Magnetization Dynamics Including Thermal Fluctuations: Basic Phenomenology, Fast Remagnetization Processes and Transitions Over High-energy Barriers

Dmitri V. Berkov
Innovent Technology Development, Jena, Germany

1 DETERMINISTIC LLG EQUATION AND TREATMENT OF THE ENERGY DISSIPATION IN MICROMAGNETICS

1.1 Origin and limitations of the standard LLG equation

An equation of motion for the magnetization of a ferromagnet, known as the Landau–Lifshitz equation

Handbook of Magnetism and Advanced Magnetic Materials. Edited by Helmut Kronmüller and Stuart Parkin. Volume 2: *Micromagnetism*. ISBN: 978-0-470-02217-7.

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = -\gamma_0 \cdot [\mathbf{M} \times \mathbf{H}^{\mathrm{eff}}] - \gamma_0 \cdot \frac{\lambda}{M_{\mathrm{S}}} \cdot [\mathbf{M} \times [\mathbf{M} \times \mathbf{H}^{\mathrm{eff}}]] \quad (1)$$

has been introduced 1935 by L.D. Landau and E.M. Lifshitz in their pioneering work (Landau and Lifshitz, 1935) devoted to the phenomenological evaluation of the permeability tensor of ferromagnets.

The first term in (1) describes the magnetization precession of the total effective field and can be derived in frames of a general phenomenological theory which is based on the assumption that for low temperatures and slowly (in space in time) varying magnetization the magnitude of the magnetization vector $\mathbf{M} = M_{\mathrm{S}}$ is conserved. The latter statement, in turn, follows from the assumption that the equilibrium value of M_{S} is fixed by the exchange interaction, which is assumed to be the strongest interaction in a ferromagnet, which is definitely the case for all 'normal' ferromagnetic materials. Comparison of the equation (1) with the precession equation for a 'free' magnetic moment in the small damping limit provides the value $\gamma_0 = g|e|/2m_{\mathrm{e}}c$ which is the so-called gyromagnetic ratio. The g-factor in this definition, being $g = 2$ for a free electron, may slightly vary around this value dependent on the concrete material. Usually at least the combination $\gamma_0 M_{\mathrm{S}}$ may be measured with a high accuracy using FMR, so that the treatment of the first term in (1) is quite straightforward.

In contrast to the precession term, the handling of the energy dissipation processes in ferromagnets turned out to be a highly complicated issue. The double vector product term in (1) was chosen in the original paper (Landau

and Lifshitz, 1935) basing on the purely phenomenological reason that energy dissipation processes (i) should drive the magnetization toward the effective field direction (in the minimal energy state the magnetization is directed along $\mathbf{H}^{\text{eff}}$), but (ii) the magnetization magnitude M_S should still remain constant. The second term in (1) obviously satisfies these both conditions, being directed toward the effective field and perpendicular to the magnetization.

It was pointed already by the authors of (Landau and Lifshitz, 1935) themselves that the damping form suggested by (1) is by neither the only possible nor the most general one. First of all, it can be seen by a sole inspection of this equation that *if* we understand the coefficient before the damping term λ in a usual way, that is, as a parameter whose value is proportional to the intensity of the energy dissipation processes in our system, then (1) cannot be used to describe the magnetization motion for moderate and large damping. The reason for this limitation can be seen immediately if we consider the overdamped regime ($\lambda \geq 1$) where the magnetic moment motion is dominated by the dissipation term. In this case, the basic equation (1) predicts that the magnetization relaxation is getting *faster* when the damping value *increases* which is in a strong contradiction with a physical picture of damping.

This circumstance was realized already by Landau and Lifshitz themselves, who have pointed out that their equation may be used in the precession-dominated regime ($\lambda \ll 1$) only. The first phenomenological equation which qualitatively, reasonably describes the magnetization motion in the whole dissipation range was suggested by Gilbert (1955):

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = -\gamma_0 \left[\mathbf{M} \times \left(\mathbf{H}^{\text{eff}} - \frac{\alpha}{M_S} \cdot \frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} \right) \right] \tag{2}$$

Gilbert has derived this equation starting with the equation of the undamped magnetization precession rewritten in the Lagrangian formalism. In this formalism, the damping can be rigorously added to the system using the so-called Rayleigh dissipation function (Landau Lifshitz, 1981). Transforming the resulting equation back to the force-torque form, Gilbert arrived at the equation (2), where the damping is represented by the second term in the round parenthesis with α being the damping constant.

The dissipation term form introduced in the Gilbert formulation of the magnetization dynamics can also be understood in the following way. First, the energy damping is supposed to slow down the precession of the magnetic moment, so the term describing the energy dissipation can be added directly to the effective field which is responsible for the moment precession and defines its frequency via $\omega = \gamma_0 H^{\text{eff}}$. Second, the magnitude of this term should be proportional to the relaxation speed, analogously to a standard (e.g., hydrodynamical) viscous damping. And finally, the added damping term should still conserve the magnetization magnitude. All these three conditions are satisfied by the dissipation term (2), which is called the *Gilbert* damping. It is straightforward to show (Kikuchi, 1956) that this formulation leads to the intuitively expected dependence of the switching time (relaxation speed) on the damping α, where the switching time is large both for small damping (precession-dominated regime, the moment performs many precession cycles before switching) and large damping (magnetic moment does not precess, but moves slowly owing to a large α value).

The Gilbert equation in its native form (2) is highly inconvenient to use, because it contains the time derivative of the magnetization on both sides. Fortunately, it can be easily cast into an explicit form ($\mathrm{d}\mathbf{M}/\mathrm{d}t$ on the left-hand side only) by multiplying both sides by the vector $\mathbf{M}$, transforming the double vector products using the standard vector algebra rule and utilizing the conservation of the magnetic moment magnitude as $(\mathbf{M}\cdot\mathbf{M}) = M_S^2$. The final result

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = -\frac{\gamma_0}{1+\alpha^2} \cdot [\mathbf{M} \times \mathbf{H}^{\text{eff}}] - \frac{\gamma_0}{M_S} \cdot \frac{\alpha}{1+\alpha^2} \cdot [\mathbf{M} \times [\mathbf{M} \times \mathbf{H}^{\text{eff}}]] \tag{3}$$

has the form of the Landau–Lifshitz equation (1) if we replace the gyromagnetic ratio γ_0 by $\gamma_0/(1+\alpha^2)$ and put $\alpha = \lambda$. It is also evident that in the small damping limit both equations coincide if $\alpha = \lambda$, as it should be.

Summarizing, we point out that the two most widely used equations for the description of the damped magnetization motion are formally equivalent. However, the physics involved in the definition of the damping coefficient is quite different: Whereas in the Gilbert form (2) the value of the damping parameter α can be assumed to be proportional to the intensity of the energy dissipation processes which slow down the magnetic moment motion, for the Landau–Lifshitz form it is true only in the small damping limit. Although this issue has been discussed in the past several times (see in addition to the original papers of Landau *et al*. and Gilbert, e.g., also comments in Kikuchi (1956) and Mallinson (1987)), everybody who is familiar with the contemporary literature will surely agree that recalling them once more will at least not harm. The difference mentioned in the preceding text tends to be forgotten simply because for the overwhelming majority of applications (including, but limited to the magnetic random access memory (MRAM) technology and high density storage media) the case of the small damping is of major interest. However, this circumstance detracts no way from the importance of the physical meaning of the parameters entering into these equations.

We conclude this subsection with the following important remark. As soon as we have introduced the restriction $\mathbf{M} = M_S$, the magnetization is allowed only to rotate so that *any* magnetization change $\Delta\mathbf{M}$ should be perpendicular to $\mathbf{M}$ itself. Hence, one can gain the impression, that the equation (1) (or its equivalent forms) is the most general form of the equation of motion in this situation: from a purely mathematical point of view (1) can be considered as an expansion of the magnetization time derivative (which should lie in the plane perpendicular to the magnetization) over the two vectors $[\mathbf{M} \times \mathbf{H}]$ and $[\mathbf{M} \times [\mathbf{M} \times \mathbf{H}]]$ which form an orthogonal basis in this plane. Coefficients γ and λ can then be viewed simply as the expansion coefficients of an arbitrary allowed magnetization change in this basis. However, from the physical point of view the equation (1) and its analogues contain a much stronger assumption than the trivial statement about the existence of such an expansion. Namely, the form (1) means that the expansion coefficients are *time-independent scalar* variables, so that the magnetization motion including damping can be described on a phenomenological level with two *scalar* parameters whose values do *not* depend on the instantaneous magnetization configuration. As we shall see in the subsequent text, this is not necessarily true. However, in many cases the combined Landau–Lifshitz–Gilbert (LLG) equation (3) provides a convenient and at least semiquantitatively adequate description of the damped magnetization precession.

Concluding this subsection, we would like to briefly mention possible extensions of the standard LLG-equation (3) to the case, when the magnetization magnitude is *not* conserved, which may be particularly interesting for temperatures not very far from the Curie point. One of the possibilities to account for a change of the magnetization *value* on the mesoscopic level is the insertion into the LLG equation of the *longitudinal* relaxation term. This question was discussed already by Aharoni in the 1980s. For the recent development, including a more detailed discussion of possible forms of such an additional term, other physical problems concerning this topic (e.g., calculation of the corresponding relaxation time) and some simulation results obtained with such a generalized LLG equation, we refer the reader to the publications (Garanin, 1997; Smith, 2002; Grinstein and Koch, 2003; Garanin and Chubykalo-Fesenko, 2004).

1.2 To the possibility of alternative forms of the damping term

Rigorous *evaluation* of the damping constant entering the LLG equation is one of the most complicated problems of the modern solid-state magnetism not only due to a large variety of the damping mechanisms in ferromagnets, but also because in most cases it is really difficult to separate the contributions of these mechanisms to the energy dissipation rate measured experimentally (e.g., using FMR). To the mechanisms mentioned in the preceding text belong the spin-lattice relaxation (magnon-phonon interaction), two-and many-magnon processes, magnon-impurity interactions, magnon scattering on the surface and interface defects and so on, we refer the interested reader to the contributions in the Volume 1 of this Handbook.

For this reason, aphenomenological approach which could provide nontrivial conclusions about the form of the damping term and supply methods for the evaluation of the damping value beyond the simplest LLG phenomenology, but still valid for a relatively broad class of the energy dissipation processes is in principle highly desirable. However, a development of such an approach is a delicate matter, requiring careful analysis of each step of a corresponding 'general' procedure.

As an example of a such an attempt we would like to analyze a recently developed and meanwhile widely cited approach of Safonov and Bertram (2003) and references therein, who suggested to use the *normal mode analysis* of a magnetic system coupled to a thermal bath. Up to a certain level such an analysis can be performed without specifying the concrete energy dissipation mechanism, but merely employing general assumptions concerning the Hamilton function, which describes the interaction between the magnetization and thermal reservoir.

In the simplest case of a uniformly magnetized ferromagnetic particle, the normal mode analysis starts with the Taylor expansion of the magnetic energy density E/V over small magnetization variations M_x and M_y

$$\frac{E}{V} = \frac{H_{0x}}{2M_S} M_x^2 + \frac{H_{0y}}{2M_S} M_y^2 \tag{4}$$

near the equilibrium moment orientation $\mathbf{M}_0$, whereby the $0z$-axis is directed along $\mathbf{M}_0$. Dimensionless coefficients before the squares of magnetization components $M_{x(y)}$ in (4) can be written in the form $H_{0x(y)}/M_S$, where parameters $H_{0x(y)}$ (which describe the curvature of the energy surface for the equilibrium magnetization state) have the dimension of a field [1].

Any equation of motion which describes the magnetization precession in terms of components M_x and M_y leads to the coupling of their time-dependencies and is thus unsuitable for the identification of the system normal modes. For this reason, Safonov and Bertram proceed by introducing the (complex-valued) functions $a(t) \sim M_x + iM_y$ and $a^*(t) \sim M_x - iM_y$. These functions are fully

analogous to the corresponding quantum-mechanical operators which increase/reduce the z projection of the angular momentum, but the formalism developed by Safonov and Bertram is a purely classical one. The functions $a(t)$ are required to describe the interaction of a magnetic system with a thermal bath, because they appear in the corresponding interaction term of a Hamiltonian; this is the reason why this intermediate transformation should be explicitly introduced. The second transformation $a(t) \sim uc(t) + vc^*(t)$ and $a^*(t) \sim uc^*(t) + vc(t)$ diagonalizes the Hamilton function, which becomes H: $\omega_0 c^* c$, with $\omega_0 = \gamma\sqrt{H_{0x}H_{0y}}$ being the FMR frequency of our system. Equation of motion for these new 'coordinates' are uncoupled, that is, $\dot{c}(t) = -i\omega_0 c(t)$ and $\dot{c}^*(t) = i\omega_0 c^*(t)$.

At the next step the interaction between the *normal modes* of the magnetization oscillations and a thermal bath described *as a set of harmonic oscillators* is introduced via the insertion of a linear coupling between functions $a(t)$ and the thermal bath normal modes b_k into the system Hamiltonian (see Safonov and Bertram, 2003 for details):

$$H \sim E_{\text{mag}} + E_{\text{bath}} + \sum\nolimits_k \left[G_k(ab_k^* + a^* b_k) + F_k(ab_k + a^* b_k^*)\right] \tag{5}$$

where E_{mag} and E_{bath} are the energies of a magnetic system and thermal bath only.

Transformation of the interaction term to the variables $c(t)$ and $c^*(t)$, solution of the resulting dynamic equations for the thermal bath modes b_k and substitution of the expression for b_k into the dynamic equations for $c(t)$ leads to the equation

$$\frac{\mathrm{d}c}{\mathrm{d}t} = -i(\omega_0 + \Delta\omega)c - \eta c + f(t) \tag{6}$$

which describes a damped harmonic oscillator under the influence of a random force (noise) $f(t)$. The damping constant η in this formalism

$$\eta = \pi \left|\tilde{G}_k\right|^2 D(\omega_0) \tag{7}$$

is proportional to (i) the squared interaction coefficient $\tilde{G}_k = uG_k + vF_k$ of a magnetic system with the bath normal mode having the same frequency ω_0 as the undisturbed magnetic system and (ii) the density of states $D(\omega_0)$ of a thermal bath at this frequency.

An important conclusion which can be drawn from the consideration in the preceding text is that in this phenomenological model the magnetization damping present in the Gilbert or Landau–Lifshitz equation of motion for the components of the system magnetization $\mathbf{M}(t)$

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = -\gamma_0 \left[\mathbf{M} \times \left(\mathbf{H}^{\text{eff}} - \frac{\hat{\alpha}}{M_{\text{S}}} \cdot \frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t}\right)\right] \tag{8}$$

cannot be reduced anymore to a scalar variable, but is a tensor

$$\hat{\alpha} = \frac{\eta}{\omega_0^2} \begin{pmatrix} H_{0x} & 0 & 0 \\ 0 & H_{0y} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{9}$$

where the diagonal components of this tensor are proportional to the parameters $H_{0\text{x(y)}}$ characterizing the energy surface curvature near the equilibrium magnetization state (expression (9) may be verified by rewriting M_x and M_y via the functions $c(t)$ and $c*(t)$ and substituting the resulting expressions into (8), thus arriving at the equation (6)). This tensor reduces to a scalar only for a symmetric energy minimum with $H_{0\text{x}} = H_{0y}$, which is rather an exception.

The formalism outlined in the preceding text can be extended to magnetic systems with arbitrarily nonhomogeneous equilibrium magnetization configuration (Bertram, Safonov and Jin, 2002). After the discretization of a ferromagnetic body into N finite elements the eigenmodes of its magnetization state can be identified using the equation of motion *without* damping $\dot{\mathbf{M}}_i = -\gamma[\mathbf{M}_i \times \mathbf{H}_i^{\text{eff}}]$. Expanding the system state magnetization vector $\mathbf{T} = (\mathbf{M}_1, \ldots, \mathbf{M}_i, \ldots, \mathbf{M}_N)$ the near the equilibrium state one obtains a set of coupled equations $\dot{T} = -\gamma\hat{\mathbf{H}}\cdot\mathbf{T}$ where tensor $\hat{\mathbf{H}}$ consists of corresponding effective field components. Diagonalization of this tensor provides the eigenvectors $\mathbf{c}_i$, which give the spatial distribution of the i-th eigenmode and eigenvalues, which are proportional to the mode oscillation frequencies ω_i. Further analysis proceeds then as in the simplest case discussed in the preceding text.

However, the procedure sketched in the preceding text, which leads to the conclusion that the damping in the LLG equation *should* be a tensor, is incorrect, as shown by Smith (2002). Taking into account that the question about the form of the phenomenological damping term is crucially important for micromagnetic calculations, we shall briefly reproduce here the argumentation of Smith, replacing his simple illustrative example by an even simpler one.

To find out where is the flaw in the argumentation of Bertram and Safonov, we consider the simplest possible mechanical system with more than one eigenmode, that is, two particles of equal masses m, each of which is attached to a wall via a spring with the stiffness constant k. Particles are placed in the *isolated* reservoirs filled with a fluid with

different viscosities (this is the only difference between the particles) and coupled via a spring with the stiffness k_c. If particle coordinates x_1 and x_2 in the mechanical equilibrium are set to zero, the corresponding equations of motion are

$$m\ddot{x}_1 = -\eta_1\dot{x}_1 - kx_1 + k_c(x_2 - x_1) + F_1^L \tag{10}$$

$$m\ddot{x}_2 = -\eta_2\dot{x}_2 - kx_2 - k_c(x_2 - x_1) + F_2^L \tag{11}$$

Here, the friction coefficients η_1 and η_2 are different owing to different fluid viscosities. Random thermal forces F^{L} responsible for the Brownian motion of the particles are obviously *uncorrelated* (because the reservoirs with the particles are isolated from each other), δ correlated in time and their mean-square values are proportional to the system temperature T.

The determination of the system normal modes should be performed in a standard way, using the equations (10) and (11) without dissipative and fluctuation forces. In this very simple case, it is enough to add and subtract these equations in order to find out that the normal modes are $x_+^{\mathrm{n}} = (x_1 + x_2)/\sqrt{2}$ and $x_-^{\mathrm{n}} = (x_1 - x_2)/\sqrt{2}$, so that the transformation matrix U between the vector of the particle coordinates $\mathbf{x}$ and the normal modes of our system $\mathbf{x}^{\mathrm{n}}$ (i.e., $\mathbf{x} = U\cdot\mathbf{x}^{\mathrm{n}}$) is

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{12}$$

Next, we apply this transformation to the initial equations of motion system, (10) and (11), written in the matrix form

$$M\ddot{\mathbf{x}} = -\mathrm{H}\dot{\mathbf{x}} + K\mathbf{x} + F^{\mathrm{L}} \tag{13}$$

where the mass and friction matrices are diagonal ($M = diag(m, m)$, $\mathrm{H} = diag(\eta_1, \eta_2)$), and the stiffness matrix K has both diagonal $K_{11} = K_{22} = -(k + k_c)$ and off-diagonal elements $K_{12} = K_{21} = k_c$. The transformation to the normal modes proceeds in the usual way: we insert the unit matrix in the form $I = U\cdot U^{\mathrm{T}}$ into the matrix-vector products into the equation (13) and multiply it by the matrix U^{T} from the left. Using the relation $U^{\mathrm{T}}\mathbf{x} = \mathbf{x}^{\mathrm{n}}$ and performing the matrix multiplications $\tilde{M} = U^{\mathrm{T}}MU = M$, $\tilde{\mathrm{H}} = U^{\mathrm{T}}\mathrm{H}U$, $\tilde{K} = U^{\mathrm{T}}KU$, we arrive at the *equations of motion for the normal modes*:

$$\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}\cdot\begin{bmatrix} \ddot{x}_+^{\mathrm{n}} \\ \ddot{x}_-^{\mathrm{n}} \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} \eta_1 + \eta_2 & \eta_1 - \eta_2 \\ \eta_1 - \eta_2 & \eta_1 + \eta_2 \end{bmatrix}\cdot\begin{bmatrix} \dot{x}_+^{\mathrm{n}} \\ \dot{x}_-^{\mathrm{n}} \end{bmatrix} + \begin{bmatrix} -k & 0 \\ 0 & -(k + 2k_c) \end{bmatrix}\cdot\begin{bmatrix} x_+^{\mathrm{n}} \\ x_-^{\mathrm{n}} \end{bmatrix} + \begin{bmatrix} F_+^{\mathrm{n,L}} \\ F_-^{\mathrm{n,L}} \end{bmatrix} \tag{14}$$

where the vector of fluctuation forces $\mathbf{F}^{\mathrm{n,L}} = U^{\mathrm{T}}\cdot\mathbf{F}^{\mathrm{L}}$ now contains the corresponding forces for the normal modes defined as $F_+^{\mathrm{n,L}} = (F_1^{\mathrm{L}} + F_2^{\mathrm{L}})/\sqrt{2}$, $F_-^{\mathrm{n,L}} = (F_1^{\mathrm{L}} - F_2^{\mathrm{L}})/\sqrt{2}$.

It can be immediately seen from equation (14) that the equations of motion for the normal modes which include the damping term are *correlated* owing to the non-diagonal character of the transformed damping matrix $\tilde{\mathrm{H}} = U^{\mathrm{T}}\mathrm{H}U$. This means that the simple phenomenological addition of the *uncoupled* damping terms like the term $-\eta\cdot c$ in (6) is incorrect. This term in (6) results from writing the equations of motion for the coupled system 'magnetic body + thermal bath' directly in terms of the normal modes (see (5)), which hence turns out to be incorrect at least when the coupling of these modes and correlation properties of corresponding random (fluctuation) forces are important. The conclusion (9) about the obligatory tensor form of the damping in the LLG equation, being a direct consequence of the *uncoupled* damping of *normal modes*, is for this reason erroneous. A detailed extension of this discussion to the case of micromagnetic equations can be found in (Smith, 2002).

Another very important point is that the addition of *uncoupled* fluctuation forces into the equations of motion for *the normal modes* (instead of using corresponding equations for real particles) as it is done in equation (6) is also physically incorrect. Namely, if thermal forces acting on the *normal modes* would be indeed uncorrelated, then the corresponding forces acting on *real* particles (obtained via the backward transformation as $\mathbf{F}^L = U\cdot\mathbf{F}^{\mathrm{n,L}}$) would be correlated, which is completely unphysical. This is evident at least in the case considered in the preceding text where particles are placed into *separate isolated* reservoirs. Taking into account that the random (Brownian) forces are caused by the chaotic motion of the fluid molecules, any correlations between these forces in two separate isolated fluid volumes are obviously absent.

The discussion in the preceding text should remind the reader, that although normal modes of the system are doubtlessly a very useful physical concept, one should not forget that initially dynamical equations must be written for real particles (magnetic moments) where the physical meaning of various terms in such equations can be directly examined.

The last comment in order is that the argumentation line presented here does *not* mean that the tensor form of the damping term is *forbidden*. As already mentioned by Smith (2002), this only means that any conclusion about the specific

form of the damping should be based on the consideration of the corresponding physical mechanism responsible for the energy dissipation.

2 STOCHASTIC LLG EQUATION: CHOICE OF THE STOCHASTIC CALCULUS

2.1 General introduction to the solution of SDEs

Equation-of-motion simulations of the remagnetization dynamics *including* thermal fluctuations require, in contrast to the case $T = 0$, the solution of stochastic differential equations (SDE), which is by far more difficult. A simple example of such stochastic (Langevin) equations arising in the theory of stochastic processes is the equation of motion for a particle in a viscous medium under the influence of a deterministic force F_{det} and thermal fluctuations represented via the random (or Langevin) force ξ_{L}(η denotes the friction coefficient)

$$\dot{x} = \frac{1}{\eta} F_{\mathrm{det}}(t) + a(x,t)\cdot\xi_{\mathrm{L}}(t) \tag{15}$$

$\xi_{\mathrm{L}}(t)$ is usually assumed to be a random Gaussian variable, δ correlated in time:

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(0)\cdot\xi(t) \rangle = 2D\cdot\delta(t) \tag{16}$$

with the noise power $D \sim T$. The 'good' function $a(x,t)$ contains the coordinate- and time-dependencies of the noise characteristics.

The problem with the equations of this kind is that they can *by no means* be interpreted as the 'usual' differential equations (DE). Namely, any attempt to integrate equation (15) as an ordinary DE leads to the integral

$$W(t) = \int_0^t \xi(t')\cdot \mathrm{d}t' \tag{17}$$

which obviously represents a random process, because its integrand is a random variable. From the correlation properties (16) it follows that $W(t)$ is the standard Wiener process (Brownian motion) Gardiner (1997). This process is *not* differentiable – the ratio $(W(t+\Delta t) - W(t))/\Delta t$ diverges in the limit $\Delta t \to 0$ almost surely. Because the derivative $\mathrm{d}W/\mathrm{d}t = \xi(t)$ does not exist, in the usual sense the equation (16) including this derivative *does not exist* also and hence cannot be interpreted as a 'normal' ordinary DE.

The proper way to assign a mathematically correct meaning to the equations like (15) is to introduce the differential of the Wiener process $\mathrm{d}W$ (which is usually viewed by a physicist as a replacement of the product $\xi(t)\mathrm{d}t$) and to define the corresponding integral

$$I = \int a(x,t)\cdot \mathrm{d}W(t) \tag{18}$$

analogously to the standard Riemann–Stieltjes integrals as the limit of partial sums

$$\begin{aligned} I &= \lim_{n\to\infty} \sum_{i=1}^{n} a(x(\tau_i), \tau_i)\cdot \Delta W(\Delta t_i) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} a(x(\tau_i), \tau_i)\cdot [W(t_i) - W(t_{i-1})] \end{aligned} \tag{19}$$

with the points τ_i where the integrand values are evaluated lie inside the interval $[t_{i-1}, t_i]$.

This limit, being understood in the mean-square sense (see any handbook on stochastic calculus) is convenient enough to develop the complete analysis of such *stochastic* integrals. The real problem is that (in a heavy contrast to the standard analysis) this limit itself – and not just the values of partial sums in (19) – *depends on the choice of the intermediate points* τ_i (see Chapter 3 in Gardiner (1997) for a simple but impressive example).

The only way to cope with this problem is to introduce some standard choices of the intermediate points and to find the best choice from the physical point of view. The two standard choices – (i) $\tau_i = t_{i-1}$ coincide with the *left* points of the intervals (*Ito* stochastic calculus) and (ii) $\tau_i = (t_{i-1} + t_i)/2$ are in the *middle* of the intervals (*Stratonovich* interpretation) – lead to *different* solutions if the noise in a stochastic equation is *multiplicative* – that is, the random term is *multiplied* by some function of the system variables. In this case, usually the *Stratonovich* interpretation provides physically correct results, recovering, for example, some important properties of physical random processes obtained using more general methods (Gardiner, 1997).

In micromagnetics, the most common way to include thermal fluctuation effects into the consideration is the addition of the so-called 'fluctuation field' to the deterministic effective field in equation (LLG). This leads to the *stochastic* LLG equation (Brown, 1963b) for the magnetic moment motion.

$$\begin{aligned} \frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} &= -\gamma\cdot[\mathbf{M}\times(\mathbf{H}^{\mathrm{det}} + \mathbf{H}^{\mathrm{fl}})] - \lambda\cdot\frac{\gamma}{M_{\mathrm{S}}} \\ &\quad \cdot[\mathbf{M}\times[\mathbf{M}\times(\mathbf{H}^{\mathrm{det}} + \mathbf{H}^{\mathrm{fl}})]] \end{aligned} \tag{20}$$

Here, the *deterministic* effective field $\mathbf{H}^{\mathrm{det}}$ acting on the magnetization includes all the standard micromagnetic contributions (external, anisotropy, exchange and magnetodipolar interaction field). Analogous to the random force in

the mechanical equation (15), Cartesian components of the fluctuation field $\mathbf{H}_i^{\mathrm{fl}}$ are usually assumed to be δ correlated in space and time (Brown, 1963b)

$$\langle H_{\xi,i}^{\mathrm{fl}} \rangle = 0 \quad \langle H_{\xi,i}^{\mathrm{fl}}(0) \cdot H_{\psi,j}^{\mathrm{fl}}(t) \rangle = 2D \cdot \delta(t) \cdot \delta_{ij} \cdot \delta_{\xi\psi} \tag{21}$$

where i, j are the discretization cell (magnetic moment) indices and $\xi, \psi = x, y, z$. The noise power D evaluated using the fluctuation-dissipation theorem (see subsequent text for a detailed discussion) is proportional to the system temperature T and depends on γ and the damping constant λ:

$$D = \frac{\lambda}{1+\lambda^2} \cdot \frac{kT}{\gamma M_{\mathrm{S}} \Delta V} \tag{22}$$

We note in passing that the fluctuation field $\mathbf{H}^{\mathrm{fl}}$ in the dissipation term of (20) can be omitted by rescaling correspondingly the noise power D (Garcia-Palacios and Lazaro, 1998; Braun, 2000). We shall use this possibility below by comparing the Ito and Stratonovich interpretations of the LLG equations.

The noise in the Langevin equations 20 is *multiplicative*, because in the vector products the random field projections are *multiplied* by the magnetic moment projections. This fact was pointed out already in the pioneering paper of Brown (1963b), who suggested to use for this reason the Stratonovich interpretation of the equation.

Analytic solution of (20) is possible only in a few simplest cases, so that really interesting magnetic systems can be studied only numerically. For such simulations, the choice of the stochastic calculus is, in principle, of *primary* importance, because different numerical integration schemes converge to different stochastic integrals: The Euler and the simple implicit methods converge to the Ito solution, Heun and Milstein schemes – to the Stratonovich limit (McShane, 1974) and Runge-Kutta methods can converge to both types of the stochastic integrals (including the in-between cases) depending on their coefficients (Rümelin, 1982). Most authors (see, e.g., Garcia-Palacios and Lazaro, 1998; Scholz, Schrefl and Fidler, 2001; Berkov, Gorn and Görnert, 2002) and commercial micromagnetic packages (advanced recording model (ARM), LLG, MicroMagus) use the Heun, Runge-Kutta or modified Bulirsch–Stoer methods converging to the Stratonovich solution, but several groups employ the Ito-converging Euler (Zhang and Fredkin, 2000; Lyberatos and Chantrell, 1993) method and implicit schemes (Nakatani, Uesaka, Hayashi and Fukushima, 1997). These last papers were criticized in Garcia-Palacios and Lazaro (1998) where it has been claimed that only the Stratonovich interpretation ensures the physically correct solution of (20).

Fortunately, we could show that for *standard* micromagnetic models (where $|\mathbf{M}| = M_{\mathrm{S}} = Const$) both Ito and Stratonovich stochastic calculi provide identical results, so that the only criterion by the choice of the integration method is its efficiency by the solution of the LLG stochastic equation for the particular system under study. In the next subsection, we address this question in more detail because of its methodical importance.

2.2 Equivalence of Ito and Stratonovich stochastic calculus for standard micromagnetic models

In this subsection, we shall prove that for the system in which the dynamics is described by the stochastic equation (20) the Ito and Stratonovich versions of the stochastic calculus are equivalent *if the magnetization magnitude (or the magnitude of the discretization cell/single particle) is assumed to be constant* [2]. This is true for many magnetic system models including the classical Heisenberg and related models, spin glasses, fine magnetic particle systems (Dotsenko, 1993; Hansen and Morup, 1998), and in standard micromagnetics (Brown, 1963a).

First, we repeat that the fluctuation field in the dissipation term of (20) that can be omitted by rescaling correspondingly the noise power D (Garcia-Palacios and Lazaro, 1998; Braun, 2000). Thus we can restrict ourselves to the study of a simpler equation

$$\begin{aligned} \frac{\mathrm{d}\mathbf{m}_i}{\mathrm{d}\tau} = &- \left[\mathbf{m}_i \times (\mathbf{h}_i^{\mathrm{eff}} + \mathbf{h}_i^{\mathrm{fl}})\right] \\ &- \lambda \cdot \left[\mathbf{m}_i \times \left[\mathbf{m}_i \times \mathbf{h}_i^{\mathrm{eff}}\right]\right] \end{aligned} \tag{23}$$

To proceed, we recall that by the transition between the Ito and Stratonovich forms of a stochastic differential equation the additional drift term appears: if one adds to a SDE system

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = A_i(\mathbf{x}, t) + \sum_k B_{ik} \xi_k \tag{24}$$

the *deterministic* drift $D \sum_{jk} B_{jk} (\partial B_{ik} / \partial x_j)$, then the *Ito* solution of this new system

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = A_i(\mathbf{x}, t) + D \sum_{jk} B_{jk} \frac{\partial B_{ik}}{\partial x_j} + \sum_k B_{ik} \xi_k \tag{25}$$

is equivalent to the *Stratonovich* solution of the initial system (24) (Kloeden and Platen, 1995).

For the LLG equation written in Cartesian coordinates the matrix $\mathbf{B}$ is $B_{ik} = \sum_j \varepsilon_{ijk} m_j$, and the additional drift term reduces to $\mathrm{d}\mathbf{m}_i / \mathrm{d}\tau = -2D\mathbf{m}_i$. This drift is directed *along* the magnetic moment $\mathbf{m}_i$ trying to change its *magnitude*, which is forbidden by the model. Hence, this term *must* be discarded, which leads to the equivalence of the Ito and Stratonovich schemes.

The same result can be obtained (and understood) in a much simpler way rewriting the LLG equation using spherical coordinates (θ, ϕ) of the moment unit vector $\mathbf{m}$ (Cartesian coordinates of magnetic moments are *not* independent: owing to the conservation of a moment magnitude they are subject to the restriction $m_{x,i}^2 + m_{y,i}^2 + m_{z,i}^2 = 1$). In spherical coordinates the part containing the fluctuation (stochastic) field part of (23) which is of interest for us reads (Brown, 1963b; Braun, 2000).

$$\frac{\mathrm{d}\theta}{\mathrm{d}\tau} = h_{\varphi}^{fl}, \quad \frac{\mathrm{d}\varphi}{\mathrm{d}\tau} = -\frac{1}{\sin\theta} h_{\theta}^{fl} \tag{26}$$

so that the matrix $\mathbf{B}$ responsible for the drift mentioned in the preceding text is

$$B = \begin{pmatrix} B_{\theta\theta} & B_{\theta\varphi} \\ B_{\theta\varphi} & B_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1/\sin\theta & 0 \end{pmatrix} \tag{27}$$

and this drift is exactly zero: $D\sum_{jk} B_{jk}(\partial B_{ik}/\partial x_j) = 0$ ($i, j, k = 1, 2$ and $x_1 = \theta, x_2 = \phi$). Hence, we arrive at the same result that Stratonovich and Ito stochastic integrals are equivalent in this case, which means that *for stochastic dynamics of models with rigid dipoles* (*dipoles with constant magnitudes*) *there is no difference between the Ito and Stratonovich solutions of corresponding stochastic differential equations* (Berkov and Gorn, 2002).

It is interesting to see why the opposite statement made in Garcia-Palacios and Lazaro (1998) is incorrect. Using the Fokker–Planck equation (FPE) which describes the evolution of the probability distribution of the magnetization orientation $P(\mathbf{m}, t)$, the authors of Garcia-Palacios and Lazaro (1998) have shown that an additional drift term $\partial(\mathbf{m}P)/\partial\mathbf{m}$ appears in the FPE derived from the *Ito* interpretation of the Langevin equation. But this term should be excluded from the FPE because it leads to the probability density drift *along* the magnetization vector which would change the moment magnitude (this can be most easily seen in *spherical* coordinates (m, θ, ϕ) where this term reduces to $\partial(mP)/\partial m$).

To support our conclusion about the equivalence of the Ito and Stratonovich integrals for models with constant magnetic moment magnitudes, we have performed numerical experiments simulating equilibrium (density of states) and nonequilibrium (time dependent magnetization relaxation) properties of a disordered system of magnetic dipoles. We have solved the stochastic LLG equation (20) using methods converging either to its Ito (Euler scheme) or Stratonovich (drift-modified Euler and Heun schemes) solutions. Results obtained by all these methods coincide within the statistical accuracy, confirming that Ito and Stratonovich calculi lead, for these systems, to the same physical results despite that the noise in the stochastic LLG equation is *multiplicative*.

However, we point out that the proof in the preceding text *heavily relies* on the *conservation* of the moment magnitude. Hence for models where this is *not* the case – for example, by simulations of the heat assisted magnetic recording (HAMR) or for models attempting to relax the local restriction $\mathbf{M} = M_S = \text{Const}$ (see preceding text) – one should pay close attention to the choice of a numerical method used to solve equation (20).

3 STOCHASTIC LLG EQUATION: THERMAL NOISE CORRELATIONS

3.1 Thermal fluctuations for a single magnetic moment

3.1.1 Introduction

As it was stated in Section 2, the 'standard' way to take into account the thermal fluctuations of the magnetization in micromagnetic simulations is the inclusion of the 'fluctuation field' $\mathbf{H}^{fl}$ into the basic equation (20), which we write out here for a single magnetic moment $\boldsymbol{\mu}$ once more to have it at hand:

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = -\gamma\cdot\left[\boldsymbol{\mu}\times(\mathbf{H}^{det}+\mathbf{H}^{fl})\right] - \lambda\cdot\frac{\gamma}{\mu}\cdot\left[\boldsymbol{\mu}\times\left[\boldsymbol{\mu}\times(\mathbf{H}^{det}+\mathbf{H}^{fl})\right]\right] \tag{28}$$

Components of the this field $\mathbf{H}_i^{fl}$ are supposed to have δ correlation in space and time

$$\langle H_{\xi,i}^{fl}\rangle = 0 \;\langle H_{\xi,i}^{fl}(0)\cdot H_{\psi,j}^{fl}(t)\rangle = 2C\cdot\delta(t)\cdot\delta_{ij}\cdot\delta_{\xi\psi} \tag{29}$$

where i, j are the discretization cell (or magnetic moment) indices and $\xi, \psi = x, y, z$. The proportionality coefficient before the δ functions in (29) is the noise power C, which can be evaluated as (when the random field is present in both terms on the right-hand side of (20))

$$C = \frac{\lambda}{1+\lambda^2}\cdot\frac{kT}{\gamma\mu} = \frac{\lambda}{1+\lambda^2}\cdot\frac{kT}{\gamma M_S V_p} \tag{30}$$

In the last equality we have used the relation $\mu = M_S V_p$ between the particle magnetic moment μ, saturation magnetization of the particle material M_S and the particle volume V_p.

The question of main interest is whether the properties (29) and (30), introduced by Brown (1963b) for a single-domain particle – actually for a single magnetic moment surrounded by a thermal bath – would survive for a typical micromagnetic system, where complicate interactions between the magnetic moments of the finite elements used to discretize a continuous problem are present.

We start the discussion of this principal problem with the statement that it actually contains two separate question:

1. Whether the correlations between random field components *can be really treated as δ -functional ones* and
2. Whether *the noise power can be evaluated using the universal expression like* (30) which contains *only* the system temperature T, the gyromagnetic ratio γ, the damping constant λ and the magnitude of the cell magnetic moment $\mu = M_S \Delta V$, and thus does not depend on any interaction details and other features of the concrete system under study.

To answer these questions for a micromagnetic system, we first recall how the δ correlations (29) of the random noise are introduced for a single magnetic particle, how the power of this noise (30) can be obtained for this simplest case, and discuss a physical sense of the expression (30).

First of all, in case of a single particle (single magnetic moment) we are not interested in *spatial* correlation properties of the noise. The statement that temporal correlations of the random field components are δ functional is an *assumption* based explicitly on the properties of that physical component of a thermal bath which is responsible for the appearance of thermal fluctuations. The most common point of view is that thermal fluctuations are mainly due to the interactions with phonons. In this case, we are interested in the correlation time associated with typical phonons, which contribute to thermal bath fluctuations. If we study the properties of our system at room temperature, which is of the same order of magnitude as the Debye temperature of typical materials, then we are speaking about phonons with the wavelength about several interatomic distances; typical life time of such room-temperature phonons is about a picosecond or less, so for remagnetization processes on time scales much larger than that we can safely accept the temporal correlation function $\delta(t)$ in (29). We note in passing that for the description of the magnetization dynamics on a much shorter time scales the phenomenology (20) is not valid anyway, because the equilibrium between various subsystems of a magnetic body (electrons, phonons, magnons etc.) cannot be reached. We also point out that the logic presented above to justify the δ-functional character of temporal noise correlations fails for systems at low temperatures, where the decay time of the characteristic (long-wave) phonons may be more than several nanoseconds, especially in pure materials.

3.1.2 *Derivation of the relation between the noise power and system properties for a mechanical Brownian motion*

Let us now examine the derivation of the relation (30). Usually it is quoted as a consequence of a so-called fluctuation-dissipation theorem (see subsequent text), but in many cases it can be obtained in a much simpler way, which also make the physical sense of this relation more transparent. To illustrate this point, we start with the mechanical Brownian motion of a 'normal' free particle described by the Langevin equation

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\eta\frac{\mathrm{d}x}{\mathrm{d}t} + F^{\mathrm{fl}}(t) \tag{31}$$

where the first term on the right is the friction force and the second term represents the Langevin (fluctuating) force, which has by the assumptions outlined above zero-mean value and the correlation properties $\langle F^{\mathrm{fl}}(0)\cdot F^{\mathrm{fl}}(t)\rangle = 2C\cdot\delta(t)$. Introducing the particle velocity $v = \mathrm{d}x/\mathrm{d}t$, we obtain for v from the equation above a simple first-order differential equation

$$\frac{\mathrm{d}v}{\mathrm{d}t} + \frac{\eta}{m}v = \frac{1}{m}F^{\mathrm{fl}}(t) \tag{32}$$

which explicit solution

$$v(t) = v(0)e^{-\eta t/m} + \frac{1}{m}\int_0^t e^{-\eta(t-t')/m}F^{\mathrm{fl}}(t')\,\mathrm{d}t' \tag{33}$$

allows the straightforward evaluation of the mean-square velocity: writing the velocities for two different time moments $v(t_1)$ and $v(t_2)$ using (33), multiplying these quantities and taking the thermal average, we obtain the expression for the product $\langle v(t_1)v(t_2)\rangle$ which contains only the correlation function $\langle F^{\mathrm{fl}}(t_1)\cdot F^{\mathrm{fl}}(t_2)\rangle = 2C\cdot\delta(t_1 - t_2)$ (all other terms are zero because the velocity values are not correlated with the values of the random force). Putting $t_1 = t_2$ and using the basic property of the δ function, we finally obtain the desired result $\langle v^2\rangle = C/2\eta m$. On the other hand, in the *thermal equilibrium* the average kinetic energy of the particle is $\langle E_K\rangle = m\langle v^2\rangle/2 = kT/2$, so that $\langle v^2\rangle = kT/m$. Equating these two expressions for the mean square of the particle velocity, we obtain the desired result

$$C = \eta\cdot kT \tag{34}$$

which connects the noise power C with the friction coefficient η and system temperature T. Note that by derivation of this relation we have used only the fact that the system (particle) is in a thermodynamical equilibrium with the surrounding thermal bath.

For a particle which moves in an external potential and hence possesses the energy $V(x)$ the situation is more complicate, because no general analytical solution of the corresponding Langevin equation

$$\eta\frac{\mathrm{d}x}{\mathrm{d}t} = -\nabla V(x) + F^{\mathrm{fl}}(t) \tag{35}$$

is available in (35) we have neglected the inertial term for simplicity, which means that we are interested in times much larger than the velocity equilibration time $t_p = m/\eta$. In this situation, the relation between the noise power C and the system features can be established using another principal equation of the theory of stochastic processes – the FPE.

The FPE describes the temporal and spatial evolution of the probability density $w(x,t)$ which gives the probability $w(x,t)\cdot dx\cdot dt$ to find the particle inside the region $[x; x+dx]$ during the time interval $[t; t+dt]$. This equation can be derived in many ways; the most transparent general method to obtain the FPE starts from the so-called Chapman–Kolmogorov (also called *Smoluchovski*) equation. From a physical point of view, this latter equation simply states that the conditional probability $P(x_1,t_1|x_3,t_3)$ to find the particle at x_1 at the time t_1, *if* its position at the time t_3 was x_3, can be obtained by integrating the product of conditional probabilities for the transitions $(x_1,t_1)\to(x_2,t_2)$ and $(x_2,t_2)\to(x_3,t_3)$ over all intermediate particle positions x_2. The derivation of the FPE from the Chapman–Kolmogorov equation is conceptually very simple (see, e.g., Gardiner, 1997) and assumes only the existence of the limits

$$A(x,t) = \lim_{\Delta t\to 0}\frac{\langle x(t+\Delta t)-x(t)\rangle}{\Delta t}$$
$$B(x,t) = \frac{1}{2}\lim_{\Delta t\to 0}\frac{\langle (x(t+\Delta t)-x(t))^2\rangle}{\Delta t} \qquad (36)$$

The resulting FPE reads in a general case

$$\frac{\partial w(x,t)}{\partial t} = -\frac{\partial}{\partial x}[A(x,t)\cdot w(x,t)] + \frac{\partial^2}{\partial x^2}[B(x,t)\cdot w(x,t)] \qquad (37)$$

Here, the first term on the right describes the systematic particle *drift* due to the potential force $F_{pot}(x) = -\nabla V(x)$ and the second term is responsible for the particle *diffusion* due to thermal fluctuations. For each concrete system the limits (36) can be evaluated from the corresponding Langevin equation (Coffey, Kalmykov and Waldron, 2004) in conceptually the same manner as the mean square of the particle velocity was derived from (32). For the simple system described by the equation (35) the result is

$$A = -\nabla V(x)/\eta, \quad B(x) = Const = C/\eta^2 \qquad (38)$$

(C is the noise power from the correlation function $\langle F^{fl}(t_1)\cdot F^{fl}(t_2)\rangle = 2C\cdot\delta(t_1-t_2)$ of the random force) so that the FPE (37) has the form of a standard diffusion equation

$$\frac{\partial w(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[\frac{1}{\eta}\frac{\partial V}{\partial x}\cdot w(x,t)\right] + \frac{C}{\eta^2}\cdot\frac{\partial^2 w(x,t)}{\partial x^2} \qquad (39)$$

which means that the noise power C is equal to the diffusion coefficient.

To establish now the required relation between the noise power and other system properties, we use the statement that in the equilibrium state (where the time derivative $\partial w/\partial t = 0$ the probability distribution function should be given by the (normalized) Boltzmann exponent $w(x) = N\exp(-V(x)/kT)$. Substitution of this expression into the right-hand side of the FPE (39) gives the equation, which is worth to be explicitly written out:

$$\frac{d^2V}{dx^2}\left(\eta - \frac{C}{kT}\right) - \frac{1}{kT}\cdot\left(\frac{dV}{dx}\right)^2\left(\eta - \frac{C}{kT}\right) = 0 \qquad (40)$$

The equation (40) should be satisfied for an arbitrary potential $V(x)$ which leads exactly to the same relation $C = \eta\cdot kT$ between the random noise power and system parameters as the result (34) for a free particle. Now we have proved that this relation *does not* depend on the concrete potential $V(x)$, and thus represent a very general statement. Again, the only physical assumption used to derive this formula was the statement that *the system is in a thermodynamical equilibrium*, so that the Boltzmann distribution for the probability density $w(x,t)$ could be used.

3.1.3 FPE and noise power evaluation for a single magnetic moment

The FPE can be derived for a single magnetic moment in the same way as in the preceding text for a mechanical particle. The resulting equation is, however, much more complicate due to the following reasons: (i) we deal here with a *rotational* diffusion and (ii) the precession term (which is the counterpart of the inertial term in mechanics) in the corresponding Langevin equation (20) cannot be neglected, because this would lead to a *qualitatively* incorrect description of the magnetization dynamics for the overwhelming majority of physically interesting systems.

The first derivation of the corresponding FPE was presented (up to our knowledge) by Brown (1963b), who used the 'physical' or 'intuitive' method, employing the continuity equation for the distribution density $w(\theta,\phi)$ of the magnetic moment directions. This equation relates the time derivative of $w(\theta,\phi)$ and its flux on the (θ,ϕ) sphere, whereby the diffusion term is added to the flux in a phenomenological way using the similarity with the equation describing the mechanical rotational diffusion. The rigorous derivation of this FPE, which uses the functional analysis methods can be found in the Appendix of Garanin (1997). Another derivation of the same FPE from the Langevin equation (20) which employs the relations (known from the general rules of stochastic calculus) between the coefficients of the Langevin

equation and corresponding terms from FPE is contained in Garcia-Palacios and Lazaro (1998). The structure of the resulting equation

$$\frac{\partial w(\mathbf{m},t)}{\partial t} = -\nabla_{\mathbf{m}}\Big\{\gamma[\mathbf{m}\times\mathbf{H}^{\text{eff}}]w - \gamma\cdot\lambda\cdot[\mathbf{m}\times[\mathbf{m}\times\mathbf{H}^{\text{eff}}]]w + D_{\text{rot}}[\mathbf{m}\times[\mathbf{m}\times\nabla_{\mathbf{m}}]]w\Big\} \tag{41}$$

is obviously inherited from the Langevin equation in the LLG form (20) which the FPE (41) has been derived from: the first term in the curved parenthesis describes the drift of the probability density due to the precession term in (20), the second term corresponds to the drift due to the damping torque and the third (diffusion) term also has the structure of the damping term in the original LLG. The operator $\nabla_{\mathbf{m}}$ acting on the vectors on the right-hand side of (41) means the divergence over the components of the vector $\mathbf{m}$, so that, for example, in Cartesian coordinates $\nabla_{\mathbf{m}}\mathbf{a} \equiv \sum_i \partial a_i/\partial m_i (i = x, y, z)$. The coefficient before the rotational diffusion term, expressed in terms of the quantities entering into the LLG equation (20), is

$$D_{\text{rot}} = \gamma^2(1+\lambda^2)\cdot C \tag{42}$$

where C is the noise power of the fluctuating field $\mathbf{H}^{\text{fl}}$.

The procedure for the calculation of the noise power using this FPE is fully analogous to that described in the preceding text for the mechanical translational diffusion. Namely, we substitute the equilibrium probability density $w(x) = N\exp(-V(\theta,\varphi)/kT)$ into the equation (41) with zero left-hand side ($\partial w/\partial t = 0$). Taking into account the effective field definition $\mathbf{H}^{\text{eff}} = -\partial V/\partial\mathbf{m}$ we arrive after a very tedious, but straightforward differentiation at the relation similar to (40), which in this case can be satisfied only if

$$D_{\text{rot}} = kT\cdot\frac{\gamma\lambda}{\mu} \tag{43}$$

Equating the two expression (42) and (43) for D_{rot}, we arrive at the final result which establish the relation (30) between the noise power C and the system parameters:

$$C = \frac{\lambda}{1+\lambda^2}\cdot\frac{kT}{\gamma\mu} = \frac{\lambda}{1+\lambda^2}\cdot\frac{kT}{\gamma M_{\text{S}} V_{\text{p}}} \tag{44}$$

As in the corresponding relation (34) for the mechanical translational Brownian motion, the noise power is proportional to the system temperature kT and to the friction constant λ. Also, fully analogous to the mechanical noise power (34), the value of C in (44) does *not* depend on the concrete potential $V(\theta,\phi)$ acting on the magnetic moment, which can be understood as a first hint that this result will remain unchanged for a system of *interacting* magnetic moments also. The appearance of the gyromagnetic ratio γ in the denominator is due to its presence as a common factor for both drift terms in the initial FPE (41), so that after the substitution of an equilibrium Boltzmann probability density and differentiation it appears in the equality (43). The factor $(1+\lambda^2)$ in the denominator reflects the special structure of the LLG (28), where the fluctuating field $\mathbf{H}^{\text{fl}}$ has been added *both* to the precession and dissipation terms. It is also possible to use an alternative form of this equation – with $\mathbf{H}^{\text{fl}}$ added to the precession term *only*, in which case the factor $(1+\lambda^2)$ in the relation (45) is absent (Garcia-Palacios and Lazaro, 1998).

One aspects of this relation deserve a special discussion, namely, the inverse proportionality of the noise power C to the total particle magnetic moment μ, or – taking into account that the saturation magnetization of the particle material M_{S} is constant – to the particle volume V_{p}. In other words, the noise power, or the dispersion of the fluctuation field $\sigma_{\text{fl}}^2 = \langle(\mathbf{H}^{\text{fl}})^2\rangle decreases$ linearly with the increasing particle volume. This dependence can be understood on an intuitive level in the following way. Let us consider a small magnetic particle consisting of N atoms with magnetic moments $\boldsymbol{\mu}_i$. For each atomic magnetic moment the equation of motion (we neglect damping to simplify the discussion) can be written as

$$\frac{\mathrm{d}\boldsymbol{\mu}_i}{\mathrm{d}t} = -\gamma\cdot\big[\boldsymbol{\mu}_i\times(\mathbf{H}^{\text{det}} + \mathbf{H}_i^{\text{fl}})\big] \tag{45}$$

Here, we have assumed that the particle is so small deterministic effective fields are approximately equal for all elementary moments. Thermal fluctuation field $\mathbf{H}_i^{\text{fl}}$ randomly varies from one atom to another. To obtain the equation of motion for the total particle magnetic moment $\boldsymbol{\mu}_{\text{tot}} = \sum_{i=1}^{N}\boldsymbol{\mu}_i$, we have to sum the equations (45) over all particle atoms:

$$\sum_i \frac{\mathrm{d}\boldsymbol{\mu}_i}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{\mu}_{\text{tot}}}{\mathrm{d}t} = -\gamma\cdot\sum_i\big[\boldsymbol{\mu}_i\times\mathbf{H}^{\text{det}}\big] - \gamma\cdot\sum_i\big[\boldsymbol{\mu}_i\times\mathbf{H}_i^{\text{fl}}\big] \tag{46}$$

Taking into account the independence of $\mathbf{H}^{\text{det}}$ on the elementary moment index, the first sum on right-hand side immediately transforms to $[\boldsymbol{\mu}_{\text{tot}}\times\mathbf{H}^{\text{det}}]$. Situation with the second (random field) term is more complicate, because $\mathbf{H}_i^{\text{fl}}$ is a random quantity. Using again the assumption that the particle is small enough to ensure that due to the exchange interaction atomic magnetic moments are nearly parallel, we

can write the elementary moment on the second sum as $\boldsymbol{\mu}_i = \boldsymbol{\mu}_{\text{tot}}/N$ and factor the i-independent total moment out of the sum. Then the equation (46) takes the desired form of the Langevin equation for the total moment

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{\mu}_{\text{tot}}}{\mathrm{d}t} &= -\gamma\cdot\left[\boldsymbol{\mu}_{\text{tot}}\times\mathbf{H}^{\text{det}}\right] - \gamma\cdot\left[\frac{\boldsymbol{\mu}_{\text{tot}}}{N}\times\sum_i \mathbf{H}_i^{\text{fl}}\right] \\ &= -\gamma\cdot\left[\boldsymbol{\mu}_{\text{tot}}\times(\mathbf{H}^{\text{det}}+\mathbf{H}_{\text{tot}}^{\text{fl}})\right] \end{aligned} \tag{47}$$

if we *define* the fluctuation field $\mathbf{H}_{\text{tot}}^{\text{fl}}$ acting on the total moment as

$$\mathbf{H}_{\text{tot}}^{\text{fl}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{H}_i^{\text{fl}} \tag{48}$$

The relation (48) between the total and elementary thermal fields means that the dispersion of the total fluctuation field σ_{tot}^2 can be evaluated as

$$\sigma_{\text{tot}}^2 = \frac{1}{N^2}\cdot\sum_{i=1}^{N}\sigma_i^2 = \frac{1}{N^2}\cdot N\sigma_{\text{at}}^2 = \frac{1}{N}\cdot\sigma_{\text{at}}^2 \tag{49}$$

where we have used the assumption that all fluctuation fields are random *independent* variables with *equal* dispersions $\sigma_i^2 = \sigma_{\text{at}}^2 \forall i = 1,\ldots,N$. The dispersion of the fluctuating field on a single atom σ_{at}^2 does not depend on the system volume, so the equality (49) explains why the dispersion of the total fluctuation field $\mathbf{H}_{\text{tot}}^{\text{fl}}$ which appears in the LLG equation (28) is inversely proportional to the number of elementary magnetic moments in the system, that is, to the system volume.

The relation (30) can also be understood as the statement that with increasing particle volume the importance of thermal fluctuation decreases, in accordance with an intuitive picture of Brownian motion.

Already here we would like to point out, that the increase of the noise power (30) with the decreasing particle volume just discussed has serious consequences for numerical micromagnetic simulations. The random field $\mathbf{H}^{\text{fl}}$ present in the LLG equation of motion is the *main factor*, which limits the simulation time step when an algorithm with the built-in adaptive step-size control is used (which should always be the case). The reason is quite simple: fluctuation field, being a random process, is *not* a smooth function of time and space, which naturally strongly diminishes the efficiency of any numerical integrator. It can be even shown that the order of a numerical integration scheme applied to a stochastic equation is usually a *square root* of the order of the same scheme applied to an ordinary differential equation (Kloeden and Platen, 1995). Hence, the growth of the mean fluctuation field amplitude with the decreasing discretization cell volume (entering into (30) instead of the particle volume) enforced by the relation (30) leads to the *decrease* of the integration time step in micromagnetic simulations on finer grids, which should be always kept in mind by estimating simulation time basing on the data obtained on coarser grids.

3.2 Noise correlations for an interacting system: general theory

The most interesting question concerning the random field concept used to simulate magnetization dynamics under the influence of thermal fluctuations is the following: can the random field components on different spatial locations (different discretization cells) still be considered as *independent* (uncorrelated) random variables, despite the strong interactions between magnetic moments? The hand-waving argument that this interaction should not influence the correlation properties of random fields, because all the interaction kinds are already included in the deterministic part of the effective field $\mathbf{H}^{\text{det}}$ cannot be considered as fully satisfactory. For example, in a system of interacting particles moving in a viscous media, random forces acting on different particles should be treated as correlated ones to ensure correct statistical properties of such systems (see, e.g., Ermak and McCammon, 1978). For this reason, we have to resort to a general theory which allows to evaluate noise correlations in interacting many-particle systems in a rigorous way.

This theory operates with the so-called *thermodynamically* conjugate variables (Landau and Lifshitz, 1980) (which should not be confused with the conjugate variables known from quantum mechanics) and we repeat here briefly the major points of this concept to make this review self-contained. In short, we consider a system which state is fully characterized by a set of N variables $\{\mathbf{x} = x_1, x_2, \ldots, x_N\}$ chosen so that their values at equilibrium are zero: $\mathbf{x}_0 = 0$. If the system fluctuates, that is, the values of system variables deviate from these equilibrium values and change with time, then the time derivatives $\mathrm{d}x_i/\mathrm{d}t$ can be expressed as functions of instantaneous values of $\mathbf{x}(t)$ as $\dot{x}_i = f_i(\mathbf{x})$. If the deviations from the equilibrium are small, we can expand the functions f_i around $\mathbf{x} = 0$ and maintain only the first order terms in small quantities x_i (at equilibrium in the absence of thermal fluctuations $\dot{x}_i = 0$), obtaining a system of first-order differential equations describing the relaxation of the system variables to their equilibrium values as

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \sum_k \Lambda_{ik}x_k, \text{ where } \Lambda_{ik} = \frac{\partial f_i}{\partial x_k} \tag{50}$$

To account for thermal fluctuations, we introduce into (50) random forces $\xi_i(t)$ which are assumed to be responsible for

the fluctuations of the system variables:

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = \sum_k \Lambda_{ik} x_k(t) + \xi_i(t) \tag{51}$$

Assuming that the correlation times of these forces are much less than all characteristic system relaxation times, we can write the correlation functions $K_{\mathrm{ik}}(t)$ of the random forces as $\langle \xi_i(0)\xi_k(t)\rangle = 2C_{ik}\cdot\delta(t)$. The problem is to calculate the correlation coefficients C_{ik} in order to obtain from the system (51) correct statistical properties of the system fluctuations.

It turns out that the correlation matrix C_{ik} can be expressed in the simplest way when the system (51) is rewritten in terms of *thermodynamically conjugate variables* $\{\mathbf{X} = X_1, X_2, \ldots, X_N\}$ defined as derivatives of the system entropy S over the 'initial' variables $\mathbf{x}$: $X_i = -\partial S/\partial x_i$. Near the equilibrium the difference between the system entropy S and its maximum (equilibrium) value $S_{\max}$ can be expanded over small deviations x_i: $S - S_{\max} = -\frac{1}{2}\sum_{i,k}\beta_{ik}x_i x_k$, so that thermodynamically conjugate variables X_i are linear functions of $\{\mathbf{x}\}$:

$$X_i = -\frac{\partial S}{\partial x_i} = \beta_{ik}x_k, \quad \text{or} \quad \mathbf{X} = \hat{\boldsymbol{\beta}}\mathbf{x} \tag{52}$$

where $\hat{\beta} = \{\beta_{ik}\}$. Substituting $\mathbf{x} = \hat{\beta}^{-1}\mathbf{X}$ into the sum on the right-hand side of (51), we obtain the stochastic equations for thermal fluctuations of our system near its equilibrium state in the form

$$\begin{aligned}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} &= -\hat{\Lambda}\mathbf{x} + \boldsymbol{\xi} = -(\hat{\Lambda}\cdot\hat{\beta}^{-1})\mathbf{X} + \boldsymbol{\xi} \\ &= -\hat{\Gamma}\cdot\mathbf{X} + \boldsymbol{\xi}, \ \text{or} \end{aligned} \tag{53}$$

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -\sum_k \Gamma_{ik}X_k + \xi_i \tag{54}$$

with the matrix $\hat{\Gamma} = \{\Gamma_{ik}\}$ defined as $\hat{\Gamma} = \hat{\Lambda}\hat{\beta}^{-1}$.

The usefulness of these transformations becomes apparent when we express the correlation coefficients of the random forces C_{ik} in terms of matrices $\hat{\Lambda}$, $\hat{\Gamma}$, and $\hat{\beta}$. Corresponding derivation for the many-variable case is quite tedious, so we restrict ourselves to the system characterized by a single variable x and its conjugate $X = -\partial S/\partial x = \beta x$. Such a system is described by the relaxation equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\Lambda x(t) + \xi(t) \tag{55}$$

The temporal correlation function of x defined as $\varphi(t - t') = \langle x(t)x(t')\rangle$ may depend on the difference $t - t'$ only (stationary fluctuations) and can be easily found multiplying (55) by $x(t')$, performing statistical averaging $\langle\cdots\rangle$ and taking into account that the values of $x(t')$ and $\xi(t)$ are uncorrelated. The result is $\varphi(t) = \langle x^2\rangle\exp(-\lambda|t|) = (1/\beta)\exp(-\Lambda|t|)$, where the mean square of x was evaluated from its probability distribution $w(x)\ \exp\{S(x)\}$ using the above mentioned quadratic expansion $S - S_{\max} = -(\beta/2)x^2$ near its maximum.

Denoting the Fourier transform (FT) of $x(t)$ as $\tilde{x}(\omega)$, we rewrite the definition of $\phi(t)$ as

$$\begin{aligned}\varphi(t - t') &= \langle x(t)x(t')\rangle \\ &= \frac{1}{(2\pi)^2}\iint \langle\tilde{x}(\omega)\tilde{x}(\omega')\rangle e^{-i(\omega t + \omega' t')}\mathrm{d}\omega\mathrm{d}\omega' \end{aligned} \tag{56}$$

The statement that the CF $\varphi(t - t')$ may depend on the time *difference* only (see above) requires that the FT product in (56) has the form $\langle\tilde{x}(\omega)\tilde{x}(\omega')\rangle = 2\pi\cdot P_x(\omega)\cdot\delta(\omega + \omega')$. Here $P_x(\omega)$ is the spectral power of $x(t)$ and is, according to (56), the Fourier transform of its correlation function $\varphi(t) = \langle x(t)x(0)\rangle$ (Wiener–Khinchin theorem). For $\varphi(t) = (1/\beta)\exp(-\Lambda|t|)$ the simple integration gives $P_x(\omega) = 2\Lambda/\beta(\omega^2 + \Lambda^2)$.

Now we apply the same method to calculate the correlation function of the random force. Expressing $\xi(t)$ via its FT $\tilde{\xi}(\omega)$, (due to (55) we have $\tilde{\xi}(\omega) = (\Lambda - i\omega)\tilde{x}(\omega)$), and using the definition $K(t - t') = \langle\xi(t)\xi(t')\rangle$, we obtain

$$\begin{aligned}K_\xi(t - t') &= \frac{1}{(2\pi)^2}\iint \langle\tilde{\xi}(\omega)\tilde{\xi}(\omega')\rangle e^{-i(\omega t + \omega' t')}\mathrm{d}\omega\mathrm{d}\omega' \\ &= \frac{1}{(2\pi)^2}\iint (\lambda - i\omega)(\lambda - i\omega')\langle\tilde{x}(\omega) \\ &\quad \times \tilde{x}(\omega')\rangle e^{-i(\omega t + \omega' t')}\mathrm{d}\omega\mathrm{d}\omega' \\ &= \frac{1}{2\pi}\int (\lambda^2 + \omega^2)P_x(\omega)e^{-i\omega t}\mathrm{d}\omega \end{aligned} \tag{57}$$

where by the last transformation the property $\langle\tilde{x}(\omega)\tilde{x}(\omega')\rangle = 2\pi\cdot P_x(\omega)\cdot\delta(\omega + \omega')$ was used. Equation (57) means that the spectral power of the random noise fluctuations $P_\xi(\omega)$ is related to the spectral power $P_x(\omega)$ of the x-fluctuations via $P_\xi(\omega) = (\omega^2 + \Lambda^2)P_x(\omega)$. Hence the random force spectrum $P_\xi(\omega) = 2\Lambda/\beta$ is frequency independent, as it should be for the quantity with the CF $K(t) = \langle\xi(t)\xi(0)\rangle = (2\lambda/\beta)\cdot\delta(t)$.

In terms of the conjugate variable the relaxation equation for a single-variable system reads $\mathrm{d}x/\mathrm{d}t = -\Gamma\cdot X(t) + \xi(t)$, so that $\tilde{\xi}(\omega) = \Gamma\cdot\tilde{X}(\omega) - i\omega\tilde{x}(\omega)$. The procedure identical to (57) results in the relation $P_\xi(\omega) = \omega^2 P_x(\omega) + \Gamma^2 P_X(\omega)$ between the spectral powers of the random noise, $x(t)$ and its conjugate $X(t)$. The spectrum $P_X(\omega) = 2\beta\Lambda/(\omega^2 + \Lambda^2)$ can be found from the equation $\mathrm{d}X/\mathrm{d}t = -\Lambda X(t) + \xi(t)/\beta$ for $X(t)$ (it follows from (55) and $X = \beta x$). Combining $P_x(\omega)$

and $P_X(\omega)$ into the noise spectrum $P_\xi(\omega)$ and using the relation $\Gamma = \Lambda/\beta$ (analogue of matrix relation $\hat{\Gamma} = \hat{\Lambda}\hat{\beta}^{-1}$), we obtain $P_\xi(\omega) = 2\Gamma$, so that $\langle \xi(t)\xi(0)\rangle = 2\Gamma\cdot\delta(t)$.

The whole exercise presented in the preceding text makes real sense only for the many-variable case (random noise correlations as a function of Γ can be obtained for a single-variable system immediately from the equation $\mathrm{d}x/\mathrm{d}t = -\Lambda x(t) + \xi(t)$ and the relation $\Gamma = \Lambda/\beta$). For a system described by *many* variables, however, we need a set of CF's $\varphi_{ik}(t) = \langle x_i(t)x_k(0)\rangle$ and their FT's $P_{ik}^{(x)}(\omega)$. The *system* of differential equations for these CF's $\mathrm{d}\varphi_{ik}(t)/\mathrm{d}t = \sum_l \lambda_{il}\varphi_{ik}(t)$ is transformed into the system of linear algebraic equations for $P_{ik}^{(x)}(\omega)$. Solving this and analogous system for spectral powers $P_{ik}^{(X)}(\omega)$ of the conjugates, using of the relations between the FT's of noise components, initial and conjugate variables $\tilde{\xi}_i(\omega) = \sum_k \Gamma_{ik}\tilde{X}_k(\omega) - i\omega\tilde{x}_i(\omega)$ (which follows from $\mathrm{d}x_i/\mathrm{d}t = -\sum_k \Gamma_{ik}X_k(t) + \xi_i(t)$) we obtain the result $P_{ik}^{(\xi)}(\omega) = \Gamma_{ik} + \Gamma_{ki}$. Hence the required correlation properties of the random noise components are

$$\langle \xi_i(t)\xi_k(0)\rangle = (\Gamma_{ik} + \Gamma_{ki})\cdot\delta(t) \tag{58}$$

which is a direct generalization of the single-variable relation $\langle \xi(t)\xi(0)\rangle = 2\Gamma\cdot\delta(t)$.

The simplicity of the relation (58) which gives the correlation coefficients of the random noise matrix directly in terms of the elements of $\hat{\Gamma}$-matrix is the reason to use equation (54) where the stochastic motion of a system near its equilibrium state is described using the *conjugate* variables $\{\mathbf{X}\}$.

The relation (58) can be used to derive correlation properties of the random noise in a very important particular case, which includes also micromagnetic models. To do this, we begin with the important remark that our starting point, namely the relaxation equations (50) for the system variables are *not* the equation of motion derived from some physical formalism (like Newton laws or Lagrange mechanics), but are merely a direct consequence of the mathematical assumption that the relaxation rates $\mathrm{d}x_i/\mathrm{d}t$ of the system variables near its equilibrium state can be expanded in terms of small deviations $\{\mathbf{x}\}$ from their equilibrium values. Equations (50) maintain only the first-order terms of this expansion.

For many physical system the expansion coefficients Γ_{ik} may be given in a more specific form, which allows further progress by evaluating the correlation matrix $\langle \xi_i\xi_k\rangle = (\Gamma_{ik} + \Gamma_{ki})$. First we note, that the entropy change by the system deviation from its equilibrium state may be expressed via the *minimal work* $A_{\min}$ required to transfer the system from equilibrium into the state with the entropy S as $S - S_{\max} = -A_{\min}/kT$ (see, e.g., Landau and Lifshitz, 1980). This allows us to calculate conjugate variables as derivatives of this work: $X_i = -\partial S/\partial x_i = (1/kT)\cdot(\partial A_{\min}/\partial x_i)$.

Further, for a wide class of physical systems this minimal work $A_{\min}$ is equal to the energy difference $E - E_0$ between the system energy in the equilibrium state E_0 and in the given state E, which can be also expanded near the equilibrium state as

$$E - E_0 = \frac{1}{2} + \sum_{i.k} a_{ik}x_i x_k \tag{59}$$

so that

$$X_i = \frac{1}{kT}\frac{\partial A_{\min}}{\partial x_i} = \frac{1}{kT}\frac{\partial E}{\partial x_i} = \frac{1}{kT}\sum_k a_{ik}x_k \tag{60}$$

which means that the matrix $\{\beta_{ik}\}$ from the definition (52) in this case is $\hat{\beta} = \hat{a}/kT$.

If the deterministic motion of the system can be described by the Newtonian equations with the inertial term neglected, that is, in the form $\eta_i\cdot\mathrm{d}x_i/\mathrm{d}t = F_i$, then, evaluating the forces as $F_i = -\partial E/\partial x_i = -\sum_k a_{ik}x_k$, introducing particle mobilities as inverses of the corresponding friction coefficients $\kappa_i = 1/\eta_i$, and adding random forces to the right-hand side of the equation of motions, we obtain these equations in the form

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -\kappa_i\sum_k a_{ik}x_k + \xi_i \tag{61}$$

where the random forces ξ_i are related to the forces F_i used in (35) via $\xi_i = F_i/\kappa_i$. Comparing this system to the relaxation equations (51), we find that in this case $\hat{\Gamma} = \hat{\kappa}_{\mathrm{diag}}\hat{a}$, where the diagonal matrix $\hat{\kappa}_{\mathrm{diag}}$ contains mobilities κ_i on its main diagonal. Substituting $\hat{\Gamma} = \hat{\kappa}_{\mathrm{diag}}\hat{a}$ and $\hat{\beta} = \hat{a}/kT$ into the definition of the matrix $\hat{\Gamma} = \hat{\Lambda}\hat{\beta}^{-1}$, we arrive at the important result

$$\hat{\Gamma} = kT\cdot\hat{\kappa}_{\mathrm{diag}} \tag{62}$$

This means that for a system where (i) the minimal work required to bring it out of the equilibrium is equal to the corresponding energy change and (ii) the relaxation of the system coordinates can be expressed via the damped equations of motion, the correlation matrix of random forces acting on different variables

$$\begin{aligned}\langle \xi_i(t)\xi_k(0)\rangle &= (\Gamma_{ik} + \Gamma_{ki})\cdot\delta(t)\\ &= 2\kappa_i\cdot kT\cdot\delta_{ik}\cdot\delta(t)\end{aligned} \tag{63}$$

is *diagonal*. This property is *independent* on the specific expression of the interaction energy $E\{\mathbf{x}\}$ between the particles, which means that for such a system random noise components are uncorrelated despite the presence of an interparticle interaction.

3.3 Noise correlations for an interacting system: application to micromagnetic simulations

Using the formalism developed above, we shall rigorously demonstrate in this subsection that *physical* correlations between the random fields on *different* cells or between the random field components on *one and the same* cell are absent.

The formalism from the previous subsection may be applied directly to micromagnetic Langevin dynamics simulations of magnetization fluctuations in a thermodynamic equilibrium. Although the situation in micromagnetics is slightly more complicate than for systems discussed in Section 3.2 (due to the presence of a *precessional* term), a complete description of random field correlation properties is nevertheless possible.

For a micromagnetic system the variables $\{\mathbf{x}\}$, which determine its state are the magnetization projection. Here we consider a system which is already discretized into finite elements (cells) and denote the projections of the magnetization inside the i-th cell as $M_i^\alpha, \alpha = x, y, z$. The most important step now is the establishing of the corresponding conjugate variables. Comparing the definition of the deterministic effective field

$$\mathbf{H}_i^{\text{det}} = -\frac{1}{\Delta V_i}\frac{\partial E}{\partial \mathbf{M}_i}, \text{ or}$$
$$H_{i,\alpha}^{\text{det}} = -\frac{1}{\Delta V_i}\frac{\partial E}{\partial M_i^\alpha}(\alpha = x, y, z) \tag{64}$$

appearing in the LLG equation for a discretized system

$$\frac{\mathrm{d}\mathbf{M}_i}{\mathrm{d}t} = -\gamma\cdot\Big[\mathbf{M}_i \times (\mathbf{H}_i^{\text{det}} + \mathbf{H}_i^{\text{fl}})\Big] - \lambda\cdot\frac{\gamma}{M_{\text{S}}}\cdot\Big[\mathbf{M}_i \times \Big[\mathbf{M}_i \times (\mathbf{H}_i^{\text{det}} + \mathbf{H}_i^{\text{fl}})\Big]\Big] \tag{65}$$

with the definition (60) of the conjugate variables $X_i = (1/kT)(\partial E/\partial x_i)$, we immediately see that the variable X_i^α conjugate to the projection M_i^α is simply proportional to the corresponding effective field projection:

$$X_i^\alpha = -\frac{\Delta V_i}{kT}\cdot H_{i,\alpha}^{\text{det}} \tag{66}$$

The direct consequence of this proportionality is the absence of correlations between the random field projection on *different* discretization cells, because in the LLG equation of motion for the magnetization of the i-th cell only the effective field projections *for the same cell* (and hence – conjugate variables with the same index i) do appear. This means that the matrix elements $\Gamma_{ik}^{\alpha\beta}$ from the system (54) with different cell indices $i \neq k$ are automatically zero, ensuring the absence of intercell correlations according to (58).

To find out, whether any correlations between the random field components on *one and the same* cell do exist, some technical work should be done. To simplify the treatment, we assign to each cell its own coordinate system with the $0z$-axis parallel to the equilibrium direction of the cell moment. The quadratic energy expansion around the equilibrium magnetization state

$$E - E_0 = \frac{1}{2}\sum_{i,j}\sum_{\alpha,\beta=x,y} a_{ik}\cdot\Delta M_i^\alpha \Delta M_j^\beta \tag{67}$$

will then include small deviations $\Delta M_i^\alpha \ll M_i^z \approx M_{\text{S}}$ (from zero) of the x and y magnetization components only, because the magnitude of the cell magnetic moment should be conserved. The deterministic effective field evaluated according to (64) will have on each cell also only x and y components $H_i^{x(y)}$ which will be of same order of magnitude as ΔM_i^α. Writing the LLG equations of motion (65) for ΔM_i^x and ΔM_i^y, neglecting the terms $\Delta M_i^{x(y)}\cdot H_{i,z}^{\text{fl}}$ which are small compared to $M_i^z\cdot H_{i,x(y)}^{\text{fl}}$ (because $\Delta M_i^\alpha \ll M_i^z$) and linearizing the resulted equations with respect to small deviations $\Delta M_i^x \equiv M_i^x$ and $\Delta M_i^y \equiv M_i^y$, we obtain the system

$$\frac{1}{\gamma M_{\text{S}}}\frac{\mathrm{d}M_i^x}{\mathrm{d}t} = +(H_{i,\text{det}}^y + H_{i,\text{fl}}^y) + \lambda(H_{i,\text{det}}^x + H_{i,\text{fl}}^x) \tag{68}$$

$$\frac{1}{\gamma M_{\text{S}}}\frac{\mathrm{d}M_i^y}{\mathrm{d}t} = -(H_{i,\text{det}}^x + H_{i,\text{fl}}^x) + \lambda(H_{i,\text{det}}^y + H_{i,\text{fl}}^y) \tag{69}$$

It is evident from these equations that the coefficients $\Gamma_{ik}^{\alpha\beta}$ for $\alpha \neq \beta$ obey the relation $\Gamma_{ii}^{xy} = -\Gamma_{ii}^{yx}$, so that the cross-correlation coefficients of the fluctuation field projections on the given cell are identically zero: $C_{ii}^{xy} = C_{ii}^{yx} \sim (\Gamma_{ii}^{xy} + \Gamma_{ii}^{yx}) = 0$. This property, being obtained in our specific coordinate system, should remain the same in any other coordinates due to the space isotropy. *Hence there are no physical correlations neither between the random fields on different cells nor between the random field components on one and the same cell*. This result was obtained also in Chubykalo *et al*. (2003) with a somewhat more complicated method.

Note that this statement does not apply to artificial (having nonphysical nature) correlations between the random fields on different cells that appear due to the finite-element discretization of an initially continuous problem. This topic is discussed in the next subsection.

* * *

A very interesting theme is the discussion of the correlation properties of the random thermal fields in micromagnetics

from a point of view of the fluctuation-dissipation theorem (FDT). Due to the space limitations we preferred not to include this topic into the current review, because it requires a careful and detailed discussion supplemented by a substantial amount of the readers theoretical knowledge. A very transparent and clear discussion of this topic at a high scientific level can be found in Smith (2001).

4 DISCRETIZATION EFFECTS IN DYNAMIC MICROMAGNETIC SIMULATION

4.1 Discretization effects for $T = 0$

In this subsection, we discuss the influence of a finite-element representation of the continuous micromagnetic problem (discretization) on the magnetization dynamics observed in numerical simulations performed *without* taking into account thermal fluctuations, that is, for $T = 0$.

To demonstrate the importance of the discretization effects we have chosen the following problem: we study the switching dynamics of a nanoelement with lateral sizes $L_x \times L_z = 400 \times 600\,\text{nm}$, thickness $h = 5\,\text{nm}$ (our $0xz$-plane coincides with the element plane), $M_S = 1000\,\text{G}$ and exchange stiffness $A = 10^{-6}\,\text{erg cm}^{-1}$. To simplify our task, we have set the magnetocrystalline anisotropy to zero. The switching of this element is simulated integrating the LLG equation (1) using an optimized Bulirsch–Stoer algorithm with the adaptive step-size control. We start from the S-type remanent state applying at $t = 0$ the external field $\mathbf{H} = H_z\mathbf{e}_z$ with $H_z = -200\,\text{Oe}$; this field is well beyond the corresponding quasistatic switching field $H_{sw} \approx -80\,\text{Oe}$. To study the discretization effects *s*imulations were done for five sequentially refined grids ($N_x \times N_z = 40 \times 60, 60 \times 90, 80 \times 120, 120 \times 180, 200 \times 300$) with the same (1:1) aspect ratio of the grid cell.

The switching process for the most interesting low damping case $\lambda = 0.01$ is shown in Figure 1(a) (grid $N_x \times N_z = 120 \times 180$): it starts with the reversal of closure domains near the short element borders, proceeds via the reversal of the central domain and is completed by the 'flip' of narrow domains near the long sides.

To emphasize the importance of the discretization effects we have compared results for several grids listed above. In Figure 1(b) we present corresponding $m_x(t)$ dependencies, because in our geometry the influence of the discretization effects can be most clearly seen on this projection. It can be clearly seen that the remagnetization curves do *not* converge to any limiting curve up to the finest grid $N_x \times N_z = 200 \times 300$. The effect is even *qualitative*, as it can be seen from the comparison of final states (m_x-gray-scale maps on the right in Figure 1(b)) for all discretizations $N_x \times N_z \leq 120 \times 180$ and for $N_x \times N_z = 200 \times 300$.

This discrepancy can not be attributed to an insufficient discretization of the interaction (energy) terms, because already for a moderate grid $N_x \times N_z = 80 \times 120$ the cell sizes $\Delta_x = \Delta_z = 5\,\text{nm}$ are smaller than our characteristic micromagnetic length $l_{\text{dem}} = (A/M_S^2)^{1/2} = 10\,\text{nm}$. We have also verified starting from the grid $N_x \times N_z = 60 \times 90$ quasistatic hysteresis loops did not change when the grid was refined further.

The reason for a significant modification of the switching process by the grid refinement is a strong influence

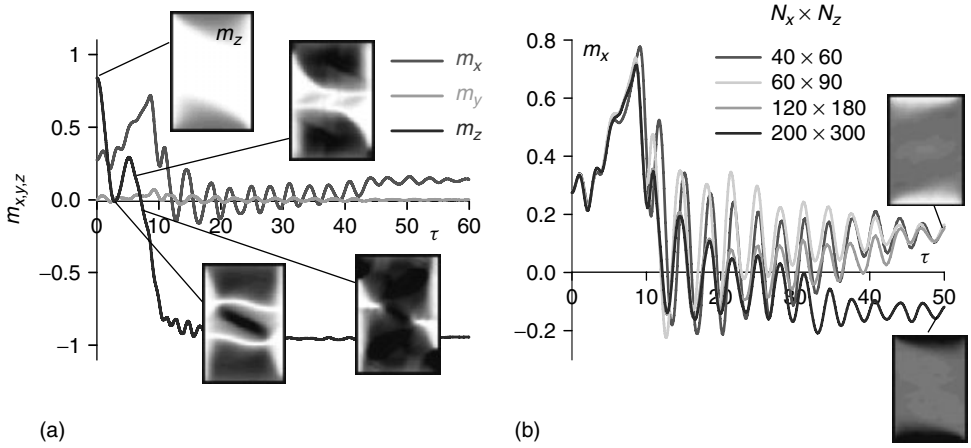

Figure 1. Switching of a thin 'soft' magnetic element ($400 \times 600 \times 5\,\text{nm}$) with a *low* damping ($\lambda = 0.01$) in a field $H_z = -200\,\text{Oe}$ starting from the S-type remanent state: (a) time-dependencies of all magnetization projections for the discretization $N_x \times N_z = 120 \times 180$($m_z$-grey-scale maps for several times $\tau = t\gamma M_S$ are shown). (b) m_x time dependencies simulated using various grids as shown in the legend. (Reproduced from D.V. Berkov *et al*., 2002. © 2002 with permission from IEEE.)

of magnetic excitations with a short wavelength. For the static case, it is sufficient to discretize the system using the mesh size somewhat smaller than the characteristic (exchange or demagnetizing) magnetic length of the material. For dynamics it is, generally speaking, not true. Even if the mesh is fine enough to represent all the features of the starting magnetization state, during the remagnetization process magnons with the wavelengths shorter than the grid cell size may play an important role, so that magnons with decreasing wavelengths appear when the remagnetization proceeds. As soon as the grid is unable to support these magnons, simulations became inadequate (Berkov, 2002a). This means that in such situations *dynamical* simulations on a given lattice are valid up to some *maximal* time duration.

The problem emphasized in this subsection requires further thorough investigation, because recently the so-called spin-injection driven magnetization dynamics was predicted theoretically and discovered experimentally (see the review of Miltat, Albuquerque and Thiaville, 2001 in this Volume). It has been shown that when the magnetization dynamics is excited by a spin polarized current, both (i) large deviations of the magnetization from its equilibrium position and (ii) quasichaotic magnetization dynamics (where spatial variations of the magnetization are very fast) are possible and even represent quite common features of corresponding remagnetization processes. This rapidly developing topic with potentially very rich applications makes the study of numerical artifacts due to the interplay of the discretization grid with short-wavelength magnons really important.

For systems with *larger* dissipation the effect demonstrated above may be absent due to a much smaller decay times of the short-wave magnons (so that they do not play any significant role by switching).

4.2 Influence of the discretization on the random field correlations

We proceed with the consideration of discretization effects on the magnetization dynamics simulated at finite temperatures, that is, with the fluctuation field $\mathbf{H}^{\mathrm{fl}}(\mathbf{r}, t)$ included into the LLG equation. The standard assumption (29) that this field is δ correlated in space and time may become invalid due to the following effect: as pointed above, by discretizing a continuous magnetic film we exclude all magnons with the wavelength smaller than the grid cell size $\Delta = \min(\Delta x, \Delta z)$. However, these magnons can still have a mean free path *much larger* than the grid cell size, thus causing substantial correlations especially of the exchange fields on neighboring cells. Although these excitations cannot be included into simulations on the given grid explicitly, it is possible to take them into account as an additional contribution to the fluctuation field $\mathbf{H}^{\mathrm{fl}}$ with the corresponding correlation properties.

To compute the correlation function (CF) of the effective field produced by such short-wave magnons, we have first performed simulations (with the white noise only) solving LLG equations at the grid which was finer than the 'actual' grid intended to be used for final simulations. Then from the total effective field produced at this fine grid all contributions with the wavelength *larger* than the cell sizes of the 'actual' grid Δ were cut out, so that only magnons with short wavelengths $\lambda_{\mathrm{mag}} \leq \Delta$ remained. Afterwards magnetic field generated by these short-wave magnons was calculated and averaged over all subcells inside the given cell of the actual simulation grid. Finally, the correlation properties of *this averaged field* were evaluated (Berkov and Gorn, 2004).

The resulting CF has a quite complicated form both in space (Figure2a) and time (Figure 2b) and can be roughly described as exponentially decaying oscillations. Both the

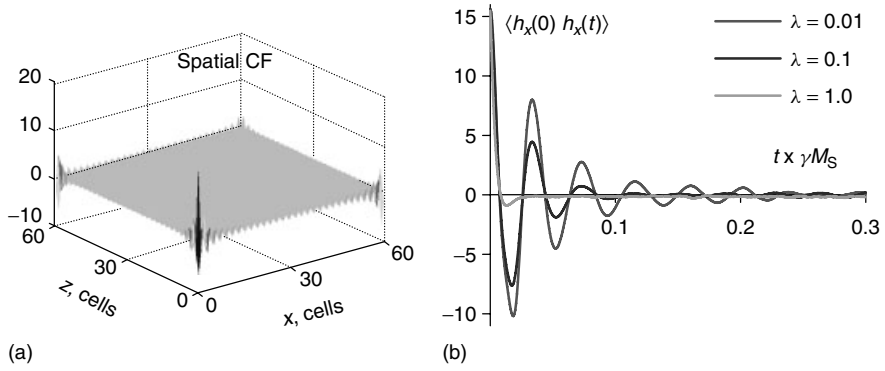

Figure 2. Correlation functions (CF) of the x component of the random field $\mathbf{H}^{\mathrm{fl}}$ resulting from the short-wave magnons. (a) 2D spatial CF, (b) Temporal CF on one and the same cell for various dissipation constants λ. (Reproduced from Berkov *et al*., 2004, with permission from Elsevier. © 2004.)

time and space oscillation periods are determined by the corresponding properties of the magnons with the shortest wavelength available for the 'actual' grid. The *space* decay length depends mainly on the 'quasistatic' magnetic system parameters like the saturation magnetization and exchange constant, whereas the decay *time* is determined by the damping λ.

To take into account the influence of the short-wave magnons which are cut off by the given grid, one should, strictly speaking, perform simulations on this grid by solving the SDE system (20) with both the standard white noise term and an additional colored noise with correlations imposed by *the field of these short-wave magnons*. For this purpose it is necessary to implement an algorithm, which could generate a Gaussian random noise with any given correlation function.

Several methods for generation of such a noise are available. Matrix methods (James, 1980) can generate random sequences with arbitrary given CF, but are very time and storage consuming. Linear Langevin equation (Garcia-Ojalvo and Sancho, 1994) or the so-called 'physical' methods (simulation of a simplified system without long-range interactions) are fast, but can generate only noise with monotonously or regularly oscillating exponentially decreasing CFs. These methods may be in principle applied to micromagnetic systems, because correlations caused by the short-wave magnons are mostly oscillating and exponentially decreasing.

However, correlation functions like those shown above still exhibit significant irregularities, which cannot be reproduced by 'physical' method. In this case, some version of a spectral method (Romero and Sancho, 1999) should be used, because this method can generate arbitrarily correlated random numbers. The method is based on the usage of the temporal and spatial FT of the required correlation function which after the multiplication with the so-called anticorrelated (in the Fourier space) random numbers gives the Fourier image of the random number field with requested correlations. Their inverse FT provides the requested random numbers themselves (Romero and Sancho, 1999). The important disadvantage of the method is that the complete correlation matrix should be stored (for a spatially 2D system our the dimension of the this matrix would be $N_x \times N_z \times L_t$, where L_t is the number of time steps), requiring large memory resources. For this reason, in most cases the usage of so-called 'external storage' of the Fourier transform is unavoidable (Press, Teukolsky, Vetterling and Flannery, 1992).

4.3 Discretization and the density of magnon states

Finite-element discretization of a micromagnetic system *qualitatively* affects the spectral power its thermal excitations, so this question should be addressed here in detail. To present the main point as transparent as possible, we consider a simplest model: a square region of an extended thin film (periodic boundary conditions are assumed) in an external field perpendicular to the film plane and neglect anisotropy and magnetodipolar interactions.

A spectrum of *equilibrium* thermal magnetic excitations can be efficiently computed using the Langevin dynamics (Berkov and Gorn, 2005). Starting from the saturated state along the external field we integrate the SDEs (20) till the total energy does not change systematically with time (equilibrium is reached). From this time moment we save the trajectories of every cell magnetization during a sufficiently long time (which depends on the desired accuracy and

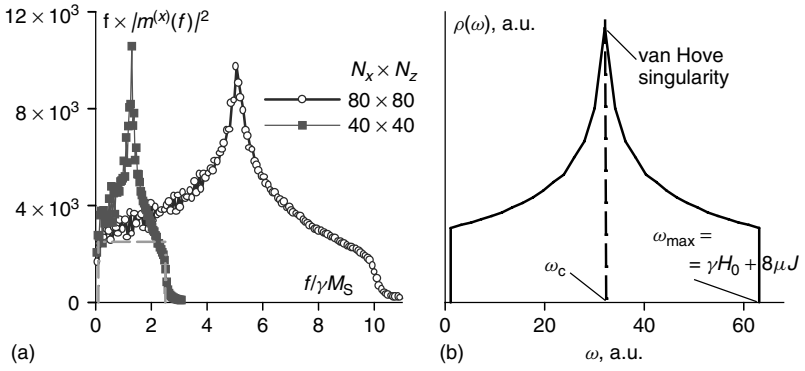

Figure 3. (a) Power spectrum of m_x-oscillations for various discretizations. Cusps are the manifestations of the magnon DoS singularity (right panel). The dashed rectangle represents the spectrum expected in the limit of very fine discretization for a low frequency region available for the 40×40 discretization; (b) density of magnon states for the model (70). The cusp by $\omega_c = \omega_{min} + \omega_{max})/2$ is the van Hove singularity common for all 2D models with the cosine like $\omega(\mathbf{k})$ dependence. (Reproduced from Berkov *et al*., 2005, with permission from Elsevier. © 2005.)

frequency resolution of the spectral power finally obtained). Finally we perform the temporal FT of these trajectories and averaging over several thermal noise realizations.

Corresponding results for a thin-film region with lateral sizes 400×400 nm, thickness $h = 5$ nm, $M_S = 1000$ G, $A = 10^{-6}$ erg cm^{-1} placed in an external field $H_{ext} = 100$ Oe are presented in Figure 3(a), where oscillation power spectra of m_x projection at $T = 10$ K are shown for two different discretizations. The most striking features of these spectra are (i) sharp cusps in the middle and (ii) a shift of this cusp towards higher frequencies when the grid is refined.

This result can be easily understood as follows. After the in-plane discretization of the film into $N_x \times N_y$ cells with the sizes Δx and Δy and volumes ΔV the system energy $E = E_{\text{ext}} + E_{\text{exch}}$ is converted into the sum over cells i, j

$$E = -\mu \sum_{i,j=1}^{N_{x(y)}} \mathbf{m}_{ij} \mathbf{H}_{ij}^{\text{ext}} - \frac{1}{2} J \mu^2 \sum_{\langle i,j \rangle} (\mathbf{m}_i \mathbf{m}_j) \qquad (70)$$

where μ is the cell magnetic moment. The exchange constant J in (70) depends on the exchange stiffness A and the grid cell parameters as

$$J = \frac{A}{M_S^2 \cdot \Delta V} \left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} \right) \qquad (71)$$

As usual, the total oscillation power for the given frequency ω in a thermal equilibrium is directly proportional to the number of modes contributing to this frequency, that is, to the magnon density of states $\rho(\omega)$. For a typical *lattice* model described by the energy (70) this density of states is well known. Namely, the quadratic expansion of (70) over small magnetization deviations from the ground state (small temperature or large external field limit) leads to the eigenfrequencies $\omega_{pq} = \gamma \cdot (H_{\text{ext}} + \mu J \cdot f_{pq})$ depending on the eigenmodes wave vector indices p and q via the sum of cos functions as $f_{pq} = 2 \cdot (2 - \cos(2\pi p / N_x) - \cos(2\pi q / N_z))$, which is common for all 2D lattice models with the nearest-neighbors harmonic interaction. If the eigenfrequencies depend on the wave vectors in a cosine like manner, then the density of states (Figure 3b) contains the famous van Hove singularity in its middle, which is clearly visible in both spectra in Figure 3(a) as a cusp. The spectrum shift toward higher frequencies when the discretization is refined follows simply from the fact that the eigenfrequencies $\omega_{pq} = \gamma \cdot H_{\text{ext}} + \mu J \cdot f_{pq})$, being proportional to the exchange constant J, increase according to (71) as an inverse square of a cell size when a mesh is refined.

The actual excitation spectrum of a real system (which we attempt to simulate) also contains such a cusp (a real system is discrete at the atomic level) but for frequencies determined by the interatomic distances and thus absolutely unavailable for simulations. This means that the correct spectrum of the continuous thin film model in the frequency region available for micromagnetic simulations is nearly flat as shown by the dashed rectangle in Figure 3(a). Hence, in order to obtain adequate results for equilibrium system properties using such simulations, one should either work in the frequency region where the spectrum is still approximately flat ($\omega \ll \omega_c$) or use a *colored* noise to correct the excitation spectrum of the corresponding lattice model.

5 MAGNETIZATION RELAXATION OVER HIGH ENERGY BARRIERS

For system with high energy barriers $\Delta E \gg kT$ direct simulation of the magnetic moment trajectories using the Langevin dynamics is fairly impossible. Such simulations simply mimic the time-dependent system behavior so that the simulation time necessary to overcome the barrier *exponentially* growths with its height following the Arrhenius–Van't Hoff law (probability to overcome the barrier is $p \sim \exp(-\Delta E / T)$)–exactly as for real systems.

Nevertheless, methods for numerical simulations allowing to study transitions over large barriers are highly desirable from the practical point of view: they are the only way to predict the long-time stability of the information storage devices. To evaluate the transition probability p over such barriers we must in the first place find the *lowest* saddle point between the two metastable state of interest. Its height gives us the corresponding energy barrier ΔE between these states, allowing to estimate p from the Arrhenius–Van't Hoff law. Analytical methods for the saddle-point search exist only for relatively simple magnetic systems (Braun, 1994, 2000; Klik and Gunther, 1990). In principle, such a saddle point can be found by solving a system of nonlinear equations $\partial E / \partial x_i = 0$ (where x_i denote the variables of a system configuration space), because at a saddle point *all* energy derivatives $\partial E / \partial x_i$ should be zero, but neither an energy maximum nor a minimum should be achieved. However, general methods for the solution of such systems are not available, and there exist even arguments that there will never be any (Press, Teukolsky, Vetterling and Flannery, 1992). For this reasons numerical methods based on other principles are required.

5.1 Time-temperature scaling method

The time-temperature scaling method (Xue and Victora, 2000) quantifies the rough idea that in some cases simulations of the transition over high energy barriers involving

macroscopically long waiting times at *low* temperatures can be replaced by simulations over the same energy barrier, but at a much *higher* temperature so that the transition time (and hence – the simulation time) is much smaller and is accessible for simulations.

To make this idea applicable in numerical simulations, we need the quantitative relation between the time and temperature scales. To establish such a relation, we start from the simplest version of the Arrhenius–Van't Hoff law which states that the average transition time τ_{av} depends on the energy barrier height ΔE and the system temperature T mainly exponentially via their ratio as

$$\tau_{\mathrm{av}} = \frac{1}{\nu_0} \exp\left(\frac{\Delta E}{kT}\right) \tag{72}$$

where the prefactor is defined using the so-called 'attempt frequency' ν_0. If the temperature dependence of this frequency is weak compared to the exponent $\exp(-\Delta E/kT)$ then the product $kT{\cdot}\log(\tau_{\mathrm{av}}{\bullet}\nu_0)$ remains constant for the transition over this barrier. This means that if a transition over some barrier takes on average a long time τ_{long} at a (low) temperature T_{long}, then in order to observe the same transition during a desired short time τ_{short} we need to increase the temperature up to the value $T_{\mathrm{s/l}}$ which is related to the quantities introduce above via

$$T_{\mathrm{s/l}} \log(\nu_0 \tau_{\mathrm{short}}) = T_{\mathrm{long}} \log(\nu_0 \tau_{\mathrm{long}}) \tag{73}$$

To calculate from this equation, the scaled temperature $T_{s/l}$ which we should use in simulations if we would like to reduce our simulation time from the inaccessible value τ_{long} down to τ_{short}, we need to determine the attempt frequency ν_0 for the system under study. Analytical formulae for ν_0 are available only for the simplest systems like a single-domain particle (Brown, 1963b). For this reason Xue and Victora (2000) have proposed the following trick. They have introduced a new time τ_{ref} which is much larger than the short time τ_{short} (which we would like to use for final simulations), but still small enough so that simulations during this time are possible and the desired transition occurs during the time τ_{ref} at some intermediate temperature T_{ref}. The first step for the determination of the attempt frequency ν_0 (and hence – the temperature $T_{\mathrm{s/l}}$) is the simulation of a system at the temperature T_{ref} during the time τ_{ref}, whereby some physical property of the system is determined or some dynamical process in the system is recorded. Then one should perform several simulation runs at different temperatures during the short time τ_{short} and *find the temperature* $T_{\mathrm{s/r}}$ for which the process recorded at T_{ref} during the time τ_{ref} *proceeds as similar as possible* to the process observed during the short time τ_{short}. This means, that the corresponding times and temperatures are connected via the same relation as (73), namely

$$T_{\mathrm{s/r}} \log(\nu_0 \tau_{\mathrm{short}}) = T_{\mathrm{ref}} \log(\nu_0 \tau_{\mathrm{ref}}) \tag{74}$$

This latter relation can be used to extract the attempt frequency ν_0, because all other quantities here are known. Determination of ν_0 from (74) and its substitution into (73) leads to the following expression for the required high simulation temperature $T_{\mathrm{s/l}}$:

$$T_{\mathrm{s/l}} = T_{\mathrm{long}} + (T_{\mathrm{s/r}} - T_{\mathrm{ref}}) \cdot \frac{T_{\mathrm{long}} \log(\tau_{\mathrm{long}}/\tau_{\mathrm{short}})}{T_{\mathrm{ref}} \log(\tau_{\mathrm{ref}}/\tau_{\mathrm{short}})} \tag{75}$$

Simulations at this temperature during the time τ_{short} should now reproduce the behavior of the system under study at the low temperature T_{long} during the time τ_{long} which is exponentially larger than τ_{short} due to the relation (73).

Xue and Victora have applied their algorithm to the simulations of the hysteresis loops at various field sweep rates R (which served as inverse time scales τ_{short} etc.). A remarkable agreement between the two numerically calculated loops for the sweep rates 0.5 and 50 Oe nsec^{-1} was obtained and the loop measured experimentally at $R = 50\,\mathrm{Oe\,sec^{-1}}$ (i.e., nine orders of magnitude slower) could be successfully predicted (Xue and Victora, 2000). They have also simulated the process of a bit decay in magnetic recording media (Xue and Victora, 2001) over a macroscopically long time scale, which is highly interesting for the development of high-density magnetic storage.

Concluding this subsection, we note that the method outlined above probably is not able to reproduce correctly the magnetization dynamics for a system with *low* dissipation, when the precession term in the LLG equation is really important (results of Xue and Victora were obtained on systems with moderate damping). The reason is that actual simulations in this method are performed at temperatures much higher that the actual system temperature, so that the relation between the random (fluctuation) field and the deterministic field is wrong. Another limitation of this formalism is the usage of the relation (72), which is valid only when the entropic contribution to the transition probability (curvature of the energy landscape in the vicinity of a saddle point and energy minima) can be neglected.

5.2 Rigorous evaluation methods for the energy barrier height

In this subsection, we describe general numerical methods for the evaluation of the energy barrier height between the two metastable states. All methods aim to evaluate some

kind of an optimal trajectory between these states and the required energy barrier can then be calculated as the barrier along this trajectory. The transition probability between the energy minima in question can then be evaluated using the general transition-rate theory (Hoenggi, Talkner and Borkovec, 1990), whereby the features of the energy surface near critical points may be also taken into account.

5.2.1 *Minimization of a thermodynamical action (Onsager–Machlup functional)*

General idea

This method is based on the search for the *most probable transition path* between the two energy minima by minimizing the corresponding *thermodynamical action* derived from the path-integral formulation of the problem. The underlying idea (Onsager and Machlup, 1953) can be explained considering a system of N particles with coordinates $x_i (i = 1, \ldots, N)$ and the interaction energy $V(\mathbf{x})(\mathbf{x} = (x_1, \ldots, x_N))$ in a viscous fluid. Langevin equations for this system are

$$\dot{x}_i = -\frac{\partial V(\mathbf{x})}{\partial x_i} + \xi_i(t), \quad i = 1, \ldots, N \tag{76}$$

where we have neglected the inertial term for simplicity and absorbed the friction constant into the time scaling. Langevin forces ξ_i are again assumed to be independent Gaussian δ-correlated random variables: $\langle \xi_i(0)\xi_j(t)\rangle = 2D\delta_{ij}\delta(t)$.

Due to these simple correlation properties the probability of some *particular* noise realization $\{\xi_i(t)\}, i = 1, \ldots, N$ for the time period $[0, t_\mathrm{f}]$ is (Onsager and Machlup, 1953); (Bray and McKane, 1989)

$$P[\boldsymbol{\xi}(t)] = A \exp\left[-\frac{1}{4D}\int_0^{t_\mathrm{f}} \sum_i \xi_i^2(t)\mathrm{d}t\right] \tag{77}$$

Rewriting the system (76) as $\xi_i(t) = \mathrm{d}x_i/\mathrm{d}t + \partial V(\mathbf{x})/\partial x_i$ and introducing the Jacobian $J[\mathbf{x}(t)]$ of the transformation $\mathbf{x} \to \boldsymbol{\xi}$, we immediately obtain that the probability to observe a *given trajectory* $\mathbf{x}(t)$ for the transition $A \to B$ during the time t_f $(\mathbf{x}_A(0) \to \mathbf{x}_B(t_\mathrm{f}))$ is

$$P[\mathbf{x}(t)] \sim J[\mathbf{x}] \exp\left[-\frac{S(\mathbf{x}(t), t_\mathrm{f})}{4D}\right] \tag{78}$$

where the *thermodynamical action* $S(\mathbf{x}(t))$ is defined as

$$S(\mathbf{x}(t), t_\mathrm{f}) = \int_0^{t_\mathrm{f}} \mathrm{d}t \sum_i \left(\frac{\mathrm{d}x_i}{\mathrm{d}t} + \frac{\partial V(\mathbf{x})}{\partial x_i}\right)^2 \tag{79}$$

It is obvious that the trajectory which minimizes the action $S(\mathbf{x}(t))$ provides the most probable (optimal) transition trajectory $\mathbf{x}_\mathrm{opt}(t)$, along which the energy barrier for this transition can be found: $\Delta E(A \to B) = E_\mathrm{max}(\mathbf{x}_\mathrm{opt}) - E_A$. The minimization of the functional $S(\mathbf{x}(t))$ can be performed using various numerical methods which employ the discretization of the transition path $\mathbf{x}(t)$ thus reducing the task of minimizing (79) to the problem of the minimization of a many-variable function (see Berkov, 1998 for detail).

Unfortunately, the minimization of the functional (79) by itself, being technically quite complicate, does *not* represent a main problem when searching for an optimal physical transition path. The main problem is the presence of many 'false' local minima of the functional (79), that is, the existence of many trajectories between the states A and B *which minimize* (79) *but do not provide any information about the corresponding energy barriers*.

To explain why this is almost always the case we note that for *any* path for which the conditions $\mathrm{d}x_i/\mathrm{d}t = \pm\partial V(\mathbf{x})/\partial x_i$ are fulfilled (the plus/minus sign correspond to the downhill/uphill trajectory parts) provides an extremum to the action (79) (Bray and McKane, 1989). This means that the extremal trajectories for the action functional (79) go *along the gradient lines* of the energy surface.

On a very simple 2D energy landscape shown in Figure 4(a) *both* the solid line $M_1 \to M_2$ and the dashed line $M_1 \to P_2 \to M_2$ deliver local extrema to the action for the transition $M_1 \to M_2$, because both paths proceed along the gradient lines of the energy surface. Moreover, these both extrema are local *minima* of the action; in fact, they were obtained by minimizing (79) with the potential shown in Figure 4 as a gray-scale map simply starting from different initial trajectories. However, the solid line trajectory passes through the saddle point, giving the correct energy barrier height (the 'true' optimal trajectory), whereas the dashed line $M_1 \to P_2 \to M_2$ goes via the energy *maximum* supplying no useful information whatsoever ('false' optimal trajectory). The next example shown in Figure 4(b) demonstrates that even the *value* of the action (given by the sum of heights which an optimal trajectory has to climb over) along the 'false' optimal path $(M_1 \to P_2 \to M_2)$ may be smaller than the corresponding value along the 'true' optimal trajectory. Hence without a *reliable* algorithm able to distinguish between these two kinds of optimal paths the whole method is absolutely useless, because the number of 'false' optimal trajectories exponentially growth with the complexity of system.

An apparently straightforward possibility to discriminate between these two cases is the analysis of the *curvature tensor* of the energy surface at the points where the energy along the optimal trajectory has local *maxima*: if the corresponding matrix of the second energy derivatives has exactly *one*

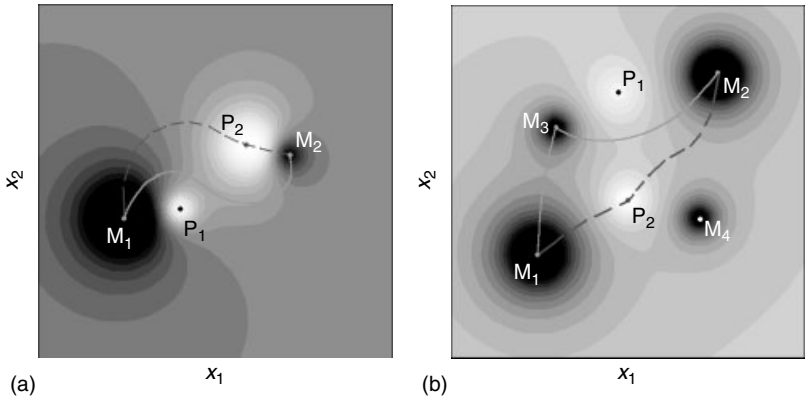

Figure 4. (a) 'True' (solid line) and 'false' (dashed line) optimal trajectories for a simple energy landscape. (b) An example of an energy landscape, where the action (79) along the 'false' optimal path $M_1 \to P_2 \to M_2$ may be even smaller than along the 'true' path $M_1 \to M_3 \to M_2$ (see text for details). (Reproduced from Berkov *et al.*, 1998, with permission from Elsevier. © 1998.)

negative eigenvalue, then this point indeed corresponds to a transition saddle. This method, however, is not sufficiently reliable due to a discrete representation of the continuous trajectory and a finite accuracy by the determination of an 'optimal' path. An alternative algorithm based on small 'jumps' away from the trajectory point with the highest energy in a random direction and subsequent minimization of the system energy starting from this new position, is described in detail in (Berkov, 1998).

Another problem arises due to the presence of the transition time t_f in the action (79) as the upper integral limit, which should be known in advance to set the time step and/or the number of time slices in the discretized action version; this transition time is of course not known. Fortunately, the barrier *height* determined from the discretized action turned out to depend on the t_f-value only slightly. For this reason sufficiently accurate results could be obtained by minimizing the discretized action using the small constant time step and simply doubling the number of time slices until the relative difference between the two barriers heights obtained for the subsequent action minimizations becomes less than a certain small threshold.

Implementation for magnetic systems

To apply this method to systems of interacting magnetic moments we have to start with the magnetic counterpart to the Langevin equation (76), namely, with the stochastic Landau–Lifshitz–Gilbert equation of motion for magnetic moments (20). The precession term in this equation affects, of course, the optimal transition *trajectory*, but does not change the system energy and hence it is reasonable to assume that it does not change the barrier *height* for this transition (see also our discussion of the string method below). For this reason we neglect the precession term in (20) and obtain the equation of motion for the magnetization unit vectors $\mathbf{m}_i$

$$\frac{\mathrm{d}\mathbf{m}_i}{\mathrm{d}t} = -\Big[\mathbf{m}_i \times \Big[\mathbf{m}_i \times (\mathbf{h}_i^{\mathrm{det}} + \mathbf{h}_i^{\mathrm{fl}})\Big]\Big] = -\mathbf{m}_i \cdot (\mathbf{m}_i \cdot \mathbf{h}_i^{\mathrm{tot}}) + \mathbf{h}_i^{\mathrm{tot}} \tag{80}$$

where all constants are again absorbed in the time unit, the total field is $\mathbf{h}^{\mathrm{tot}} = \mathbf{h}^{\mathrm{det}} + \mathbf{h}^{\mathrm{fl}}$ and the normalization $m_i = 1$ was used by the last transformation.

The conservation of the magnetic moment magnitude enforces the transition to spherical coordinates (θ, ϕ) of $\mathbf{m}$, because only the random field components perpendicular to $\mathbf{m}$ should be taken into account. Transforming all vectors to the new coordinates with the z'-axis along $\mathbf{m}$ and the x'-axis in the meridian plane of the initial spherical coordinates (so that in the initial system $m_x = \sin\theta\cos\varphi$, $m_y = \sin\theta\sin\varphi$, $m_z = \cos\theta$), we obtain equations of motion for the magnetization angles

$$\frac{\partial\theta_i}{\partial t} = -\frac{\partial E\{\Omega\}}{\partial\theta_i} + h^{\mathrm{fl}}_{i,x'},$$
$$\sin\theta_i \cdot \frac{\partial\phi_i}{\partial t} = -\frac{1}{\sin\theta_i} \cdot \frac{\partial E\{\Omega\}}{\partial\phi_i} + h^{\mathrm{fl}}_{i,y'} \tag{81}$$

where $h^{\mathrm{fl}}_{x'}$ and $h^{\mathrm{fl}}_{y'}$ are Cartesian components of the fluctuation field $\mathbf{h}^{\mathrm{fl}}$ in the new coordinate system. Deterministic effective field $\mathbf{h}^{\mathrm{det}}$ is already contained in corresponding angular derivatives of the magnetic energy E{Ω} (where {Ω} denotes the set of all angles (θ_i, φ_i)), which may include also the interaction energy of any kind (i.e., exchange, dipolar, RKKY, etc.).

The system (81) is fully analogous to (76) so that under the same assumptions (Langevin field components are independent Gaussian δ-correlated random variables) the magnetization path in the Ω space which minimizes the

thermodynamical action S for a magnetic system

$$S[\Omega(t)] = \int_0^{t_f} \mathrm{d}t \sum_i \left[\left(\frac{\mathrm{d}\theta_i}{\mathrm{d}t} + \frac{\partial E\{\Omega\}}{\partial \theta_i} \right)^2 + \left(\sin\theta_i \cdot \frac{\mathrm{d}\phi_i}{\mathrm{d}t} + \frac{1}{\sin\theta_i} \cdot \frac{\partial E\{\Omega\}}{\partial \phi_i} \right)^2 \right] \qquad (82)$$

provides the information about the energy barrier separating the states Ω_A and Ω_B.

Application for magnetic nanocomposites

With this method we have calculated the distribution of the energy barriers in a system of single-domain magnetic particles (embedded in a nonmagnetic matrix) with the uniaxial anisotropy and magnetodipolar interaction between the particles. The most intriguing question for this system is the influence of the magnetodipolar interaction on the distribution density of the energy barriers $\rho(E)$ (Hansen and Morup, 1998; Dormann, Fiorani and Tronc, 1999), which controls both the reversible and irreversible thermodynamics of the system. To solve this question, we have computed $\rho(E)$ for various volume concentrations of the magnetic phase, thus varying the interaction strength.

Calculations were performed for systems with high ($\beta = 2K/M_S^2 = 2.0$), moderate ($\beta = 0.5$) and low ($\beta = 0.2$) single-particle anisotropies. The energy barrier distributions were accumulated from $N_{conf} = 8$ realizations of the particle disorder; for each configurations about $N_{trans} = 200$ transitions between metastable states were analyzed. Corresponding results are shown in Figure 5 were the distribution of the reduced energy barriers $\varepsilon = E/M_S^2 V$ are presented.

First of all, it can be seen that for low particle concentrations ($\leq 1\%$)$\rho(\varepsilon)$ consists of the relatively narrow peak positioned at the value corresponding to the energy barrier $\varepsilon_{sp} = \beta/2$ for a single particle moment flip, as it should be for a weakly interacting system. The position of this single-particle flip barrier is shown both in Figure 5 with the dashed line. As expected, with increasing concentration the energy barrier density broadens, but for the systems with the low and high anisotropy this broadening occurs in a qualitatively different ways. For the high-anisotropy case (Figure 5, right column) the broadening of $\rho(\varepsilon)$ with increasing concentration is accompanied by its shift toward *lower* energy barriers, so that already for moderate particle concentration ($\geq 4\%$) almost all barriers lie below the value for a single particle.

For the system of particles with the low anisotropy (Figure 5, left column) barriers both higher *and* lower than a single particle barrier arise. However, the overall

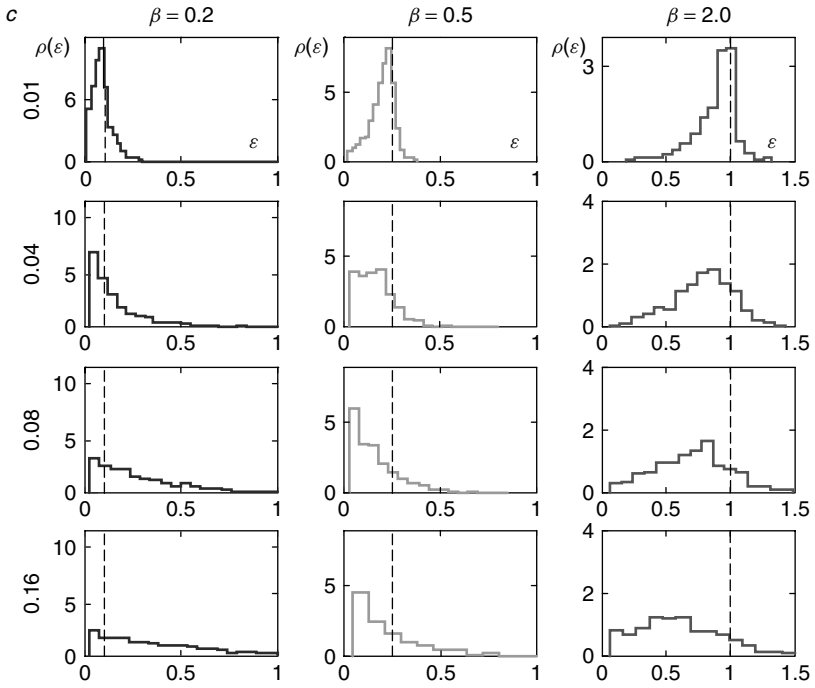

Figure 5. Density of the energy barriers for transitions between randomly chosen energy minima in a disordered magnetic particle system with the low ($\beta = 0.2$, left column), moderate ($\beta = 0.5$, middle column) and high ($\beta = 2.0$, right column) single-particle anisotropies for various particle volume fractions c. Dashed lines show positions of the energy barriers for a single particle with the corresponding anisotropy.

energy barrier spectrum clearly shifts toward *higher* energies with increasing particle concentration. Detailed physical explanation of this behavior can be found in Berkov (2002b).

It is also important to keep in mind that different transitions cause different moment changes. The key question is whether the magnitude of the moment changes is *correlated* with the height of the corresponding energy barrier. If, for example, the moment change tends to zero when the energy barrier height for this particular transition decreases, small energy barriers would not play any significant role in the system thermodynamics, because corresponding magnetization changes would be nearly undetectable. For this reason we need a 2D *mutual* distribution of the energy barriers and moment changes $\rho(E,\Delta m)$. Corresponding contour plots for a system with the low anisotropy $\beta = 0.2$ and two different concentrations are shown in Figure 6. For the low concentration $c = 0.01$ the density $\rho(E, \Delta m)$ consists of a single sharp peak positioned near the point ($\varepsilon = 0.1$, $\Delta m = 2.0$), which corresponds to a single-particle flip. From $\rho(E, \Delta m)$ for the high concentration $c = 0.16$ it can be seen that, although the moment changes for the low barriers are concentrated at somewhat smaller values than Δm for the higher ones, they do *not* tend to zero. Hence all transitions provide approximately equivalent contributions to the system thermodynamics.

5.2.2 The string method

General idea

To explain the main idea of the string method (Ren and Vanden-Eijnden, 2002), we start with the same basic equation

$$\dot{x}_i = -\nabla V(\mathbf{x}) + \xi_i(t), \quad i = 1, \ldots, N \tag{83}$$

as for the action minimization method. It is intuitively clear, that for a system which time evolution is described by this

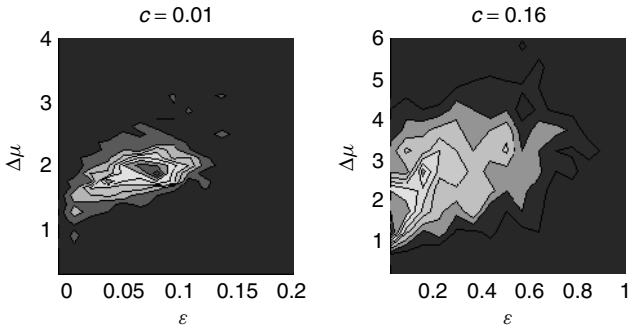

Figure 6. Mutual 2D distribution density $\rho(E, \Delta m)$ of the energy barriers and moment changes for dilute ($c = 0.01$, left panel) and concentrated ($c = 0.16$, right panel) system of magnetic particles with $\beta = 0.2$. (Reproduced from D.V. Berkov *et al.*, 2002. © 2002 with permission from IEEE.)

equation, the path $\phi_{opt}(\mathbf{x})$ which connects the two given metastable states A and B of the potential $V(\mathbf{x})$ and goes via a saddle point with the minimal energy barrier, satisfies the condition

$$(\nabla V)_\perp(\phi_{\text{opt}}) = 0 \tag{84}$$

In this notation $(\nabla V)_\perp(\phi)$denotes the component of the energy gradient, which is perpendicular to a curve ϕ. The physical sense of the statement (84) is that at any point of the optimal transition path the energy gradient is perpendicular to this path, with other words, the transition path proceeds along the energy gradient lines (see also our discussion of the action minimization method given above). A rigorous proof of this statement can be found in (Freidlin and Wentzell, 1998).

Although the equation (84) is only the necessary, but by no means the sufficient condition that the path ϕ goes through the lowest energy barrier between A and B (see, e.g., Figure 4a, where for the dashed line curve the condition $(\nabla V)_\perp(\phi) = 0$ is also fulfilled everywhere), this equation provides a useful hint how to find ϕ_{opt} starting from some arbitrary path ϕ: one can simply 'move' this path with the 'velocity'$\mathbf{u} = (\nabla V)_\perp(\phi)$which is normal to the path curve, until the stationary state of the system (given by the condition $\mathbf{u} = 0$) is reached.

To implement this idea, it is necessary to introduce some parameterization of a path ϕ, so that the coordinates of a point along the curve ϕ (in the N-dimensional configuration space of our system) are represented as functions of some parameter α: $x_1 = x_1(\alpha), \ldots, x_N = x_N(\alpha)$. Then, treating the evolution of ϕ with the velocity $\mathbf{u} = (\nabla V)_\perp(\phi)$ as a 'motion' in a fictitious time t, so that the instantaneous position of the path is given by the functions $\phi(\alpha, t) = (x_1(\alpha, t), \ldots, x_N(\alpha, t))$, we can write the corresponding 'dynamical' equation for $\phi(\alpha, t)$ as (Ren and Vanden-Eijnden, 2002)

$$\frac{\partial \phi(\alpha, t)}{\partial t} = -(\nabla V)_\perp(\phi) = -[\nabla V(\phi) - (\nabla V, \mathbf{e}_\tau)\mathbf{e}_\tau] \tag{85}$$

where the vector $\mathbf{e}_\tau$ is the unit tangent vector ϕ and thus its components at the curve point characterized by the parameter value α are

$$e_\tau^{(l)} = \frac{1}{\|\mathbf{x}_\alpha\|} \cdot \frac{\partial x^{(l)}(\alpha, t)}{\partial \alpha}, \quad \|x_\alpha\| = \sqrt{\sum_{l=1}^{N} \left(\frac{\partial x^{(l)}}{\partial \alpha} \right)^2} \tag{86}$$

so that the expression in the square brackets on the right-hand side of (85) is, as required, the component of the energy gradient vector normal to ϕ.

Numerical solution of (85) requires the discretization of the path ϕ using some concrete parametrization. A significant advantage of the string method is that this parametrization can be chosen arbitrarily basing on the considerations of either simplicity, or stability of a numerical method used to integrate (86), or required accuracy for the energy barrier value etc. The simplest choice is the 'natural' parametrization of a curve using the parameter α equal to its normalized arclength. In this case for a path starting at the point A with the coordinates $(x_A^{(1)}, \ldots, x_A^{(N)})$ and ending at $B(x_B^{(1)}, \ldots, x_B^{(N)})$ we have $\alpha(A) = 0$ and $\alpha(B) = 1$. Starting from some smooth curve between A and B, we discretize it into $K+1$ points equally spaced along the curve (for this kind of parametrization!) where the k-th point along the curve with coordinates $\mathbf{x}_k$ is characterized by the parameter value $\alpha_k = k/K (k = 0, \ldots, K)$. Calculating the derivatives of the coordinates as functions of α by some finite-difference approximation method, we compute the components of the unit tangent vector (86) for the given position of a string (at the given 'time' t) and make a 'time' step integrating numerically equations (85).

Application to micromagnetic systems

By implementation of this formalism for a micromagnetic system we encounter the same question as in the action minimization method: the dissipation term in the stochastic LLG-equation indeed represents the gradient of the micromagnetic energy (or, to be more precise, projection of this gradient into the hyperplane normal to all magnetic moment vectors), but the precession term does not. In their paper (Ren and Vanden-Eijnden, 2003) devoted to the usage of the string method in micromagnetism E *et al*. mention that they could prove that the local minima and saddle points remain the same after neglecting the precession term (unfortunately, this result is cited in (Ren and Vanden-Eijnden, 2003) as 'unpublished').

E *et al*. have demonstrated the applicability of the string method to micromagnetism in the papers (Ren and Vanden-Eijnden, 2002, 2003), where they have studied the remagnetization of a thin Permalloy nanoelement ($200 \times 200 \times 10\,\mathrm{nm}^3$ nanoelement, discretized in-plane only) choosing the two remanent S states with opposite magnetization orientations as initial and final states of the transition over a barrier (an external field was assumed to be absent). They could show that there exist at least two possible paths for this transition. The first path correspond to the magnetization switching via the intermediate S states (rotated by 90° relative to the initial and final S states) flower state, flower states (which were identified as saddle points) and the C state which was the lowest energy minimum visited during the transition process. The second transition path corresponded to the formation of the two vortices, which propagation through the nanoelement governed the switching process. The energy barriers for this second path were found to be significantly higher than for the first one. Another example briefly considered in (Ren and Vanden-Eijnden, 2003) deals with the remagnetization of a rectangular prism with a square cross-section ($200 \times 50 \times 50\,\mathrm{nm}^3$, discretized in 3D), where also two possible transition paths for the switching between the two states with the magnetization oriented (on average) along the two opposite directions of the long prism axis have been found.

From the methodical point of view, the string method has two significant advantages compared to the action minimization. First, a thermodynamical action itself already contains the first derivatives of the system energy (forces or torques). Hence its minimization with any method employing the *derivatives* of the function to be minimized (and only such methods provide a reasonable convergence speed) requires the evaluation of second derivatives of the system energy, that is, its Hessian matrix. By the string method which is based on the 'equation of motion' like (85), only the first energy derivatives are required, so that computational cost should be lower and the stability of the method higher than for the action minimization. The second important issue is the appearance of many undesired local maxima of the action, as discussed in the previous subsection. In the string method these maxima will probably play no significant role, because each point of the string is moved according to the equation (85) in the direction toward *lower* values of the system energy, so that it is highly unlikely that the string gets stuck at some 'false' metastable state like that shown in Figure 4, because such a 'false' path always goes through at least one energy *maximum*.

5.2.3 The elastic band method

Description of the method

The elastic band method, belonging to the so-called 'chain-of-states' method for searching the saddle points in complicate energy landscapes, is closely related to the string method discussed above. The main initial idea (exactly as in the discretized version of a string method) was to represent the continuous path in the configuration space of a system under study as a number of discrete states $\{\mathbf{S}_k\}(kd = 0, \ldots, K)$ and to build up an 'object function' of a type

$$Q = \sum_{k=1}^{K} V\{\mathbf{S}_k\} + \kappa \sum_{k=1}^{K-1} (\mathbf{S}_k - \mathbf{S}_{k-1})^2 \qquad (87)$$

Keeping the initial ($k = 0$) and final ($k = K$) states fixed and minimizing this function with respect to the set of states $\{\mathbf{S}_k | k = 1, \ldots, K-1\}$ should, on the one hand, lead to the decrease of the energies of the states involved (given by the terms $V\{\mathbf{S}_k\}$ in the first sum). On the other hand, the second sum should prevent the neighboring states (along the path) to

'run away' from each other (because the terms $(\mathbf{S}_k - \mathbf{S}_{k-1})^2$ give the distances between the states in the configuration space), thus keeping the discretized path reasonably smooth. Together these two tendencies should provide the sequence of closely positioned states (second sum) with the energy being as low as possible (first sum), which from the qualitative point of view obviously corresponds to the path over a saddle point between the fixed initial and final states. The method was named an 'elastic band' method, because the second term in (87) exactly corresponds to the energy of elastic bands (springs) with zero natural length and elastic constant κ 'built in between' the neighboring states of the chain $\{\mathbf{S}_k\}$.

In practice, this idea does not work really well: if the elastic constant κ is chosen too large (elastic term dominates) then the chain of states tends to 'round the corners' of the energy landscape, trying to reduce the length of the chain (distances between the neighboring states) on the cost of increasing the potential energy of the chain states. If κ is too small, then the elastic term is not able to prevent the states from sliding into the potential minima (initial and final states), so that the saddle point can be located with a reasonable accuracy. The region of intermediate κ values where the saddle point position of a continuous transition path is reproduced by the chain (87) well enough, is usually very narrow or may even not exist (see an excellent review contained in Jónsson, Mills and Jacobsen, 1998). For this reason the elastic band method, introduced in the middle of the 1980th, was considered as unreliable a decade long.

The solution of both problems was suggested by Mills and Jonsson (1994) (see Jónsson, Mills and Jacobsen, 1998 for a detailed explanation), who noted that both effects were due to the 'too physical' understanding of the model (87). Namely, by the minimization of the object function (87) both the potential (due to the first sum) and elastic (second sum) forces were fully taken into account using the straightforward differentiation of the object function Q. By its constructed Q fully mimics the energy of a set of 'particles' (states) connected via springs with elastic constants κ and moving in a potential landscape V. But the actual purpose of this function is quite different from simply imitating the behavior of the physical system just described: Q should be constructed so that the chain of 'particles' (states) reproduces as good as possible the *optimal continuous* path between the two given energy minima, which goes from the starting to the final state *along the gradient lines* of the potential $V\{\mathbf{x}\}$. Hence, Q fulfills its purpose well enough if (i) the first term would move the states *perpendicular* to the gradient lines (grad(V) is normal to the optimal path, see above) and (ii) the second term would produce only the force *parallel* to the path (to ensure that the states stay close to each other, it is sufficient to apply a force along the line connecting the states). For this reason we can *separate* the effects of the first (potential) and the second (elastic) terms by taking into account (i) only that projection of the potential force which is *perpendicular* to the path and (ii) only that projection of the elastic force which is *parallel* to the path.

This leads to the 'nudged elastic band'(NEB) model with the 'equation of motion' for the state i in the form similar to that of (85)

$$\frac{\partial \mathbf{S}_k}{\partial t} = -[\nabla V - (\nabla V, \mathbf{e}_\tau)\mathbf{e}_\tau]_{\mathbf{S}_k} + (\mathbf{F}_k^{\mathrm{el}}, \mathbf{e}_\tau)\mathbf{e}_\tau \qquad (88)$$

where the first term in square brackets is fully analogous to the corresponding term in equation (85) thus 'moving' each state in the direction normal to the transition path and the second term represents the tangential projection of the elastic force $\mathbf{F}^{\mathrm{el}}$ (derivative of the second sum in (87)) which takes care that the states remain close to each other in the configuration space. During to the fact that the two terms on the right-hand side of (88) are perpendicular to each other, that is, fully decoupled, there exist now a wide range of the elastic constants κ where the position of the saddle point along the transition path can be reproduced with a nearly arbitrary accuracy just by increasing the number of states used to discretize a path.

Finally, we note that the quality of decoupling of the two force contributions in (88) and thus – the quality of the saddle point determination and the stability of the method as a whole – crucially depend on the calculation accuracy of the tangent vector direction $\mathbf{e}_\tau$. For this reason large effort has been devoted to the development of improved method for the tangent determination for discretized curves (Jonsson and Henkelmaan, 2000).

Micromagnetic simulations using the elastic band method

Up to our knowledge, first application of the NEB method to micromagnetic simulations is due to Dittrich *et al*., (2002,2003a,b), Dittrich, 2003 and Dittrich, Thiaville, Miltat and Schrefl (2003). In their first paper, Dietrich *et al*. (Dittrich *et al*., 2002) describe their concrete implementation of the general NEB algorithm for micromagnetics, which involves a transition to the spherical coordinates of magnetic moment (as really independent variables, see the discussion above) and a proper finite-difference approximation of the tangent vector for the discretized transition path.

In this first study Dittrich *et al*. (2002) noted that for some simple systems the method works well even without the spring force, that is, without the second term in (88). This means that one can sometimes obtain a good approximation to a saddle-point path simply by moving the states of the discretized initial guess for the transition path along the energy gradient projection perpendicular to the instantaneous trajectory configuration (we note in passing that this simplified method has nothing to do anymore with the

'NEB'). In particular, in Dittrich *et al*. (2002) the energy barriers calculated numerically for the system of two interacting single-domain particles with uniaxial anisotropies were found to agree well with the analytical results available for this system. The ability of the method to find energy barriers for (i) a coherent switching of a small ($5 \times 5 \times 1\,\mathrm{nm}^3$) rectangular nanoelement, (ii) a switching of an elongated slab (typical cross-section size 13 nm, length 70 nm) via the domain wall motion, and (iii) magnetization reversal of a piece of a granular magnetic media was also demonstrated.

Later Dittrich *et al*. have implemented also a complete elastic band method (Dittrich *et al*., 2003a), including the second (elastic) term into their 'equation of motion' for the states in the chain. As expected, they have observed that for every concrete problem there exist a broad range of the elastic constant values where the height and position of a saddle point is reproduced with a sufficiently high accuracy, although for each new problem this region must be found afresh. With this improved algorithm Dittrich *et al*. could rigorously evaluate energy barriers arising due to the shape anisotropy in triangular and square nanoplatelets, investigate the increase of the energy barrier in a ferromagnetic grain coupled to an antiferromagnet (a promising candidate for a high-density recording media with improved thermal stability) and identify several possible reversal modes in a MRAM cell (Dittrich *et al*.,2003a,b; Dittrich, Thiaville, Miltat and Schrefl, 2003).

More details about micromagnetic simulations performed with this method can be found in the contribution **Numerical Methods in Micromagnetics (Finite Element Method), Volume 2** of T. Schrefl to this volume.

An interesting topic which has been pursued by several research groups in the last few years is the applicability of the Monte-Carlo (MC) methods (Binder, 1986) for dynamic micromagnetic simulations at finite temperatures. The major *advantage* of modern MC schemes when applied to the magnetization transition between various metastable states is evident: The corresponding computation time does *not* depend exponentially on the height of the energy barrier separating these states, as it is the case for the Langevin dynamics. However, there exist also several principal problems by the dynamical application of the MC methods. The two most serious of them are: (i) the difficulty to establish a relation between a MC step and a physical time and (ii) proper inclusion of the magnetization precession, which is also a highly nontrivial task, because the precession does not lead to the change of the system energy and hence does not affect the probability to accept a MC step. Recent methodical progress on this area can be found in Nowak, Chantrell and Kennedy (2000), Chubykalo *et al*. (2003), and Cheng, Jalil, Lee and Okabe (2005, 2006) and is reviewed in the contribution of U. Nowak to this volume.

NOTES

[1] Note however, that the effective field itself $\mathbf{H}^{\mathrm{eff}} = (1/V)\cdot\partial E/\partial\mathbf{M}$ is of course zero at equilibrium – up to the component along the magnetization vector which can be neglected because the magnetization magnitude is assumed to be constant

[2] The *deterministic* Landau–Lifshitz–Gilbert equation conserves the moment magnitude anyway. However this is *not* automatically the case for its stochastic analogue.

ACKNOWLEDGMENTS

The author greatly acknowledges many useful discussions on the topics treated in this chapter with J. Miltat, N. Smith, M. Stiles, A. Slavin, T. Schrefl, and U. Nowak.

REFERENCES

Berkov, D.V. (1998). Numerical calculation of the energy barrier distribution in disordered many-particle system: the path integral method. *Journal of Magnetism and Magnetic Materials*, **186**, 199.

Berkov, D.V. (2002a). Fast switching of magnetic nanoparticles: simulation of thermal noise effects using the Langevin dynamics. *IEEE Transactions on Magnetics*, **MAG-38**, 2489.

Berkov, D.V. (2002b). Density of energy barriers in fine magnetic particle systems. *IEEE Transactions on Magnetics*, **MAG-38**, 2637.

Berkov, D.V. and Gorn, N.L. (2002). Thermally activated processes in magnetic systems consisting of rigid dipoles: equivalence of the Ito and Stratonovich stochastic calculus. *Journal of Physics: Condensed Matter*, **14**, L1.

Berkov, D.V. and Gorn, N.L. (2004). Micromagnetic simulations using Langevin dynamics: studying thermal field correlations. *Journal of Magnetism and Magnetic Materials*, **272–276**, Part 1, 687.

Berkov, D.V. and Gorn, N.L. (2005). Stochastic dynamic simulation of fast remagnetization processes: recent advances and applications. *Journal of Magnetism and Magnetic Materials*, **290–291**, 442.

Berkov, D.V., Gorn, N.L. and Görnert, P. (2002). The Langevin-dynamics simulation of interacting fine magnetic particle system. *Journal of Magnetism and Magnetic Materials*, **226**, 1936.

Bertram, H.N., Safonov, V.L. and Jin, Z. (2002). Thermal magnetization noise, damping fundamentals, and mode analysis: application to a thin film GMR sensor. *IEEE Transactions on Magnetics*, **38**, 2514.

Binder, K. (Ed.) (1986) *Monte-Carlo Methods in Statistical Physics*, Springer-Verlag: Berlin.

Braun, H-B. (1994). Kramer's rate theory, broken symmetries and magnetization reversal. *Journal of Applied Physics*, **76**, 6310.

Braun, H-B. (2000). Stochastic magnetization dynamics in magnetic nanostructures: from Neel-Brown to soliton-antisoliton creation. In *Structure and Dynamics of Heterogeneous Systems*, Entel, P. and Wolf, D. (Eds.), World Scientific, Singapore.

Bray, A.J. and McKane, A.J. (1989). Instanton calculation of the escape rate for activation over a potential barrier driven by colored noise. *Physical Review Letters*, **62**, 493.

Brown, W.F. (1963a). *Micromagnetics*, Wiley: New York.

Brown Jr., W.F. (1963b). Thermal fluctuations of a single domain particle. *Physical Review*, **130**, 1677.

Cheng, X.Z., Jalil, M.B.A., Lee, H.K. and Okabe, Y. (2005). Time-quantifiable Monte Carlo method for simulating a magnetization-reversal process. *Physical Review*, **B72**, 094420.

Cheng, X.Z., Jalil, M.B.A., Lee, H.K. and Okabe, Y. (2006). Mapping the Monte Carlo scheme to Langevin dynamics: A Fokker-Planck approach. *Physical Review Letters*, **96**, 067208.

Chubykalo, O. , Smirnov-Rueda, R., Gonzalez, J.M., *et al*. (2003). Brownian dynamics approach to interacting magnetic moments. *Journal of Magnetism and Magnetic Materials*, **266**, 28.

Chubykalo, O., Nowak, U., Smirnov-Rueda, R., *et al*. (2003). Monte Carlo technique with a quantified time step: Application to the motion of magnetic moments. *Physical Review*, **B67**, 064422.

Coffey, W.T., Kalmykov, Yu.P. and Waldron, J.T. (2004). *The Langevin Equation (With Applications to Stochastic Problems in Physics, Chemistry and Electrical Engineering)*, World Scientific.

Dittrich, R. (2003). *Finite Element Computation of Energy Barriers in Magnetic Systems*, Ph.D. Thesis, Vienna University of Technology.

Dittrich, R., Schrefl, T., Forster, H., *et al*. (2003a). Energy barriers in magnetic random access memory elements. *IEEE Transactions on Magnetics*, **39**, 2839.

Dittrich, R., Schrefl, T., Suess, D., *et al*. (2003b). Thermally induced magnetization reversal in AFC media, *Journal of Applied Physics*, **93**, 7405.

Dittrich, R., Schrefl, T., Suess, D., *et al*. (2002). A path method for finding energy barriers and minimum energy paths in complex micromagnetic systems. *Journal of Magnetism and Magnetic Materials*, **250**, L12.

Dittrich, R., Thiaville, A., Miltat, J. and Schrefl, T. (2003). A rigorous micromagnetic computation of configurational anisotropy energies in nanoelements. *Journal of Applied Physics*, **93**, 7891.

Dotsenko, V.S. (1993). Physics of the spin-glass state. *Physics - Usp. (USA)*, **36**, 455.

Dormann, J.L., Fiorani, D. and Tronc, E. (1999). On the models for interparticle interaction in nanoparticle assemblies: comparison with experimental results. *Journal of Magnetism and Magnetic Materials*, **266**, 28.

E, W., Ren, W. and Vanden-Eijnden, E. (2002). String method for the study of rare events. *Physical Review*, **B66**, 052301.

E, W., Ren, W. and Vanden-Eijnden, E. (2003). Energy landscape and thermally activated switching of submicron-sized ferromagnetic elements. *Journal of Applied Physics*, **93**, 2275.

Ermak, D.L. and McCammon, J.A. (1978). Brownian dynamics with hydrodynamic interaction. *Journal of Chemical Physics*, **69**, 1352.

Freidlin, M.I. and Wentzell, A.D. (1998). *Random Perturbations of Dynamical Systems*, Springer: New York.

Garanin, D.A. (1997). Fokker-Planck and Landau-Lifshitz-Bloch equations for classical ferromagnets. *Physical Review*, **B55**, 3050.

Garanin, D.A. and Chubykalo-Fesenko, O. (2004). Thermal fluctuations and longitudinal relaxation of single-domain magnetic particles at elevated temperatures. *Physical Review*, **B70**, 212409.

Gardiner, S. (1997). *Handbook on Stochastic Processes*, Springer-Verlag: Berlin.

Garcia-Ojalvo, J. and Sancho, J.M. (1994). Colored noise in spatially extended systems. *Physical Review*, **E49**, 2769.

Garcia-Palacios, J.L. and Lazaro, F.J. (1998). Langevin-dynamics study of the dynamical properties of small magnetic particles. *Physical Review*, **B58**, 14937.

Gilbert, T.L. (1955). A Lagrangian formulation of gyromagnetic equation of the magnetization field. *Physical Review*, **100**, 1243.

Grinstein, G. and Koch, R.H. (2003). Coarse graining in micromagnetics. *Physical Review Letters*, **90**, 207201.

Hansen, M.F. and Morup, S. 1998. Models for the dynamics of interacting magnetic nanoparticles. *Journal of Magnetism and Magnetic Materials*, **184**, 262.

Henkelmaan, G. and Jonsson, H., (2000). Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *Journal of Chemical Physics*, **113**, 9978.

Hoenggi, P., Talkner, P. and Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers. *Reviews of Modern Physics*, **62**, 251.

James, F. (1980). Monte Carlo theory and practice. *Reports on Progress in Physics*, **43**, 1145.

Jónsson, H., Mills, G. and Jacobsen, K.W. (1998). Nudged elastic band method for finding minimum energy paths of transitions. In *Classical and Quantum Dynamics in Condensed Phase Simulations*, Berne, B.J., Ciccotti, G. and Coker, D.F. (Eds.), World Scientific.

Kikuchi, R. (1956). On the minimum of magnetization reversal time. *Journal of Applied Physics*, **27**, 1352.

Klik, I. and Gunther, L. (1990). Thermal relaxation over a barrier in single domain ferromagnetic particles. *Journal of Applied Physics*, **67**, 4505.

Kloeden, P.E. and Platen, E. (1995). *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag: Berlin.

Landau, L.D. and Lifshitz, E.M. (1935). On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. *Physikalische Zeitschrift der Sowjetunion*, **8**, 153.

Landau, L.D. and Lifshitz, E.M. (1980). *Statistical Physics (Part 1)*, Butterworth-Heinemann.

Landau, L.D. and Lifshitz, E.M. (1981). *Mechanics*, Butterworth-Heinemann.

Lyberatos, A. and Chantrell, R.W. (1993). Thermal fluctuations in a pair of magnetostatically coupled particles. *Journal of Applied Physics*, **73**, 6501.

Mallinson, J.C. (1987). On damped gyromagnetic precession. *IEEE Transactions on Magnetics*, **23**, 2003.

McShane, E.J. (1974). *Stochastic Calculus and Stochastic Models*, Academic Press: New York.

Mills, G. and Jonsson, H. (1994). Quantum and thermal effects in H_2 dissociative adsorption: evaluation of free energy barriers in multidimensional quantum systems. *Physical Review Letters*, **72**, 1124.

Miltat, J., Albuquerque, G. and Thiaville, A. (2001). An introduction to micromagnetics in the dynamical regime. In *Spin Dynamics in Confined Magnetic Structures I*, *Topics in Applied Physics*, Hillebrands, B. and Ounadjela, K. (Eds.), Springer-Verlag: Berlin, Vol. 83.

Nakatani, Y., Uesaka, Y., Hayashi, N. and Fukushima, H. (1997). Computer simulation of thermal fluctuation of fine particle magnetization based on Langevin equation. *Journal of Magnetism and Magnetic Materials*, **168**, 347.

Nowak, U., Chantrell, R.W. and Kennedy, E.C. (2000). Monte Carlo simulation with time step quantification in terms of Langevin dynamics. *Physical Review Letters*, **84**, 163.

Onsager, L. and Machlup, S. (1953). Fluctuations and irreversible processes. *Physical Review*, **91**, 1505.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992). *Numerical Recipes in Fortran: The Art of Scientific Computing*, Cambridge University Press.

Romero, A.H. and Sancho, J.M. (1999). Generation of short and long range temporal correlated noises. *Journal of Computational Physics*, **156**, 1.

Rümelin, W. (1982). Numerical treatment of stochastic differential equations. *SIAM Journal on Numerical Analysis*, **19**, 604.

Safonov, V.L. and Bertram, H.N. (2003). Linear stochastic magnetization dynamics and microscopic relaxation mechanisms. *Journal of Applied Physics*, **94**, 529.

Scholz, W., Schrefl, T. and Fidler, J. (2001). Micromagnetic simulations of thermally activated switching in fine particles. *Journal of Magnetism and Magnetic Materials*, **233**, 296.

Smith, N. (2001). Modelling of thermal magnetization fluctuations in thin-film magnetic devices. *Journal of Applied Physics*, **90**, 5768.

Smith, N. (2002). Fluctuation–dissipation considerations for phenomenological damping models for ferromagnetic thin films. *Journal of Applied Physics*, **92**, 529.

Xue, J. and Victora, R.H. (2000). Micromagnetic predictions for thermally assisted reversal over long time scales. *Applied Physics Letters*, **77**, 3432.

Xue, J. and Victora, R.H. (2001). Micromagnetic predictions of bit decay caused by thermal fluctuations over long time scale. *Journal of Applied Physics*, **89**, 6985.

Zhang, K. and Fredkin, D.R. (2000). Thermal relaxation of clusters of interacting Stoner-Wohlfarth particles. *Journal of Applied Physics*, **87**, 4795.

FURTHER READING

Advanced Recording Model, package for micromagnetic simulations, Euxine technologies, http://www.euxine.com.

Brown, G., Novotny, M.A. and Rikvold, P.A. (2001). Thermal magnetization reversal in arrays of nanoparticles. *J. Appl. Phys.*, **89**, 7588.

Chikazumi, S. (1997). *Physics of Ferromagnetism*, Clarendon Press.

Feng, X. and Fischer, P.B. (2001). Coarse-graining Landau-Lifshitz damping. *J. Appl. Phys.*, **80**, 6988.

Feynman, R.P. and Hibbs, A.R. (1965). *Quantum Mechanics and Path Integrals*, McGraw-Hill: New York.

Hubert, A. and Schäfer, R. (1998). *Magnetic Domains: The Analysis of Magnetic Microstructures*, Springer-Verlag.

Kittel, C. (1953). *Introduction to Solid State Physics*, John Wiley & Sons: New York.

Landau, L.D. and Lifshitz, E.M. (1975). *The Classical Theory of Fields*, Pergamon Press: Oxford.

Landau, L.D. and Lifshits, E.M. (1985). *Electrodynamics of Continuous Media*, Elsevier Science.

Lee, K.J., Park, N.Y. and Lee, T.D. (2001). Numerical study of spin relaxation by thermal fluctuation: effect of shape anisotropy. *J. Appl. Phys.*, **89**, 7460.

Martinez, E., Lopez-Diaz, L. Torres, L. and Alejos, O. (2003). Temperature dependence of spontaneous magnetization using a continuous model, *IEEE Transactions on Magnetics*, **39**, 2522.

McKane, A.J., Luckock, H.C. and Bray, A.J. (1990). Path integrals and non-Markov processes I and II. *Phys. Rev.*, **A41**, 644.

O'Grady, K. and Laider, H. (1999). The limits of magnetic recording – media considerations. *J. Magn. Magn. Mat.*, **200**, 616.

Safonov, V.L. and Bertram, H.N. (2002). Thermal magnetization noise in a thin film. *Phys. Rev.*, **B65**, 172417.

Scheinfein, R. Landau-Lifshitz-Gilbert Micromagnetics Simulator, http://www.llgmicro.com.

Suhl, H. (1998). Theory of the magnetic damping constant. *IEEE Transactions on Magnetics*, **MAG-34**, 1834.

Thiaville, A., Miltat, J. and Ben Youssef, J. (2001). Dynamics of vertical Bloch lines in bubble garnets: experiments and theory. *Europ. Phys. J. B*, **23**, 37.

Wohlfarth, E.P. and Stoner, E.C. (1948). A mechanism of magnetic hysteresis in heterogeneous alloys. *Phil. Trans. Roy. Soc. (London), Ser. A*, **240**, 599.