julliere is that, in the parallel case, more electrons can tunnel through the spin-down channel where there is a high DOS in both leads; while in the antiparallel case, each channel has low DOS in one of the two interfaces, so the current is reduced. In terms of the spin polarisation P at E_F the magnetoresistance ratio is expressed as

$$\frac{\Delta g}{g_P} = \frac{\Delta I}{I_P} = \frac{2PP'}{1+PP'}, \text{ with } P = \frac{n_\uparrow - n_\downarrow}{n_\uparrow + n_\downarrow} \tag{1.16}$$

(and similarly for P').

The model of Julliere is attractive due to its simlicity, and in many cases can explain the experimental trends. In the last decade, since the re-discovery of TMR with much higher ratio by Miyazaki and Tekuza, and Moodera and co-workers[12], the prospect of applications, particularly in non-volatile Magnetic Random Access Memosies (MRAMs), fueled the research in this field. In most experiments, the model of Julliere, or somewhat improved models such as the one of Sloncewski [14], were employed for the interpretation. However, there have been first-principles calculations [13] showing that in some cases the model of Julliere must be inapplicable. We turn our attention to these cases now, offering a simple way to understand the physics involved [15].

If a junction consists of perfectly or almost perfectly ordered materials and interfaces, then $\mathbf{k}_\parallel$ is conserved during the scattering at the interface, and one must examine the transmission probability for each $\mathbf{k}_\parallel$ separately. The situation is pretty much analogous to what we saw earlier for free electrons incident on a square barrier, only that now the band structure must be taken into account, and $\mathbf{k}$ and $\mathbf{k}_\parallel$ refer to Bloch crystal momentum. Since we have a tunneling current, we should examine the complex band structure of the insulator at E_F (which is in the gap region), and find for which $\mathbf{k}_\parallel$ the decay parameter $\kappa(\mathbf{k}_\parallel)$ takes its minimum value, $\kappa_{\min}$. When our junction becomes a little thicker, all other states will decay much faster than this. For this reason it is not relevant to examine the whole DOS, $n_\uparrow(E_F)$ and $n_\downarrow(E_F)$, as in Julliere's model, but we must concentrate our study to the $\mathbf{k}_\parallel$ for which $\kappa(\mathbf{k}_\parallel) = \kappa_{\min}$.

If we fix some $\mathbf{k}_\parallel$, then in the case of free electrons the complex band $E(\kappa)$ in the barrier has the form of an inverse parabola, given if we solve eq. (1.11) for E. Furthermore, if we insist in giving a periodic lattice structure to free space and confine our study in the first Brillouine zone, then for a certain $\mathbf{k}_\parallel$ we get more solutions, because we also have the inverse parabolas that correspond to $(\mathbf{k}+\mathbf{G})_\parallel$, where $\mathbf{G}$ is a vector of the inverse lattice. But in the case of a real material, the structure is more complicated. For example, in a semiconductor we have a band gap at the center of the Brilloune zone Γ that opens due to the periodic potential. If we imagine the strength of this periodic potential varying continuously from zero to its normal value, then the band gap opens up gradually, by lifting the degeneracy at Γ. In this case, the two states that were previously degenerate are connected via a complex band, which thus forms a loop rather than an inverse parabola. The complex band structure consists mainly of such loops, plus inverse parabolas starting from the positions where they would be in a free electron

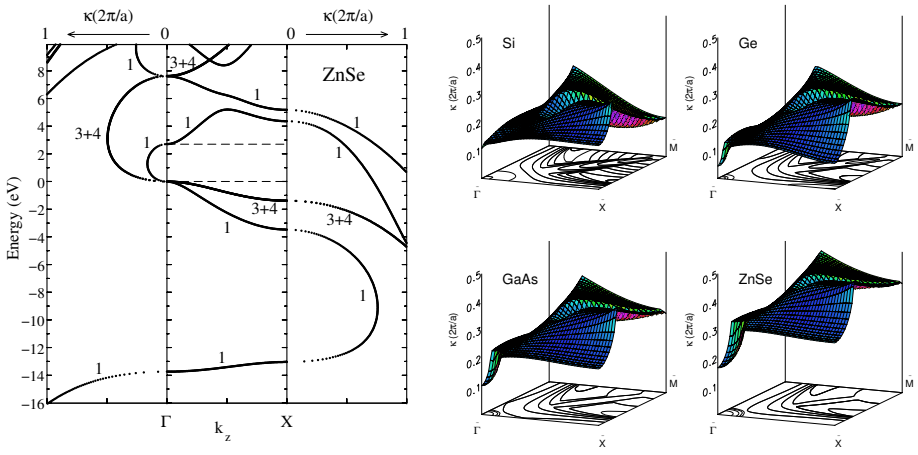

Fig. 1.6: Left: Real and complex band structure of ZnSe along k_z for $\mathbf{k}_\parallel = 0$. In the middle pannel the real bands are shown, while left and right the complex bands $E(\kappa)$ appear for the real part $k_z = 0$ and $k_z = 2\pi/\mathrm{a}$, respectively. Complex bands with real part k_z other than 0 and $2\pi/\mathrm{a}$ are not included. Right: Constant energy surface in $(\mathbf{k}_\parallel, \kappa)$ space for $E = E_F$, showing that $\kappa_{\min}$ indeed appears for $\mathbf{k}_\parallel = 0$.

system. In order to calculate it, one can use standard band-structure methods, but with the possibility that the Bloch k_z-vector takes also complex values (the energy must always be real). Then the imaginary part of the k_z vector is just the decay parameter κ.

As an example, we present in fig. 1.6 the real and complex band structure of ZnSe along the k_z direction, for $\mathbf{k}_\parallel = 0$ (the $\bar{\Gamma}$-point). At E_F, in the middle of the gap, we see more than one complex bands: one loop connecting the top of the valence band with the bottom of the conduction band with a real part of $k_z = 0$ in the left pannel, another similar loop (doubly degenerate) connecting again the valence band with higher bands, plus free-electron-like inverse parabolas with a real part of $k_z = 2\pi/\mathrm{a}$ in the right pannel (a is the lattice constant). Evidently, the smallest decay parameter is that of the small loop. But we must also scan at other $\mathbf{k}_\parallel$ vectors to see if this is indeed the absolute minimum. This scanning gives us a constant-energy surface in $(\mathbf{k}_\parallel, \kappa)$ space, the analogoue of the Fermi surface. It is shown in fig. 1.6, right. As we can see, $\kappa_{\min}$ appears indeed at the $\bar{\Gamma}$-point. Although this property is very common among semiconductors, the case could perhaps be different, especially in an indirect-gap semiconductor such as Si, if the bottom of the valence band (which is off $\bar{\Gamma}$) is close to E_F.

Now that we have located $\kappa_{\min}$ at $\bar{\Gamma}$, we can proceed to the case study of spin-dependent tunneling in the model system Fe/ZnSe/Fe (001). We must know what states are incident from Fe at $\mathbf{k}_\parallel = 0$, because these will couple to the ZnSe state with $\kappa_{\min}$. This can be done most easily by examining the symmetry character of the states. If we examine the band structure of Fe at $\bar{\Gamma}$ along k_z (see fig. 1.7) around E_F, we see that, for spin-up we have one band of symmetry Δ_1, one of symmetry $\Delta_{2'}$, and one of symmetry Δ_5. The symmetry characterises the behaviour of the corresponding wavefunctions under rotation around the z-axis (we mean the fourfold rotation group C_{4v} that preserves the Fe crystal stucture). In particular, Δ_1 is the rotationally invariant state, while the others change sign under some of the rotations. On the other hand, for spin-down we find no Δ_1 band at E_F, but only $\Delta_{2'}$ and Δ_5. The following table shows the irreducible representations of the symmetry groups C_{4v} and C_{2v} needed here, and the compatibility with local angular-momentum orbitals.

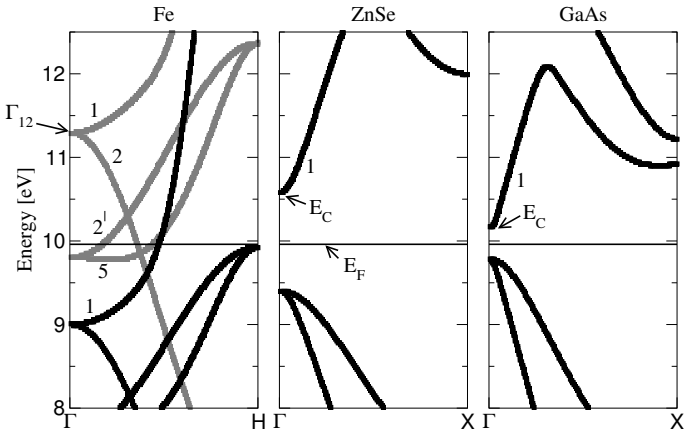

Fig. 1.7: Energy bands of Fe (left), ZnSe (middle) and GaAs (right), along k_z (the Δ-direction) near E_F. The majority-spin bands of Fe are represented by black lines and the minority ones by gray lines. The numbers denote the corresponding irreducible representation of the group of rotations around the Δ-direction, which is C_{4v} for Fe and C_{2v} for the semiconductor.

C_{4v} (Fe (001))		C_{2v} (ZnSe (001))	
Δ_1	s, p_z, d_{z^2}	Δ_1	s, p_z, d_{xy}, d_{z^2}
$\Delta_{2'}$	d_{xy}		
Δ_2	$d_{x^2-y^2}$	Δ_2	$d_{x^2-y^2}$
Δ_5	p_x, p_y, d_{xz}, d_{yz}	Δ_{3+4} (Δ_5)	p_x, p_y, d_{xz}, d_{yz}

Next we must examine the symmetry character of the decaying state with $\kappa_{\min}$ in ZnSe. We find that it is of symmetry Δ_1 (rotationally invariant). We conclude then that the incident spin-up Δ_1 state of Fe can couple to this Δ_1 state of ZnSe and tunnel through, while the rest of the majority states, plus all the minority states, are incompatible with the $\kappa_{\min}$-state.[4] As a consequence of the above, we realise that the spin-up current can tunnel best though the Δ_1 state, while the minority current decays much faster, since the minority incoming states couple to other ZnSe decaying states with larger decay parameter.

Having established that in the parallel moment configuration the spin-up states conduct much better than the minority, we examine the antiparallel configuration by reversing the moment (and the spin directions) in the second lead. As before, only spin-up wavefunctions will be able to reach the second lead. But there they will confront bands of incompatible wrong symmetry ($\Delta_{2'}$ and Δ_5), because the two spin directions are reversed. As a result they cannot couple to outgoing wavefunctions, and they are blocked at the second interface. Thus the transmission probability, and by the Landauer formula the conductance, is much lower than in the parallel case, and the TMR effect presents itself. In fact, Ref. [13] shows a large TMR ratio for these junctions, close to the ideal 100%.

These considerations and results ar completely at odds with the Julliere model. There, only the DOS and the spin polarisation at E_F in the vicinity of the interface play a role. Here, the

[4]To be more exact, the symmetry of ZnSe is reduced compared to Fe, and it follows the C_{2v} group around the z axis. Therefore its Δ_1 symmetry is also compatible with the $\Delta_{2'}$ symmetry of Fe. But the coupling there is really very weak, because the Fe $\Delta_{2'}$ states consist locally of d_{xy} orbitals and these are very locallised and directed in-plane.

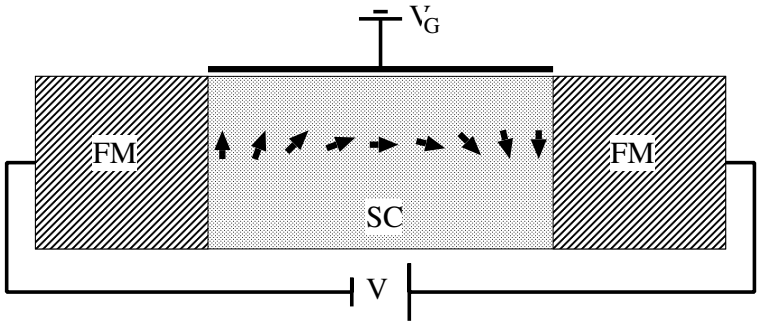

Fig. 1.8: The device proposed by Datta and Das. Current is injected from the left ferromagnetic lead, manipulated within the semiconductor and detected by the right ferromagnetic lead.

detailed band structure is taken into account. Actually the two approaches are valid in different regimes. Julliere's model works well in the case of rough interfaces or amorphous barriers, where one should really average over all incident states by some method. Then we can imagine that the symmetry and other details of the incoming state is of no importance, since it will suffer random diffuse scattering at the interface. On the other hand, the second approach is valid in the ballistic regime, when we have completely ordered structures.

In the present day, most applications of TMR use alumina (Al_2O_3) as a tunneling barrier, which is amorphous; also, no special care is taken for a clean interface. Therefore Julliere's model is applicable. On the other hand, many experiments continue to be oriented towards the construction of clean, ordered stuctures, where the ballistic approach is applicable. As we have seen, one advantage of succeeding in creating such structures would be an extremely high TMR ratio.

1.5 Spin Injection

In 1990, Datta and Das [3] proposed a spin-filter transistor where a spin-polarized current could be created, manipulated and detected. For details we refer the reader to the original publication, but in short, we can say that the device (shown here schematically in Fig. 1.8) would consist of two ferromagnetic leads sandwiching a semiconductor region. The semiconductor part should be constructed in such a way (by doping or other techniques) that the Fermi level would be slightly within the conduction band. Under a small bias voltage in such a structure, spin-polarised current could be injected from the ferromagnetic lead into the semiconductor. Then this spin current could be manipulated, *i.e.*, its magnetisation direction could be rotetated willingly by utilising the Rashba spin-orbit effect via an applied field (the strength of the field would determine the degree of rotation). Finally it would be detected by a second ferromagnetic lead, and the conductance should depend on the degree of rotation; for example, if no rotation was performed one would expect a high conductance, while if the magnetisation axis was reversed, a very low conductance. Thus a continously varying magnetoresistance ratio could be achieved, depending on the strength of the external field. As a model the conception was ideal, but many difficulties had (and still have) to be overcome before is could be realised.

It seems that the bottleneck is the success in injecting a spin-polarised current from the fer-

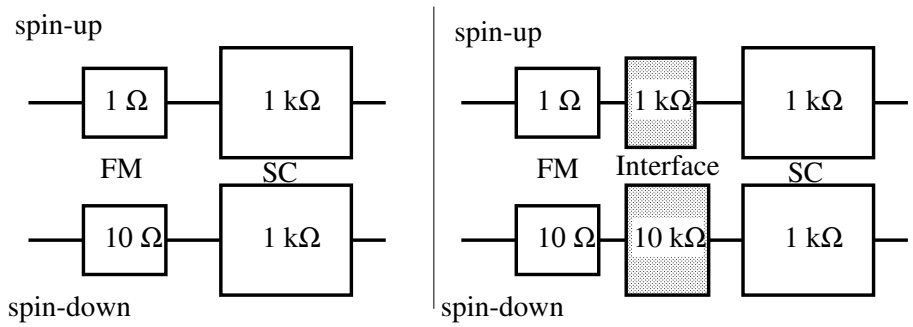

Fig. 1.9: The origin of the problem in spin injection (left) and its solution (right).

romagnet (FM) into the semiconductor (SC), the so-called *spin injection*. Many experimental trials have been unsuccessful, and a theoretical model that could explain the failure was proposed by Schmidt and co-workers [4]. The idea is presented in Fig. 1.9 (left). In short, although there is a difference between the resistance of the spin-up and spin-down channel, the resistivity of the semiconductor itself is so large that it drowns every other effect. The only possibility to measure some effect would be that either the current is initially almost 100% spin-polarised, as would be the case in a half-metallic lead, or that the interface resistance is so large and different for the two spin directions that it dominates the effect (as shown in Fig. 1.9, right).

This solution was pointed out by Rashba, and Fert and Jaffres [5], who proposed that one should include a tunneling barrier between the ferromagnet and the semiconductor. Such a barrier would present the desired high resistance and also spin-selectivity as in the TMR junctions. In fact, there exists a natural way to obtain such a barrier: it is the Schottky barrier[5] at the ferromagnet/semiconductor contact, the height and thickness of which can be manipulated by appropriate doping and interface engineering. Spin injection experiments including tunneling have indeed been successful [16].

The problem pointed out by Schmidt *et al.* [4] manifests itself in the diffusive regime, when the resistivity of the semiconductor is very high. However, if one would work in the ballistic regime, with a defect-free semiconductor, then the scattering at the interface itself would be the main source of resistance without even the need of a tunnel barrier. If one could then achieve spin-selective interface scattering, *e.g.* high transmission for spin up and high reflection for spin down electrons, spin injection would be possible [17, 18]. That such an option exists in principle has been shown by *ab initio* calculations [19–21]. Here we shall describe briefly the ideas and conclusions of this approach.

The systems under study here are Fe/SC(001) junctions, where SC stands for ZnSe and GaAs. The lattice mismatch between Fe and the SC is less than 2%, making a coherent growth of the two materials possible. In fact, the lattice constant of Fe is practically half the one of the SC, so that one has just to consider a double-sized Fe unit cell. Having this in mind, we can assume in-plane periodicity, so that $\mathbf{k}_{\|}$ is conserved during scattering.

We consider thermal injection, that is, injection of electrons at E_F directly into the SC conduction band. This is of course only possible if the conduction band minimum E_c is lower than

[5] We remind the reader that the Schottky barrier is a region near the metal-SC interface, extending into the SC, where no free charges are present even if such charges exist within the SC due to doping.

E_F. In reality this can be achieved via a gate voltage, by n-doping, or by formation of a quantum well. Here we emulate such a situation by rigidly shifting the SC potential from the third monolayer (ML) and on, until E_c falls slightly under E_F; we use $E_F - E_c$ up to 100meV which is related to typical carrier concentrations in similar experimental situations [17], and then we vary E_c to find the conductance as a function of the gate voltage shift.

In Fig. 1.7 we show the energy bands of Fe, ZnSe, and GaAs along the k_z-direction, for $\mathbf{k}_\parallel := (k_x, k_y) = 0$, *i.e.* for normal incidence on the (001) interface. As the energy E_c is lowered, the injection will start at the threshold $E_c = E_F$, at which there will be available states to carry the current deep into the SC bulk. Close to E_c, the Fermi sphere of the SC is very small, due to the almost parabolic dispersion relation; for this reason only electron states with $\mathbf{k}_\parallel$ almost zero will be transmitted. Therefore we shall base our discussion on what happens exactly at $\mathbf{k}_\parallel = 0$ and argue that because of continuity the results will be similar for $\mathbf{k}_\parallel \simeq 0$. We have also compared the results at $\mathbf{k}_\parallel = 0$ with integrated results over the whole surface Brillouin zone, and our results for the spin current polarisation remained unchanged within a few percent.

We can argue directly that the spin injection is largely determined by the symmetries of the band structures. For $\mathbf{k}_\parallel = 0$, the irreducible representation of the incoming Fe state must be compatible with the one of the outgoing SC state, if there is to be any transmission from the former to the latter. Now we can do a symmetry analysis, just as in the case of TMR earlier. In both GaAs and ZnSe, the conduction band along the Δ-direction belongs to the irreducible representation Δ_1 of C_{2v}, which consists locally of s, p_z, d_{z^2}, and d_{xy}-like states,[6] and is invariant under all point group operations of C_{2v} around the z-axis. One must see which Fe bands are compatible with this representation. In the majority-spin direction, in the energy range of interest Fe has only one band, which belongs to the irreducible representation Δ_1 of C_{4v}, consisting locally of s, p_z, and d_{z^2} orbitals, and which is compatible with the SC Δ_1 of C_{2v}. On the other hand for the minority spin Fe states this Δ_1 band starts much higher, at the energy denoted as Γ_{12} in Fig. 1.7, due to the exchange splitting and the s-d hybridisation gap of Fe. Up to there, there exist bands belonging to the irreducible representations Δ_5, Δ_2, and $\Delta_{2'}$ of C_{4v}. The last one has d_{xy} character and can couple to the SC conduction band, while the rest are incompatible because they have the wrong symmetry (they consist mainly of d_{xy} and p_x and p_y orbitals) and are orthogonal to the SC states. As we shall see from the results later on, even this $\Delta_{2'}$-state couples only weakly to the SC conduction band, presumably because it is more localised than the Δ_1 Fe band which contains the extended s and p_z orbitals. As a result, although Fe has a higher density of states in the minority spin band than in the majority, most of the candidate conductance channels are blocked for symmetry reasons and the majority-spin states prevail. The situation is expected to change above the energy level Γ_{12} where the Fe minority Δ_1-state comes into play.

In Fig. 1.10 (left) we see the spin-resolved conductance in the case of thermal injection in Fe/ZnSe(001) and Fe/GaAs(001), for both SC terminations in each system, as a function of $E_F - E_c$. Evidently, the majority-spin conductance is orders of magnitude higher than the minority spin conductance; the latter had to be magnified by the inset number in each case. This leads to an extremely high spin polarisation. The reason is, again, the extremely poor

[6] The d-states in question are not inherent to the SC atoms, but rather describe the distortion of the s-p states by the neighbours in the tetrahedral positions, if one uses the language of local orbitals.

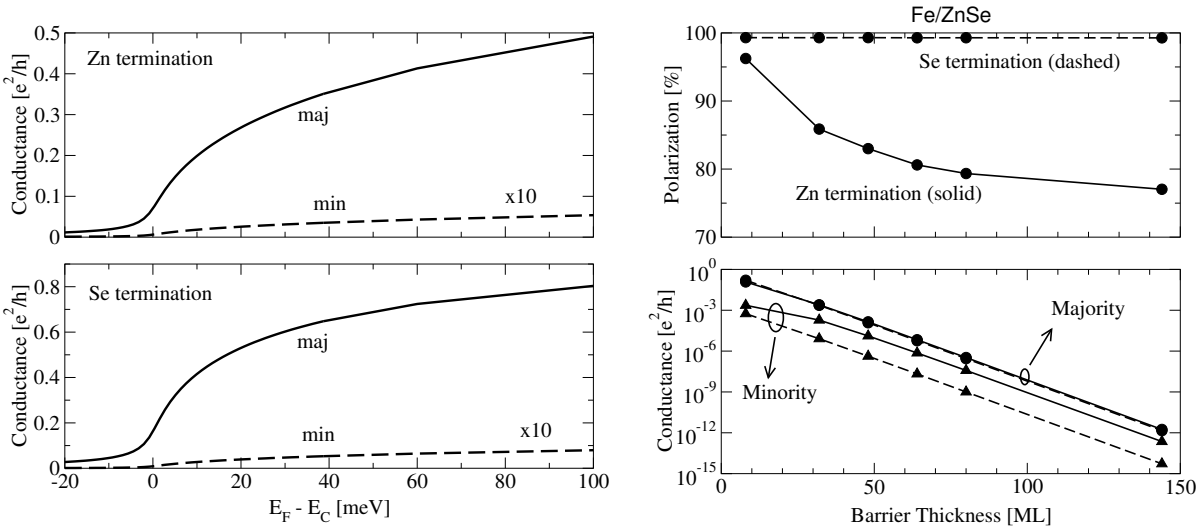

Fig. 1.10: Left: Thermal injection in Fe/ZnSe(001) without an intermediate tunneling barrier, for $\mathbf{k}_{\|} = 0$. Both SC terminations are shown. The full lines represent the majority-spin conductance, and the dashed lines the minority one. The minority spin conductance is orders of magnitude lower than its majority counterpart and has been magnified by the inset factor each time in order to be viewable. Right: Thermal injection in Fe/ZnSe(001), with a simulated intermediate Schottky tunneling barrier, at $\mathbf{k}_{\|} = 0$. In the upper panel the current polarisation is shown as a function of the barrier thickness, and in the lower panel the majority and minority conductance is shown. Both SC terminations are considered: the full lines represent the Zn termination, while the dashed lines represent the Se termination. The minority spin conductance is again orders of magnitude lower than its majority counterpart; note the logarithmic scale of the conductance. The results are similar for the Fe/GaAs system.

coupling of the Fe minority spin $d_{x^2-y^2}$-like $\Delta_{2'}$ state to the SC conduction band states, as opposed to the majority Δ_1 state.

The conductance as a function of energy behaves as $g \propto \sqrt{E_F - E_c}$; this is understandable since the transmission probability is always proportional to the group velocity of the outgoing states. As a trivial example, one can consider the one known from elementary quantum mechanics, where an incoming free electron in one dimension with energy $E = \hbar^2 k^2/2m$ (and velocity $\hbar k/m$) encounters an infinitely long potential step of height $V_0 < E$; if the outgoing wavenumber is then $k' = \sqrt{2m(E - V_0)/\hbar}$, the transmission probability is $4kk'/(k+k')^2 \propto \sqrt{E - V_0}$ for $k' \to 0$. The tails of the conductance below $E_F - E_c = 0$ in Fig. 1.10 are due to a small but nonzero imaginary part in the energy, necessary for numerical reasons.

Next we simulate the effect of a Schottky barrier; to this end we perform the shift of the SC potential (to bring $E_c < E_F$) not abruptly after two ML, but gradually over a thickness of N monolayers. This would represent an N-monolayer-thick Schottky barrier, through which the electrons would have to tunnel before reaching the SC conduction band. This tunneling procedure is also spin-selective with a mechanism similar to the one encountered in epitaxial magnetic tunnel junctions [13, 15]. The evanescent states within the barrier, determined by the complex band structure of the SC, have different coupling properties with the incoming Fe states: the majority-spin electrons are able to couple well to the states of Δ_1 symmetry which

show the slowest decay, in contrast to the minority spin electrons. Thus such a barrier promotes also a majority-spin selection in the transmitted current.

In Fig. 1.10 (right), the conductance for both spins is demonstrated as a function of barrier thickness, together with the resulting spin polarisation, at $\mathbf{k}_{\|} = 0$. Evidently the conductance decays exponentially, with a decay parameter that is characteristic of the evanescent states within the barrier. The junctions are still highly spin-polarising. However, in the case of Zn termination in Fe/ZnSe increasing the barrier thickness seems to reduce the current spin polarisation. This effect can be explained by the existence of interface resonances, as metal-induced gap states, in the Fe/SC(001) system among the minority spin states. These resonance states are of Δ_1 symmetry and therefore the minority Fe $\Delta_{2'}$ states as well as the semiconductor Δ_1 states can couple to them. Such resonances strongly enhance the electron tunneling, and can even be the main channels of conductance. This requires, however, that the resonance energy of the interface states lies very close to the Fermi level, which is not the case for the considered systems. For large thicknesses there is of course a saturation of the polarisation, since these resonances decay exponentially as well. Finally we note that calculations of the $\mathbf{k}_{\|}$-resolved densities of states have verified our suspicion for the existence of the above-mentioned interface resonances.

We must stress that the origin of the very high spin polarisation of the current here, as well as in the TMR Fe/ZnSe/Fe (001) junction, is basically the symmetry of the incident wavefunctions and of the SC conduction band. Thus we can speak of *symmetry enforced spin polarisation* of the current. This means that other interfaces, such as the (111) or (110), cannot show as strong an effect, as one can easily see by inspection of the symmetry properties of those bands. Also, strong interface disorder can kill the effect, since the various incoming channels will be mixed at the interface [21]. Therefore good quality interfaces are required.

Finally we note the possibility to use other materials for spin injection, such as half-metals or diluted magnetic semiconductors; such experiments are in the center of interest at the present day.

1.6 Diluted Magnetic Semiconductors

It has been attempted to make use of not only the charge but also the spin degree of freedom in modern semiconductor electronics for information processing [22, 23]. This new developing field is called spintronics. In order to establish spintronics as a practical technology, it is indispensable to fabricate a new material, because currently used materials in conventional electronics are usually non-magnetic and only charges are controllable. In order to introduce magnetic properties into semiconductors, diluted magnetic semiconductors (DMS) were proposed and fabricated. Actually, the discovery of the ferromagnetism in (In, Mn)As [24] and (Ga, Mn)As [25] promoted DMS to be fundamental elements for spintronics because of a compatibility with semiconductors used in present electronics. However, their Curie temperatures T_{C}, *e.g.*, 110 K for (Ga, Mn)As [26], are not high enough for real applications. Therefore, many experiments were performed to search for DMS with room-temperature ferromagnetism and recently high T_{C}-values were reported for several systems ((Ga, Mn)N [27–29], (Ga, Cr)N [30] and (Zn, Cr)Te [31]).

The ferromagnetism in DMS has been investigated theoretically either by a model Hamiltonian

[32–34] or by *ab initio* methods [35–39]. Despite of their basic differences, both methods gave similar predictions for the ferromagnetism. Nevertheless, no consensus has been reached about the origin of the ferromagnetism. In fact, Dietl *et al.* proposed Zener's *p-d* exchange interaction to describe the magnetism [32, 33]. This model predicts room-temperature ferromagnetism in (Ga, Mn)N [32, 33] and explained many physical properties of (Ga, Mn)As successfully, as is shown by MacDonald *et al.* [34]. On the other hand, Akai pointed out by first-principles calculations that Zener's double exchange mechanism is responsible for the ferromagnetism in (In, Mn)As [35]. Similar arguments were also given by Sato *et al.* [36, 37], who predicted high $T_{\rm C}$-values for (Ga, Mn)N, (Ga, Cr)N and (Ga, Cr)As. Thus, despite the fact that ferromagnetism of DMS is one of the most important topics in spin-electronics, this is still a controversial issue. In this lecture, we show that both mechanisms are important for understanding ferromagnetism in DMS. We calculate the electronic structure of (Ga, Mn)X, where X refers to N, P, As or Sb, and estimate their $T_{\rm C}$'s from first-principles within the mean field approximation. We discuss the relation between the electronic structure and the dominant exchange mechanism and show that the two mechanisms lead to very different concentration dependences of $T_{\rm C}$.

The electronic structures of DMS are calculated based on the local spin density approximation (LSDA). In DMS, Mn impurities are introduced randomly into cation sites of the host semiconductor. The substitutional disorder in DMS is described by using the Korringa-Kohn-Rostoker coherent-potential-approximation (KKR-CPA). In the CPA, configuration-averaged properties of alloys are calculated within a sigle-site approximation. By this method, it is also possible to take magnetic disorder into account. For example, we can simulate the disordered local moment (DLM) state by considering a multi-component alloy consisting of $\rm Mn^{\uparrow}$, $\rm Mn^{\downarrow}$ and host atoms at the cation site, where up and down arrows indicate the direction of the local moment of the Mn. It is known that this approach gives a good description of the paramagnetic state of ferromagnetic (FM) materials above $T_{\rm C}$. Therefore, to discuss the magnetism of DMS, the total energies (TE) of both the FM state, described as $(\rm Ga_{1-\mathit{x}}, Mn_{\mathit{x}}^{\uparrow})X$, and the DLM state, described as $(\rm Ga_{1-\mathit{x}}, Mn_{\mathit{x}/2}^{\uparrow}, Mn_{\mathit{x}/2}^{\downarrow})X$, are calculated and in particular the energy difference $\Delta E = TE(\text{DLM state}) - TE(\text{FM state})$ is estimated. Throughout the present calculations, we used the KKR-CPA package (MACHIKANEYAMA2000) developed by Akai [40]. Local lattice distortions due to Mn impurities are neglected and the experimental lattice constants of host semiconductors are used in the calculations [41]. Relativistic effects are considered in the scalar relativistic approximation.

Since the LSDA describes only the ground state properties, we need an additional procedure to estimate $T_{\rm C}$. To describe the magnetic properties at finite temperatures we use the Heisenberg model, which can be written as $H = -\Sigma_{i\neq j} J_{ij} \vec{S}_i \vec{S}_j$, where $\vec{S}_i$ is the spin at site i and J_{ij} is the exchange coupling constant between sites i and j. We can calculate the total energy difference $\Delta E^{\rm H}$ in this Heisenberg model within the mean field approximation as $\Delta E^{\rm H} = S^2 c^2 \Sigma_{n\neq 0} J_{n0}$, where c is the concentration of the magnetic ions and n sums over all sites of the cation sublattice. This $\Delta E^{\rm H}$ can be directly identified with the energy difference ΔE calculated from first principles, since the CPA represents also a mean field theory. On the other hand, within the mean field theory of the Heisenberg model we can estimate $T_{\rm C}$ by using the Brillouin function expression, known as the molecular field theory, leading to $k_{\rm B} T_{\rm C} = \frac{2}{3} S^2 c \Sigma_{n\neq 0} J_{n0}$, where $k_{\rm B}$ is Boltzmann constant. The result $k_{\rm B} T_{\rm C} = \frac{2}{3} \Delta E / c$ allows to evaluate $T_{\rm C}$ from first principles.

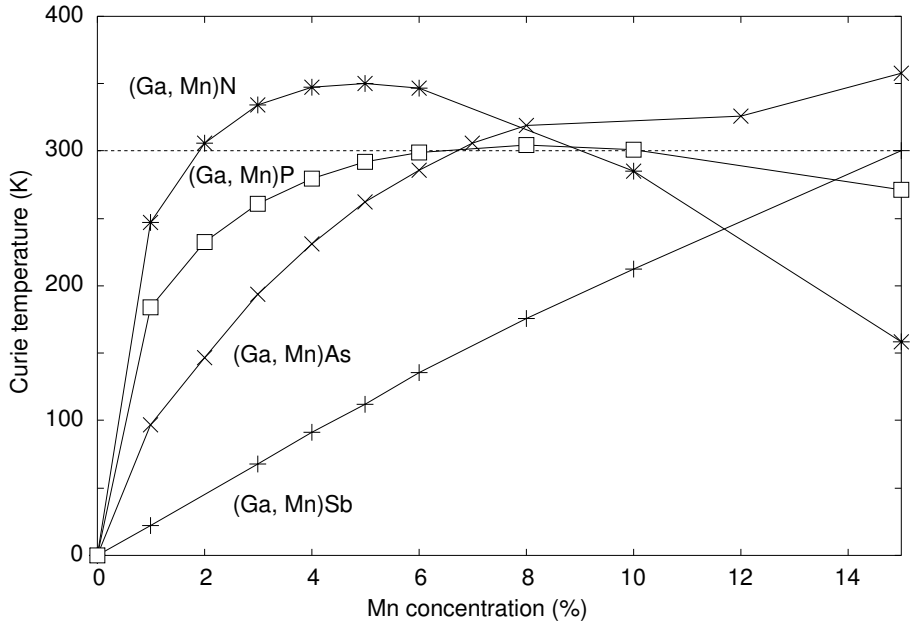

Fig. 1.11: Curie temperatures of (Ga, Mn)N, (Ga, Mn)P, (Ga, Mn)As and (Ga, Mn)Sb calculated from first-principles in the mean field approximation. The dashed line indicates room temperatures (300K).

Fig. 1.11 shows the Curie temperatures of (Ga, Mn)N, (Ga, Mn)P, (Ga, Mn)As and (Ga, Mn)Sb calculated in the mean field approximation from the total energies. For example, T_{C} is calculated as 260 K for 5% Mn-doped GaAs, which is much higher than the experimental value of 110 K [26]. However, when comparing with experiment one has to remember that the mean field approximation usually somewhat overestimates the Curie temperature of ferromagnets, which is in our case presumably not a large effect due to the long range nature of the interaction. Moreover, due to compensation effects the effective Mn concentration is not well known. In the present calculation the neutral charge state of Mn is considered. In case of (Ga, Mn)N, we have a maximum T_{C} of approximately 350 K, which seems to be inconsistent with recent experimental T_{C} of 940 K [27]. The ferromagnetism in (Ga, Mn)P [42] and (Ga, Mn)Sb [43] were also observed experimentally.

For low concentrations, the T_{C}'s of (Ga, Mn)N, (Ga, Mn)P and (Ga, Mn)As scale roughly proportional to the square root of the Mn concentration. This behavior is also observed in Cr-doped III-V DMS [37]. In contrast to this, the T_{C} of (Ga, Mn)Sb shows an almost linear concentration dependence. As a whole, we find clear chemical trends in the concentration dependences. In (Ga, Mn)N, T_{C} goes up very sharply and reaches a maximum value of 350 K at approximately 5 % of Mn, then goes down. In (Ga, Mn)P, T_{C} increases sharply up to a saturated value of about 300 K. In (Ga, Mn)As, T_{C} shows similar dependences to (Ga, Mn)P, but T_{C} increases more moderately for low concentrations and still increases for high concentrations. Finally, T_{C} of (Ga, Mn)Sb shows a linear dependence with nearly no curvature. As a result, the concentration dependence shows in the sequence (Ga, Mn)N $\rightarrow$ (Ga, Mn)Sb a dramatic transition from a $\sqrt{c}$ dependence to a linear behavior.

In order to discuss the origin of the ferromagnetism, we show in Fig. 1.12 the total density of states (DOS) and local density of d-states at a Mn site in the FM state. In III-V compounds,

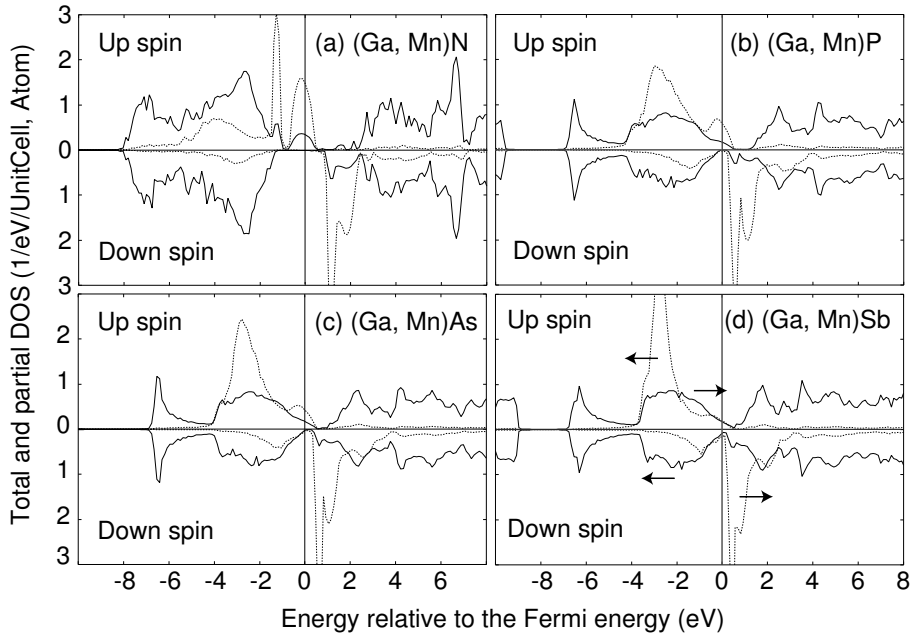

Fig. 1.12: Total density of states per unit cell (solid lines) and local density of d-states at Mn site (dotted lines) in (a) (Ga, Mn)N, (b) (Ga, Mn)P, (c) (Ga, Mn)As and (d) (Ga, Mn)Sb in the ferromagnetic state for 5% Mn. In (d), the arrows indicate the relative shift of the p- and d-states due to hybridization.

the host valence band originates from anion p-states. Therefore, the DOS reflects the increase of the atomic p level in the series N $\rightarrow$ Sb, resulting in the gradual change of relative position of Mn d-states with respect to the host valence band. In case of (Ga, Mn)N, the host valence band is lower in energy than the Mn d-states and clear impurity bands appear in the band gap. The Fermi level (E_{F}) is located in the t_2 impurity band, being induced by the majority d-states of Mn. Since the minority d-states are much higher in energy, (Ga, Mn)N shows a half-metallic behavior. On the other hand, in case of (Ga, Mn)Sb, the majority d-states of Mn are located deep in the host valence band and the local DOS at E_{F} agrees well with the host DOS. Thus in (Ga, Mn)Sb the majority d-states are localized representing a d^5 configuration, and the hole states above E_{F} consist of host valence states. The DOS is no longer half-metallic. This is presumably an LDA-error since in the LDA the band gap of GaSb vanishes. (Ga, Mn)P and (Ga, Mn)As are apparently intermediate cases. Their local DOS still show small peaks around the E_{F}, slightly larger in GaP than in GaAs. However, the gap states are almost merged into valence bands showing rather broad resonances (Fig. 1.12-(b) and (c)). Both systems are half-metallic and the total magnetic moment per Mn atom is just 4 μ_{B} as in (Ga, Mn)N.

The chemical trends seen in the DOS allow to explain the drastic differences in the concentration dependence of T_{C}. Here we will concentrate on the extreme cases, (Ga, Mn)N and (Ga, Mn)Sb. The electronic structure of (Ga, Mn)N is characterised by the sharp e impurity band and the broader t_2 impurity band. Of the 7 valence electrons of Mn, three are accommodated in the valence band, two in the majority e band and the remaining two in the threefold degenerate t_2 band, so that the Fermi energy falls in the upper part of this band, leaving one state per Mn empty. With increasing concentration the impurity bands broaden, and it is just the broadening of the partially filled t_2 band which stabilized the ferromagnetism. This is the double exchange interaction in DMS [35, 44]. The energy gain due to double exchange is proportional to the band

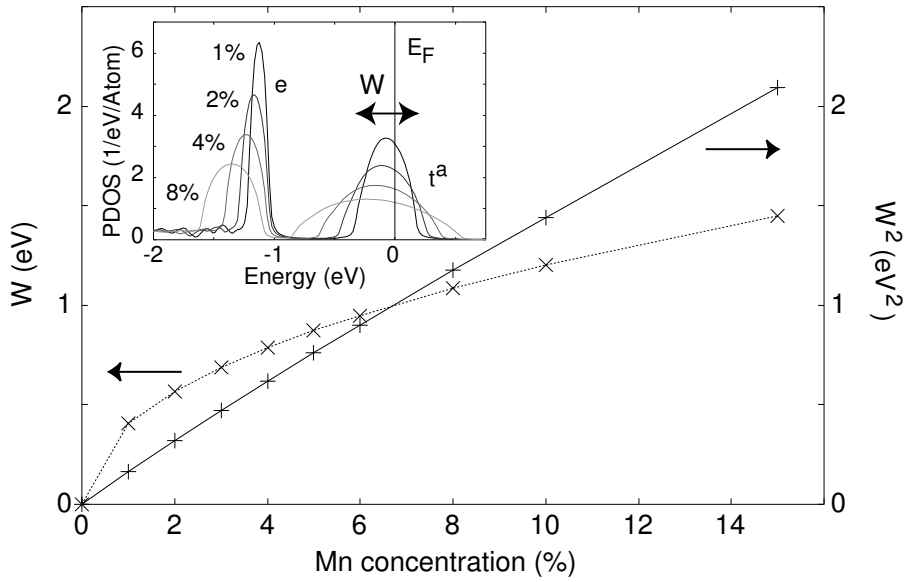

Fig. 1.13: The width (W) and its square (W^2) of the impurity t_2-band of (Ga, Mn)N as a function of Mn concentration. The inset shows the local density of Mn gap-states.

width (W), being defined as the deviations of the energy eigenvalues E from the mean value $\bar{E}$, *i.e.*, $W^2 = \langle (E - \bar{E})^2 \rangle = \Sigma_{m \neq 0} |H_{0m}|^2$. The last step follows from a tight-binding description and H_{0m} is the hopping integral between the sites 0 and m. The impurity band is formed due to the hopping of electrons between Mn atoms. Suppose a certain configuration of Mn atoms in GaN, with one of them at site 0. In this configuration, H_{0m} has a finite value, say t_{0m}, if another Mn atom sits at site m, otherwise $H_{0m} = 0$. In the CPA, we make a configurational average over all sites m $\neq 0$. The probability to find a Mn atom at site m is given by c, the atomic concentration of Mn. Thus the configurational average gives: $\langle W^2 \rangle_{\text{conf.}} = c \Sigma_{m \neq 0} |t_{0m}|^2$. Therefore the effective band width W is proportional to $\sqrt{c}$ [45]. This is in fact found in our KKR-CPA calculations, as is shown in Fig. 1.13. Here the effective quantities W and W^2 of the t_2 impurity band of (Ga, Mn)N are plotted as a function of the Mn concentration. The linear behavior of W^2 proofs the $\sqrt{c}$-scaling of W, in agreement the above arguments.

Now we discuss the stability of the FM state with respect to the DLM state. In the latter configuration, half of Mn neighbors have moments parallel to the Mn moment at site 0. Therefore the energy gain due to double exchange is always by a factor of $1/\sqrt{2} \sim 0.71$ smaller than the one of the FM state. The other half of the Mn neighbors, being anti-parallel aligned, gain energy by the super-exchange interaction [35, 44]. For each spin direction, the majority d-states of the one Mn and the minority d-states of the other one hybridize covalently, *i.e.*, the majority states at one site are pushed down, in this way stabilizing the anti-ferromagnetic alignment, if the Fermi level lies between the two peaks or coincides with one of them. This super-exchange interaction is expected to scale with the concentration $c/2$ of anti-aligned Mn-pairs. Therefore for small concentrations always the double exchange interaction wins due to the $\sqrt{c}$ behavior and stabilizes the ferromagnetism. This explain the behavior of T_{C} for (Ga, Mn)N shown in Fig. 1.11. The decrease at large concentrations can be explained by a rather strong super exchange interaction in this system, favoring the DLM state. However, if the t_2-impurity band is full or completely empty, as it is *e.g.* the case for (Ga, Fe)N or (Zn, Mn)Te, the double exchange interaction

vanishes and the super-exchange interaction stabilizes the DLM state. This is confirmed by our calculations giving negative $\Delta\epsilon$ or T_C values scaling linearly with c.

For (Ga, Mn)Sb, the other extreme case, the majority d-states can be regarded as localized and holes exist in the majority valence band of GaSb. This behavior is well described by Zener's p-d exchange model used by Dietl *et al.* [32, 33] and MacDonald *et al.* [34]. As in the discussion of Kanamori [44], in the FM state the hybridization between the Mn d levels and the Sb p-states pushes the lower levels down and the higher levels up, as is indicated for both spin directions by the arrows in Fig. 1.12-(d). As a result, holes appear in the majority host band, so that the GaSb host becomes polarized with a moment anti-parallel to the local Mn moment. This moment is close to 1 μ_B, since the system is nearly half-metallic. By configurational averaging, a homogeneous host polarization of $c\mu_B$ is obtained, which favors the ferromagnetic coupling of the Mn moments, by an energy proportional to the host polarization, scaling linearly with c. In the DLM state the average host polarization vanishes, so that this state is unfavorable. In conclusion, ferromagnetism is stabilized and T_C increases linearly with c.

In (Ga, Mn)P and (Ga, Mn)As, the impurity bands fuse with the host valence bands and the lower Mn d-states are not yet fully localized. In this sense, these systems are intermediate cases being influenced by both mechanisms, so that one might expect a superposition of $\sqrt{c}$ and a linear c-dependence. This is in line with Fig. 1.11, indicating that (Ga, Mn)As shows a weaker $\sqrt{c}$ and a somewhat stronger c dependence than (Ga, Mn)P.

In this lecture, we have discussed the origin of the ferromagnetism in DMS based on *ab initio* calculations for Mn-doped III-V compounds. We find that double exchange dominates if impurity bands in the gap are formed and that then T_C increases proportional to $\sqrt{c}$, where c is the Mn concentration. A typical example for this is (Ga, Mn)N. On the other hand, the p-d exchange mechanism dominates, if the d-states of the impurity are nearly localized, as it is the case for (Ga, Mn)Sb, and then a linear c-dependence of T_C is obtained. (Ga, Mn)P and in particular (Ga, Mn)As are intermediate cases. Actually, (Ga, Mn)As is on the border between the two mechanisms, and the results depend sensitively on the position of Mn d-states.

Acknowledgements

We gratefully acknowledge the support of M. Lezaic in carefully reviewing the manuscript.

References

[1] M. N. Baibich *et al.*, Phys. Rev. Lett. **61**, 2472 (1988).

[2] G. Binash, P. Grünberg, F. Saurenbach, and W. Zinn, Phys. Rev. B **39**, 4828 (1989).

[3] S. Datta and B. Das, Appl. Phys. Lett. **56**, 665 (1990).

[4] G. Schmidt, D. Ferrand, L. W. Molenkamp, A. T. Filip, and B. G. van Wees, Phys. Rev. B **62**, R4790 (2000).

[5] E. I. Rashba, Phys. Rev. B **62**, R16267 (2000); A. Fert and H. Jaffres, Phys. Rev. B **64**, 184420 (2001).

[6] J. C. Slonczewski, J. Magn. Magn. Matter **159**, L1 (1995); L. Berger, Phys. Rev. B **54**, 9353 (1996).

[7] G. Tatara (private communication).

[8] P. Bruno and C. Chapert, Phys. Rev. Lett. **67**, 1602 (1991).

[9] P. Zahn, I. Mertig, M. Richter, and H. Eschrig, Phys. Rev. Lett. **75**, 2996 (1995).

[10] S. Zhang and P. M. Levy, J. Appl. Phys. **69**, 4786 (1991).

[11] M. Julliere, Phys. Lett. A **54**, 225 (1975).

[12] T. Miyazaki and N. Tekuza, J. Magn. Magn. Mater. **139**, L231 (1995); J. S. Moodera, L. R. Kinder, T. M. Wong, and R. Meservey, Phys. Rev. Lett. **74**, 3273 (1995).

[13] J. M. MacLaren, X.-G. Zhang, W. H. Butler, and Xindong Wang, Phys. Rev. B **59**, 5470 (1999).

[14] J. C. Slonczewski, Phys. Rev. B **39** 6995 (1989).

[15] Ph. Mavropoulos, N. Papanikolaou, and P. H. Dederichs, Phys. Rev Lett. **85**, 1088 (2000).

[16] H. J. Zhu, M. Ramsteiner, H. Kostial, M. Wassermeier, H.-P. Schönherr, and K. H. Ploog, Phys. Rev. Lett. **87**, 016601 (2001); R. P. Hammar and M. Johnson, Appl. Phys. Lett. **79**, 2591 (2001); A. T. Hanbicki, B. T. Jonker, G. Itskos, G. Kioseoglou, and A. Petrou, *ibid.* **80**, 1240 (2002).

[17] D. Grundler, Phys. Rev. B **63**, 161307 (2001).

[18] G. Kirczenow, Phys. Rev. B **63**, 054422 (2001).

[19] O. Wunnicke, Ph. Mavropoulos, R. Zeller, P. H. Dederichs, and D. Grundler, Phys. Rev. B **65**, 241306(R) (2002).

[20] Ph. Mavropoulos, O. Wunnicke, and P. H. Dederichs, Phys. Rev. B **66**, 024416 (2002).

[21] M. Zwierzycki, K. Xia, P. J. Kelly, G. E. W. Bauer, and I. Turek, cond-mat/020442 (2002).

[22] H. Ohno, Science **281** 951 (1998)

[23] S. A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. von Molnár, M. L. Roukes, A. Y. Chtchelkanova and D. M. Treger, Science **294** 1488 (2001)

[24] H. Ohno, H. Munekata, T. Penney, S. von Molnár and L. L. Chang, Phys. Rev. Lett. **68** 2664 (1992)

[25] H. Ohno, A. Shen, F. Matsukura, A. Oiwa, A. Endo, S. Katsumoto and Y. Iye, Appl. Phys. Lett. **69** 363 (1996)

[26] F. Matsukura, H. Ohno, A. Shen and Y. Sugawara, Phys. Rev. B **57** R2037 (1998)

[27] S. Sonoda, S. Shimizu, T. Sasaki, Y. Yamamoto and H. Hori, J. Crys. Growth **237-239** 1358 (2002)

[28] N. Theodoropoulou, A. F. Hebard, M. E. Overberg, D. R. Abernathy, S. J. Pearton, S. N. G. Chu and R. G. Wilson, Appl. Phys. Lett. **78** 3475 (2001)

[29] M. L. Reed, N. A. El-Masry, H. H. Stadelmaier, M. K. Ritums, M. J. Reed, C. A. Parker, J. C. Roberts and S. M. Bedair, Appl. Phys. Lett. **79** 3473 (2001)

[30] M. Hashimoto, Y. K. Zhou, M. Kanamura and H. Asahi, Solid State Comm. **122** 37 (2002)

[31] H. Saito, V. Zayets, S. Yamagata, Y. Suzuki and K. Ando, J. Supercond. to be published.

[32] T. Dietl, H. Ohno, F. Matsukura, J. Cibert and D. Ferrand, Science **287** 1019 (2000)

[33] T. Dietl, Semicond. Sci. Technol. **17** 377 (2002)

[34] T. Jungwirth, J. König, J. Sinova, J. Kučera and A. H. MacDonald, Phys. Rev. B **66** 012402 (2002)

[35] H. Akai, Phys. Rev. Lett. **81** 3002 (1998)

[36] K. Sato and H. Katayama-Yoshida, Semicond. Sci. Technol. **17** 367 (2002)

[37] K. Sato, P. H. Dederichs and H. Katayama-Yoshida, Europhys. Lett. **61** 403 (2003).

[38] J. Kudrnovský, I. Turek, V. Drachal. F. Máca, J. Mašek and P. Weinberger, submitted to Phys. Rev. B.

[39] L. M. Sandratskii and P. Bruno, Phys. Rev. B **66** 134435 (2002)

[40] H. Akai, http://sham.phys.sci.osaka-u.ac.jp/kkr/

[41] R. W. G. Wyckoff, *Crystal Structures*, Wiley, New York (1986)

[42] N. Theodoropoulou, A. F. Hebard, M. E. Overberg, C. R. Abernathy, S. J. Pearton, S. N. G. Chu and R. G. Wilson, cond-mat/0201492

[43] E. Abe, F. Matsukura, H. Yasuda, Y. Ohno and H. Ohno, Physica E **7** 981 (2000)

[44] J. Kanamori, Transac. Mag. Soc. Jpn. **1** 1 (2001)

[45] B. Velický, S. Kirkpatrick and H. Ehrenreich, Phys. Rev. **175** 747 (1968)