Theory of Spin Transfer Torque

M. D. Stiles
Electron Physics Group
National Institute of Standards and Technology
Gaithersburg MD 20899-8412, USA mark.stiles@nist.gov

October 17, 2005

Abstract

In magnetic multilayers, the spin polarized currents flowing between the magnetic layers can exert torques on the magnetizations of the layers when the magnetizations are not collinear. The theory of these spin transfer torques is developed in terms of semiclassical transport calculations coupled with quantum mechanical calculations of the behavior of spins at interfaces. The result is an expression for the torque in terms of the geometry of the devices and the magnetic configuration.

1 Introduction

The observation of giant magnetoresistance [1, 2] in multilayers of magnetic metals separated by nonmagnetic metals established that electric current in these magnetic multilayers is spin polarized. Two years later, Slonczewski [3] and Berger [4] showed that the angular momentum carried by the spin polarized current between the magnetic layers exerts torques on the magnetizations. They predicted that when the current is large enough, these spin transfer torques can cause reversal and precession of the magnetization. This prediction has been confirmed in a number of laboratories and in a number of different sample geometries including mechanical point contacts [5, 6], lithographically defined point contacts [7, 8], electrochemically grown nanowires [9], manganite junctions [10], lithographically defined nanopillars [11, 12, 13, 14, 15, 16, 17, 18, 19], tunnel junctions [20, 21, 22, 23], and semiconductor structures [24].

There has been a great deal of research both exploring the variety of behaviors exhibited by these systems and testing the possible explanations that have been offered. Two chapters in the volume by Rippard and by Buhrman and Ralph describe some of these efforts in detail. Spin transfer torques are not only interesting scientifically but are also potentially important in some commercial applications. In existing implementations of magnetic random access memory (MRAM), bits are switched by the magnetic fields caused by current pulses, see the chapter by Engel and Rizzo for details. This method of switching severely constrains the fabrication of the bits because the magnetic fields are not well localized. The possibility that the bits could be directly addressed and switched by a polarized current is one promising application of spin transfer torques. Another possible application is based on using the rapid precession observed in these multilayers under suitable conditions. The high frequency oscillating resistance suggests the use of these systems as current controlled high frequency oscillators, see the chapter by Rippard for more details. On the other hand, in read heads with perpendicular current flow, precession interferes with
the operation of the heads by creating an unwanted noise source. In the end, research is necessary to both optimize and minimize these effects.

The giant magnetoresistance is the dependence of the resistance of a magnetic multilayer on the orientation of the magnetizations in neighboring ferromagnetic layers. Thus, it provides a mechanism for inferring the relative orientations of the magnetizations in a multilayer simply by measuring the resistance. When spin transfer torques change the magnetic configuration of the structure, the resistance of the structure changes as well. To detect these torques, typically one layer, called the free layer, responds to the spin transfer torques, and the other, called the fixed layer, is constructed in a way so that it does not. Thus, when the current rotates the magnetization of the free layer, the resistance of the multilayer changes. In the precessing state, a constant input current gives a rapidly varying voltage due to the time variation in the resistance.

All of the devices studied share some common geometrical attributes, starting with the fact that the current flows perpendicularly to the magnetic layers. Spin transfer torques are interfacial (see below) so the free layer is kept thin, typically 3 nm to 6 nm thick, because the current necessary to move the magnetization from its ground state is proportional to this thickness. Typically the fixed layer is much thicker than the free layer so it responds less strongly to the current induced torques. Typical cross sectional dimensions are on the order of 100 nm for two reasons related to the high current density necessary to move the magnetization. First, the heating generated by this current density needs to be tied to a relatively large heat sink. Second, the torques due to the current-induced magnetic field become less important as the cross sectional area decreases.

Tsoi et al. observed spin-transfer torques using a point contact to a magnetic multilayer in a high applied magnetic field [5]. The signature of a spin-transfer torque in their measurement was the observation of peaks in the differential resistance, peaks that appeared only for one direction of current flow. These peaks were interpreted as evidence of a precessing state. Shortly thereafter, measurements on lithographically defined samples, point contacts by Myers et al. [7] and nanopillars by Katine et al. [11], showed both a similar peak, asymmetric in the current, in high applied magnetic fields, and hysteretic switching between parallel and antiparallel states at lower field. The crossover field between these behaviors is roughly the zero-current coercive field. The switching was attributed to spin transfer torques because the behavior was asymmetric. For large currents of electrons flowing from the fixed layer to the free layer, the parallel state was stable, whereas for large currents in the opposite direction, the antiparallel state was stable. If the current induced switching were due to the magnetic fields produced by the current, the stable state in high currents would be independent of the direction of current flow.

High frequency measurements of resistance provided more data to interpret the origin of the peak in the differential resistivity at high fields and currents. Urazhdin et al. [14] showed that rapid fluctuations between two states with different resistances give rise to peaks in the differential resistance. The dwell times in each state depended strongly on current, so that with increasing current there was a rapid transition from spending most of the time in the low resistance state to spending most of the time in the high resistance state. This rapid and reversible rise in the resistance leads to a peak in the differential resistance. Shortly thereafter, measurements at still higher frequencies by Kiselev et al. [25] and Rippard et al. [26] demonstrated the existence of precession for currents and fields close to the peak, but not clearly associated with it. They observed sharp peaks in the power spectrum density, a measure of the frequency dependent resistance in the devices, characteristic of precession. In some cases, these peaks were extremely sharp, with Q values around 18 000 [27]. Krivorotov et al. [28] observed the precession through real time measurements of the resistance and even measured the few precession periods the system undergoes on the way to reversal in lower applied fields.
Finite sample temperatures add additional complications. For low applied fields, the hysteretic switching depends on the measurement time, or the rate at which the field or current is swept [29]. The observed two level switching is also strongly temperature dependent. There have been a number of studies of this and the associated low frequency noise [5, 29, 14, 30, 31]. Sometimes the power spectrum density has the Lorentzian form expected for two level switching, but frequently the low frequency noise has a more general form.

A number of experiments have tested qualitative and quantitative predictions of the models used to describe these systems. The dependence of the switching currents on the thicknesses of the layers confirm that the spin transfer torque is interfacial [32]. The dependence of the switching current on the angle of the applied field relative to the easy axis of the free layer shows the expected increases as the field becomes perpendicular [33]. Comparisons between calculated and measured “phase diagrams” [11, 34, 35] in which the behavior of the system is measured as a function of current and applied field show that the models qualitatively reproduce the measurements. For devices that are symmetric, the asymmetric behavior that typically is a signature of the spin transfer torque becomes symmetric in current, as expected [36]. Studies of the material dependence show the expected reduction of switching currents with decreasing saturation magnetization [37]. By choosing appropriate and different materials for each magnetic layer, it is possible to make a device in which the state with antiparallel magnetizations has the lowest resistance, the so called inverse giant magnetoresistance. Studies of spin transfer torques using these material systems [38] show that the stability of the parallel and antiparallel states depends mainly on the properties of the fixed layer, and only weakly on those of the free layer. Again, this result is consistent with the theoretical models described below.

In the rest of the chapter, I describe the theoretical developments put forth to predict and explain the observed behavior. Section 2, reviews spin transport in systems with collinear magnetizations. These calculations show how the current in the non-magnetic layers becomes spin polarized and what determines how large the polarization is in different devices. However, the magnetization must be non-collinear for spin transfer torques to play a role. To understand these torques it is necessary to calculate the behavior of individual electrons scattering from interfaces with their spin non-collinear with the magnetization. I describe these calculations in Sec. 3. Sec. 4 combines the transport calculations and the scattering calculations to give calculations of the torque as a function of the device geometry and the orientations of the magnetization. I briefly discuss the consequences of these torques on the dynamics in Sec. 5. More details of this behavior are given in the chapter by Thiaville.

2 Collinear Transport

The theory of spin transport near interfaces started with Aronov [39] in 1976. A macroscopic theory was described by Johnson and Silsbee [40], and van Son et al. [41], and derived from the Boltzmann equation by Valet and Fert [42]. The history of these developments is discussed by in a review article by Žutić et al. [43]. In this section, I discuss charge and spin transport in magnetic multilayers with collinear magnetizations, starting with a single interface between a ferromagnet and a non-magnet. The results for single interfaces combine to give results for more complicated structures, including the structure of principle interest in this chapter, two ferromagnetic layer separated by a non-magnetic layer embedded in a non magnetic host. The differences in the transport for parallel and antiparallel magnetizations of the two layers gives the giant magnetoresistance.

The negative charge of the electron is a potential source of confusion. For example, charge currents move oppositely to number (and spin) currents. Additionally, the electron’s magnetic moment is opposite
to its spin, leading to several more possibilities for confusion [44]. In this chapter, I choose to address these issues by avoiding charge and magnetization and working instead with number densities and currents and spin densities and currents. Multiplying the number current by \(-e\) gives the charge current; \(e\) is positive by convention. Similarly, multiplying the spin density by \(g\mu_B/\hbar\) gives the magnetization density. For free electrons, \(g_e \approx -2.002319\). In transition metal ferromagnets, in which the orbital moment is largely quenched, the \(g\)-values are close to \(-2\). In the examples discussed below, one of the magnetic layers will have its magnetization in the \(-z\) direction, so that the spin density is in the \(z\) direction. Majority electrons, referred to as “spin up,” have their spins aligned parallel to the spin density and minority electrons, referred to as “spin down,” antiparallel. The symbols \(\uparrow\) and \(\downarrow\) refer to majority and minority electrons respectively.

The main quantities of interest in this section are the spin accumulation and the spin current. The spin accumulation is the excess spin density above the equilibrium amount for each material. For the non-magnetic material, this is just the spin density itself. The spin current is the net flow of spins. Consider an isolated interface between a non-magnet and a ferromagnet with a constant and uniform current consisting of electrons flowing from the non-magnetic layer perpendicularly to the interface into the ferromagnetic layer. Fig. 1 shows the behavior of the spin accumulation and the spin current. Far from the interface, both behave as they would in bulk materials. There is no spin accumulation in either material and the spin current is zero in the non-magnetic material. In the ferromagnet on the other hand, the conductivity for majority electrons, \(\sigma_\uparrow\) is greater than that for minority electrons, so the current density is spin polarized. More current is carried by majority than by minority electrons, \(j_\uparrow > j_\downarrow\). As a reminder, \(j\) here is a number current density, not a charge current density. The difference in the spin dependence of the bulk conductivities between the ferromagnet and non-magnet means that more majority electrons are extracted from the interface region into the ferromagnet than are delivered to it from the non-magnet. Thus, a deficit of majority electrons builds up near the interface. Similarly, there is an accumulation of minority electrons near the interface because fewer are extracted into the ferromagnet than are delivered from the non-magnet.

When the density of spins varies spatially, as it does here, with an excess of minority spins and a deficit of majority spins accumulating near the interface, the spins diffuse in a direction to reduce this variation. Thus, majority electrons diffuse into the interface region and minority electrons diffuse away from it. The diffusion of these spins gives rise to the spatially varying parts of the spin current seen in Fig. 1.
material, the spin current for the majority electrons is given by

\[
j_{\uparrow} = (\sigma_{\uparrow}/e)E - D_{\uparrow}\nabla\delta n_{\uparrow},
\]

where \(\delta n_{\uparrow}\) is the excess density of majority electrons, and \(D_{\uparrow}\) is the spin dependent diffusion constant, which, like the conductivity \(\sigma_{\uparrow}\) is different in each material. The diffusion constants and the conductivities are related through the Einstein relation \(\sigma_{\uparrow}/\tau_{\uparrow} = D_{\uparrow}\mathcal{N}_{\uparrow}\) where \(\mathcal{N}_{\uparrow}\) is the spin-dependent density of states. The density of states also relates the accumulation of each spin to a spin-dependent chemical potential \(\delta n_{\uparrow} = \mathcal{N}_{\uparrow}\mu_{\uparrow}\).

The transport equations for majority and minority spins can be combined into an equation describing the net spin current and spin accumulation. The screening length in metals is on the atomic scale, so there is no charge accumulation in the interior of a conductor. This justifies setting \(\delta n_{\downarrow} = -\delta n_{\uparrow}\) except across interfaces where dipole layers are possible. Summing the transport equations for majority, Eq. (1), and minority spins, solving for the electric field \(E\) in terms of the current \(j\), and inserting the result into the difference of the two transport equations gives

\[
j_s = \gamma^P j - D_s \nabla \delta n_s,
\]

where \(j_s = j_{\uparrow} - j_{\downarrow}\) and \(\delta n_s = \delta n_{\uparrow} - \delta n_{\downarrow}\). The first term on the right hand side of Eq. (2) is zero in nonmagnets, and gives the current polarization in the bulk of the ferromagnet with \(\gamma^P = (\sigma_{\uparrow} - \sigma_{\downarrow})/(\sigma_{\uparrow} + \sigma_{\downarrow})\). The second term describes the net diffusion of spins with \(D_s = (D_{\uparrow}\sigma_{\uparrow} + D_{\downarrow}\sigma_{\downarrow})/(\sigma_{\uparrow} + \sigma_{\downarrow})\).

The system reaches the steady state shown in Fig. 1 because spin-flip scattering, which reduces the number of excess spins by coupling their angular momentum with the lattice, cuts off a limitless increase. The net spin accumulation \(\delta n_s\) decays with a characteristic spin flip scattering time \(\tau_{sf}\)

\[
\frac{d\delta n_s}{dt} = -\nabla \cdot j_s - \frac{\delta n_s}{\tau_{sf}}.
\]

The first term on the right hand side describes the net flow of spins into and out of a region and the second describes the reduction in the spin density due to spin flip scattering. In steady state, the left hand side vanishes so that Eq. (2) and Eq. (3) can be combined to give a diffusion equation for spins. The spin diffusion length \(l_{sf} = \sqrt{\tau_{sf}D_s}\) is the characteristic length scale in this diffusion equation. In Fig. 1, the spin accumulation in both materials decays exponentially away from the interface with the spin diffusion length appropriate for each material.

Spin flip scattering provides an important source of angular momentum for the current carrying electrons. As noted above, angular momentum does not flow in from the nonmagnet but does flow out into the ferromagnet. The angular momentum coupled from the lattice through spin-flip scattering is the source of the difference between the angular momentum flowing in the two leads.

Equations (2) and (3) describe all of the results in Fig. 1 except the discontinuity in the spin accumulation at the interface. This discontinuity arises because the interface has a spin-dependent resistance associated with it \([45, 46, 47, 48]\). For each spin, the current across the interface is related to a discontinuity in the spin chemical potential, described above, across the interface

\[
R_{\uparrow} \cdot j_{\uparrow} = \mu_{\uparrow}^{NM} - \mu_{\uparrow}^{FM}.
\]

As there is a “natural” polarization to the bulk current given by \(\gamma^P\), so is there a natural polarization to the spin current through the interface given by \(\gamma^I = (R_{\downarrow} - R_{\uparrow})/(R_{\uparrow} + R_{\uparrow})\). The apparent sign reversal comes
Figure 2: Spin accumulation ($\delta n_s$) and spin current ($j_s$) for electron flow through two ferromagnetic layers embedded in a non-magnetic host and separated by a thin non-magnetic spacer layer. The magnetizations of the two layers are parallel along the $z$ direction on the left side and antiparallel on the right. $\gamma^P$ is the polarization of the current in the bulk ferromagnet far from any interfaces.

from the definition in terms of resistances rather than conductances. Whenever the polarization of the current, $j_s/j$, in the bulk of a material is different from the polarization of the conductivity $\gamma^P$, there must be a spatially varying spin density to generate a diffusive contribution to the spin density to compensate. Similarly, whenever the polarization of the current crossing an interface differs from the polarization of the interface conductance, $\gamma^l$, there must be a discontinuity in the spin chemical potential across the interface.

The right hand panel of Fig. 1 shows that when the current direction changes, the spin accumulation and spin currents both change signs. In this case, more majority and fewer minority electrons enter the interface region than leave it. Hence, there is a positive spin accumulation and a net diffusion of majority electrons away from the interface and minority spins into it.

These ideas describing spin transport through and near interfaces provide the basis for understanding more complicated structures. For example, consider the case of two finite ferromagnetic layers separated by a non-magnetic spacer layer and embedded in a non magnetic host as depicted in Fig. 2. Far from the layers, there is no spin current or spin accumulation in the non-magnetic host. For parallel magnetizations, as depicted to the left, there are more majority electrons flowing through the layers than minority electrons. This flow of electrons leads to minority electrons accumulating to the left of the layers and diffusing to the left and majority electrons accumulating to the right and diffusing to the right. The result is a positive spin current in both cases. The polarization of the current is less than the polarization of the interface conductance, so there are substantial discontinuities in the spin accumulation across the interfaces.

On the other hand, when the two magnetizations are antiparallel, the current flowing through the layers is largely unpolarized, provided the layers are thin compared to the appropriate spin diffusion lengths. The unpolarized current flowing through the ferromagnetic layers requires substantial gradients in the spin accumulation to provide the diffusive spin currents to compensate the bulk-like spin currents. Also, since the current is much less polarized than the interface conductances, there are substantial discontinuities in the spin accumulation as well. This situation leads to an interesting reversal. In the case of parallel magnetizations, the current flowing through the spacer layer is highly polarized, but the spin accumulation is close to zero. On the other hand, for antiparallel alignment, the spin current in the spacer layer is close to zero but there is substantial spin accumulation.

Figure 2 provides an explanation [1] for giant magnetoresistance. When the magnetizations are parallel, the current carrying electrons are polarized so that more of them flow through the low resistance majority
channel. This polarization of the current lowers the average resistance of the structure and provides a short circuit as compared to the antiparallel case. In the antiparallel case the current remains largely unpolarized so that close to half of the electrons are forced to flow through the higher resistance minority channel in each layer.

It should be clear from Fig. 2 that the spin current will depend on the sample geometry, including the leads. In most samples, the leads only have the same cross sectional area as the rest of the sample for a length much shorter than the spin diffusion length. Then, the structure widens out quickly. In one dimensional drift-diffusion calculations, this widening is typically modelled by asserting that the spin accumulation goes to zero at the widening point, but the spin current is finite there. This approximation is motivated by the large density of states in the wide part of the sample. Three dimensional calculations of the transport [49, 50] bear out this picture qualitatively, but quantitative differences remain.

There are several key points from this discussion that are important for understanding spin transfer torques. Spin-dependent conductivities in ferromagnetic layers and spin dependent conductances at interfaces with ferromagnetic layers lead to spin-polarized current flowing through the non-magnetic layers in these structures. Spin-flip scattering couples angular momentum from the lattice into the electron system, allowing for apparent non-conservation of angular momentum within the electron system itself. Finally, the spin polarization of the current at particular points in the structure is not a local property. It depends on everything upstream and downstream within a few spin diffusion lengths. It also depends strongly on the alignment of the magnetizations in close-by layers.

3 Spin-dependent interfacial scattering.

The last section describes transport in devices with collinear magnetizations, but torques only occur when the magnetizations are non-collinear. Treating the non-collinear case requires understanding the behavior of spins currents at interfaces when the spins are not collinear with the magnetization of the ferromagnet. In this section, I first describe the scattering of individual electrons at interfaces as a function of their spin direction and then describe the behavior of the collection of electrons carrying current. These descriptions provide boundary conditions that can be used to generalize the results of the previous section to the case of non-collinear magnetizations. This generalization is taken up in the next section.

When the magnetizations are collinear it is possible to treat the spin accumulation as a scalar and the spin current as a vector. In the presence of non-collinear magnetizations, it is necessary to treat the spin accumulation as a vector to account for variations in the direction of the spins and the spin current as a tensor to account both the direction of the spins and the direction they are moving. Classically the spin current carried by an individual electron is the outer product of its spin and velocity \((\hbar/2)\mathbf{s} \otimes \mathbf{v}\) as illustrated in Fig. 3(a). More generally, these quantities are given by the expectation values of the spin operator \(\mathbf{S}\) and the velocity operator \(\mathbf{v}\):

\[
\mathbf{s}(\mathbf{r}) = \sum_{i\sigma \sigma'} \psi_{i\sigma}(\mathbf{r}) S_{\sigma,\sigma'}(r) \psi_{i\sigma'}(\mathbf{r})
\]

\[
\mathbf{Q}(\mathbf{r}) = \sum_{i\sigma} \text{Re} [\psi_{i\sigma}(\mathbf{r}) S_{i,\sigma,\sigma'}(r) \otimes \mathbf{v} \psi_{i\sigma'}(\mathbf{r})].
\]

(5)

In these sums, the index \(i\) refers to occupied states, and the index \(\sigma\) to the spinor components of the states. In general, the total spin current need not factor into an outer product of a spin direction and a vector current, but in systems like those considered in the previous section, it does

\[
\mathbf{Q} = (\hbar/2) \mathbf{s} \otimes \mathbf{j}, \quad (\hbar/2) \mathbf{P} \mathbf{s} \otimes \mathbf{j},
\]

(6)

7
Figure 3: (a) An electron moving in one direction with its spin in another illustrating a tensor spin current. (b) A pillbox around interface for computing the interfacial torque.

where \( \hat{s} \) is a unit vector in the direction of the spin density. The last form in Eq. (6) writes the difference of currents carried by each spin as a polarization \( P \) times the number current.

The equation of motion for the spin density can be derived in the same way as the familiar continuity equation for the number density

\[
\frac{\partial n}{\partial t} = -\nabla \cdot \mathbf{j}, \tag{7}
\]

which is derived from the commutator of the operator for the number density with the Hamiltonian. The number operator commutes with all terms except the kinetic energy, which gives the right hand side. The expression shows that the time rate of change of the number of electrons in a volume is given by the net rate of flow of electrons into the volume. For the spin density this procedure gives

\[
\frac{\partial \mathbf{s}}{\partial t} = -\nabla \cdot \mathbf{Q} + \mathbf{n}_{\text{ext}} - \alpha \hat{s} \times \frac{\partial \mathbf{s}}{\partial t} - \frac{\delta s}{\tau_{sf}}, \tag{8}
\]

where \( \nabla \cdot \mathbf{Q} = \partial_t Q_{ik} \), summing over the repeated index \( k \). The first term on the right hand side is similar to the right hand side of Eq. (7). It is a contribution to the time rate of change of the spin density given by the net flow of spin into that volume. Unlike the case for the number density, there are other terms that can change the spin density. The second term on the right hand side of Eq. (8) comes from the terms in the Hamiltonian in which the energy depends on the orientation of the spin density, terms like the Zeeman interaction with an external field, the magnetocrystalline anisotropy, and the magnetostatic interaction between the magnetizations at different points. These terms tend to cause precession of the spin density around its equilibrium direction and are combined into the effective torque. The last two terms are phenomenological representations of more complicated interactions. The first of these is the damping that tends to reduce the magnitude of the precession and the second is the spin flip scattering described in Eq. (3). With just the second and third terms on the right hand side, Eq. (8) is the familiar Landau-Lifshitz-Gilbert equation written in terms of the spin density instead of the magnetization. The last term is an addition for cases in which the magnitude of the spin density is allowed to vary due to flowing current. The first term is the gives the spin transfer torque of interest in this chapter.

There are two contributions to the spin transfer term in Eq. (8), as there are two contributions to the spin current. One contribution to the spin current is carried by all of the electrons and exists whenever
the spin density (or magnetization) varies spatially even if there is no current flow. The gradient of this contribution to the spin current gives the micromagnetic exchange interaction discussed in the chapter by Krommüller and Miltat. The second contribution is carried by non-equilibrium carriers near the Fermi energy. It only exists in the presence of a current and is the contribution of interest in this chapter. For the rest of the chapter, I ignore the contribution to the spin current due to the spatially varying magnetization and include its contribution to the torque in $\mathbf{n}_{\text{ext}}$ as is generally done in micromagnetics and only consider the spin current carried by the non-equilibrium electrons close to the Fermi energy.

The remainder of the $-\nabla \cdot \mathbf{Q}$ term in Eq. (8) is the spin transfer torque and has two important manifestations. The manifestation of principle interest in this chapter is an interfacial contribution in multilayers that is discussed below. There is an additional contribution for spatially varying spin densities in bulk materials originally studied by Berger [51, 52]. When the spin density direction varies spatially it exerts a torque on the spins carried along with the current as they move through the variation. These spins tend to rotate to stay aligned with the magnetization. If they stayed perfectly aligned, the spin current would have the form $(\hbar/2)P \hat{\mathbf{u}}(\mathbf{r}) \otimes \mathbf{j}$, where $\hat{\mathbf{u}}(\mathbf{r})$ is the local direction of the spin density. Then, the reaction torque on the spin density is $-\left(\hbar/2\right)P(\mathbf{j} \cdot \nabla)\hat{\mathbf{u}}(\mathbf{r})$ [53, 54, 55, 56] for a constant current. If the spins carried by the current do not adiabatically follow the direction of the magnetization, there can be an additional torque of the form $\xi(\hbar/2)P \hat{\mathbf{u}}(\mathbf{r}) \times (\mathbf{j} \cdot \nabla)\hat{\mathbf{u}}(\mathbf{r})$ [57, 58, 59]. However, calculations of this later quantity are still in a preliminary state.

The rest of this chapter focuses on the interfacial spin transfer torque of importance in magnetic multilayers [3, 60, 61, 62]. To understand the origin of this torque, it is useful to integrate the spin transfer torque $-\nabla \cdot \mathbf{Q}$ over the volume of a pillbox surrounding the interface between a non-magnetic metal and a ferromagnetic metal, as illustrated in Fig. 3(b). The volume integral gets converted into the difference in the flux between the two surfaces. For a spin current incident from the non-magnetic metal, the net flux can be written as a torque on the enclosed magnetization

$$
\mathbf{N}_e = (Q^\text{in} - Q^\text{tr} + Q^\text{ref}) \cdot \mathbf{A} \mathbf{x} \approx Q^\text{in}_{\perp} \cdot \mathbf{A} \mathbf{x}.
$$

(9)

where in, tr, and ref refer to the incident, transmitted and reflected spin currents respectively. The equality in this equation is the difference in the fluxes through the two interfaces and is an exact result. The approximate result, which states that the transverse component of the incident spin current is absorbed by the ferromagnet at the interface, will be developed in the rest of this section.

### 3.1 Torque due to individual electrons.

The behavior of individual electrons spin in magnetic multilayers is determined by their interaction with the spin polarized electronic structure of the ferromagnetic metals. Modelling the details of the spin transport requires a model for this electronic structure, which is complicated [63, 64, 65] because it is a balance of atomic-like intra-atomic exchange and correlation effects with interatomic hybridization. In isolated atoms, electrons occupy states in partially filled levels in such a way so as to maximum the total spin (Hund’s first rule). Maximizing the spin reduces the Coulomb repulsion between the electrons because parallel spins are naturally kept away from each other by the Pauli exclusion principle. Within partially filled atomic levels, the orbital energies of the different states are the same so there is no penalty for developing this polarization. In solids, the atomic levels hybridize with levels on neighboring atoms forming bands and removing the degeneracy of the states within a level. Since the degeneracy is lifted, there is now a cost for developing a spin polarization as each flipped spin has to be put into a higher energy state. Typically, the Coulomb energy gain from developing a spin polarization is less than the cost, so most solids remain
Figure 4: The band structures and Fermi surfaces of face-centered cubic Co. The red (blue) curves give the majority (minority) bands along two high symmetry directions through the Brillouin zone center, Γ. The dotted black curve shows the s-p band if it were not hybridized with the d bands. The bars to the right of the bands show the width of the d bands and the exchange splitting between the majority and minority bands. The red and blue arrows in the band structure plots give the width of the gap caused by the hybridization between the s-p and d bands of the same symmetry. The Fermi surfaces are by permission from Choy et al., [72].

unpolarized. However, in Fe, Co, and Ni, among others, the energy gain is sufficient for ferromagnetism to develop.

Two approaches are commonly used to describe the electronic structure of ferromagnets in the context of magnetic multilayers. One is based on models used to describe magnetic impurities in non-magnetic hosts [66]. In these models, the d-levels on the impurity are treated as localized states that form a local moment. This model ignores the hybridization of the levels forming the moment. In the most common treatment of this model, the local moment becomes a well defined object in its own right. The moment is then weakly coupled to the moment of the conduction electron spins. This model is referred to as the “s-d” model or the local moment model.

The other approach, which I will adopt in this chapter, is based on the local spin density approximation (LSDA) [67, 68, 69, 70]. This approach was developed for computing the total electronic energy of systems. In metallic solids it works very well for calculations of properties such as cohesive energies, equilibrium lattice constants, and the magnetic moments [71]. Alternatively, this method can be thought of as an approximation for the electronic structure in which the interatomic hybridization is treated exactly, but the interactions between the electrons are treated in mean field theory. I adopt this approach because it does a reasonable job describing the Fermi surfaces of the transition metals and the transport behavior is dominated by the properties of the electrons at the Fermi energy. Calculated band structures and Fermi surfaces for majority and minority electrons in face-centered cubic Co are shown in Fig. 4. The fact that the band width of the d-derived states as well as the s-d hybridization are larger than the exchange splitting highlight the importance of treating the hybridization of the d states.

Either approach can describe much of the physics of these systems, largely because so much is unknown about the details of the systems. For example, the spin-dependent conductivity would be difficult to calculate directly even if the details of all of the defects causing the scattering were known. Since these
details are not known, the conductivities are usually taken from experiment, allowing either model to correctly describe that aspect of the behavior. However, there are some processes where the details of the Fermi surfaces matter and the two models give different results. As discussed below, the spin transfer process is an example where this difference is important. Fig. 4 shows that when the hybridization of the $d$-electrons is taken into account, the Fermi surfaces of the majority and minority electrons are quite different. In local moment models, the two Fermi surfaces are almost identical.

Two different aspects of the exchange interaction are discussed in this chapter, the exchange interaction that gives rise to the moments and the micromagnetic exchange. The latter aspect is the additional energy that arises when the direction of the moment varies in space. In this chapter, the magnetization of the layers is generally assumed to be uniform, so references to the exchange interaction refer to the aspect giving rise to the moment. I use micromagnetic exchange to refer to the spatially varying aspect.

When electrons scatter from interfaces between a non-magnetic metal and a ferromagnetic metal, the exchange interaction in the ferromagnet leads to scattering that depends on the spin of the electron. For electrons with moments parallel to the magnetization in the ferromagnet, the reflection amplitude, $R_\uparrow$, is different from the reflection amplitude for electrons with moments antiparallel, $R_\downarrow$. For multilayers with collinear magnetizations, spin dependent reflection leads to a spin dependent interface resistance \cite{45, 46, 47, 48}. The transmission amplitudes, $T_\uparrow$ and $T_\downarrow$ are similarly spin dependent.

The reflection amplitudes for electrons with moments misaligned with respect to the magnetization can be computed directly from the reflection amplitudes for collinear moments. As discussed above, I assume that the magnetization is in the $-\hat{z}$-direction so that the ferromagnetic spin density is in the $\hat{z}$-direction. Then, the spin state of an electron spin pointing in the direction specified by the polar angle $\theta$ and azimuthal angle $\phi$ is given by a coherent superposition of majority and minority spin states

$$|\theta, \phi\rangle = \cos(\theta/2) e^{-i\phi/2} |\uparrow\rangle + \sin(\theta/2) e^{i\phi/2} |\downarrow\rangle. \quad (10)$$

So, an electron in the non-magnet with a spin pointing the the $\theta$, $\phi$ direction with wavevector $k$ can be described by the wave function $e^{ikx} |\theta, \phi\rangle$. Since quantum mechanics is linear, the majority and minority components of the wave function reflect as they would in the absence of any coherence. The reflected wave function is thus

$$e^{-ikx} \left[ R_\uparrow \cos(\theta/2) e^{-i\phi/2} |\uparrow\rangle + R_\downarrow \sin(\theta/2) e^{i\phi/2} |\downarrow\rangle \right]. \quad (11)$$

The reflected spin is rotated with respect to its incident direction and is pointing in the direction specified by $\tan(\theta'/2) = |R_\downarrow / R_\uparrow| \tan(\theta/2)$ and $\phi' = \phi + \text{Im} \ln(R_\uparrow R_\downarrow)$. Such rotations are sketched in Fig. 5.

Physically, the reflected electron spin precesses during its interaction with the exchange field present in the ferromagnet. Similar effects occur for transmitted electrons with one additional complication. When the electrons go through the ferromagnet they continue to interact with the exchange field and precess around it. This precession manifests itself through the different wave vectors for the two components of the wave function

$$e^{-ikx} T_\uparrow \cos(\theta/2) e^{-i\phi/2} |\uparrow\rangle + e^{-ikx} T_\downarrow \sin(\theta/2) e^{i\phi/2} |\downarrow\rangle \quad (12)$$

As they propagate, the different components accumulate additional relative phase from the factor $\exp[i(k_x - k_x)\Delta]$. This phase factor describes the precession around the magnetization as the spin propagates through the ferromagnet. In a local moment model, the spatial frequency of this precession is much slower than it is when more realistic Fermi surfaces are considered. Realistic Fermi surfaces give spatial precession periods on the order of several lattice constants.
Figure 5: Mechanisms contributing to absorption of incident transverse spin current. Electrons incident from the non-magnet (lower left) are distributed over a distribution of states represented by three different incident directions. All of these electrons are in the same spin state, which is transverse to the ferromagnetic spin density (blue arrow). The reflected electron spins have predominantly minority character and their transverse components are distributed over many directions (random spin rotation) because of the variation over the Fermi surface of the phases of the reflection amplitudes. The transmitted electron spins precess as they go into the ferromagnet because the wave vectors for the majority and minority components are different. Electrons with different initial conditions precess at different rates, leading to classical dephasing (differential precession).

When the interaction of the spin with the magnetization causes the spin to precess, there is a reaction torque from the spin on the magnetization. This reaction torque is the current induced torque described in Eq. 9. The second part of that equation expresses an approximate numerical result for the size of the torque. One of the mechanisms for this result is seen from the following simple argument based on a particular limit [60], in which the reflection probability for majority electrons is zero and for minority electrons is one. While this simple behavior is not general, it does occur for some electrons [73]. For a ferromagnetic spin density along $-\mathbf{z}$ and an interface normal in the $\mathbf{x}$ direction, consider an incident electron with its spin in the $\mathbf{x}$ direction. The spin currents on the faces of the pillbox of Fig. 3 as in Eq. (9) are

$$Q_{\text{in}} = \frac{\hbar}{2} \mathbf{\dot{x}} \otimes \mathbf{v} \mathbf{x}$$

$$Q_{\text{refl}} = \frac{\hbar}{4} (-\mathbf{\dot{z}}) \otimes (-\mathbf{v} \mathbf{x})$$

$$Q_{\text{trans}} = \frac{\hbar}{4} \mathbf{\dot{z}} \otimes \mathbf{v} \mathbf{x}. \quad (13)$$

This combination gives a current induced torque of $\mathbf{N}_c = \mathbf{\dot{x}} A \mathbf{v} \hbar / 2$. There are two features of this result. First, there is no torque along the direction of the magnetization. This is a general result and simply reflects the fact that the component of the spin along the magnetization is conserved during the scattering process. The second feature is that the component of the spin transverse to the magnetization is absorbed by the magnetization. The mechanism for the torque in this simple model is the spatial separation of the majority and minority components of the wave function. The transverse component of the spin current arises from the interference of these two components of the wave function. When they cease overlapping, the transverse...
component of the spin current vanishes. The two components of the wave function are separated by the exchange interaction, the ultimate cause for the spin dependence of the reflection amplitudes and hence the torque.

The result that the transverse spin current is absorbed at the interface is exact only for this special limit of complete spatial separation. However it holds approximately for many transition metal interfaces. Explaining this approximate result requires a discussion of the distribution of electrons that carry a current.

3.2 Torque due to an ensemble of electrons

Collinear transport is described in Sec. 2 using the language of the drift-diffusion approximation in terms of densities and currents, ignoring the behavior of individual electrons. Determining the boundary conditions at the interfaces between different materials requires summing over the behavior of the individual electrons. The electrons carrying the current are spread over the entire Fermi surface and have different properties.

Equation (11) shows that the electrons spins rotate when they reflect. First principles calculations [61, 62, 74] show that these rotations vary rapidly over the Fermi surface. A schematic of this distribution of rotations is shown in Fig. 5. For most interfaces that have been studied, there is a tendency for a polarization to develop along the direction of the magnetization; either majority or minority reflection is greater on average. However, in all cases, the transverse component of the reflected spin current sums to a value close to zero. Thus, \( Q_{\text{ref}} \) in Eq. (9) has only a small transverse component.

These same calculations show a similar, but not as complete cancellation of the transmitted transverse spin current. However, the transmitted electrons are rapidly precessing due to the evolving phase difference \( \exp[i(k_\uparrow - k_\downarrow)x] \) for the two spin components of the wave function for each electron. The difference in wave vector varies rapidly over the Fermi surface. Thus, there is an increase in the dephasing and concomitant decrease in the transverse component as a function of penetration into the ferromagnet. This precession is illustrated schematically in Fig. 5 and the net precession and decay of the spin current is illustrated in Fig. 6. If the spin current is considered several layers into the ferromagnet, the the transmitted current has only a small transverse component. The small transverse part of the transmitted and reflected spin currents means that the transverse part of the incident spin current is largely transferred to the ferromagnetic magnetization. This result is the approximate second half of Eq. (9).

Quantitative calculations like those in Fig. 6 require keeping track of the distributions of electrons on either side of the interface. A generalization of the Boltzmann equation is appropriate for this purpose. The Boltzmann equation describes the evolution of the distribution function for electrons when the spatial
variations in the system are slow enough that the electrons in a small region can be described as if they were in an infinite bulk region with similar material properties. Coherence between electrons at different wave vectors or positions is ignored. In magnetic multilayers this approximation can only be valid away from interfaces. However, distributions in two regions away from the interfaces can be joined together through boundary conditions calculated quantum mechanically.

In the ferromagnet, the distribution function $f_s(r, k)$, depends on position $r$, wave vector $k$, and band index (not indicated). The magnetization provides the natural quantization direction and there is a separate function for each spin $s = \uparrow, \downarrow$. Any spin component transverse to this direction tends to rapidly vanish due to dephasing effects similar to those discussed above for spin currents injected into a ferromagnet.

In the non-magnetic layers on the other hand, there is neither a natural quantization direction, nor the rapid dephasing that is present in ferromagnets. In this case, it is necessary to generalize the Boltzmann equation to keep track of the coherence between electrons in different spin states on each part of the Fermi surface. This can be done by generalizing the original distribution functions to $2 \times 2$ distribution matrices

$$f_{s, s'}(r, k, t) = f(r, k, t)\delta_{s, s'} + \sum_{\alpha=x, y, z} f_\alpha(r, k, t)[\sigma_\alpha]_{s, s'}$$

where $\sigma_\alpha$ are Pauli spin matrices ($\alpha = x, y, z$). This construction allows the electron spins to point in arbitrary directions. The four functions that characterize the distribution matrices give the densities and currents

$$n_0 + \delta n = \sum_k f(k) \quad j_\beta = \sum_k f(k)v_\beta(k)$$

$$s_\alpha = \frac{\hbar}{2} \sum_k f_\alpha(k) \quad Q_{\alpha\beta} = \frac{\hbar}{2} \sum_k f_\alpha(k)v_\beta(k)$$

where $v_\beta$ is a component of the velocity of the state at $k$, and $n_0$ is the equilibrium density. Moments of the Boltzmann equation give the drift-diffusion equations in the non-magnet as shown for the collinear case by by Valet and Fert [42].

The construction in Eq. (14) is based on including the coherence between majority and minority spin states, but only for states at the same point on the Fermi surface. Thus, the possibility of such a construction in ferromagnets depends on the model for the electronic structure of the ferromagnet. Calculations based on local moment models [75, 76, 77, 78, 79], typically include coherence between majority and minority states because there is little difference between the Fermi surfaces of the majority and minority transport electrons. On the other hand, for realistic models of the electronic structure, such a construction is not possible because states with the same energy do not generally have the same wave vector. Thus Boltzmann equation and drift-diffusion calculations based on realistic band structures do not allow for transverse spins in ferromagnets. It is possible to extend the Boltzmann equation so as to include the coherence that is left out of the Boltzmann equation; this is done in the first principles calculations [62, 74]. However, it is difficult to include scattering in calculations with coherence between non-overlapping Fermi surfaces.

In the Boltzmann equation, the boundary conditions for each electronic state are determined by the spin-dependent reflection and reflection amplitudes shown in Eqs. (11-12). The boundary conditions for reflection can be expressed in terms of four real reflection parameters

$$R_\uparrow = |R_\uparrow|^2 \quad R_\downarrow = \text{Re}[R_\downarrow R_\uparrow^*]$$

$$R_\downarrow = |R_\downarrow|^2 \quad R_\times = \text{Im}[R_\downarrow R_\uparrow^*]$$

(16)
The two parameters on the left are the reflection probabilities for the majority and minority components respectively. They determine the reflected number current and the reflected longitudinal spin current. The two quantities on the right, which are not probabilities as they can be negative, determine the reflected transverse spin current. $R_\perp$ describes reflection of a spin along the same azimuthal axis as the incident state. A negative value implies rotation by $180^\circ$. $R_x$ describes reflection along an axis rotated by $\pm 90^\circ$. The boundary conditions for transmission are described by the transmission probabilities for majority and minority electrons. Since these models do not include transverse spins in the ferromagnet, any such spins that are transmitted are assumed to be absorbed by the magnetization at the interface. Similarly, the incident electrons from the ferromagnet are assumed to be collinear with the magnetization, so those process are described the majority and minority transmission and reflection probabilities alone.

The transformation from the Boltzmann equation to the drift-diffusion equation requires appropriate averages over these reflection parameters. The result that the incident transverse spin current is absorbed at the interface is equivalent to stating that $R_\perp$ and $R_x$ average to zero. Then, the boundary conditions for transport in multilayers with non-collinear magnetizations are that the transverse spin current and accumulation in the ferromagnet are both zero and the transverse spin accumulation and current in the non-magnet are proportional to each other

$$Q_\perp \cdot \mathbf{n} = B s_\perp. \quad (17)$$

The $\perp$ subscript indicates the transverse component. The constant of proportionality is a constant with units of velocity $B = A_{FS}/(8\pi^3 h N)$, where $A_{FS}$ is the projected area of the Fermi surface in the non-magnetic material.

The spin transfer torque described here is due to the spin current carried by non-equilibrium carriers at the Fermi energy. While the spin transfer torque is interfacial, it is typically treated as spread uniformly throughout the free layer because the free layer is thin enough that the micromagnetic exchange interaction tends to keep the magnetization aligned. It is interesting to see how this comes about in detail. The interfacial spin transfer torque needs to be balanced by another interfacial torque so as to prevent a diverging response. This balancing torque comes from the micromagnetic exchange. Normally, the micromagnetic exchange has the boundary condition that the normal gradient of the magnetization vanishes ($\mathbf{n} \cdot \nabla \mathbf{s}(r) = 0$). This condition is simply the requirement that the interfacial torque vanish. In the presence of an interfacial spin transfer torque, the gradient is non-zero at the interface. Since the magnetization is differentiable, a finite derivative at the interface gives a finite derivative into the layer. This rotation of the magnetization effectively “spreads” the interfacial spin transfer torque over the thin layer. As mentioned after Eq. (8), the micromagnetic exchange torque can be understood as the gradient of a spin current carried by all of the electrons in the Fermi sea. A finite normal gradient corresponds to a discontinuous spin current. At the interface, both the transport spin current and the micromagnetic spin current are discontinuous. In fact, the total spin current is continuous. At the interface, the incident spin current carried by the non-equilibrium carriers at the Fermi energy is converted into a “quasi-equilibrium” spin current carried by all of the electrons. This conversion is what effectively spreads the interfacial spin transfer torque over the thickness of the film.

4 Non-Collinear Transport and Torque

With the boundary condition derived in the previous section, it is now possible to compute the transport shown in Fig. 2 for multilayers with non-collinear magnetizations. Such calculations have been undertaken
with a variety of different approaches, including the Keldysh formalism [80], the Boltzmann equation 
[81, 78], the drift-diffusion approximation [82, 83, 84, 85, 86], random matrix theory [60], and circuit 
theory [87, 88]. While the calculations differ in detail, they all find qualitatively similar results. Here we 
describe the physics in the language of the drift-diffusion approach.

A calculation of the spin current and spin accumulation in a magnetic multilayer with perpendicular 
magnetizations is shown in Fig. 7. The behavior is closely related to that shown in Fig. 2. Consider the 
component of spin along \( \hat{z} \), aligned with the spin density in the left layer. Far to the left of the sample, 
the spin current is unpolarized, but it becomes polarized due to the spin dependent conductivity of the 
ferromagnetic layer as described in Sec. 2. More majority electrons go through the left layer leading to 
backward diffusion of minority spins in the left lead and a positive spin current. This spin current continues 
unchanged through the thin spacer layer but goes to zero at the left interface of the right ferromagnet. At 
this interface, the spins are transverse to the spin density in that layer and the transverse spin current is 
absorbed as described in Sec. 3. The abrupt change in the spin current gives a current induced torque on 
the right ferromagnet at its left interface.

When the component of the spin along \( \hat{z} \) goes to zero at this interface, components along \( \hat{x} \) are generated 
by the spin-dependent reflection. The reflected electrons are predominantly antiparallel to the spin density 
in the right ferromagnetic layer and move away from the interface toward the left ferromagnet. When 
they hit the right interface of the left ferromagnet, they are again transverse to the spin density and that 
component of the spin density gets absorbed at that interface. Again, this discontinuity in the spin current 
give rise to a current induced torque.

An interesting result of this calculation is that the torques tend to make the two magnetizations 
pinwheel after each other, apparently violating conservation of angular momentum. This apparent violation 
is resolved by the role of spin flip scattering, which couples angular momentum from the lattice into the 
electron system, which is then absorbed by the magnetization. Another interesting feature of this figure is 
that in the spacer layer, the spin current direction is not aligned with the spin accumulation, but rather is 
almost perpendicular. Such a result is not entirely unexpected because the spin current is related to the 
gradient of the spin accumulation rather than the accumulation itself.

The calculation shown in Fig. 7 gives the torque and the resistance of the multilayer for a particular 
magnetic configuration. Similar calculations give the torque and magnetoresistance as a function of angle. 
Related calculations done by a number of authors [89, 90, 81, 91, 92, 93, 94, 95, 96] give similar results. If 
the spin current incident on the free layer were independent of the relative orientation of the magnetizations 
of the two layers and given by \((\hbar/2)\mathcal{P} \hat{S}_0 \otimes \hat{j}\), then the torque would be

\[
\frac{N_c}{A} = \frac{\hbar}{2} (\hat{J} \cdot \hat{x}) \mathcal{P} \left[ \hat{S}_0 - (\hat{S}_0 \cdot \hat{s}) \hat{s} \right].
\]  

(18)

where \( \hat{s} \) is the directions of the ferromagnetic spin density for the free layer. The last factor is simply 
the transverse part of the spin current. It is more frequently written in terms of the equivalent triple 
product form, \( \hat{s} \times (\hat{S}_0 \times \hat{s}) \). The magnitude of this vector is \( \sin \theta \), where \( \theta \) is the relative angle between 
the magnetization \( \hat{s} \) and the direction of the spin current \( \hat{S}_0 \). This form of the torque is frequently used in 
simulations of the magnetic dynamics.

However, calculations show that the spin current does depend on the relative angle of the magnetizations. 
However, the changes to Eq. (18) are simple. The direction of the torque remains the same if \( \hat{S}_0 \) is 
taken to be the direction of the fixed layer magnetization. This replacement works because the direction of 
spin current is generally some combination of \( \hat{s} \) and \( \hat{S}_0 \). Since \( \hat{s} \) does not have a component that is 
transverse to itself, only the contribution along \( \hat{S}_0 \) remains. After that replacement, the rest of the changes
Figure 7: Spin accumulation ($\delta s_\alpha$) and spin current ($Q_{\alpha x}$) for electron flow through two ferromagnetic layers embedded in a non-magnetic host and separated by a thin non-magnetic spacer layer. The magnetizations of the two layers are perpendicular to each other, for purposes of illustration both are in the plane of the figure. On the left side the current flows in the $\hat{x}$ direction and on the right in the $-\hat{x}$ direction. $Q_b$ is the magnitude of the spin current in the bulk ferromagnet far from any interfaces. The transverse component of the spin current is discontinuous at each interior interface giving rise to torques on the magnetizations of each layer. The directions of the torques are indicated for each of the interfaces. The bottom panel gives a cartoon of the spin current in the spacer layer. The blue arrows give the direction of the electron spins for electrons moving in the directions given by the black arrows (recall that the spin current for an electron with a spin in one direction moving to the left is the same as the spin current from an oppositely directed spin moving to the right). The purple arrows repeat the torques from above.
in the behavior found in these calculations are captured simply by letting the polarization $\mathcal{P}$ be a function of the relative angle $\theta$ between the two magnetizations.

For multilayers with equivalent magnetic layers, Slonczewski [90] derived an analytic formula for $\mathcal{P}$ by combining a density matrix description of the spacer layer with a circuit theory [87]. Extensions of the formula for arbitrary magnetic layers

$$\mathcal{P}(\theta) = \left[ \frac{q^+}{B_0 + B_1 \cos \theta} + \frac{q^-}{B_0 - B_1 \cos \theta} \right], \quad (19)$$

have been described by a number of authors [92, 96, 84, 94]. The four parameters ($B_0$, $B_1$, $q^+$, and $q^-$) in this expression can be written in terms of layer thicknesses, resistivities, spin diffusion lengths, etc. The ratio $q^-/q^+$ depends on the asymmetry between the free layer and the fixed layer and goes to zero for symmetric systems [90]. The ratio $B_1/B_0$ goes to zero when the polarization of the current is independent of the relative angle between the magnetizations. Calculations show that this behavior only occurs in unphysical limits. Xiao et al. [96] have compared solutions of a Boltzmann equation with Eq. (19) and found very good agreement.

It is difficult to directly compare the form of the torque in Eq. (19) with experiment because the torque is not directly measurable in the experiment; the only experimental observable is the time-dependent resistance. It is possible to use Eq. (19) to predict the dynamics, and then to predict the time dependent resistance to compare with experiment, such simulations are described briefly in Sec. 5 and in more detail in the chapter by Thiaville. However, such a comparison requires correctly knowing and treating all of the other factors that contribute to the dynamics. More direct comparisons can be made between the measured and calculated resistances as a function of the relative angle between the magnetizations [97, 98]. The measurements show clear deviation from simple $\sin^2(\theta/2)$ behavior. However, the deviations tend to be smaller than those calculated [94] with the models that give results like Eq. (19).

The torque in Eqs. (18) and (19) is zero for parallel and antiparallel alignments of the magnetization. The existence of transitions out of these configurations is then due to instabilities developing in the configurations as the current is varied, see [99, 100]. For appropriate directions of current flow, the spin-transfer torque tends to amplify small deviations away from the collinear configuration and damping tends to diminish them. For large enough currents, the spin-transfer torque dominates and the system becomes unstable.
A different origin of the instability has been proposed by Berger and studied by a number of authors [4, 101, 102, 103, 104, 75, 105]. Here the longitudinal spin accumulation in the ferromagnet drives the instability rather than the transverse spin current. As the current through the structure increases, the spin accumulation eventually becomes large enough that it is possible to flip a spin from minority to majority and excite a magnon. This process is most clearly seen when the spin accumulation is described in terms of a spin-dependent chemical potential. In Berger’s model, a particular magnon mode, possibly the uniform mode, gets macroscopically populated through stimulated emission. This instability then gives rise to reversal with a particular critical current density. However, Tserkovnyak et al. [106] have shown that a model without direct magnon excitation can give the same critical current density.

When the magnetization varies in space, there are additional considerations. If the magnetization varies along the direction of the current flow, there are torques on the magnetization as the spins in the current precess around the varying magnetizations, as mentioned in Sec. 3. If the magnetization varies in the plane of the interface, lateral diffusion of spins can tend to increase or decrease these variations depending on the direction of the current flow. If the net electron flow is from a non-magnet into a ferromagnet, the spins that accumulate in the non-magnet are, on average, anti-parallel to the spin density in the ferromagnet. If the spin density is not uniform, locally, the spins tend to align opposite to the local spin density. These electrons are diffusing, so there is a good chance that some of them will diffuse laterally and scatter again from the interface at some different point. If the direction of the spin density at the point is different, the electron will, on average, exert a torque on the magnetization at that point. When the electron flow from is from the non-magnet into the ferromagnet, more minority spins are diffusing, and the torque tends to amplify the variations away from the average, see Fig. 9. On the other hand, for electron flow from the ferromagnet into the non-magnet, more of the diffusing spins are aligned with the spin density, and the resulting torque tends to reduce variations away from the average.

These torques due to lateral diffusion have been discussed [107, 85] in the context of instabilities in single ferromagnetic films and [108] in the context of instabilities in the context of trilayers. For single films, these calculations show that for the current densities studied experimentally in these systems, non-uniform modes can become unstable. Instabilities have been observed in point contact experiments with single films [6, 109] and also in lithographically defined single films [110]. The critical currents are close to those found in the calculations and the field dependence of the critical currents is also similar. However, there are still details of the experimental results that are not explained by the calculations.

Finally, before I discuss the dynamics that result from the spin transfer torque, there is a correction to the boundary condition, Eq. (17) when the magnetization is time dependent

\[ \mathbf{Q}_\perp \cdot \mathbf{n} = B s_\perp + B' \mathbf{u} \times \dot{\mathbf{u}}, \]  

(20)

where \( \mathbf{u} \) is the direction of the magnetization and \( \dot{\mathbf{u}} \) is its time derivative. This additional term describes the phenomenon called spin pumping, an effect originally proposed by Berger [4, 111], developed by Tserkovnyak et al. [112, 113] and computed in a number of different models [112, 114, 115, 116]. The two terms in Eq. (20) are closely related. The first term describes the spin current discontinuity that gives rise to a torque and causes the magnetization to move and the second is the spin current that results from a moving magnetization. Both phenomena are controlled by related parameters

\[ B' = \frac{A_{FS}}{2\pi r^3} \frac{h}{2} \frac{\partial m}{\partial \mu} B \]  

(21)

where \( A_{FS} \) is projected area of the Fermi surface in the non-metal.
Figure 9: Spin transfer torque due to lateral diffusion. Each panel shows an electron diffusing in a non-magnet and reflecting twice from the interface with a ferromagnet. After it scatters, it is oriented on average either parallel or antiparallel to the magnetization depending on the direction of the electron flow, see Fig. 1. In the first panel, the ferromagnetic magnetization is uniform laterally so that when the electron scatters the second time, it is aligned with the magnetization and there is no reorientation of either. The electron flow is from the non-magnet into the ferromagnet, so the accumulated spins are minority spins. In the second panel, there is a non-uniform magnetization, and the diffusing minority spin is not aligned with the magnetization the second time it scatters. The magnetization exerts a torque on it and it exerts a torque on the magnetization which tends to amplify the spin wave. In the third panel, the electron flow is in the opposite direction so the diffusing spins are majority spins. In this case, the torque in the second scattering event tends to reduce the amplitude of the non-uniform spin wave.

Equation (20) shows that even in the absence of a current, a moving magnetization drives a spin current in the non-magnetic material. When this spin current is absorbed through spin flip scattering or scattering from another ferromagnetic layer, the spin pumping process acts like a form of interfacial damping. These effects have been observed in linewidth measurements in ferromagnetic resonance [117, 118, 119, 120, 121, 122].

5 Discussion

This chapter describes the physics that leads to spin-transfer torques in magnetic multilayers. When the magnetizations of the layers are not collinear with each other, the spin polarized currents transfer angular momentum to the magnetizations near the interfaces, giving rise to spin-transfer torques. These torques can be computed using a combination of quantum mechanical calculations of the behavior of spins at interfaces and semi-classical transport calculations to describe the flow of spins in the multilayers.

Quantum mechanical calculations of spins at interfaces show that a combination of spatial separation of spin components coupled with classical dephasing leads to the approximate absorption of the transverse spin current at the interfaces. This result means that independent of their orientation when approaching the interface, the spins are on average collinear with the magnetization when they leave the interface. In effect, any component of the incoming spins that is transverse to the magnetization is transferred to the magnetization. The angular momentum along the magnetization direction is conserved.

The results of the quantum mechanical calculations are inserted into semiclassical transport calculations as boundary conditions. These calculations determine the degree of the spin polarization at the interfaces. They show that currents become spin polarized through the combination of spin-dependent conductivities and interface conductances together with spin-flip scattering. Spin-flip scattering couples
angular momentum between the electron system and the lattice. The spin-transfer torque couples angular
momentum between the electron system and the magnetization.

The transport calculations are done for fixed magnetizations, even when the magnetizations are time
dependent. This approximation is justified by a separation of time scales. A typical precession period for
the magnetization is about $10^{-9}$ s and the transit time for a ballistic Fermi surface electron to go through
the device (about 100 nm) is about $10^{-13}$ s. Even though the transport is not ballistic, this difference in
time scales argues that treating the electrons as moving through a fixed magnetization should be a good
first approximation.

The result of the semiclassical transport calculations is an expression for the spin transfer torque as
a function of the angle between the two magnetizations, Eqs. (18) and (19). These equations summarize
the results of the calculations described in this chapter. It is difficult to directly compare this result
with experiment. Experimentally, it is the consequences of the torque that are readily measured, for
example the critical currents for switching between parallel and antiparallel configurations. To compare
with such experiments, the calculated torque is inserted into the equation of motion for the magnetization,
the Landau-Lifshitz-Gilbert equation, and used to compute the dynamics, as described in the chapter by
Thiaville. There are several approaches to such computations.

The simplest approach is to assume that the magnetization of each layers is separately uniform. The
model based on this assumption is called the macrospin model. The simplicity of the resulting equation
of motion lends itself to analytic determination of critical currents $[123, 100, 124]$ and detailed exploration
of behavior as a function of field and current $[11, 34, 35, 125]$. However, the assumption of a uniform
magnetization is not always justified. Relaxing this approximation gives a full micromagnetic model.
Such calculations $[126, 127, 128, 16, 129, 130, 131, 132, 133]$ treat the detailed spatial variation of the
magnetization, but the results are sensitive to many of the details of the samples. Since many of these
details, like grain sizes, local anisotropies, the detailed shape, etc. are not measured, it is hard to make
definitive comparisons between these calculations and measurements.

A final complication of the comparison between theory and experiment is that experiments are typically
done at finite temperature and calculations are most simply done at zero temperature. Under the influence
of temperature, transitions between different configurations become statistical rather than deterministic.
Several different approaches have been used to analyze the statistics of the switching events $[29, 35, 134,$
$135]$. The most straightforward approach to including thermal effects is to include a fluctuating field
in macrospin simulations $[136, 125]$. Such simulations give a reasonable good account of the effects of
temperature allowing for the difference in the time scales of the experiment and those achievable in the
simulations.

Comparisons between theory and experiment generally give qualitative, but not quantitative agreement.
One of the most straightforward predictions of the torque in Eqs. (18) and (19) when inserted into the
macrospin model is that the critical currents for the transitions between parallel and antiparallel alignment
are determined by the slopes of the torque as a function of angle for $0^\circ$ and $180^\circ$ relative angles. Since
transport calculations generally give a very different polarization of the current for the two alignments,
the critical current for the parallel to antiparallel transition is expected to be larger in magnitude than
that for the reverse transition. While there is wide variation seen experimentally, the trend is that this
difference is smaller than what is expected from theory. It is not clear whether this disagreement points
to errors in the calculation of the torque or deviations of the reversal process from that expected for a
macrospin. Other disagreements between macrospin calculations and experiment appear to explainable by
micromagnetic calculations. These are discussed in more detail in other chapters in the volume.
References


[97] Pratt, W. P., Jr. (private communication).


