5-2014

Properties of Multiferroic BiFeO3 from First Principles

Dovran Rahmedov
University of Arkansas, Fayetteville

Properties of Multiferroic $BiFeO_3$
from First Principles

Properties of Multiferroic $BiFeO_3$ from First Principles

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Physics

By

Dovran Rahmedov
Turkmen State University
Bachelor of Science in Physics, 2006

May 2014
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council

Laurent Bellaiche, Ph.D. Dissertation Director	William F. Oliver III, Ph.D. Dissertation Committee
Gregory Salamo, Ph.D. Dissertation Committee	Jacques Chakhalian, Ph.D. Dissertation Committee
Huaxiang Fu, Ph.D. Dissertation Committee	Peter Pulay, Ph.D. Dissertation Committee

Abstract

In this dissertation, a first-principle-based approach is developed to study magnetoelectric effect in multiferoic materials. Such approach has a significant predictive power and might serve as a guide to new experimental works. As we will discuss in the course of this work, it also gives an important insight to the underlying physics behind the experimentally observed phenomena.

We start by applying our method to investigate properties of a generic multiferroic material. We observe how magnetic susceptibility of such materials evolves with temperature and compare this evolution with the characteristic behavior of magnetic susceptibility for pure magnetic systems. Then we focus our attention to particular multiferroic – $BiFeO_3$ – and reproduce its magnetic states with all of their essential features. Those magnetic states include (i) antiferromagnetic state, (ii) state with weak ferromagnetism resulting from canting of magnetic moments, and (iii) cycloidal magnetic structure. All of those magnetic states were also studied under external electric and magnetic fields. Under such electric fields magnetic order parameters of the systems undergo interesting transformations and sometimes take unexpected path. Finally, we study the material under strain and explore possibilities of favoring one magnetic state over another and even "creating" states that can be stable only under the strain.

Acknowledgments

I thank my friends and mentors Dorel Guzun and Murad Mashayev. Without their support my journey to the graduate school would not have been possible. I also want to express acknowledgement to my advisers Laurent Bellaiche and Gregory Salamo who guided me throughout my graduate work.

Contents

Preface 1

I Mathematical methods and computational tools 3

1 Linear algebra 3

1.1 System of linear equations and vector spaces . . . 3

1.2 Magnetic and electric susceptibilities . . . 9

1.3 Euclidean space . . . 14

1.4 Matrix multiplication . . . 15

1.5 Normal equation . . . 20

1.6 Programming tools . . . 24

1.7 Tensors . . . 34

1.8 Three dimensional vector space . . . 37

2 Fourier analysis 48

2.1 Fourier series . . . 49

2.2 Discrete Fourier transform (DFT) . . . 53

3 Statistics 56

3.1 Concepts of Probability Theory . . . 56

3.2 Phase space . . . 58

3.3 Monte-Carlo (MC) approach . . . 60

3.3.1 Molecular dynamics *vs.* Monte Carlo approach . . . 60

3.4 Monte Carlo . . . 62

3.4.1 Metropolis algorithm . . . 62

4 Quantum mechanics 65

4.1 Density functional theory 67

II First-Principle-based effective Hamiltonian study 71

5 Monte Carlo method 71

5.1 Effective Hamiltonian 71

5.2 Units used in the program 77

6 Magnetic and electric properties of matter 77

6.1 Antiferromagnetic multiferroics 80

6.2 $BiFeO_3$ under electric field 95

6.3 Weak ferromagnetism in $BiFeO_3$ 95

7 Magnetic cycloid of $BiFeO_3$ 97

7.1 Spin Density Wave 100

7.2 Anharmonicity 101

7.3 Magnetic cycloid of BFO under an epitaxial strain 114

III Conclusions and prospects 117

8 Conclusions 117

9 Future prospects 119

References 122

Appendices 126

A Copyright information 126

A.1 Copyright policies of the Nature Publishing Group (NPG) 126

A.2 Copyright policies of the American Physical Society (APS) 129

B Magnetoelectric signature in the magnetic properties of antiferromagnetic multiferroics: atomistic simulations and phenomenology 132

B.1 Numerical method 132

B.2 Results 133

B.2.1 Atomistic Simulations 133

B.2.2 Phenomenology 138

B.3 Conclusions 143

Preface

Advancement in technology throughout the ages was tightly linked to advancement in the manufacturing and usage of the materials with desirable properties. For example, improvement in production methods of steel manufacturing made industrial revolution possible. In a similar manner, in a recent times, the steady progress in manufacturing and understanding of materials with good magnetic and electric properties gave rise to modern information age. The latter progress is still continuing with a rapid pace and today there is a high demand for materials with novel magnetic and electric properties. As a result, there has been a flurry of research throughout the world to discover materials that will make technology of the future possible.

In the hunt for a materials with unusual electric and magnetic properties, one promising avenue is multiferroics. Multiferroics are materials that can exhibit electric and magnetic ordering simultaneously [1, 2]. Those coexisting electric and magnetic degrees of freedom are coupled to each other and this coupling allows some phenomena that might not be observable in pure magnetically or electrically ordered systems. For example, multiferroics can exhibit magnetoelectric effect. This effect opens possibility for new devices that can be controlled both electrically and magnetically.

Among multiferroic materials, $BiFeO_3$ (BFO) occupies a very important place. What is special about BFO is that it has unusually high antiferromagnetic Neel and ferroelectric Curie temperatures ($T_N = 640$K and $T_C = 1100$K), and high electric polarisation ($P = 100\mu C/cm^2$). These uncommon properties are essential for device applications and are indeed the main reason why so many attention was given to this material [25]. From a theoretical point of view, the rich physics of this material poses challenging problems and opens up many issues, some of which still remaining unresolved. To address some of those problems and understand the physics behind unusual properties of $BiFeO_3$ is the main goal of our study.

In this work, we will mainly concentrate on magnetic, electric, and structural properties of multiferroic $BiFeO_3$. We will study unusual properties of $BiFeO_3$ under external electric and

magnetic fields applied along different directions. We will also consider the effect of strain on properties of this material. At the same time, we are aiming to arrive at general conclusions that would be applicable to the broad class of multiferroic materials (i.e., not only $BiFeO_3$).

This work is divided into three parts. In the first part, we will describe the methodology that is used in the investigation of the properties of materials. This part contains description of some of the general mathematical and computational methods that are used throughout this work. In the first part we will also cover the relevant physics behind the models that are used in the simulations. The second part consists of presenting and describing the results that were obtained from our studies. The third part provides a summary of the studies performed, and offer possible avenues to further pursue in the future.

Part I

Mathematical methods and computational tools

This work is heavily based on methods of applied mathematics and computational science. Therefore the first part of this work will start by giving an introduction and overview to the computational tools, and mathematical methods widely used in theoretical physics and material sciences. The concepts of applied mathematics will be covered in a brief but comprehensive manner. On the other hand, all the discussions related to the computational tools will be very practical and we will go no deeper than necessary to reproduce the results of this work.

Note that Readers who are familiar with linear algebra and/or basics of computational physics can safely skip this Part I, without any losses in understanding of the subsequent parts of the dissertation. In other words, these Readers can directly go to Parts II and III, which describe results obtained from our studies and future prospects of our work.

Whenever a number of subjects comes together in the single study like this one, it is impossible to avoid conflict between terminology. Space for example has one definite meaning in mathematics and it is different from the meaning used in common language. Field in mathematics might mean the generalization of numbers and in physics it usually refers to a quantity that changes in space. The same thing is true for letters used to denote those quantities. Even the central concept of this work – multiferroic material – might be defined to mean slightly different things in different works. We will not point out this ambiguity in every case unless it is absolute necessity to avoid confusion.

1 Linear algebra

1.1 System of linear equations and vector spaces

System of linear equations. Since linear algebra will provide us with a framework and language for the subsequent discussions, its concepts will be covered first. I will introduce concepts and results of linear algebra in the process of describing a system of m linear equations with n unknown

quantities. The system of m linear equations with n unknowns has the following form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = & b_2 \\ \quad \vdots \qquad \vdots \qquad\qquad \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = & b_m \end{cases} \tag{1.1.1}$$

In the system of equation (1.1.1) symbols $x_1,\ x_2,\ \ldots,\ x_n$ represent unknowns to be determined, while $a_{11},\ a_{12},\ \ldots,\ a_{mn}$ and $b_1,\ b_2,\ \ldots,\ b_m$ are given numbers which in the linear algebra are referred to as scalars.[1]

Vector notation. In the language of linear algebra, this system can be written in a more compact form using vector notation

$$\mathcal{A}\boldsymbol{x} = \boldsymbol{b} \tag{1.1.2}$$

where

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{1.1.3}$$

In the expressions (1.1.3), the quantity $\mathcal{A}$ is called $m \times n$ matrix, while $\boldsymbol{x}$ and $\boldsymbol{b}$ are called n and m dimensional column vectors, respectively. Thus column vectors are nothing more than a column of scalars arranged in particular order and matrix is just column vectors placed side by side. The set of all possible vectors of the same dimension is called vector space or linear space of that dimension.[2]

[1] Here and in all subsequent discussions, unless otherwise specified, scalars are assumed to be from the field of real numbers $\mathbb{R}$. Later, in the context of quantum mechanics, we will also consider the case where scalars are taken from the field of complex numbers. In abstract algebra, more general cases of arbitrary fields are also studied. However, those general cases find very small applications in physics and we will not mention them at all.

[2] In "advanced" level linear algebra this definition of vector spaces are given as a particular example of linear spaces and sometimes referred to it as a coordinate space. However, our way of

Operations with vectors. To show that problems stated in expressions (1.1.1) and (1.1.2) are equivalent to each other we have to introduce some conventions.

The sum of two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of the same dimensions is naturally defined to be a vector each element of which is the sum of the corresponding elements

$$\boldsymbol{x}+\boldsymbol{y}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}+\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}=\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{pmatrix} \tag{1.1.4}$$

The difference between two vectors of the same dimensions is defined in similar manner (i.e., element by element).

The product of a scalar α with a vector, $\boldsymbol{x}$ is defined to be a vector each element of which is a product of that scalar α with corresponding elements of the vector $\boldsymbol{x}$

$$\alpha\boldsymbol{x}=\alpha\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix} \tag{1.1.5}$$

Matrix $\mathcal{A}$ times vector $\boldsymbol{x}$ is interpreted as a linear combination of columns of the matrix $\mathcal{A}$

$$\mathcal{A}\boldsymbol{x}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}+x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}+\cdots+x_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \tag{1.1.6}$$

going about building foundations of linear algebra is not less general than in most advanced level presentations. Actually one can show that any finite linear space of dimension n is isomorphic, i.e. is equivalent to the n - dimensional coordinate space that we are considering here.

Linear combinations. The concept of linear combination is very important in linear algebra. In general, a linear combination of vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\cdots$, $\boldsymbol{z}$ with scalars α, β, $\cdots$, γ is another vector obtained using the following expression

$$\alpha\boldsymbol{x} + \beta\boldsymbol{y} + \cdots + \gamma\boldsymbol{z} \tag{1.1.7}$$

Again, if we put vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\cdots$, $\boldsymbol{z}$ side by side and construct a rectangular matrix from them, then the linear combination (1.1.7) can be written as

$$\begin{pmatrix} \boldsymbol{x} & \boldsymbol{y} & \cdots & \boldsymbol{z} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \vdots \\ \gamma \end{pmatrix} \tag{1.1.8}$$

Matrix is a very convenient notation for grouping vectors together. In general, matrix is an example of linear function that acts in the vector space. From this point of view, matrix takes one vector as an argument and returns a single vector as a value in a linear fashion. Such function can be completely specified by giving the set of vectors. Therefore, by putting those vectors into columns of rectangular matrix, we so to speak make complete structure of that function available for the eyes to see. More general linear functions, i.e. tensors, can not be so conveniently denoted because the paper on which we write has only two dimensions, while generic tensor would require many more dimensions for its complete specification.

Linear dependence and independence. Of course, if all the scalars α, β, $\cdots$, γ are zeros then the operation of linear combination (1.1.7) will always return a zero vector, i.e. the vector all elements of which are zeros. If, however, there is such linear combination that can give a zero vector and not all the scalars α, β, $\cdots$, γ are zeros then the set of vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\cdots$, $\boldsymbol{z}$ are called linearly dependent. Otherwise, the set is called linearly independent. Thus, for linearly independent set of

vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\cdots$, $\boldsymbol{z}$, the relation

$$\alpha\boldsymbol{x} + \beta\boldsymbol{y} + \cdots + \gamma\boldsymbol{z} = \boldsymbol{0} \tag{1.1.9}$$

can be satisfied if and only if $\alpha = \beta = \cdots = \gamma = 0$.[3]

Basis and dimension. All the linear combinations of a single vector $\boldsymbol{x}$ define a line (i.e., one dimensional subspace of the whole n dimensional vector space). All the linear combinations of two independent vectors $\boldsymbol{x}$, and $\boldsymbol{y}$ defines a plane (i.e., two dimensional subspace of the whole n dimensional vector space). In n dimensional linear space there is always a set of n vectors whose linear combinations fills up (spans) the whole n dimensional space. Such a set is called basis. The number of vectors in basis is always the same for a given space but the choice of vectors that make up the basis is not unique. The fact that the space of vectors with n elements (i.e., n dimensional vectors) always have n basis is not trivial but we will not give its "proof" here.

Coordinate system. Whenever in n dimensional space, we have any set of n independent vectors, we can use them as a basis and every other vector can be expressed as a linear combination of those basis vectors. The coefficients of those basis are called the coordinates of the vector on that basis. Therefore, choosing a basis is equivalent to defining a particular coordinate system.[4] For a given basis (coordinate system), any vector can be uniquely identified by its coordinates. The prove of this and other basic facts about vector spaces can be found in ref. [7].

Equality of vectors. Finally, for the sake of completeness, we need to define the notion of equality between two vectors. Two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of the same dimensions are equal to each other if and

[3] Let's note that the number zero denoted by symbol 0 and zero vector denoted by bold symbol $\boldsymbol{0}$ are objects of different nature. The first is a single scalar, while the second is a column of scalars. Here, we used bold symbol for the vector. However, it is common to see both of them denoted with the same symbol 0. This ambiguous way of denoting zero rarely leads to confusion since one usually can tell from context when 0 denotes a zero vector and when it denotes the number zero.

[4] Usually, however, basis are called coordinate systems and basis vectors called coordinate axis if some explicit or implicit order is specified for the basis vectors.

only if their corresponding elements are equal as a scalars $\boldsymbol{x} = \boldsymbol{y} \Leftrightarrow x_i = y_i$ for any i. With these conventions and definitions now it is obvious that expressions (1.1.1) and (1.1.2) are equivalent to each other.

Linear combination and system of linear equations. Let us now summarize the relation between the problem of solving the system of linear equations (1.1.1) and the concept of linear combination (1.1.7). For a given set of vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\cdots$, $\boldsymbol{z}$ to compute their linear combination with a set of scalars α, β, $\cdots$, γ is trivial operation. This operation involves straightforward multiplication and addition of scalars. The system (1.1.1) or equivalently the linear equation (1.1.2) poses problem of inverting this operation of linear combination. In the language of linear algebra, the equation (1.1.2) is asking us to find coefficients of the linear combination of columns of a matrix $\mathcal{A}$ which will give the vector $\boldsymbol{b}$. These coefficients then will give us the elements of vector of unknowns $\boldsymbol{x}$.

Column space. All possible linear combinations of the columns of matrix $\mathcal{A}$ is called a column space of that matrix. The column space of matrix $\mathcal{A}$ can be equal to, i.e. span, the whole n dimensional linear space. For this to happen, the number of columns should be at least equal to the number of dimensions of the columns and consequently to the dimension of the whole space. In particular, if the matrix $\mathcal{A}$ is a square matrix with independent columns, then the columns of this matrix can be used as a basis set and by definition of basis it will span the whole space. The solution of the system (1.1.2) then will be coordinates of the right hand side vector $\boldsymbol{y}$ on that basis. That is why square matrices are so special. Provided that their columns are independent the system associated with them always have one and only one solution. Square matrix with independent columns, also called non-singular, non-degenerate, or an invertible matrix.

Overdetermined systems. The column space can also be equal to only some subspace of the whole n dimensional linear space. In this case, there will be some vectors $\boldsymbol{b}$ that will be outside the column space of the matrix $\mathcal{A}$ and the system (1.1.2) for such vectors will have no solution. For

example, if the number of columns of the matrix is smaller than the dimension of the whole space (dimension of columns), then there is no way that every right hand side vector $\boldsymbol{b}$ can be obtained just by combining those few columns.[5] This latter case happens when we have more equations than unknowns in the system (1.1.1) and it is called overdetermined system. We will consider the case of overdetermined systems below first. To make the presentation of this case more concrete, we will consider examples that plays a central role in this work.

1.2 Magnetic and electric susceptibilities

Magnetic susceptibility. Under external magnetic field $\boldsymbol{H}$, all materials produce their own magnetic field, i.e. materials become magnetized. We will discuss the physical reasons that make materials magnetic and different types of magnetic materials later. Here, we are going to develop the mathematical tools that will be used explicitly in computing certain coefficients that characterizes behavior of materials under external fields.

The degree of magnetization of the material is characterized by the physical quantity called magnetization $\boldsymbol{M}$. For a small external field $\boldsymbol{H}$, the magnetization $\boldsymbol{M}$ of the material with internal magnetization $\boldsymbol{M}^0$ can be expressed as

$$\boldsymbol{M} = \boldsymbol{M}^0 + \chi_m \boldsymbol{H} \tag{1.2.1}$$

In the expression (1.2.1) the proportionality coefficient χ_m is called the magnetic susceptibility and it will be very important in our future discussions. Its magnitude shows how easy or hard it is to magnetize the material, and its behavior with temperature will give us a clue about the magnetic structure of the material as we will see later.

[5] By taking all linear combinations of a single vector from three dimensional space, we can get only vectors lying along that vector, and by taking all linear combinations of two independent vectors from three-dimensional space, one can get only vectors from the plane that contains those two vectors. In both cases it is impossible to get every vector from the three dimensional space. We need at least three vectors to accomplish that.

Different definitions of susceptibilities. The susceptibility, in the way how we defined it, is known as a volume susceptibility and sometimes denoted as χ_m^v. In the literature, one also commonly encounters other definitions. Among those alternative definitions, the following two are important: they are mass magnetic susceptibility X_m^{mass}, and the molar magnetic susceptibility χ_m^{mol}. The conversion between these three forms of susceptibilities can be made using following relations

$$\chi_m^{mass} = \frac{\chi_m^v}{\rho}$$

$$\chi_m^{mol} = M\chi_m^{mass} = M\frac{\chi_m^v}{\rho}$$

where ρ is the density and M is the molar mass of the material.

Electric susceptibility. All the materials react to an external electric field by producing their own electric field and such materials are said to become polarized. If the magnetic response to an external magnetic field is characterized by the magnetic susceptibility, the electric response of the material to an external electric field is characterized by an electric susceptibility. We will denote the electric susceptibility as χ_e and for the material with internal polarization $\boldsymbol{P}^0$, it is defined using the following expression

$$\boldsymbol{P} = \boldsymbol{P}^0 + \varepsilon_0\chi_e\boldsymbol{E} \tag{1.2.2}$$

In equation (1.2.2), the quantity $\boldsymbol{P}$ characterizes the degree of polarization of the material in response to an external electric field $\boldsymbol{E}$ and is called electric polarization or simply polarization. The physical constant ε_0 is known by names vacuum permittivity, permittivity of free space or electric constant. Its value in SI units is $\varepsilon_0 = 8.854187817 \times 10^{-12}$ farads per meter ($\mathrm{F{\cdot}m^{-1}}$).

Mathematical Analogy. From a mathematical point of view, both definitions for electric and magnetic susceptibilities are very similar. In fact, they are so similar that just by replacing one letter with another we can copy most formulas for magnetic susceptibility and use them to describe electric susceptibility. This is not an exceptional example and in physics we often encounter

similar equations in describing different phenomena. This fact makes application of mathematics in physics very fruitful, since by solving equations once we can use their solutions many times for many different phenomena.

However, let us point out than in physics we gain a lot of insight by noticing differences between apparently similar things. For example, the magnetic susceptibility can have positive values as well as negative values. Typically, when the sign of magnetic susceptibility is negative, it is also small in magnitude in comparison with positive magnetic susceptibilities. On the other hand, electric susceptibility has only one sign, namely it is always positive. Also most materials have very small magnetic and electric susceptibilities, while there is a small family of materials that possess very large susceptibilities. These and many other experimental facts have deep physical reasons behind them and we will discuss those reasons as this work progresses.

Tensor character of electric and magnetic susceptibilities. In the discussion above, the magnetic and electric susceptibilities were treated as scalar quantities. This is the case when the material under consideration is not a single crystal or if it is an amorphous material. We also can consider the susceptibility as a single number (scalar) when we are interested in applying the field only along one direction. However, in single crystal materials, one has to take into account the fact that the matter differently reacts to an external field applied along different directions. Mathematically, this is achieved by using six numbers to characterize the susceptibility instead of one number and the corresponding equations for susceptibilities become tensor equations

$$M_i = M_i^0 + \chi_{ij}^m H_j \tag{1.2.3}$$

and

$$P_i = P_i^0 + \chi_{ij}^e H_j \tag{1.2.4}$$

We will consider mathematical treatment of tensors in the appropriate section. Below, we will ignore the tensor character of the susceptibility and consider them as scalars for convenience.

Computation of susceptibilities. One way of going about computing the magnetic and electric susceptibilities is very obvious. Let us consider the magnetic susceptibility first. We can apply external magnetic field and compute corresponding magnetization of the material. We can express the resulting series of computations in the form of the system of linear equations as follows

$$\begin{cases} M^0 + \chi_m H_1 = M_1 \\ M^0 + \chi_m H_2 = M_2 \\ \quad \vdots \qquad \vdots \qquad \vdots \\ M^0 + \chi_m H_n = M_n. \end{cases} \tag{1.2.5}$$

In this linear system of equations, unknowns are M^0 and χ_m. In the matrix form, the same system can be expressed as

$$\begin{pmatrix} 1 & H_1 \\ 1 & H_2 \\ \vdots & \vdots \\ 1 & H_n \end{pmatrix} \begin{pmatrix} M^0 \\ \chi_m \end{pmatrix} = \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_n \end{pmatrix} \tag{1.2.6}$$

As we can see, here we arrived at an overdetermined system (1.2.6) and we need to describe how to deal with such systems. As we pointed out before, in general there are no such values M_0 and χ_m that will satisfy all the n equations of the system (1.2.6). However, there are such values of quantities M_0 and χ_m that will give us numbers that are very "near" to the right hand side of equation (1.2.6). This problem is analogous to the problem of fitting the "best" line into set of scattered points as it is demonstrated in Figure 1. To deal with such problems, we need to define what does it mean for one vector to be "near" another vector or for the line to passes through scattered points and to be the "best" fitting line. This involves introduction of concepts like "length" and "distance" which we will do next.

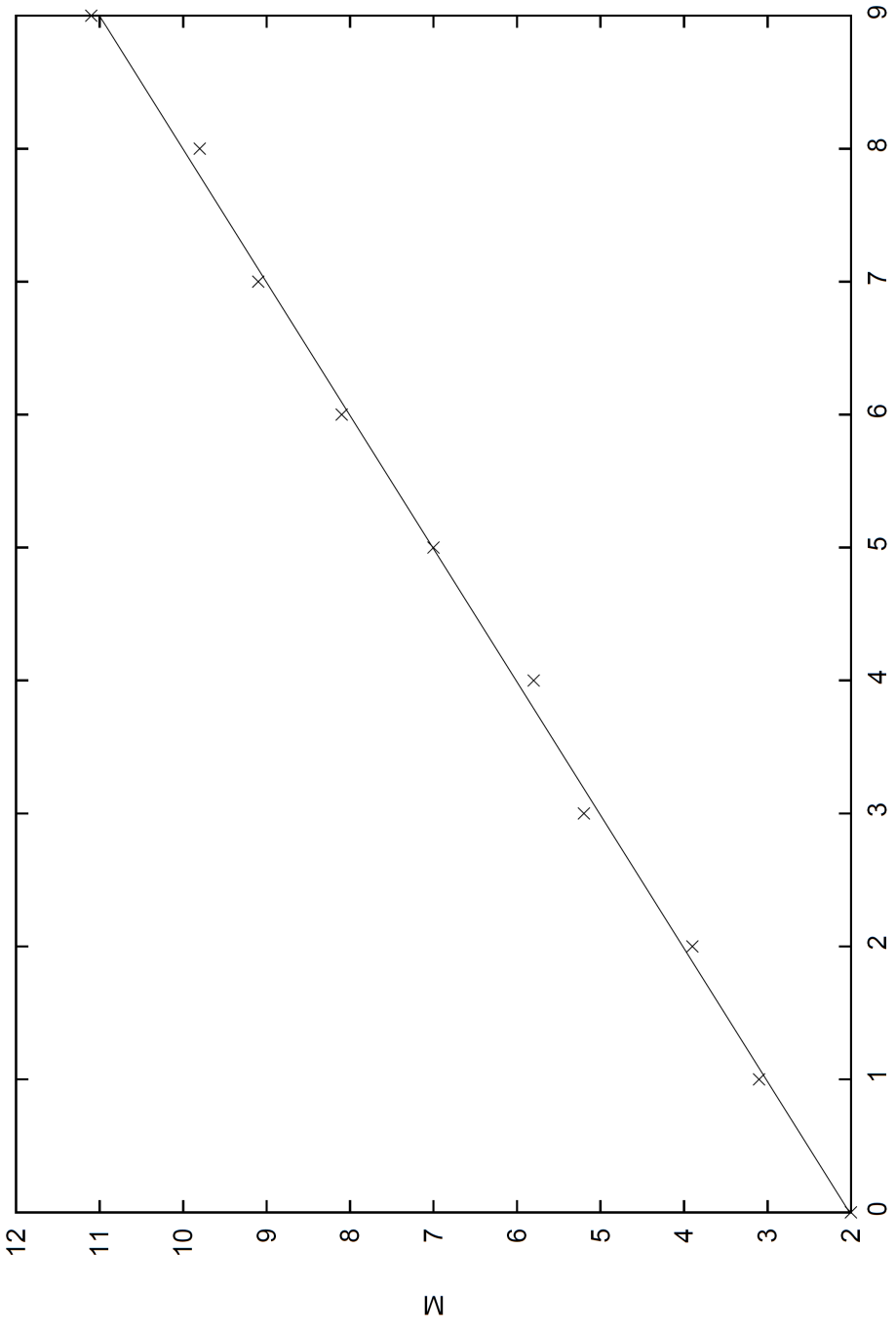

Figure 1: Problem of linear data fitting is equivalent to solving overdeterminent linear system.

1.3 Euclidean space

Scalar product. To accomplish our goal of solving overdeterminant systems and enlarge the area of application of linear algebra to other problems of physics, we need to define some additional concepts. We will start by defining the notion that will allow us to determine a distance between two n dimensional real vectors. This is achieved through the introduction of the concept of a scalar product between two vectors.

The scalar product of two real n dimensional vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is denoted by the "dot" operator and it is usually defined by the following expression

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \tag{1.3.1}$$

The length of the vector $\boldsymbol{x}$ can be defined as a squareroot of the scalar product of the vector with itself

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{\sum_{i=1}^{n} (x_i)^2} \tag{1.3.2}$$

The angle between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is then defined by the formula

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}\right) \tag{1.3.3}$$

where $\cos^{-1}$ is the inverse of the cos function and it returns angles in radians. For the definition (1.3.3) to make sense, one has to check if the following inequality is always satisfied

$$-1 \leq \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} \leq 1 \tag{1.3.4}$$

This relation (1.3.4) is called Cauchy-Schwarz inequality and its proof can be found in any standard textbooks on linear algebra.

With these definition, now we can define the distance between two vectors to be the length of

the $\boldsymbol{x}-\boldsymbol{y}$ vector

$$d(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x}-\boldsymbol{y}\| = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \tag{1.3.5}$$

The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are said to be orthogonal if their dot product is zero: $\boldsymbol{x}\cdot\boldsymbol{y}=0$. The angle between orthogonal vectors is $\theta = \frac{\pi}{2}$. For the two orthogonal vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, there is relation which can be considered as a generalization of the Pythagorean theorem from geometry

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 \tag{1.3.6}$$

To prove the generalized Pythagorean theorem, one just needs to remember that for orthogonal vectors $\boldsymbol{x}$ and $\boldsymbol{y}$: $(\boldsymbol{x}\cdot\boldsymbol{y}) = (\boldsymbol{y}\cdot\boldsymbol{x}) = 0$ and compute $\|\boldsymbol{x}+\boldsymbol{y}\|^2$ as

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = (\boldsymbol{x}+\boldsymbol{y})\cdot\ (\boldsymbol{x}+\boldsymbol{y}) = (\boldsymbol{x}\cdot\boldsymbol{x}) + (\boldsymbol{y}\cdot\boldsymbol{y}) + (\boldsymbol{x}\cdot\boldsymbol{y}) + (\boldsymbol{y}\cdot\boldsymbol{x})\ = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2, \tag{1.3.7}$$

Of course, this identity holds true for an arbitrary set of vectors provided they are pairwise orthogonal. This fact is handy when dealing with orthogonal set of basis.

As one can see from the Pythagorean theorem, the features of Euclidean geometry are emerging one by one through introduction of the scalar product. Therefore the linear vector space with this definition of length and scalar product is naturally called Euclidean space. Later, we will present a slightly different way of looking at scalar product which will allow us to introduce a new linear structure that is a tensor. But before that let us come back to the problem of solving the overdeterminant linear systems.

1.4 Matrix multiplication

Matrix time columns. We know how to multiply any $m\times n$ matrix $\mathcal{A}$ with n dimensional vector $\boldsymbol{x}$. The results is m dimensional vector $\boldsymbol{y}$

$$\mathcal{A}\boldsymbol{x} = \boldsymbol{y} \tag{1.4.1}$$

If on the right side of the matrix $\mathcal{A}$, instead of one vector $\boldsymbol{x}$ we have a set of k vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ put into one matrix $\mathcal{X}$

$$\mathcal{X} = \Big(\boldsymbol{x}_1 \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \Big) = \begin{pmatrix} (\boldsymbol{x}_1)_1 & (\boldsymbol{x}_2)_1 & \cdots & (\boldsymbol{x}_k)_1 \\ (\boldsymbol{x}_1)_2 & (\boldsymbol{x}_2)_2 & \cdots & (\boldsymbol{x}_k)_2 \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{x}_1)_n & (\boldsymbol{x}_2)_n & \cdots & (\boldsymbol{x}_k)_n \end{pmatrix} \tag{1.4.2}$$

we simply multiply the matrix $\mathcal{A}$ with every columns of $\mathcal{X}$ one by one

$$\mathcal{A}\mathcal{X} = \mathcal{A} \Big(\boldsymbol{x}_1 \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \Big) = \Big(\mathcal{A}\boldsymbol{x}_1 \mathcal{A}\boldsymbol{x}_2, \cdots, \mathcal{A}\boldsymbol{x}_k \Big) \tag{1.4.3}$$

The result is a matrix with m rows and k columns.

Formula for elements. In principle, the equation (1.4.3) is all we need to know to perform matrix multiplication. However, for different purposes some other ways of looking to the same operation are more advantageous. For direct computation purposes, for example, it is convenient to have the explicit formula for the elements of resulting matrix. Let us consider two matrices

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} & \cdots & \mathcal{A}_{1m} \\ \mathcal{A}_{21} & \mathcal{A}_{22} & \cdots & \mathcal{A}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{A}_{n1} & \mathcal{A}_{n2} & \cdots & \mathcal{A}_{nm} \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} \mathcal{B}_{11} & \mathcal{B}_{12} & \cdots & \mathcal{B}_{1p} \\ \mathcal{B}_{21} & \mathcal{B}_{22} & \cdots & \mathcal{B}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{B}_{m1} & \mathcal{B}_{m2} & \cdots & \mathcal{B}_{mp} \end{pmatrix} \tag{1.4.4}$$

By multiplying $\mathcal{A}\mathcal{B}$, we get the new matrix

$$\mathcal{A}\mathcal{B} = \begin{pmatrix} (\mathcal{A}\mathcal{B})_{11} & (\mathcal{A}\mathcal{B})_{12} & \cdots & (\mathcal{A}\mathcal{B})_{1p} \\ (\mathcal{A}\mathcal{B})_{21} & (\mathcal{A}\mathcal{B})_{22} & \cdots & (\mathcal{A}\mathcal{B})_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ (\mathcal{A}\mathcal{B})_{n1} & (\mathcal{A}\mathcal{B})_{n2} & \cdots & (\mathcal{A}\mathcal{B})_{np} \end{pmatrix} \tag{1.4.5}$$

and the elements $(\mathcal{AB})_{ij}$ of that new matrix can be computed using

$$(\mathcal{AB})_{ij} = \sum_{k=1}^{m} \mathcal{A}_{ik}\mathcal{B}_{kj} \tag{1.4.6}$$

Block matrix multiplication. There are some other important ways of looking at matrix multiplication. Most notably, block matrix multiplication is very relevant in our work because it is used by the fastest numerical algorithms to do matrix multiplication in computers. From a mathematical point of view, all ways of matrix multiplication are fundamentally the same. However, depending on how matrix is stored in the memory of computer, one way can give faster result than others. In block matrix, multiplication matrix is divided into blocks with a right size and each block is handled separately. Since the elements of every block are stored in a close proximate in memory the computer processor can access to the matrix elements in faster way and thus leads to overall fast execution. Block matrix multiplication is also the basis of fast Fourier transform algorithms. All of these technical details can not be covered in this work and I would like to refer for more information to the relevant literature on numerical computation.

Identity matrix. The matrix with ones on the main diagonal and zeros elsewhere is called identity matrix and it is denoted as I. Thus in n dimensional vector space the identity matrix has the following form

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{1.4.7}$$

When identity matrix I is multiplied with a vector $\boldsymbol{x}$, it acts as a number "one", i.e. it returns the same vector $\boldsymbol{x}$ back

$$I\boldsymbol{x} = \boldsymbol{x} \tag{1.4.8}$$

Both (1.4.7), and (1.4.8) can serve us as a definition for identity matrix since they are equivalent. Another equivalent definition of the identity matrix can be given using the following matrix equation

$$I\mathcal{A} = \mathcal{A}I = \mathcal{A} \tag{1.4.9}$$

Inverse of the matrix. Inversion is an operation that transforms a given matrix into another matrix called its inverse. It is usually denoted by placing "-1" in the superscript. The inverse $\mathcal{A}^{-1}$ of the $n \times n$ matrix $\mathcal{A}$, for example, is another $n \times n$ matrix defined according to the following relation

$$\mathcal{A}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{A} = I \tag{1.4.10}$$

If we consider a matrix as a linear function on vector space, then inverse matrix is the corresponding inverse function. In practice, most softwares use elimination methods to compute inverses of matrices. We will not go into deep discussion about inverting general $n \times n$ matrix since we will not be using them in our work. However, we will use inversion of 2×2 and 3×3 matrix, both of which we will consider next.

Inversion of 2×2 matrix. In general, inversion of a 2×2 matrix

$$\mathcal{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

can be performed using the following formula

$$\mathcal{A}^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{1.4.11}$$

Inversion of 3×3 matrices. The inverse of the matrix

$$\mathcal{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}$$

is calculated according to the following formula

$$\mathcal{A}^{-1} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}^{-1} = \frac{1}{\det(\mathcal{A})} \begin{pmatrix} A & D & G \\ B & E & H \\ C & F & K \end{pmatrix} \tag{1.4.12}$$

where the $\det(\mathcal{A})$ is a scalar number called determinant. For 3×3 matrices, it can be calculated as follows

$$\det(\mathcal{A}) = a(ek - fh) - b(kd - fg) + c(dh - eg) \tag{1.4.13}$$

Determinant is a volume of the parallelepiped made by columns of matrix $\mathcal{A}$. This fact about the determinants are true even in arbitrary dimensional spaces and can serve as a generic definition of determinants. If columns of the matrix $\mathcal{A}$ are independent then $\det(\mathcal{A}) \neq 0$ and matrix $\mathcal{A}$ is convertible.

The rest of the symbols of the equation (1.4.12) are defined as follows

$$\begin{aligned} A &= (ek - fh) \quad D = (ch - bk) \quad G = (bf - ce) \\ B &= (fg - dk) \quad E = (ak - cg) \quad H = (cd - af) \\ C &= (dh - eg) \quad F = (gb - ah) \quad K = (ae - bd). \end{aligned} \tag{1.4.14}$$

We now have enough knowledge under our belt and we can start solving an overdeterminant linear system.

1.5 Normal equation

Now, we can come back to the problem of solving overdeterminate systems. First, we will take a general approach and then we will apply our general result to the particular problem of finding susceptibilities and magnetization.

Let us introduce a new vector $\boldsymbol{\delta}$, the components of which are defined by the equations

$$\delta_i = y_i - \sum_{j=1}^{n} \mathcal{A}_{ij} x_j \tag{1.5.1}$$

If the vector $\boldsymbol{x}$ was the solution for the system (1.2.6) then this vector $\boldsymbol{\delta}$ would be a zero vector. In the case when $\boldsymbol{\delta}$ is not zero vector, it will represent the distance between "particular solution" and the vector $\boldsymbol{y}$. Therefore, our goals come to finding such a set of x_i's that would give the minimum for the length of $\boldsymbol{\delta}$ vector. If we denote the square of the length of the $\boldsymbol{\delta}$ vector by the capital letter Δ, then we can write

$$\Delta = \sum_{i=1}^{n} \delta_i^2 \tag{1.5.2}$$

To find the minimum of Δ, we can use a theorem from differential calculus. For that purpose, we should first differentiate Δ with respect to x_j's

$$\frac{\partial \Delta}{\partial x_j} = 2 \sum_{i=1}^{m} \delta_i \frac{\partial \delta_i}{\partial x_j} \tag{1.5.3}$$

According to the definition (1.5.1) of $\boldsymbol{\delta}$ vector

$$\frac{\partial \delta_i}{\partial x_j} = -\mathcal{A}_{ij} \tag{1.5.4}$$

Thus the expressions for the partial derivatives (1.5.3) becomes

$$\frac{\partial \Delta}{\partial x_j} = 2 \sum_{i=1}^{m} \left(y_i - \sum_{k=1}^{n} \mathcal{A}_{ik} x_k \right) (-\mathcal{A}_{ij}) \tag{1.5.5}$$

To find the x_j values that give the minimum for Δ and consequently the minimum for the length of

the $\boldsymbol{\delta}$ vector, we need to set partial derivatives (1.5.3) to zero

$$2\sum_{i=1}^{m}\left(y_i-\sum_{k=1}^{n}\mathcal{A}_{ik}\hat{x}_k\right)(-\mathcal{A}_{ij})=0 \tag{1.5.6}$$

Here we used $\hat{x}$ instead of just x to indicate that they represent not the solution but the closest solution. By simple manipulations, we get our final desired result

$$\sum_{i=1}^{m}\sum_{k=1}^{n}\mathcal{A}_{ij}\mathcal{A}_{ik}\hat{x}_k=\sum_{i=1}^{m}\mathcal{A}_{ij}y_i \tag{1.5.7}$$

Equation (1.5.7) is very important for us and it is sometimes called normal equation. The same equation can be written in the language of linear algebra in a more compact form

$$\boxed{(\mathcal{A}^T\mathcal{A})\hat{\boldsymbol{x}}=\mathcal{A}^T\boldsymbol{y}} \tag{1.5.8}$$

In equation (1.5.8), $\mathcal{A}^T$ represents the transpose of the $\mathcal{A}$ matrix which is achieved by reflecting the original matrix $\mathcal{A}$ over its main diagonal (i.e., over the diagonal which runs from top-left corner to bottom-right corner of the matrix $\mathcal{A}$).

The reason why it is impossible to solve overdeterminant systems in traditional way is that the right-hand side of the linear equation $\boldsymbol{y}$ is not in the columns space of the matrix $\mathcal{A}$, i.e. there is no linear combination of columns, of $\mathcal{A}$ that would give us the vector $\boldsymbol{y}$. What we did above can be summarized as follows. First, we chose the vector from the column space of matrix $\mathcal{A}$ which had a minimum distance from the vector $\boldsymbol{y}$. This step essentially represents taking projection of the vector $\boldsymbol{y}$ into the column space of $\mathcal{A}$. Second, we find a linear combination of columns of $\mathcal{A}$ that would give that projection. This is the geometrical meaning of the normal equation. There are also important statistical meanings of the normal equation, but we will not consider them. Next we will use the normal equation (1.5.8) to solve our original problem.

Computation of susceptibilities using normal equation. Now, we are ready to compute susceptibilities from the set of data for magnetic field H_i and magnetization M_i. According the formula (1.5.8), we first need to multiply both sides of the equation by the transposed matrices

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ H_1 & H_2 & \cdots & H_n \end{pmatrix} \begin{pmatrix} 1 & H_1 \\ 1 & H_2 \\ \vdots & \vdots \\ 1 & H_n \end{pmatrix} \begin{pmatrix} M^0 \\ \chi_m \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ H_1 & H_2 & \cdots & H_n \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_n \end{pmatrix} \tag{1.5.9}$$

After multiplication of matrices in both sides of the equations, we get the following resulting matrix relation

$$\begin{pmatrix} n & \sum_{i=1}^n H_i \\ \sum_{i=1}^n H_i & \sum_{i=1}^n H_i^2 \end{pmatrix} \begin{pmatrix} M^0 \\ \chi_m \end{pmatrix} = \begin{pmatrix} \sum_i^n M_i \\ \sum_i^n H_i M_i \end{pmatrix} \tag{1.5.10}$$

We can transfer the matrix to another side of the equation by inverting it

$$\begin{pmatrix} M^0 \\ \chi_m \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n H_i \\ \sum_{i=1}^n H_i & \sum_{i=1}^n H_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_i^n M_i \\ \sum_i^n H_i M_i \end{pmatrix} \tag{1.5.11}$$

Using formula for the inversion of 2×2 matrices and doing matrix to vector multiplication, we get

$$\begin{pmatrix} M^0 \\ \chi_m \end{pmatrix} = \frac{1}{n\sum_{i=1}^n H_i^2 - (\sum_{i=1}^n H_i)^2} \begin{pmatrix} \sum_i^n M_i \sum_i^n H_i^2 - \sum_i^n H_i \sum_i^n H_i M_i \\ n\sum_i^n H_i M_i - \sum_i^n H_i \sum_i^n M_i \end{pmatrix} \tag{1.5.12}$$

This result has very interesting statistical meaning as we will see. However, we will first consider obvious generalizations of our result for slightly different problems.

Magnetoelectric effect. The magnetoelectric (ME) effect refers to the phenomenon of induction of magnetization by an external electric field and induction of electric polarisation by an exter-

nal magnetic field.[6] Thus, the resulting polarisation and magnetisation due to ME effect can be expressed in the following form

$$P_i = P_i^0 + \sum \alpha_{ij} H_j + \sum \beta_{ijk} H_j H_k + \dots \tag{1.5.13}$$

$$M_i = M_i^0 + \sum \alpha_{ij} E_j + \sum \gamma_{ijk} E_j E_k + \dots \tag{1.5.14}$$

In these expressions, P_i^0 represents the component of electric polarization, and M_i^0 the component of magnetization in the absence of external electric E_i and magnetic H_i fields respectively. The parameters α_{ij}, β_{ijk} and γ_{ijk} are tensor quantities and they are known as the linear (for α_{ij}) and quadratic (for β_{ijk} and γ_{ijk}) ME coefficients. In general, these parameters will depend on temperature.

Although multiferroics are the primary candidate for magnetoelectric effect, one should note that this effect can be observed in much larger class of materials. For example, at high temperature, $BiFeO_3$ loses its magnetic ordering, but its β_{ijk} quadratic magnetoelectric coefficients are not zero.

Again, we will ignore the tensor character of the coefficients and outline the way how we will compute values of those coefficients when we will present the results for those coefficients. In a scalar form (1.5.13) can be written as

$$P = P^0 + \alpha H + \beta H^2 \tag{1.5.15}$$

In the case when $\beta = 0$, we can use our previous result (1.5.12) without any modifications. If on the other hand, $\alpha = 0$ we can generalize our previous result easily by substituting H^2 in any

[6] P. Curie deserves a credit for proposing the possibility of ME effect in the year 1894. The term itself however, was introduced by P. Debye in year 1926.

place where we see H

$$\begin{pmatrix} 1 & H_1^2 \\ 1 & H_2^2 \\ \vdots & \vdots \\ 1 & H_n^2 \end{pmatrix} \begin{pmatrix} P^0 \\ \alpha \end{pmatrix} = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{pmatrix} \tag{1.5.16}$$

Actually, this later trick can be used not only in the case of square dependence of P on H but in the case of any functional dependence $f(H)$. This allows us to fit coefficients a, and b in the relation

$$g(x) = a + bf(x)$$

for arbitrary functions $g(x)$ and $f(x)$. For example, we can use this result to fit parameters into the Curie's law which we will meet later.

However, if both coefficient α and β are present we will need to use the normal equation (1.5.8) again and derive corresponding formulas for magnetoelectric coefficients. This will involve, for example, inverting 3×3 matrix instead of 2×2 matrix. The final analytical expression arrived at in that way is not simple and will not be given here. Instead, we will implement it directly in the computer program that we will present later.

1.6 Programming tools

In the following pages of this work, methods of applied mathematics and their implementation using computers will be presented together. Therefore, before going any further with our descriptions of methods, we will need to outline some computational tools. These tools will be used to implement mathematical methods to solve physical problems.

UNIX environment. All the computations that are described in this work were carried out in the UNIX environment and in this section I would like to mention things about UNIX environment that is relevant for understanding and reproducing results described here.

UNIX is a name of operating system which is no longer around us. However, today there are

many operating systems that imitate that original system and share with it fundamental philosophy of work and basic conventions. Thus UNIX became the name for the family of the operating systems. Today when one needs reliable system for handling large amount of data, UNIX is the operating system of choice. Many UNIX systems work for years without rebooting and failing.

Linux is a form of Unix which is free of charge and it is distributed as an open software. It is spreading in the computer world very rapidly and is currently dominant as a server operating system. In the future discussions we will assume this form of UNIX. However, nothing we will say will be Linux specific. It will thus be true for other flavors of UNIX as well.

AWK utility. Linux comes with additional scripting tools that helps us to perform some of the routine tasks. AWK utility in particular is very helpful in the kind of work that we commonly do in scientific computations. This utility is very flexible and can be considered as a programming language by itself. In our work, some of the mathematical methods are implemented as a AWK script and we will present them in the following sections as such.

The C programming language. During this work it will be necessary to present some portions of source code. The language of choice for this purpose will be standard C programming language. C originally was written by assembly language programmers to program computers.[7] Therefore it is the first choice for the system level programming. Linux for example was written in C and it is an excellent choice to program in the Linux and other Unix environments. However, today in many areas of science and engineering where speed of execution and control over hardware is essential, C is widely used as a language of choice.

The language specification. Initially, there was no formal definition for the language. The first standard specification for C was C90 (also known as C89). Decade later the standards committee

[7]Unix operating system first was written in assembly language. To make some periodically occurring work easy to do, the language C was created. This is the reason why C language sometimes is referred to as portable assembly language. It gives high control over hardware second only to assembly.

came up with some additions to the language and the resulting standard was called C99. Any standard C99 complaint compiler should be able to compile every code that are presented in this work. I have tested every source code using GNU C compiler with -std=c99 option specified to enforce the C99 compliance.[8]

Programming conventions. All the programs that are presented in this work behave as a standard UNIX filters. Filters are programs that expect two kinds of command line arguments. First kind of arguments is the sequence of characters that follow the " – " character. These characters modify the default behavior of the program. All the other kinds of arguments are assumed to be the names of text files to be processed. If no file names are given, the program expects its input from standard input, which is usually the keyboard but it can be modified to be a specific file or output from some other program. The output of the program by default goes to standard out which in Linux is usually the screen. However, Linux makes it easy to redirect the output of the programs to a file or to another program.

UNIX philosophy. Users of UNIX have a particular approach to system design and software development. This approach is not imposed on the users and the system can be used in any way one desires. However, this way of working with system was with UNIX culture from the very beginning. The main idea is to have tiny programs each of which is written to perform very simple task. Complicated tasks then are accomplished by combining these building block programs together. For these arrangement to make sense, every program should be able to work with any other existing program in the system and even with programs that are not written yet. UNIX makes possible to accomplish this through the mechanisms like pipes and redirected inputs and outputs. Another things that makes this work is the ubiquity of the text in the workings of the UNIX system. Programs typically accept their inputs and produce their outputs in the form of

[8]Today there is new standard for the language C11. However, currently there is no compiler that implemented all the features of the newest standard. Although there are a few features that were added, the language C remains fundamentally the same and every code that we present here can be compiled with new compilers without any modifications to them.

plain ascii text. Even system configuration files are in the form of readable plane text files. Thus text is the universal media for the programs to communicate with each other. Individual program does not need to know from where its input is coming from or which program is using its output. It just receives input in well defined format, performs its own job, and outputs the result in a well defined format. The next program start from there and this chain continues until the final desired result is obtained.

This tiny program tradition is sometimes referred to as UNIX philosophy.[9] We have covered enough mathematical ground and now I want to demonstrate how in practice this approach works.

We have covered basics of linear algebra and the basic structure in the linear algebra is a vector. Vector can be stored in simple text file as columns of numbers. The name of file will represent the name of vector. Thus the file "`M`" can be used to store values of magnetization and file "`H`" to store corresponding values of magnetic field. Even when we store matrix it is convenient to store each of its columns in a separate files. When it is necessary to combine those column files into one matrix like file, it always can be done using standard "`paste`" utility that available in every UNIX systems. For example, we can combine N column files: "`file1, file2, ... fileN`" into one file called "matrix" using the following command

```
$ paste file1 file2 ... fileN > matrix
```

The character "`$ `" represent a sell prompt and ">" instructed the shell to redirect the output to the file "`matrix `".

If the "`-s`" option is specified, `paste` utility makes column files into row files and combines them that way. The resulting file would be the transposed form of "matrix" file. Thus after the command

```
$ paste -s file1 file2 ... fileN  >  matrixT
```

the file "`matrixT`" will contain transposed version of the file "`matrix`".

We also can combine columns vertically using "`cat`" command

[9]Of course UNIX introduced to computing world many other innovations. However, most of those innovations are invisible to ordinary users.

```
$ cat file1 file2 ... fileN > matrix
```

"cat" stands for "concatenation" and it lists files one after another to its output. If only one file is specified, cat program lists its contents. Thus, we will use it to show the text programs in the following pages.

Both programs paste and cat which were written with completely different purposes in mind. However, now they are useful to us to perform basic operations of linear algebra. This is UNIX philosophy in action. This way of working is very flexible and infinitely extensible.

Let us now start to write parts of program to solve overdeterminant systems. First program is to compute $(\mathcal{A}^T)(\mathcal{A})$

```
$ cat aTa2.awk


#!/usr/bin/awk -f
# Program to compute square of matrix: (A^T)(A)
{
    a=a+$1*$1
    b=b+$1*$2
    c=c+$2*$1
    d=d+$2*$2
}
END{
    print a, b
    print c, d
}
```

The same program for 3×3 matrices can be implemented as follows

```
$ cat aTa3.awk
```

```
#!/usr/bin/awk -f
# Program to compute square of matrix: (A^T)(A)
{
    a = a + $1*$1
    b = b + $1*$2
    c = c + $1*$3
    d = d + $2*$1
    e = e + $2*$2
    f = f + $2*$3
    g = g + $3*$1
    h = h + $3*$2
    k = k + $3*$3
}
END{
    print a, b, c
    print d, e, f
    print g, h, k
}
```

The next program computes inverse of the 2×2 matrix using the formula (1.4.11) that we presented before

```
$ cat inv2.awk


#!/usr/bin/awk -f
# Program to invert 2x2 matrix
NR==1{
      a=$1
      b=$2
```

```
}
NR==2{
      c=$1
      d=$2
}
END{
# Determinant:
      det = a*d - b*c
      print  d/det, -b/det
      print -c/det,  a/det
}
```

The AWK script that computes the inverse of 3×3 matrices can be implemented as follows

```
$ cat inv3.awk

#!/usr/bin/awk -f
# Program to compute inverse of 3x3 matrix
NR==1{
      a=$1
      b=$2
      c=$3
}
NR==2{
      d=$1
      e=$2
      f=$3
}
```

```
NR==3{
      g=$1
      h=$2
      k=$3
}
END{
# Determinant:
      det = a*(e*k - f*h) - b*(k*d - f*g) + c*(d*h - e*g)

# Row 1:
      A = (e*k - f*h)/det
      D = (c*h - b*k)/det
      G = (b*f - c*e)/det
      print A, D, G

# Row 2:
      B = (f*g - d*k)/det
      E = (a*k - c*g)/det
      H = (c*d - a*f)/det
      print B, E, H

# Row 3:
      C = (d*h - e*g)/det
      F = (g*b - a*h)/det
      K = (a*e - b*d)/det
      print C, F, K
}
```

We will often need to create a vector that has a constant value. For example, when we derived normal equation for the magnetic susceptibility, we had a column that consisted of all ones. To implement that, as a computer program written in C, we do

```
$ cat const.c


#include <stdio.h>
#include <stdlib.h>


int main(int argc, char * argv[]){
    int size = atoi(argv[1]);
    double constant = atof(argv[2]);


    for (int i=0; i < size; i++) printf("%f \n", constant);
}
```

The first argument for the program is a size of a vector and the second is a value of the constant elements.

We can compile the program as follows

```
gcc -std=c99 const.c -o const
```

and we can use it to create vector with 15 elements equal to $\pi \approx 3.1415$

```
$ const 15 3.1415
```

To multiply a 2×2 matrix $\mathcal{A}^T$ with a vector we will use the following script

```
#!/usr/bin/awk -f
# Multiply A^T with a vector:
{
```

```
    a=a+$1*$3
    b=b+$2*$3
}
END{
    print a
    print b
}
```

To multiply a 2×2 matrix $\mathcal{A}^T$ with a vector

```
#!/usr/bin/awk -f
# Multiply A^T with a vector:
{
    a=a+$1*$4
    b=b+$2*$4
    c=c+$3*$4
}
END{
    print a
    print b
    print c
}
```

Now, we have all the programs that we can use to solve overdeterminant systems. Of course, once written, these programs can be used for many other purposes and combined in unexpected ways. We should remark that these programs and programs that we will present in the future are for illustrative purpose, and do not handle unexpected situations like input with a wrong format. This is done to keep the size of programs as small as possible.

1.7 Tensors

Dual space. According to our conventions, by a vector $\boldsymbol{x}$ we understand a column vector

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \tag{1.7.1}$$

Now, we would like to introduce for any such vector $\boldsymbol{x}$ a corresponding row vector $\boldsymbol{x}^T$ called its transpose

$$\boldsymbol{x}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}^T = \begin{pmatrix} x_1 \, x_2 \, \ldots \, x_m \end{pmatrix} \tag{1.7.2}$$

Thus the transposed column vector is just a row vector, and of course, by transposing the same vector $\boldsymbol{x}$ twice we get that vector $\boldsymbol{x}$ back again

$$(\boldsymbol{x}^T)^T = \boldsymbol{x}.$$

The transpose of a vector is sometimes called its dual. The set of all dual vectors makes up a new vector space called the dual space of the original vector space. When vector space and its dual are considered together, vectors of the "original" space are called *contravariant* and vectors of dual space are called *covariant* vectors.

Linear functions on vector space and scalar product. We already considered the notion of a scalar product between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ and denoted such operation using dot operator as $\boldsymbol{x} \cdot \boldsymbol{y}$. However, if we consider $\boldsymbol{x}^T$ as a $1 \times n$ matrix, we can express a dot product as $\boldsymbol{x}^T \boldsymbol{y}$, i.e. without placing the dot operator between the two vectors. In such expressions, $\boldsymbol{x}^T$ behaves as a

function that takes $\boldsymbol{y}$ as an argument and returns a scalar value. It is easy to check that the function defined in this way is linear, i.e. for the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ and for the scalar α, it has the following properties

$$f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y}) \tag{1.7.3}$$

$$f(\alpha\boldsymbol{x}) = \alpha f(\boldsymbol{x}) \tag{1.7.4}$$

In a more conventional notation, the linearity of the scalar product can be expressed in the following form

$$\boldsymbol{z}^T(\boldsymbol{x}+\boldsymbol{y}) = (\boldsymbol{z}^T\boldsymbol{x}) + (\boldsymbol{z}^T\boldsymbol{y}) \tag{1.7.5}$$

$$\boldsymbol{z}^T(\alpha\boldsymbol{x}) = \alpha(\boldsymbol{z}^T\boldsymbol{x}) \tag{1.7.6}$$

Of course, the scalar product equivalently can also be considered as a result of action of the column vector on row vector. In that alternative view, column vector plays the role of linear function and row vectors behave as an argument. Linearity of this operation then will be expressed as follows

$$(\boldsymbol{x}^T+\boldsymbol{y}^T)\boldsymbol{z} = (\boldsymbol{x}^T\boldsymbol{z}) + (\boldsymbol{y}^T\boldsymbol{z}) \tag{1.7.7}$$

$$(\alpha\boldsymbol{x}^T)\boldsymbol{z} = \alpha(\boldsymbol{x}^T\boldsymbol{z}) \tag{1.7.8}$$

Thus the whole theory is symmetric with respect to row vectors and column vectors.

Equivalence of linear functions and dual vectors. Actually, the product between dual vectors is not just an example of linear function but it can be shown that for any linear function there is an equivalent dual vector. The proof of this fact is simple and can be found in reference [7]. This way of looking at scalar product (i.e., as a function on vector space) might seem unnecessary complication at first; however, in this form it is naturally extendable to the idea of tensors which we will do next.

Tensors as a linear function on a vector space. Tensors are defined as a linear function on vector space that returns, single scalar value. However, tensors in general can accept more than one vector as an argument and those vectors can be row vectors as well as column vectors. For example, $n \times n$ matrix $\mathcal{A}$ accepts n dimensional row vector $\boldsymbol{x}^T$ and n dimensional column vector $\boldsymbol{y}$ and returns a scalar $(\boldsymbol{x}^T \mathcal{A}\, \boldsymbol{y})$. Thus matrix is a function that accepts one row vector and one column vector and returns a scalar, i.e. it is a tensor. If a tensor accepts k vectors as an argument and returns a single scalar, then we will say that this tensor has a rank k.

Again, as in the case of matrices, such generalized linear function also can be represented by a finite set of numbers. As a matter of fact, this is an alternative way of introducing tensors. Let us outline this way of looking at tensors.

So far, elements of vectors that we considered were scalars. However, it is meaningful to consider, for example, column vectors, elements of which are row vectors. Actually we already did, and called such structures as matrices. Now we can consider vectors, elements of which are matrices, or matrices elements of which are other matrices. Possibilities of such generalizations are endless, and data structures constructed with this general idea can be shown to be equivalent to the concept of tensors.

This second way of introducing tensors is common in physics and engineering. If we decide to define tensors as a set of numbers, however, we have to explicitly indicate rules for transforming tensors from one coordinate system into another. On the other hand, if we define tensors as a linear function on vector space, the rules of transformation can be deduced as consequence of that definition. Therefore, to consider tensors as linear functions is more appealing from a mathematical point of view since all of its properties are the consequence of its definition.

Index notation for tensors. Representation of a generic tensor as rows and columns (as we did in the case of matrices) is very limited when tensor has a rank three or higher. More flexible notation is an index notation. In such notation, every index will indicate a single rank of the tensor. Every rank can be written as subscript or superscript. Subscript will represent column type

argument and superscript will represent row type argument. For example, if we want to conform to this convention we should denote matrix as $\mathcal{A}^i_j$. This notation is also meaningful if we remember that a matrix is a column of rows and, according to our convention, row index is denoted as a superscript and column index denoted as a subscript.[10] However, this notation will conflict with other notations that we already adopted and we will not use it in our work.

1.8 Three dimensional vector space

Vector space of three dimensions. Three dimensional vector space is special. This is not only because the physical space has three dimensions. In three dimensions for every two noncolinear vectors, there is one perpendicular direction, and this mathematical fact allows us to define cross product. This is not true in four dimensions, for example. In four dimensions, for two vectors there is a whole perpendicular plane and every vector in that plane is perpendicular to both vectors. In other words, for any two vectors, there is no unique third vector and the concept of cross product can not be defined in a unique manner in four dimensions. This and other important "accidental" features of three dimensional vector space give us an excuse to study it on its own right.

Representation of physical space. Three dimensional physical space is studied by representing it as a set of all possible three dimensional vectors with real elements. In this context, vectors denoted as letter $\boldsymbol{r}$ or $\boldsymbol{R}$ and they are called radius vectors. We will adopt the notation in which small letter $\boldsymbol{r}$ denotes arbitrary point in three dimensional space and capital letter $\boldsymbol{R}$ denotes only points with integer coordinates. The latter set of points are also known as Bravais lattices and we will consider them later in the context of crystalline matter.

[10] Strictly speaking, $\mathcal{A}_{ij}$ would be columns of columns and it would act on two column vectors. We should also point out that in index notation we do not need to keep strict order of tensors and vectors

$$\mathcal{A}^{lmn}_{ijk} x_l y^k z_n = x_l \mathcal{A}^{lmn}_{ijk} y^k z_n = \mathcal{A}^{lmn}_{ijk} y^k x_l z_n$$

As long as the names of indexes are kept straight, it is clear on which vectors the tensor $\mathcal{A}^{lmn}_{ijk}$ is acting.

Three very special vectors deserve the special notations

$$\boldsymbol{\imath} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{\jmath} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{1.8.1}$$

Unless otherwise specified, these vectors are assumed to be chosen as basis. This basis is called standard basis, natural basis or canonical basis. These three vectors are mutually orthogonal and each of them has a unit length.[11]

When natural basis are used, elements of any vector $\boldsymbol{r}$ are also its coordinates on this special basis

$$\boldsymbol{r} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = r_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + r_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + r_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = r_1\boldsymbol{\imath} + r_2\boldsymbol{\jmath} + r_3\boldsymbol{k} \tag{1.8.2}$$

Thus, in physics, we use words like coordinates of a vector and elements of a vector interchangeably. However, one must remember that, in general, such loose use of language might be ambiguous.

In three dimensions, instead of numbering coordinates and elements by 1, 2, 3, indexes are given special names: x, y, z; and sometimes the unit basis vectors themselves are denoted by placing hats above the corresponding indexes: $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$. Actually, hat is always used to indicate that vector has a unit length.[12] For example, the unit normal vector along $\boldsymbol{n}$ can be expressed as

$$\hat{\boldsymbol{n}} = \frac{\boldsymbol{n}}{\|\boldsymbol{n}\|} \tag{1.8.3}$$

Now we have a notion of three dimensional space, and we can start considering the concept of fields defined on it.

[11] This is essentially introduction of Cartesian coordinate system without help of geometry.

[12] The hat notation in quantum mechanics will denote operators.

Fields. In general, field is a physical quantity that depends on position in space. If the quantity in question is a scalar quantity, then the corresponding field is called a scalar field. Electric charge density inside the matter is an example of a scalar field. In physics, we also often encounter vector quantity that depends on position in space. Such a quantity is called a vector field. Example of vector field is a magnetic dipole moments inside the matter. One natural generalization of scalar and vector fields is tensor field. From this point of view, a scalar field is a tensor field of rank zero, and vector field is a tensor field of the rank one.

Magnetic structure as a vector field. Three examples of the vector field will be important for us in the future sections. The first example is a constant vector field. Constant vector field is the simplest field one can imagine, and can be given by the expression

$$\boldsymbol{m}(\boldsymbol{R}) = \boldsymbol{m}_0 \tag{1.8.4}$$

where $\boldsymbol{R}$ is a lattice point and $\boldsymbol{m}_0$ is a magnetic moment at that point. This kind of field is appropriate to describe distribution of magnetic moments in ferromagnetic materials at zero Kelvin.

Another vector field can be given by slightly complicated expression as follows

$$\boldsymbol{m}(\boldsymbol{R}) = (-1)^{i+j+k}\boldsymbol{m}_0 \tag{1.8.5}$$

where i, j, and k are the coordinates of the lattice point $\boldsymbol{R}$ in some basis set $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$

$$\boldsymbol{R} = i\boldsymbol{a} + j\boldsymbol{b} + k\boldsymbol{c} \tag{1.8.6}$$

This kind of field can be used to describe magnetic structure called G-type antiferromagnetic.

More general and at the same time more complicated forms of magnetic moment distribution is described by a cycloidal field. In a cycloidal field, the end point of the vector $\boldsymbol{m}_0$ circularly rotates with a constant rate as one moves from one lattice point to another along the direction given by a

vector $\boldsymbol{k}$. In a simple case, the rotation happens in a constant plane called cycloidal plane. Such cycloidal magnetic structure can be described by the following vector field

$$\boldsymbol{m}(\boldsymbol{R}) = m_0 \left[\hat{\boldsymbol{k}} \sin(\boldsymbol{k}\cdot\boldsymbol{R}) + \hat{\boldsymbol{p}} \cos(\boldsymbol{k}\cdot\boldsymbol{R})\right] \tag{1.8.7}$$

In this expression, $\hat{\boldsymbol{k}}$ and $\hat{\boldsymbol{p}}$ represent unit vectors that define the cycloidal plane and the magnetic moments entirely lie on that plane. The propagation direction and the period of the cycloid is determined by the vector $\boldsymbol{k}$. In the case when

$$\boldsymbol{k} = 0 \Rightarrow (\boldsymbol{k}\cdot\boldsymbol{R}) = 0$$

the structure described by a cycloidal field is ferromagnetic. If on the other hand , when

$$(\boldsymbol{k}\cdot\boldsymbol{R}) = \pi n$$

where n is an integer, we get an antiferromagnetic structure. In that sense the cycloidal magnetic order is a generalization of the ferromagnetic and antiferomagnetic orders.

Functions on vector fields. Having introduced three examples of vector field, we now can consider function on those fields. Let us start with a function that return vectors. The first function that we will consider is called magnetization vector $\boldsymbol{M}$ and it is defined by the following expression

$$\boldsymbol{M} = \frac{1}{V}\sum_{\boldsymbol{R}} \boldsymbol{m}(\boldsymbol{R}) \tag{1.8.8}$$

where V is a volume. The magnetization is the quantity that represents an average value of magnetic moments. Another function that returns a vector is called antiferromagnetic vector $\boldsymbol{L}$ and it is defined using the following expression

$$\boldsymbol{L} = \frac{1}{V}\sum_{\boldsymbol{R}} (-1)^{i+j+k} \boldsymbol{m}(\boldsymbol{R}) \tag{1.8.9}$$

As one can see, if the material is pure ferromagnetic then

$$\text{PURE FERROMAGNETIC} \begin{cases} \boldsymbol{M} & = \frac{\boldsymbol{m}_0}{V_0} \\ \boldsymbol{L} & = 0 \end{cases} \tag{1.8.10}$$

where V_0 is a volume of the unit cell. On the other hand, if the material under considerations is pure antiferromagnetic then

$$\text{PURE ANTIFERROMAGNETIC} \begin{cases} \boldsymbol{M} & = 0 \\ \boldsymbol{L} & = \frac{\boldsymbol{m}_0}{V_0} \end{cases} \tag{1.8.11}$$

In general, of course both of these vectors can be zero (for example in paramagnets) or nonzero (for example in antiferromagnets with canted moments).

Characterisation of the cycloidal magnetic order is not a trivial task. If we know the propagation direction and the plane of the cycloid, one way to quantify the cycloidal order would be by computing the following quantity

$$Cyc = \frac{1}{V} \sum_{\boldsymbol{R}} \left[\hat{\boldsymbol{k}} \sin(\boldsymbol{k} \cdot \boldsymbol{R}) + \hat{\boldsymbol{p}} \cos(\boldsymbol{k} \cdot \boldsymbol{R}) \right] \cdot \boldsymbol{m}(\boldsymbol{R}) \tag{1.8.12}$$

It is easy to check that if the magnetic state is cycloidal, then $Cyc = \frac{m_0}{V_0}$. However in practice, we do not know what is the magnetic structure of the system beforehand. Therefore, we first have to try all possible $\boldsymbol{k}$ vectors and see which one of them gives us the largest value for the Cyc parameter. This is the subject of Fourier analysis which we will cover later. Next, we will outline the procedure by which we can determine the cycloidal plane if we know the propagation direction $\boldsymbol{k}$. For that purpose, we have to introduce the concept of cross product first.

Cross Product. We can define the cross product between two three–dimensional vectors by first introducing the cross product for natural basis $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$. Then, we can define this operation for arbitrary vectors.

The cross product is a binary operation denoted by symbol "$\times$" and for natural basis

$$\begin{aligned} \boldsymbol{i} \times \boldsymbol{j} &= \boldsymbol{k}, \\ \boldsymbol{j} \times \boldsymbol{k} &= \boldsymbol{i}, \\ \boldsymbol{k} \times \boldsymbol{i} &= \boldsymbol{j}, \end{aligned} \tag{1.8.13}$$

The cross product of the same base vector with itself is zero

$$\boldsymbol{i} \times \boldsymbol{i} = \boldsymbol{j} \times \boldsymbol{j} = \boldsymbol{k} \times \boldsymbol{k} = \boldsymbol{0} \tag{1.8.14}$$

The change in order of the operands leads to the change in the sign of the cross product

$$\begin{aligned} \boldsymbol{j} \times \boldsymbol{i} &= -\boldsymbol{k}, \\ \boldsymbol{k} \times \boldsymbol{j} &= -\boldsymbol{i}, \\ \boldsymbol{i} \times \boldsymbol{k} &= -\boldsymbol{j}. \end{aligned} \tag{1.8.15}$$

Since any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, in the natural basis $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ can be expressed by their coordinates, we can define the cross product between them as follows

$$\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} = (a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}) \times (b_1\boldsymbol{i} + b_2\boldsymbol{j} + b_3\boldsymbol{k}) = \tag{1.8.16}$$

$$a_1b_1\boldsymbol{i} \times \boldsymbol{i} + a_1b_2\boldsymbol{i} \times \boldsymbol{j} + a_1b_3\boldsymbol{i} \times \boldsymbol{k} + \tag{1.8.17}$$

$$a_2b_1\boldsymbol{j} \times \boldsymbol{i} + a_2b_2\boldsymbol{j} \times \boldsymbol{j} + a_2b_3\boldsymbol{j} \times \boldsymbol{k} + \tag{1.8.18}$$

$$a_3b_1\boldsymbol{k} \times \boldsymbol{i} + a_3b_2\boldsymbol{k} \times \boldsymbol{j} + a_3b_3\boldsymbol{k} \times \boldsymbol{k} = \tag{1.8.19}$$

$$c_1\boldsymbol{i} + c_2\boldsymbol{j} + c_3\boldsymbol{k} \tag{1.8.20}$$

with

$$
\begin{aligned}
c_1 &= a_2 b_3 - a_3 b_2 \\
c_2 &= a_3 b_1 - a_1 b_3 \\
c_3 &= a_1 b_2 - a_2 b_1
\end{aligned}
\tag{1.8.21}
$$

In a vector notation

$$
\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \tag{1.8.22}
$$

For analytic computations, rules for cross product can be also summarized concisely in using determinant

$$
\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \tag{1.8.23}
$$

If we know the direction of propagation of the cycloid, we can define the plane of the rotation by taking cross product between neighboring vectors along that propagation direction. Averaging all of those cross products will result in a vector that approximately defines the plane of rotation. This way of defining the rotation plane is not perfect since the normal vector changes a lot for slight deviations from the plane. However, this will be good enough for our purposes in the future. AWK program to compute such a average plane using this idea can be implemented as follows

```
cat cross.awk


#!/usr/bin/awk -f


{
    ax=$1; ay=$2; az=$3;
    if (NR > 1) {
```

```
        cx=ay*bz-az*by;
        cy=az*bx-ax*bz;
        cz=ax*by-ay*bx

        avCx=avCx+cx;
        avCx2=avCx2+cx*cx;

        avCy=avCy+cy;
        avCy2=avCy2+cy*cy;

        avCz=avCz+cz;
        avCz2=avCz2+cz*cz;

        print cx, cy, cz
    }
    bx=ax;
    by=ay;
    bz=az;
} END {
    avCx2=avCx2/(NR-1);
    avCy2=avCy2/(NR-1);
    avCz2=avCz2/(NR-1);

    avCx=avCx/(NR-1);
    avCy=avCy/(NR-1);
    avCz=avCz/(NR-1);
```

```
    print "-------------------------------------"

    devCx=sqrt(avCx2-avCx*avCx);
    devCy=sqrt(avCy2-avCy*avCy);
    devCz=sqrt(avCz2-avCz*avCz);

    print "Deviation: ", devCx, devCy, devCz;
    print "Average: ", avCx, avCy, avCz;
}
```

This program assumes that input file has three columns, each of which the component of the magnetic moments.

The angle between a vector and a plane. Any plane can be defined by its normal. Let $\boldsymbol{n}$ be the normal vector to the plane and $\boldsymbol{a}$ is a given vector. We want to find the angle between vector $\boldsymbol{a}$ and the plane given by the normal $\boldsymbol{n}$. Angle between $\boldsymbol{n}$ and the plane is sum of the angle between $\boldsymbol{n}$ and $\boldsymbol{a}$ and angle between $\boldsymbol{a}$ and the plane. By definition of the normal, this sum must be $\pi/2$. The cosine of the angle between $\boldsymbol{n}$ and $\boldsymbol{a}$ is

$$\cos(\widehat{\boldsymbol{an}}) = \cos\left(\frac{\pi}{2} - \textsc{PlaneAngle}\right) = \sin(\textsc{PlaneAngle}) = \frac{\boldsymbol{a}\cdot\boldsymbol{n}}{\|\boldsymbol{a}\|\|\boldsymbol{n}\|} \tag{1.8.24}$$

$$\textsc{PlaneAngle} = \arcsin\left(\frac{\boldsymbol{a}\cdot\boldsymbol{n}}{\|\boldsymbol{a}\|\|\boldsymbol{n}\|}\right) \tag{1.8.25}$$

Using this idea, we can find the projection of the magnetic moments onto the cycloidal plane and the angle made by the magnetic moments with the cycloidal plane. Implementation of this computation using AWK program, as follows

```
$ cat angle.awk
```

```
#!/usr/bin/awk -f
# Program computes:
{
pi = 3.14159265

# components of the moments along the x, y, z:
ax = $1
ay = $2
az = $3

# the magnitude of the moments (4 Bohr's):
absA = sqrt(ax*ax + ay*ay + az*az)

# polarisation i.e. [111] direction:
px = 1/sqrt(3)
py = px
pz = px

# cycloidal propagation direction:
kx = 0;
ky = -1/sqrt(2)
kz  =1/sqrt(2)

# normal to the cyclodal plane:
nx = -2/sqrt(6)
ny = 1/sqrt(6)
nz = 1/sqrt(6)
```

```
# components of the moments along 3-directions
mP = (ax*px + ay*py + az*pz)
mK = (ax*kx + ay*ky + az*kz)
mN = (ax*nx + ay*ny + az*nz)

# the angle between moments and the cycloidal plane:
sinus = mN / absA
cosinus = sqrt(1 - sinus*sinus)
planeAngle = (180/pi)*atan2(sinus,cosinus)

# the angle between [111] - direction and moments:
cosinusPM = mP / absA
sinusPM=sqrt(1-cosinusPM*cosinusPM)
pmAngle=(180/pi)*atan2(sinusPM,cosinusPM)

print NR, planeAngle, pmAngle, mK, mP, mN
}
```

This program assumes that the input file has a set of magnetic moments along the cycloidal propagation direction and that propagation direction is $[0\bar{1}1]$.

Voigt notation. In physics, we often encounter 3×3 symmetric tensors (matrices). Such tensors are sometimes written with a single index. The common way of achieving this is by using Voigt

notation.[13] In Voigt notation, the matrix

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{xx} & \mathcal{A}_{xy} & \mathcal{A}_{xz} \\ \mathcal{A}_{yx} & \mathcal{A}_{yy} & \mathcal{A}_{yz} \\ \mathcal{A}_{zx} & \mathcal{A}_{zy} & \mathcal{A}_{zz} \end{pmatrix} \tag{1.8.26}$$

can be written as a vector by adopting the following convention

$$\begin{pmatrix} \mathcal{A}_{xx} \\ \mathcal{A}_{yy} \\ \mathcal{A}_{zz} \\ \mathcal{A}_{yz}/2 \\ \mathcal{A}_{xz}/2 \\ \mathcal{A}_{xy}/2 \end{pmatrix} \equiv \begin{pmatrix} \mathcal{A}_1 \\ \mathcal{A}_2 \\ \mathcal{A}_3 \\ \mathcal{A}_4 \\ \mathcal{A}_5 \\ \mathcal{A}_6 \end{pmatrix} \tag{1.8.27}$$

Note that since for symmetric matrices $\mathcal{A}_{ij} = \mathcal{A}_{ji}$, we do not need to write all the $9 = 3 \times 3$ elements. By using Voigt notation, we also can write $(3 \times 3 \times 3 \times 3)$ tensor that has 81 elements as a (6×6) tensor, i.e. as a matrix.

2 Fourier analysis

There are many ways of approaching the subject of Fourier analysis[14] and every one of those approaches can be justified as being the right way of introducing the ideas of Fourier analysis for specific purposes. One mathematically attractive way is to start by Fourier transformation and represent Fourier series as a limiting case of general idea. We will, however, start by describing Fourier series and expand this idea to a Fourier transform. This is the historical path by which the subject was first developed. Then, we will introduce discrete Fourier transform (DFT) because this

[13] This convention is named after the physicist Woldemar Voigt.

[14] Fourier analysis is named after Jean-Baptiste Joseph Fourier (1768–1830) who introduced ideas of Fourier series to solve the heat equation.

is a way how computers perform Fourier analysis. It is possible to combine all of those ideas into one unified theory through abstract mathematical constructions but from an application point of view there is no gain in such theoretical elaborations, and we will not go in that direction.

Fourier series can be identified with the study of periodic[15] functions by decomposing them into complex exponential functions. Any function which is defined in a finite region can be artificially repeated infinitely, thus producing periodic function. From this point of view, periodic functions are equivalent to a function that is defined in a finite interval. Fourier transform is a generalization of the idea of Fourier series for non periodic functions or functions defined in an infinite region. Discrete Fourier transform is linearly transforming finite dimensional vectors to other vectors with the same dimension. Actually, in general, the Fourier analysis can be considered as a linear transformations in a vector space of different nature. Sometimes, this space is identified with the space of integrable functions defined in a finite interval (case of Fourier series), sometimes with the space of integrable functions defined in infinite region (case of Fourier transforms), and other times underlining space is a space of vectors of finite dimensions (this corresponds to the case of Discrete Fourier transforms).

2.1 Fourier series

Let us consider function $f(x)$ with period 2π, i.e. the function with property $f(x+2\pi) = f(x)$[16]. Coefficients defined as

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx, \quad n \geq 0 \tag{2.1.1}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx, \quad n \geq 1 \tag{2.1.2}$$

are called Fourier coefficients for the function $f(x)$.

[15] Symmetry is the word used in physics for periodicity. For example, when problem involves a periodic function depending on position, the related physical problem is said to have a translational symmetry. One can also say that periodicity of functions in physical problems is a consequence of symmetry in a corresponding physical problem.

[16] If a function $f(x)$ periodic with period $2l$ instead of 2π, our results can be extended by change of variables according to the formula $x' = \frac{lx}{\pi}$.

Again, the language of linear algebra is very powerful in expressing this concept. Functions are example of vectors, and, from this point of view, Fourier coefficients are nothing but dot product between function (vector) $f(x)$ and trigonometric functions (vectors). With this definition of coefficients a_n and b_n, the sum

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + b_n \sin(nx) \tag{2.1.3}$$

is called Fourier series for the function $f(x)$.

The Fourier series is a change of basis: the trigonometric functions are new basis vectors and Fourier coefficients are coordinates of the function in that new basis[17]. Our insistence of expressing things in terms of vectors and operations over them can seem superfluous at this point. However, such language makes it extremely clear how to extend the ideas and concepts from Fourier series to so many different areas of science and mathematics. In particular, Quantum mechanics has benefited greatly from such way of expressing things.

The way how we expressed Fourier series is convenient for performing actual analytic computations. However, for theoretical investigation, so called complex exponential form of Fourier series is used more often. To get exponential form of Fourier series, we will use one of the deepest

[17]Trigonometric functions sin and cos are not just vectors, they are orthogonal vectors. Orthogonality can be proved by checking the following integral relations

$$\int_{-\pi}^{\pi} \cos(mx)\cos(nx)\,dx = \pi\delta_{mn}, \quad \text{for} \quad m,n \geq 1,$$

$$\int_{-\pi}^{\pi} \sin(mx)\sin(nx)\,dx = \pi\delta_{mn}, \quad \text{for} \quad m,n \geq 1$$

and

$$\int_{-\pi}^{\pi} \cos(mx)\sin(nx)\,dx = 0;$$

where δ_{ij} is the Kronecker delta

$$\delta_{ij} = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}$$

relationship in mathematics – Euler's formula[18]

$$e^{inx} = \cos(nx) + i\sin(nx) \tag{2.1.4}$$

If we define a set of complex numbers as

$$c_n = \begin{cases} \frac{1}{2}(a_n - ib_n) & n > 0 \\ \frac{1}{2}a_0 & n = 0 \\ \frac{1}{2}(a_{-n} + ib_{-n}) & n < 0. \end{cases} \tag{2.1.5}$$

The coefficients a_n, b_n can be expressed in terms of these new coefficients c_n using relations

$$a_n = c_n + c_{-n} \tag{2.1.6}$$

$$b_n = i(c_n - c_{-n}) \tag{2.1.7}$$

Now using Euler's formula we arrive at the following form for Fourier series

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} \tag{2.1.8}$$

where

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx \tag{2.1.9}$$

This is the complex exponential form of Fourier series.

Fourier transform. Fourier transform is used in many areas of science and engineering and different areas adopted different conventions in defining the Fourier transform. Although all of those definitions are fundamentally the same, the final equations look slightly different depending which one of the definitions is adopted. In mathematics, one often encounters the following definition for

[18]It is better to consider "Euler's formula" not as a formula but as a definition.

this concept

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i x\xi}\, dx \tag{2.1.10}$$

Corresponding inverse transform then has the form

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{2\pi i \xi x}\, d\xi, \tag{2.1.11}$$

In quantum mechanics, we use the convention in which the Fourier transform of the function is given by

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{ix\xi}\, dx \tag{2.1.12}$$

and corresponding inverse transform has the formula

$$f(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(x)\, e^{-ix\xi}\, dx \tag{2.1.13}$$

Mathematical convention is guided by symmetry and simplicity of the formulas. In physics, along with symmetry and simplicity, we also have to take into account the meaning of the symbols and operations. Then, the slight lost in simplicity is compensated by the gain in intuition and physical feeling obtained in this way. For example, Fourier transform in the way how it is defined in quantum mechanics, is also the transformation of the same quantum state from coordinate representation to momentum representation and corresponding inverse Fourier transformation brings us back to coordinate representation again.

As one can see from the similarity between the definitions of the Fourier transform and exponential form of Fourier series in introducing the transform, the intention is the extension of the ideas of Fourier series to an arbitrary function. The question about in which cases such extension is possible is complicated with problems like convergence. These investigations were the source of many mathematical inventions and were important in shaping modern mathematics. However, we have a different goal in mind and we will not discuss them in this work. Instead, we will describe discrete Fourier transforms which is the way of performing analogous transformations on finite

dimensional vector spaces.

2.2 Discrete Fourier transform (DFT)

Computers can not handle continuous functions or any vectors with infinite number of elements. For that reason, any function is approximated with a finite dimensional vector to perform numerical computations on them. More importantly, in experimental science like physics, we get our data not in the form of continuous functions but as a discrete set of data points. We must know how to handle such a set of data efficiently and perform different operation on them fast. Discrete Fourier transform is one of those operations that we often need to perform on data and, in the following pages, we would like to outline this operation.

Let us consider function $f(x)$ in the region $[a,b]$. We can divide the region of interest into n sub-intervals: $[a = x_0, x_1], [x_1, x_2], \cdots, [x_{n-1}, x_n = b]$. Usually, intervals between points are equally spaced and the step size is denoted by h. When it comes to experimental data, however, we might not have a control over h. In that case, we can interpolate the initial data and make it equally spaced with a given h. In both cases, the vector with components $f_i = f(x_i)$ can be taken as an approximate representation of the"true" function $f(x)$.

Discrete Fourier transform is the transformation of the vector $\boldsymbol{f}$ with N complex or real elements $\{f_n\}$ to another vector with the same number of elements according to the formula

$$\hat{f}_k = \sum_{n=0}^{N-1} f_n \cdot e^{-i2\pi kn/N} \tag{2.2.1}$$

Initial vector $\boldsymbol{f}$ then can be restored by performing the inverse operation which is given as follows

$$f_n = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}_k \cdot e^{i2\pi kn/N} \tag{2.2.2}$$

The discrete Fourier transform is a linear operation on a vector space. This means that we can

express this transformation as a matrix multiplication. If we denote the N^{th} root of unity as

$$\omega_N = e^{-2\pi i/N} \tag{2.2.3}$$

we can write the matrix to perform Fourier transform in the following form

$$\mathcal{F} = \begin{pmatrix} \omega_N^{0\cdot 0} & \omega_N^{0\cdot 1} & \dots & \omega_N^{0\cdot(N-1)} \\ \omega_N^{1\cdot 0} & \omega_N^{1\cdot 1} & \dots & \omega_N^{1\cdot(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_N^{(N-1)\cdot 0} & \omega_N^{(N-1)\cdot 1} & \dots & \omega_N^{(N-1)\cdot(N-1)} \end{pmatrix} \tag{2.2.4}$$

In the language of linear algebra, Fourier transform can be expressed as a system of linear equations using matrix $\mathcal{F}$

$$\hat{\boldsymbol{f}} = \mathcal{F}\boldsymbol{f} \tag{2.2.5}$$

and the corresponding inverse transformation can be performed using the following expression

$$\boldsymbol{f} = \frac{\mathcal{F}^{-1}}{N}\hat{\boldsymbol{f}} \tag{2.2.6}$$

where $\mathcal{F}^{-1}$ represents inverse of the matrix $\mathcal{F}$.

The expressions (2.2.5) and (2.2.6) do not look symmetric. This can be fixed by changing the definition (2.2.1) and (2.2.2) slightly so that inverse transform is achieved just by transforming corresponding matrix $\mathcal{F}$. However, this is not necessary for our purposes.

Bravais lattice and periodic functions. Fourier analysis is essentially the change of the basis. We change the basis because some problems look simpler and their solutions are very intuitive in a right chosen basis.

In Fourier analysis, the basis are complex exponentials or trigonometric functions and in problems that has some periodicity, it is a good idea to use trigonometric functions as basis. Crystalline

matter provides us with a very rich set of periodic functions and Fourier analysis is widely used in solid state physics in describing such functions. The charge density $\rho(x_1,x_2,x_3)$ inside the matter is one example of such functions:

$$\rho(x_1,x_2,x_3) = \rho(\boldsymbol{r}) = \rho\left(x_1\frac{\boldsymbol{a}_1}{a_1} + x_2\frac{\boldsymbol{a}_2}{a_2} + x_3\frac{\boldsymbol{a}_3}{a_3}\right) \tag{2.2.7}$$

In this expression, $\boldsymbol{r}$ is a radius vector

$$\boldsymbol{r} = x_1\frac{\boldsymbol{a}_1}{a_1} + x_2\frac{\boldsymbol{a}_2}{a_2} + x_3\frac{\boldsymbol{a}_3}{a_3}, \tag{2.2.8}$$

of an arbitrary point inside the crystal and $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ are periods of the crystal also known as the primitive vectors. Primitive vectors are the periods of the lattice in the sense that, from any two positions separated by those vectors, the crystal looks exactly the same.

For our purpose, an important function will be magnetic moments $\boldsymbol{m}$ inside the matter. Magnetic moments are distributed continuously, but in many materials, it is a good approximation to model them as a discrete set of vectors at lattice points

$$\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3 \tag{2.2.9}$$

In this work, we will need to perform discrete Fourier transform of magnetic moments. For that purpose, we will need to address a way of generalizing our results for one dimension to a three dimensional space. This is done using the following expression

$$\hat{f}_{k_1,k_2,k_3} = \sum_{n_1=0}^{N_1-1}\left(\omega_{N_1}^{k_1n_1}\sum_{n_2=0}^{N_2-1}\left(\omega_{N_2}^{k_2n_2}\sum_{n_3=0}^{N_3-1}\omega_{N_3}^{k_3n_3}\cdot f_{n_1,n_2,\,n_3}\right)\right) \tag{2.2.10}$$

In this expression, we assumed that the number of lattice point N_i along the different directions are

not the same. If $N_1 = N_2 = N_3 = N$, then same expression can be also written as

$$\hat{f}_{k_1,k_2,k_3} = \sum_{n_1,n_2,n_3=0}^{N-1} \omega_N^{k_1 n_1 + k_2 n_2 + k_3 n_3} \cdot f_{n_1,n_2,\, n_3} \tag{2.2.11}$$

In general, the discrete Fourier transform in d dimensions is performed using the following formula

$$\hat{f}_{k_1,k_2,\ldots,k_d} = \sum_{n_1=0}^{N_1-1} \left(\omega_{N_1}^{k_1 n_1} \sum_{n_2=0}^{N_2-1} \left(\omega_{N_2}^{k_2 n_2} \cdots \sum_{n_d=0}^{N_d-1} \omega_{N_d}^{k_d n_d} \cdot f_{n_1,n_2,\ldots,n_d} \right) \right) \tag{2.2.12}$$

3 Statistics

In physics of matter, we are dealing with very large number of atoms and molecules. To describe the whole system, by giving information about positions and velocities of its particles, is just not realistic and, even if it were realistic, such detailed information would be overwhelming and quite useless. For that reason, we deal with statistically averaged quantities. In physics of matter, we do not know what every particle is doing at every moment of time. However, we can tell about average quantities using theorems of statistics. This is the idea behind statistical physics which we will introduce in the following pages. We also will overview the results from mathematical statistics and theory of probability that will be useful to us in many different ways.

3.1 Concepts of Probability Theory

In order to introduce the subject of statistics in a proper way, one has to start with a set theory and then introduce ideas and concepts of theory of probability. Since our aim is practical, we will not follow such a path and we will limit ourselves to basic concepts and important definitions. In particular, we will not prove any of theorems of probability theory.

The set of all possible outcomes of an experiment or an observation is called the sample space. Particular subset of sample space is called an event. For each event from the sample space, we associate a number between zero and one. Such a number is called the probability of that particular

event. The three basic properties of probability function can be stated in the following equations

$$P(\text{EMPTY SET}) = 0 \tag{3.1.1}$$

$$P(\text{WHOLE SAMPLE SPACE}) = 1 \tag{3.1.2}$$

For the pairwise disjoint subsets of the sample space $A_1, A_2, \cdots, A_n$, the following identity holds:

$$P(\text{THE UNION OF ALL } A_i) = \sum_i^n P(A_i) \tag{3.1.3}$$

The equations (3.1.1), (3.1.2), and (3.1.3) are called axioms of probability or Kolmogorov's axioms and they can be considered as a mathematical basis on which whole theory of probability is constructed. Deduction of the results of theory of probability starting from these axioms is the way how this subject is covered in the most of the standard courses.

Random variable. A random variable is a function from sample space into the set of real numbers. Thus to define a random variable, one needs to define two things: the region from the set of real numbers, and the function that assigns probability for every number from that set. The possible values of a random variable and their associated probabilities is described by a function called a probability distribution.

Here, we should make distinction between countable and uncountable sample spaces. The reason for such distinction is that the probability distribution is specified in different ways, depending which kind of sample space we are talking about. In practice, we are dealing only with the space of countable nature. Furthermore, in computation we only consider spaces that are finite. Therefore one might argue that, in applications of statistics to physics, there is no need for considering anything but finite sample spaces. However, some formulas for uncountable sample spaces are much simpler to derive, since in these spaces, we can use integration and differentiation operations. Thus, sometimes we approximate our finite sample space with the uncountable sets in order to use some of the results from differential and integral calculus. To make our discussion concrete,

we will continue with examples from statistical physics.

3.2 Phase space

Basic concepts and theorems of statistical physics can be presented from the point of view of classical physics or quantum physics. However, the theory arrived by both ways are fundamentally the same. Therefore, in the subsequent pages, we will state our arguments in the language of classical physics exclusively.

Macroscopic mechanical system consisting of n degrees of freedom provides us with a good example of random variable. To uniquely define the configuration of such system, one must specify at least n quantities. For example, the state of system consisting of N particles at every moment of time is characterized by specifying its $3N$ coordinates (3 coordinates for every particle). Experience shows us that, to determine subsequent configuration of this mechanical system, one also has to know derivatives of its coordinates with respect to time (i.e., velocities). Set of coordinates and velocities of all particles at given time is called phase point. All possible phase points of the system define its phase space.

Let us divide the phase space of the system into pieces, each of which has a volume $\Delta q_i \Delta \dot{q}_i$. If we wait for a time T and, during the ΔT portion of that total time T, the system under consideration spend in the volume $\Delta q_i \Delta \dot{q}_i$ then

$$P_i = \frac{\Delta T}{T} \tag{3.2.1}$$

will give us probability distribution for the system to be found in that region of phase space.

However, if we consider the phase space as a continuous set of points, i.e. when $\Delta q_i \Delta \dot{q}_i \to 0$, then we will need to characterize the system not with the probability function P_i but by the continuous function $\rho = \rho(q_i \dot{q}_i)$ called the probability density. Thus, we have to define what we mean by the probability density.

Probability density. The probability of the continuous random variable x to take value from the interval $[a,b]$ is given by the expression:

$$P[a \leq x \leq b] = \int_a^b \rho(x)\,dx \tag{3.2.2}$$

The expression (3.2.2) can serve us as a definition of the probability density $\rho(x)$ of the random variable x.

The restatement of the second axiom (3.1.2) of probability in terms of density function ρ is called normalization condition. The normalization condition for the phase space can be given by the following integral relation:

$$\int_\Omega \rho dq d\dot{q} = 1 \tag{3.2.3}$$

where the region of integration Ω is the entire phase space.

Probably, the most famous probability distribution is a normal or Gaussian distribution which has the following functional form

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3.2.4}$$

For our purposes, however, the most important distribution will be the following

$$\rho(E) = \frac{\exp(-\beta E)}{\int_E \exp(-\beta E)} \tag{3.2.5}$$

In this expression, the quantity β is defined as $\beta = \frac{1}{k_B T}$, where T represents temperature. The constant k_B is known as a Boltzmann constant and its value in SI units is $k_B = 1.3806503 \times 10^{-23}\ \mathsf{m}^2 \cdot \mathsf{kg} \cdot \mathsf{s}^{-2}$.

The distribution given by probability density (3.2.5) characterizes the probability that the system with constant volume V, temperature T, and number of particles N will occupy the state with energy E. There is no strict "proof" of this relation in a mathematical sense and it must be considered as an experimental fact. However, there is many physical reasoning that leads us to this

relation. We will not give them here but we will describe how to use this result in the Monte-Carlo method to find the state of different physical systems at different temperatures.

For the system with discrete states, the corresponding relation for the probability is given by

$$P(E = E_i) = \frac{\exp(-\beta E_i)}{\sum_i \exp(-\beta E_i)} \tag{3.2.6}$$

and this is the form of the distribution that is used to perform computations.

3.3 Monte-Carlo (MC) approach

In the following sections, we will start introducing computational methods in physics. This will mainly include molecular dynamics and Monte-Carlo methods. Those two methods allow us to conduct simulation at a finite temperature. Molecular dynamics is not heavily used in the current work but one has to refer to them to understand methods that are used. Then, we will model properties of the models that we are investigating. In the section devoted to the summary of the results, we will apply all of these methods to study properties of multiferroic materials, in particular $BiFeO_3$.

3.3.1 Molecular dynamics *vs.* Monte Carlo approach

In this work, we will not be conducting molecular dynamic simulations. However, to understand the ideas behind some of the methods that we will be using, it is instructive to compare and contrast them with molecular dynamic approaches. Therefore, some discussion of molecular dynamics techniques will be necessary.

Any physical quantity f that characterizes the system, in general, will be defined as a function of coordinates q_i and momenta p_i of the particles that make up that system (i.e., as a function of the phase points). Coordinates and momenta of the particles evolve with time t. Consequently, the value of the quantity $f = f(q_i, p_i)$ will also depend on time. The first question that needs to be addressed therefore is the law by which the quantities q_i and p_i evolve with time, i.e. to find

functions: $q_i = q_i(t)$, and $p_i = p_i(t)$. In classical mechanics this laws expressed as a system of differential equations. It is a rule rather than an exception that the differential equations governing the motion of particles can not be solved analytically for any realistic system that we encounter in physics. However, we can solve them approximately using computers. Then, in molecular dynamics, an average property of physical quantities are calculated by integrating them over time. The average value of a quantity $f = f(t)$, for example on a time interval $[t_1, t_2]$, can be calculated by

$$\overline{f} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} f(t)\, dt \tag{3.3.1}$$

This way of computing the averages is a characteristic aspect of molecular dynamics and it is called time averaging. In statistical methods like the Monte Carlo method, a different procedure is used for averaging and it is called ensemble average.

Time scale problem. To calculate averages like (3.3.1) in practice, one must trace the entire path of the system in a phase space in details. Typical oscillation of atoms at room temperature is about 10-100 femto seconds (10-100 10^{-15}s), and if we want to simulate phenomena that take fraction of seconds there is a huge time scale to bridge. Thus if the simulation time is not long enough, some of the properties are not captured by molecular dynamics. This problem is called the time scale problem.

The time scale problem is of course a computational problem and in principle can be solved with faster processors. From a physical point of view, however, there is a bigger problem. Sometimes, the dynamics of the system might not be known or it might be very complicated. This is particularly true in the case of magnetic systems. We might not know the detailed dynamics of magnetic spins for a particular system. However, general theorems of statistical mechanics are still applicable and they can provide us with the knowledge about the average quantities. Why one should follow the system through all of its steps to calculate an average quantity which is insensitive to a detailed dynamics of the system? If the average value is the only thing we need, can it be obtained by more direct methods using statistical properties of the system? Monte-Carlo method

provides us with such approach.

3.4 Monte Carlo

Monte Carlo (MC) methods are a family of computational experiments that are based on generation of random numbers to compute certain quantities. They are used in physics to simulate systems with many degrees of freedom such as solids and fluids. In mathematics, to compute multidimensional integrals for regions with irregular boundaries, MC methods provide fast and accurate results. Use of MC methods are not limited to science and mathematics. In business, for example, they are used to calculate risk in uncertain conditions in market place. Such a wide application of the MC methods makes it impossible to give a simple definition which does not end up to be too wide or too narrow. For our purpose, MC method will be equivalent to its particular implementation – a Metropolis algorithm – which we will describe in the next section.

3.4.1 Metropolis algorithm

Let us consider some random variable x with probability distribution $P(x)$. We want to generate sets of random numbers $\{x_0, x_1, x_2, \ldots\}$ which are distributed with probability $P(x)$. Metropolis algorithm provides us with one way of achieving this. Of course, we have a particular probability distribution in our mind, i.e. the distribution given by the expression (3.2.5) and we want to describe the algorithm in relation to that distribution. However, we would like to note that Metropolis algorithm works independently on particular forms of probability distributions.

We start with some initial state. One can choose this initial state in a random fashion or the initial state can be the result of some previous computation. We then generate new state by making a perturbation of the initial state. Let us denote corresponding energy change as ΔE. If the energy decreased as a result of such perturbation and consequently $\Delta E < 0$, the moves is always accepted. On the other hand if $\Delta E > 0$, we accept the move with probability $\exp(-\beta \Delta E)$.

In computers, it is usually easy to generate random variables between 0 and 1. If we call function for generating such sequence of numbers as $\mathsf{rand}(0,1)$ we can write the acceptance criteria

as

$$\text{if} \quad \text{rand}(0,1) < \exp(-\beta \Delta E) \quad \text{then accept the move}$$

If we generate a sequence of states with the procedure outlined above, the energy of the state will be distributed according to the distribution (3.2.5). If we have n of such states, we then can calculate the average value of the property a according to a simple formula:

$$(\text{Average of the property } a) = \frac{1}{n}\sum_{i=1}^{n} a_i = \frac{a_1 + a_2 + \cdots + a_n}{n}. \tag{3.4.1}$$

This property a can represent the energy, magnetic moment, or electric polarization of the system.

In order to implement this algorithm in practice, there are many additional questions that should be addressed: what is a nature of the states and what kind of moves are allowed? What is the energy corresponding to different moves? Those questions should be answered in basis of the physics of the problem under consideration. Depending on what is included in the model and what we ignored, some of the phenomena will be reproduced or overlooked by that model.

Expressions for the energy and generalization of the scalar product. From the point of view of modeling, the system is completely defined by giving the expression for the energy that describes it. The expression for a particular energy will be given in appropriate place in the results section of this work. In this section, however, I would like to talk a little bit about the general forms of the energies that are used in this simulation. All of them have very similar form and they can be considered as a generalized scalar product between order parameters.

Let us consider the dipole-dipole interaction between magnetic moments for example. This interaction is usually modeled using the following energy expression:

$$\mathcal{W}_{mag-dipol-dipol} = -\sum_{i,j} \mathcal{J}_{ij} \left(\boldsymbol{m}_i \cdot \boldsymbol{m}_j \right) \tag{3.4.2}$$

and all the other energy terms have very similar form to (3.4.2).

As one can see, (3.4.2) is analogous in form to the scalar product, but instead of expression like $\boldsymbol{x}^T\boldsymbol{y}$ we have expression of the form $\boldsymbol{x}^T J\boldsymbol{y}$ where J is a matrix of coefficients. Actually, in general, we could define the scalar product to be expression of the form

$$\boldsymbol{x}\cdot\boldsymbol{y} \equiv \boldsymbol{x}^T J\boldsymbol{y}$$

with some coefficient matrix (tensor) J, provided the following conditions are satisfied.

The first condition can be called the symmetry condition and it requires the following relation to hold

$$\boldsymbol{x}\cdot\boldsymbol{y} = \overline{\boldsymbol{y}\cdot\boldsymbol{x}}$$

this condition will be satisfied provided that the matrix J is symmetric i.e. the relations $J^T = J$ is satisfied for it. In physics, energy coefficients are symmetric and this condition is easily satisfied.

The next two properties are elementary consequences of the definition of matrices and tensors:

$$\begin{aligned}(a\boldsymbol{x})\cdot\boldsymbol{y} &= a(\boldsymbol{x}\cdot\boldsymbol{y})\\ (\boldsymbol{x}+\boldsymbol{y})\cdot\boldsymbol{z} &= (\boldsymbol{x}\cdot\boldsymbol{z})+(\boldsymbol{y}\cdot\boldsymbol{z})\end{aligned} \tag{3.4.3}$$

The last property that is required from the dot product is the following:

$$\begin{aligned}&\boldsymbol{x}\cdot\boldsymbol{x} \geq 0\\ \text{and}\quad &\boldsymbol{x}\cdot\boldsymbol{x} = 0 \quad \text{if and only if} \quad \boldsymbol{x} = \boldsymbol{0}.\end{aligned} \tag{3.4.4}$$

This property is satisfied if $\boldsymbol{x}^T J\boldsymbol{x} \geq 0$ for any vector $\boldsymbol{x}$, and $\boldsymbol{x}^T J\boldsymbol{x} = 0$ if and only if $\boldsymbol{x} = \boldsymbol{0}$. Matrix J with such property is called positive definite matrix and matrices that describe the energy in physics are indeed positive definite. The whole point of this short discussion is that the energy coefficient have to satisfy some symmetry conditions and they can not be arbitrary set of numbers. We will not discuss all of those conditions in this work.

The only question that we need to ask ourselves is how to find out the values of the relevant

coefficients that describe different energy terms. In principle, those coefficient should be obtainable if we are familiar with all the particles that matter is made of and which the fundamental laws that govern motion of those particles. Discussion of these topics is the subject of the next section.

4 Quantum mechanics

Standard view of matter. Matter can be studied as consisting of a large number of interacting particles. By their electric properties, particles of matter can be divided into three kinds. Conventionally, they are called positively charged particles, negatively charged particles and particles of no charge. Particles of no charge are also called neutral particles. An example of positively charged particle is a proton, i.e. atomic hydrogen without its electron. Electron is a particle of negative charge and neutron has no charge. The charges of electron and proton are equal to each other in magnitude but opposite in signs. The charge of proton is called the elementary charge, and we will denote it as e. In SI units the value of elementary charge is:[19]

$$e = 1.602176565(35) \times 10^{-19} \text{ Coulombs} \tag{4.0.5}$$

Three main particles. These three particles: neutron, proton, and electron are responsible for almost all properties of matter. Therefore, complete understanding of laws that govern motions of these particles, in principle, would give us all the information that we want to know about the matter. However, there are many difficulties with such direct approaches for understanding properties of matter. First of all, when it comes to protons and neutrons, our understanding of these particles is still incomplete. Second, even in the domains where we know all the laws, any realistic problem involves very complicated situations and one has to rely on approximations. In this section we will start to overview those approximations.

[19] This was first measured by Robert A. Millikan in a famous oil drop experiment which he conducted in 1909.

Nucleus. Neutrons and protons form intimate mixture which is called nucleus. We do not need to know details involved in the structure of nucleus except basic facts that will allow to us make some approximations. First of all, nucleus is very heavy in comparison with the mass of electrons. The total charge of nucleus is equal to the number of protons that it contains times elementary charge e. Any nuclei with the same number of protons will have the same chemical properties. Nucleus can be considered as a classical point particles that obey laws of classical mechanics. Nuclei move very slowly and, at every configuration of nuclei, electrons always have enough time to reach their ground state. This is the idea behind the Born-Oppenheimer approximation.[20] This is all we will need to know about the nucleus and all the subsequent discussions will involve electrons. Electrons move around nuclei as a quantum particles and to understand their properties one has to use ideas of quantum mechanics.

Quantum mechanics. Historically, quantum mechanics was developed as a response to unexpected experimental results. The final theoretical framework was simple and elegant. However, in "real world", quantum mechanics is almost always used in the situations that involve many particles in complicated interactions which makes it a very difficult subject. Thus applied quantum mechanics is reduced to a complex system of approximations in varying degrees of validity and complexity. We will outline those approximations in the consequent sections since they are basis of some of our methods. We will start by outlining fundamental equation of quantum mechanics which is called Schrodinger's equation.

Linear algebra and quantum mechanics. The language of linear algebra is particularly convenient in relation to quantum mechanics. The main equation of quantum mechanics is a linear equation and, in compact form, it is often written as

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi \tag{4.0.6}$$

[20]This approximation is named after Max Born and J. Robert Oppenheimer who proposed it in 1927. This approximation is also known as adiabatic approximation.

In the equation (4.0.6), Ψ is called the probability amplitude, and it is a vector with complex numbers as its elements. The probability amplitude is sometimes called wave function and is all we need to know about the system. All the properties of the system will be expressed as a function of that probability amplitude. Also as equation (4.0.6) shows, the future of the system is also completely determined in terms of wave function. Thus the state of quantum system is determined by giving its wave function.

4.1 Density functional theory

The state of a system of N quantum particles (i.e., electrons) is specified by a vector $\Psi(\vec{r}_1,\ldots,\vec{r}_N)$, and in stationary case it obeys the following Schrodinger's equation:

$$\hat{H}\Psi = \left[\hat{T}+\hat{V}+\hat{U}\right]\Psi = \left[\sum_i^N -\frac{\hbar^2}{2m_i}\nabla_i^2 + \sum_i^N V(\vec{r}_i) + \sum_{i<j}^N U(\vec{r}_i,\vec{r}_j)\right]\Psi = E\Psi \tag{4.1.1}$$

In this equation $V(\vec{r}_i)$ represents the external potential that electrons feel from the nuclei. The fact that $V(\vec{r}_i)$ does not depend on time reflects the Born-Oppenheimer approximation that we adopted. The electron-electron interaction energy is given by the energy term $U(\vec{r}_i,\vec{r}_j)$.

Variational method. Functional is a function that takes a vector (in particular a function) as an argument and returns a scalar as a value. For any wave function $|\Psi\rangle$ for example, and for given Hamiltonian $\hat{H}$ we can define the following functional:

$$\varepsilon[\Psi] = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{4.1.2}$$

If it happens that $|\Psi\rangle$ is a ground state then the functional will return a ground state energy. However, if $|\Psi\rangle$ is not a ground state then the value that it returns will be higher than the ground state energy. Thus by trying many forms of $|\Psi\rangle$ and choosing the one which gives the lowest value for the functional, we can approximate the ground state by that value. Usually, one chooses a parametric form of $|\Psi\rangle$ function and finds such parameters which will give the smallest possible value

for ε.

Hartree-Fock method. Now we will consider one of the earliest and very influential methods for electronic structure computation, mainly the Hartree-Fock method [28]. This method is based on very simple ideas. It is reformulation of Schrodinger equation in the language of the Variational method. We will end up using density functional theory in our actual computations but in practice ideas from Hartree-Fock method is used in some density function calculations as well.

Hartree-Fock method uses the following generic function to use in variational method

$$\Psi(\vec{r}_1,\vec{r}_2,\ldots,\vec{r}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1(\vec{r}_1) & \chi_2(\vec{r}_1) & \cdots & \chi_N(\vec{r}_1) \\ \chi_1(\vec{r}_2) & \chi_2(\vec{r}_2) & \cdots & \chi_N(\vec{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\vec{r}_N) & \chi_2(\vec{r}_N) & \cdots & \chi_N(\vec{r}_N) \end{vmatrix} \equiv \begin{vmatrix} \chi_1 & \chi_2 & \cdots & \chi_N \end{vmatrix}, \tag{4.1.3}$$

The wave function as we know is a function of all N electrons. On the other hand, the Hartree-Fock method works with functions χ_i of a single electron. So Hartree-Fock method replaces a single wave function with N wave functions.

Like any variational method, this method will approach the ground state energy from above and will not go below it. Since the whole wave function is written as determinant (known as a Slater determinant), any symmetry rules for fermions are also automatically satisfied.

Another important feature of the Hartree-Fock method is that it is based on very solid theory and is easily extensible to more accurate methods. Thus today we have many methods from the Hartree-Fock family. Of course, the computational cost of those improved versions is also very high and this puts some limit on their practicality.

Density functional theory [29]. Hartree-Fock method was developed to serve the need of quantum chemistry. On the other hand, density functional theory came fromthe solid state community.

The fundamental variable in this approach is the charge density

$$n(\vec{r}) = N \int \mathrm{d}^3 r_2 \int \mathrm{d}^3 r_3 \cdots \int \mathrm{d}^3 r_N \Psi^*(\vec{r}, \vec{r}_2, \ldots, \vec{r}_N)\Psi(\vec{r}, \vec{r}_2, \ldots, \vec{r}_N) \tag{4.1.4}$$

i.e. it is not the wave function. The wave function has all the information about the system. However, it is computationally complicated to deal with. At the end, we want to compute the charge density of the system. Therefore, the main idea behind the density functional theory is to find equation that will give us the charge density directly without going through wave function.

It turns out that when it comes to the ground state of the quantum system, for a given charge density there is only one possible external potential that can produce such charge density. That means that, as far as ground state is concerned, the charge density can be considered as fundamental as a wave function in the sense that it has all the information about the system. Thus, energy of the system in principle can be computed from the charge density. The only problem is that we do not know the equation that will give us the energy from the charge density. If we know such equation, then it would be equivalent to the Schrodinger equation.

There were many attempts to approximately compute the energy of the system from the charge density. The problem is not trivial and it is not likely that there is a simple expression that would achieve that goal, and density functional theory will remain as set of approximations that would give better and better results without being theoretically exact.

To show the nature of the problem, let us discuss how one might get the kinetic energy from the charge density. To find the kinetic energy of the system from the charge density is not trivial. The kinetic energy is the second derivative of the wave function. How one can recover the second derivative of the wave function from the charge density? One approach is called the Local Density approximation. In this approach, we divide the region of interest into small sub regions. In every subregion, the charge is assumed to be homogeneous and non-interacting. For such electron gas, we can compute the charge density. By summing all those kinetic energies, we calculate the kinetic

energy, of the whole system:

$$T_{TF}[n] = C_F \int n(\vec{r}) n^{2/3}(\vec{r}) d^3r = C_F \int n^{5/3}(\vec{r}) d^3r$$

where

$$C_F = \frac{3h^2}{10m_e} \left(\frac{3}{8\pi} \right)^{2/3}$$

This is also the idea behind the Thomas-Fermi approximation [23]. This works fine for metals and other systems that have very smooth charge distribution. On the other hand, in systems like atom and molecules where electrons are distributed very inhomogeneously, the local density approximation is not well suited.

Density functional theory is less expensive computationally than the Hartree-Fock methods and, when it comes to solids, it is also overall more accurate. In practice, the application of these methods comes down to using appropriate software that implements them. It is unrealistic to implement all of those methods from scratch by any person.

Part II

First-Principle-based effective Hamiltonian study

5 Monte Carlo method

5.1 Effective Hamiltonian

From the point of view of Monte Carlo simulations, a system is defined by giving the expression for its energy as a function of the degrees of freedom. The system that we are interested in can be modeled using the following degrees of freedom:

$$\begin{aligned} &\boldsymbol{u} \text{ – amplitude of local mode relative to cubic structure} \\ &\boldsymbol{\omega} \text{ – angle of rotation of oxygen octahedra} \\ &\boldsymbol{m} \text{ – magnetic moment} \\ &\eta_l \text{ – strain tensor} \end{aligned} \tag{5.1.1}$$

In addition to those degrees of freedom, we have two external fields: electric $\boldsymbol{E}$ and magnetic $\boldsymbol{H}$.

Local mode $\boldsymbol{u}$ characterizes the polar displacements, occurring in one unit cell from cubic ideal positions of the different ions, and it is proportional to the electric dipole moment $\boldsymbol{p}$ of that cell and that is induced due to these displacements. The proportionality constant between $\boldsymbol{u}$ and $\boldsymbol{p}$ has the units of charge and it is called Born effective charge Z^*. The value of the Born effective charge Z^* is obtained from first-principle calculations. In case of $BiFeO_3$, for example, its value in atomic units equals $Z^* = 5.868$. Thus the polarization $\boldsymbol{P}$ can be expressed in terms of Z^*, and $\boldsymbol{u}$ as follows

$$\boldsymbol{P} = \frac{Z^* \boldsymbol{u}}{V_0} \tag{5.1.2}$$

where V_0 is a volume of the unit cell.

The electric field is coupled only with local dipole moments $\boldsymbol{p}_i$ through the energy term:

$$E_e = -\sum_i \boldsymbol{E} \cdot \boldsymbol{p}_i \tag{5.1.3}$$

Whenever we do trial change of the local mode $\boldsymbol{u}_i$ and consequently the local electric dipole moments $\boldsymbol{p}_i$ in MC simulation, we have to take into account this term in the calculation of the energy change.

As in the case of the classic harmonic spring–mass system, there is an energy cost associated with the displacement of the unit cell i. Such energy can be called self energy since its calculation involves local mode at a given site alone. The self energy of isolated local mode at cell i with amplitude $\boldsymbol{u}_i$ (relative to cubic structure) is given as follows

$$E_{\boldsymbol{u}} = \sum_i k_2 u_i^2 + \alpha u_i^4 + \gamma(u_{ix}^2 u_{iy}^2 + u_{iy}^2 u_{iz}^2 + u_{iz}^2 u_{ix}^2) \tag{5.1.4}$$

The expression (5.1.4) include only even terms because of the symmetry considerations [46]. Also to reproduce ferroelectricity, one has to have higher order than harmonic terms. Parameters k_2, α, γ and all the other energy parameters that are introduced below are obtained from first-principle calculations.

One characteristic thing about many perovskites like $BiFeO_3$ is that they have very rigid oxygen octahedra. The movements of oxygen atoms can be summarized as a rotation of this octahedra rigidly along some axis – so called antiferrodistortive (AFD) motion. This rotation is linked to many important phenomena in perovskites and it has to be taken into account to reproduce those phenomena. We will characterize this rotation by a vector $\boldsymbol{\omega}$ and we will express it in radians. The direction of the vector $\boldsymbol{\omega}$ is chosen to be the direction of the axis about which the oxygen octahedra rotates.

As it is the case with local modes $\boldsymbol{u}$, there is a self energy associated with antiferrodistortive motion. For analogous (symmetry) reasons as in the case of local modes, we write the self energy

for the antiferrodistortive motion in the following form [47]:

$$E_{\boldsymbol{\omega}} = \sum_i k_A \omega_i^2 + \alpha_A \omega_i^4 + \gamma_A(\omega_{ix}^2\omega_{iy}^2 + \omega_{iy}^2\omega_{iz}^2 + \omega_{iz}^2\omega_{ix}^2) \tag{5.1.5}$$

Magnetic moments $\boldsymbol{m}_j$ in $BiFeO_3$ are assumed to be localized at lattice sites i with constant magnitude of $4\mu_B$ This is consistent with first-principle calculations (see reference [37]). Assumption about the localized magnetic moments are not true in general. However, in case of insulators like $BiFeO_3$, such approximation about localized moments is well justified. If we make this assumption, then the energy of the system due to external magnetic field $\boldsymbol{H}$ can be calculated according to the following expression

$$E_h = -\sum_i \boldsymbol{H} \cdot \boldsymbol{m}_i \tag{5.1.6}$$

The change in this energy has to be calculated every time we make trial change in $\boldsymbol{m}_i$.

We will write expression for the elastic energy by dividing it into homogeneous E_H and inhomogeneous E_I parts [46]. Homogeneous part of the elastic energy is given by the following expression

$$\begin{aligned} E_H =& B_{11}(\eta_{H,1}^2 + \eta_{H,2}^2 + \eta_{H,3}^2) \\ &+ B_{12}(\eta_{H,1}\eta_{H,2} + \eta_{H,2}\eta_{H,3} + \eta_{H,3}\eta_{H,1}) \\ &+ B_{44}(\eta_{H,4}^2 + \eta_{H,5}^2 + \eta_{H,6}^2) \end{aligned} \tag{5.1.7}$$

where the variable $\eta_{H,l}$ is called homogeneous strain and it constitutes part of the total strain. The other part of the strain is called inhomogeneous strain. If we denote displacement vectors at the corner of the unit cell i as $\boldsymbol{v}(i)$, the elastic energy due to in homogeneous strain can be expressed

as follows:

$$\begin{aligned} E_I = \sum_i \{ & \gamma_{11}[v_x(i) - v_x(i \pm 1_x)]^2 \\ & + \gamma_{12}[v_x(i) - v_x(i \pm 1_x)][v_y(i) - v_y(i \pm 1_y)] \\ & + \gamma_{44}[v_x(i) - v_x(i \pm 1_y) + v_y(i) - v_y(i \pm 1_x)]^2 \\ & + \text{cyclic permutations} \} \end{aligned} \tag{5.1.8}$$

In the expression for E_I, the 1_x refers to the neighboring cell along the x axis.

If we define the following quantities

$$\Delta v_{xx} = \sum_{\boldsymbol{d}=0,1_y,1_z,1_{y+z}} [v_x(i-d-1_x) - v_x(i-d)] \tag{5.1.9}$$

$$\Delta v_{xy} = \sum_{\boldsymbol{d}=0,1_y,1_z,1_{y+z}} [v_y(i-d-1_x) - v_y(i-d)] \tag{5.1.10}$$

then the diagonal part of the inhomogeneous strain is defined as

$$\eta_{I,1}(i) = \frac{\Delta v_{xx}}{4} \tag{5.1.11}$$

Non-diagonal components of the strain are defined as

$$\eta_{I,4}(i) = \frac{\Delta v_{yz} + \Delta v_{zy}}{4} \tag{5.1.12}$$

Whenever we talk about strain $\eta_l(i)$ we mean the sum of two components of the strain:

$$\eta_l(i) = \eta_{H,l}(i) + \eta_{I,l}(i) \tag{5.1.13}$$

Next, we will describe the energy terms that are introduced to model couplings between the different degrees of freedom and interactions between two different lattice sites. They will have similar analytic form and can be thought of as the smallest possible order of the Taylor expansion

of the total energy as a function of the degrees of freedom. Despite these mathematical similarities between different terms, they have different physical origins behind them. So introduction of particular term has to be guided by physics. For example, one has to make sure that the symmetry requirements are met. Also as we mentioned already, those parameters has to be obtained from first-principle calculations.

The interactions of the local mode $\boldsymbol{u}_i$ at site i with local modes $\boldsymbol{u}_j$ at other sites and strain η_l are modeled using the following energy

$$
\begin{aligned}
E_{\text{LocalMode}} = & \\
& + \quad \textstyle\sum_{i,j}\sum_{\alpha,\beta} H_{ij,\alpha\beta} u_{i,\alpha} u_{j,\beta} && (5.1.14)\\
& + \quad \textstyle\sum_{i,j}\sum_{\alpha,\beta} Z_{ij,\alpha\beta} u_{i,\alpha} u_{j,\beta} && (5.1.15)\\
& + \quad \textstyle\sum_{i}\sum_{\alpha,\beta} Y_{i,\alpha\beta} \eta_l(i) u_{i,\alpha} u_{i,\beta} && (5.1.16)
\end{aligned}
$$

where indices α, and β, - run over Cartesian components. The expression (5.1.14) describes the long-range interaction between local modes of sites i and j. The j index for it runs over all the other sites of a given site i. The j of the expression (5.1.15) models the short-range exchange interactions and index j for it runs over all the first, second, and third nearest neighbors of the site i. This later interaction in calculated by excluding dipole-dipole interactions and can be considered as a result of overlapping between wave functions of the neighboring sites and/or cutting bonds. Naturally this interaction will have short range in insulators like $BiFeO_3$. Interaction between local modes $\boldsymbol{u}$ and the strain η_l are expressed as smallest possible order of the Taylor expansion (allowed by symmetry) of the energy in terms of $\boldsymbol{u}$ and η_l. For symmetry reasons there are only few independent coefficient $Y_{i,\alpha\beta}$ of this interaction.

The interaction of the $\boldsymbol{\omega}$ with other degrees freedom (excluding magnetic ones) is modeled as

follows

$$E_{\mathrm{AFD}} =$$

$$+ \quad \sum_{i,j}\sum_{\alpha,\beta} R_{ij,\alpha\beta}\omega_{i,\alpha}\omega_{j,\beta} \tag{5.1.17}$$

$$+ \quad \sum_{i}\sum_{\alpha,\beta,\gamma,\delta} S_{i,\alpha\beta\gamma\delta}\omega_{i,\alpha}\omega_{i,\beta}u_{i,\gamma}u_{i,\delta} \tag{5.1.18}$$

$$+ \quad \sum_{i}\sum_{\alpha,\beta} X_{i,\alpha\beta}\eta_{l}(i)\omega_{i,\alpha}\omega_{i,\beta} \tag{5.1.19}$$

In this expression α, β, γ, and δ run over Cartesian components, index i runs over all the lattice sites, and j run over the six nearest neighbors of site i. The energy term (5.1.17) can be considered as energy gain or cost due to interaction between two rotating oxygen octahedra. Thus we need to include only interactions between nearest neighbors for this energy term. The term (5.1.18) describes interaction between local mede and antiferrodistortive motion at a given site i. In analogous to expression (5.1.16) we introduce the term (5.1.19). This last term expresses the interaction between strain η_l and antiferrodistrotive vector ω_i at site i.

Magnetic dipoles at side i are coupled with all the other degrees of freedom according to the following expression

$$E_{\mathrm{MAG}} =$$

$$+\sum_{i,j,\alpha,\gamma} A_{ij,\alpha\gamma}m_{i,\alpha}m_{j,\gamma} \tag{5.1.20}$$

$$+\sum_{i,j,\alpha,\gamma} D_{ij,\alpha\gamma}m_{i,\alpha}m_{j,\gamma} \tag{5.1.21}$$

$$+\sum_{i,j,\alpha,\gamma,\nu,\delta} E_{ij,\alpha\gamma\nu\delta}m_{i,\alpha}m_{j,\gamma}u_{i,\nu}u_{i,\delta} \tag{5.1.22}$$

$$+\sum_{i,j,\alpha,\gamma,\nu,\delta} F_{ij,\alpha\gamma\nu\delta}m_{i,\alpha}m_{j,\gamma}\omega_{i,\nu}\omega_{i,\delta} \tag{5.1.23}$$

$$+\sum_{i,j,l,\alpha,\gamma} G_{ij,l,\alpha\gamma}\eta_{l}(i)m_{i,\alpha}m_{j,\gamma} \tag{5.1.24}$$

$$+\sum_{i,j} K_{ij}(\boldsymbol{\omega}_i - \boldsymbol{\omega}_j)\cdot(\boldsymbol{m}_i \times \boldsymbol{m}_j) \tag{5.1.25}$$

$$-\sum_{i,j} C_{ij}(\boldsymbol{u}_i \times \boldsymbol{e}_{ij})\cdot(\boldsymbol{m}_i \times \boldsymbol{m}_j) \tag{5.1.26}$$

In this expression α, γ, ν, and δ run over Cartesian components and index i runs over all the lattice

sites. The j of the expression (5.1.20) runs over all the sites. It models long-range interaction between magnetic dipoles. The j of the expression (5.1.21) models the short-range exchange interactions. The j of the second, third, fourth, and fifth energies runs over the first, second, and third nearest neighbors of the site i. The j index of the expression (5.1.25) runs over the first nearest neighbors of the side i. The index j expression (5.1.26) runs over the second-nearest neighbors of the site i, and $\boldsymbol{e}_{ij}$ is a vector from site i to j with the length $\sqrt{2}$. The purpose of the energy terms (5.1.22), (5.1.23), and (5.1.24) are to model interactions between magnetic moments and local modes, antiferrodistortive motion, and strain respectively. The energy terms (5.1.25) and (5.1.26) have very important consequences for our study and we will describe the physics behind them in a later section.

5.2 Units used in the program

For numerical calculation, it is important to use units that makes all quantities of the same order. For example, it is convenient to express lengths and displacements in the units of the lattice constant of the cubic phase, and energy in Hartree. The temperature also has to be expressed in the units of the energy as follows

$$T \longrightarrow \frac{T(\text{Kelvin})}{3.15778 \times 10^5} \tag{5.2.1}$$

Electric and magnetic fields are expressed in atomic units using the following transformation

$$E \longrightarrow E(\text{Volt/m}) \cdot 1.9447 \times 10^{-12} \tag{5.2.2}$$

$$H \longrightarrow H(\text{Tesla}) \cdot 2.1271912 \times 10^{-6} \tag{5.2.3}$$

6 Magnetic and electric properties of matter

Matter consists of charged particles that interact with each other through the laws of electromagnetism. These laws are summarized in Maxwell's equations and Lorentz's force law. There are

many equivalent ways of presenting Maxwell's equations. We will present them in a form that is most commonly encountered in the literature and, unless otherwise specified, SI units will be assumed throughout.

The first law – Gauss's law – indicates that there is a source of electric field (also called electric charge) in nature so that the following relation holds

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon_0} \tag{6.0.4}$$

On the other hand, the magnetic field has no analogous source, i.e. there is no magnetic charge from which the magnetic field can be produced. This is expressed in the second equation

$$\nabla \cdot \boldsymbol{B} = 0 \tag{6.0.5}$$

The first law of electromagnetism that describes the relation between electric and magnetic fields is the Faraday's law or law of induction

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{6.0.6}$$

Ampere's law with Maxwell's generalization historically was the last law that was discovered as a result of set of excellent experiments and reasoning by Ampere and later correction made by Maxwell

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J} + \mu_0 \varepsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} \tag{6.0.7}$$

Together with Lorentz force law

$$\boldsymbol{F} = q[\boldsymbol{E} + (\boldsymbol{v} \times \boldsymbol{B})] \tag{6.0.8}$$

Maxwell's equations (6.0.4),(6.0.5), (6.0.6), and (6.0.7) give us description of the whole classical electrodynamics.

Two of the Maxwell's equations i.e. Equations (6.0.6) and (6.0.7) show us that in a dynamic

situations, electric and magnetic fields are coupled to each other. They tell us that changing electric field can produce a magnetic field and conversely changing magnetic field can produce an electric field. This fact is the basis of all modern inventions that uses radio-waves, and main means by which electricity is produced.

At any time when we have ability to control or produce one field by the other, it opens for us a lot of possibilities for new applications and making improvement in old ones. In applications and for research purposes however, it is much easier to produce, apply and control constant electric and magnetic fields. If we will not consider dynamic conditions and limit ourselves only to static electric and magnetic fields then we have

$$\frac{\partial \boldsymbol{E}}{\partial t} = 0 \quad \text{and} \quad \frac{\partial \boldsymbol{B}}{\partial t} = 0$$

If we put these latter equations into Maxwell's equations, we get the corresponding equations for the static fields

Electrostatics:

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon_0} \tag{6.0.9}$$

$$\nabla \times \boldsymbol{E} = 0 \tag{6.0.10}$$

Magnetostatics:

$$\nabla \cdot \boldsymbol{B} = 0 \tag{6.0.11}$$

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J} \tag{6.0.12}$$

As one can see that for static conditions, electric and magnetic fields are decoupled and they do not affect each other as they do in dynamical cases. However, in some materials, static electric field can magnetize the matter and consequently create a magnetic field; and magnetic field can polarize the matter electrically and consequently create an electric field. This effect is called magnetoelectric effect and it opens a possibility to convert an electric field to a magnetic field and a magnetic field

to an electric field in static conditions.

The class of materials that are considered as primary candidates for exhibiting magnetoelectric effect is *multiferroics*. These are materials that exhibit electric and magnetic orderings simultaneously. One important example of multiferroics is $BiFeO_3$. The goal of this work is to study magnetoelectric effect in multiferroics. In particular, we will focus our attention to $BiFeO_3$.

In a study like this, we start by describing the system energetically. This involves the description of matter by giving its energy as a function of its different degrees of freedom and energy parameters. Then we use ab-initio methods, i.e. methods based on quantum mechanics to compute energy parameters. Ab-initio methods provide us with the starting point for our simulation. The next step is to compute relevant quantities that characterize the system at a finite temperature using Monte Carlo methods (see section 3.4). We did not describe all of those steps in details, but we did mention important things about them in previous pages. In the remaining pages, we will mainly focus on results that were obtained as a consequence of application of the our method to particular problems.

6.1 Antiferromagnetic multiferroics

In early days of experiments on magnets, materials were discovered which act like paramagnets in all respect but had atypical dependence of their susceptibilities on temperature. Later it was discovered that this unusual behavior is due to the difference in the magnetic structure of such materials.[21] Today those materials are considered as a separate class of magnetic materials and they are called antiferromagnetic. Now we know that antiferromagnetic materials are more common than ferromagnets and they have a rich set of interesting properties. Table 1 lists some of the typical antiferromagnetic materials with their properties.

In magnetic materials with antiferromagnetic structure, magnetic moments on neighboring lattice sites point in opposite directions. Lattice points gathering magnetic moments that point along

[21]The contributions of French physicist Louis Neel (22 November 1904 – 17 November 2000) is very important in this respect. He received a Nobel Prize for Physics in 1970 for his work on magnetic properties of solids.

Material	Metal Ion Arrangement	$T_N(K)$	$\Theta(K)$	$\frac{\Theta}{T_N}$	$\frac{\chi_\perp(0)}{\chi_\perp(T_N)}$
MnO	fcc	122	610	5.0	0.69
FeO	fcc	198	570	2.9	0.78
CoO	fcc	293	280	1.0	—
NiO	fcc	523	3000	5.7	0.67
α-MnS	fcc	154	465	3.0	0.82
β-MnS	fcc	155	982	6.3	—
α-Fe_2O_3	r	950	2000	2.1	—
Cr_2O_3	r	307	1070	3.5	0.76
$CuCl_2 \cdot 2H_2O$	r	4.3	5	1.2	—
FeS	hl	613	857	1.4	—
$FeCl_2$	hl	24	−48	−2.0	<0.2
$CoCl_2$	hl	25	−38	−1.5	∼0.6
$NiCl_2$	hl	50	−68	−1.4	—
MnF_2	bct	67	80	1.2	0.76
FeF_2	bct	79	117	1.5	0.72
CoF_2	bct	40	53	1.3	—
NiF_2	bct	78	116	1.5	—
MnO_2	bct	84	—	—	0.93
Cr	bcc	310			
α-Mn	cc	100			

Table 1: List of some antiferromagnetic materials [22] with their properties. The θ in this Table refers to a parameter of the equation $\chi = \frac{C}{(T+\theta)}$, T_N is the Neel temperature, and fcc – face-centered cubic, r – rhombohedral, hl – hexagonal layers, bct – body-centered tetragonal, bcc – body-centered cubic, cc – complex cubic. $\chi_\perp$ is the component of the magnetic susceptibility that is perpendicular to the antiferromagnetic vector.

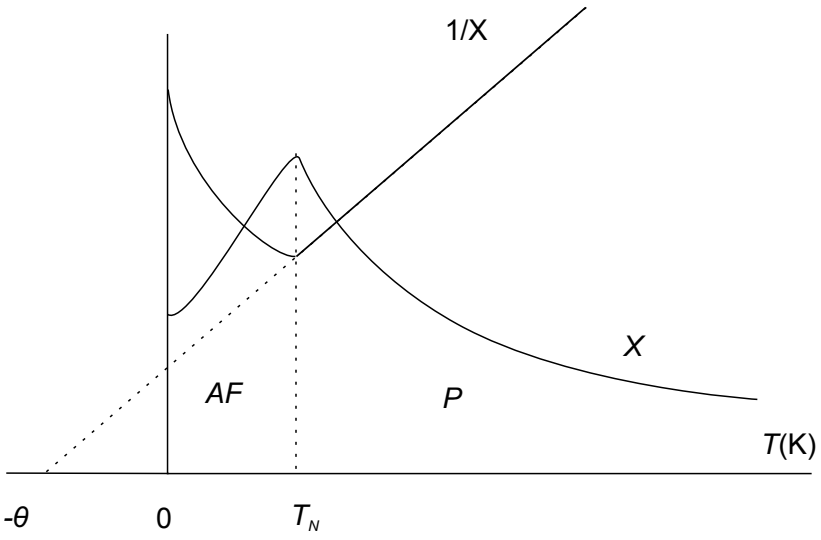

Figure 2: Temperature dependence of susceptibility χ and inverse susceptibility $1/\chi$ for an antiferromagnetic material (schematic). AF – antiferromagnetic, P – paramagnetic (see [22]).

the same direction form a magnetic sub-lattice. Thus antiferomagnets consist of two magnetic sublatices for which the magnetic moments point along opposite directions. As we discussed before, G-type antiferromagnets can be characterized quantitatively using the antiferromagnetic vector $\boldsymbol{L}$ defined as follows

$$\boldsymbol{L} = \frac{1}{\mathrm{V}} \sum_{i,j,k} (-1)^{i+j+k} \, \boldsymbol{m}(i,j,k) \tag{6.1.1}$$

where V is the volume over which the summation is conducted.

Antiferromagnetic substances have a small positive susceptibility and typically their susceptibilities depend on the temperature as shown in Figure 2. The transition temperature of an antiferromagnet is called the Neel temperature and above that temperature the antiferromagnet becomes a paramagnet. This Neel temperature is denoted as T_N in Figure 2. Above the Neel temperature, the dependence of magnetic susceptibility on temperature can be expressed using Curie-Weiss law

$$\chi = \frac{C}{T + \theta} \tag{6.1.2}$$

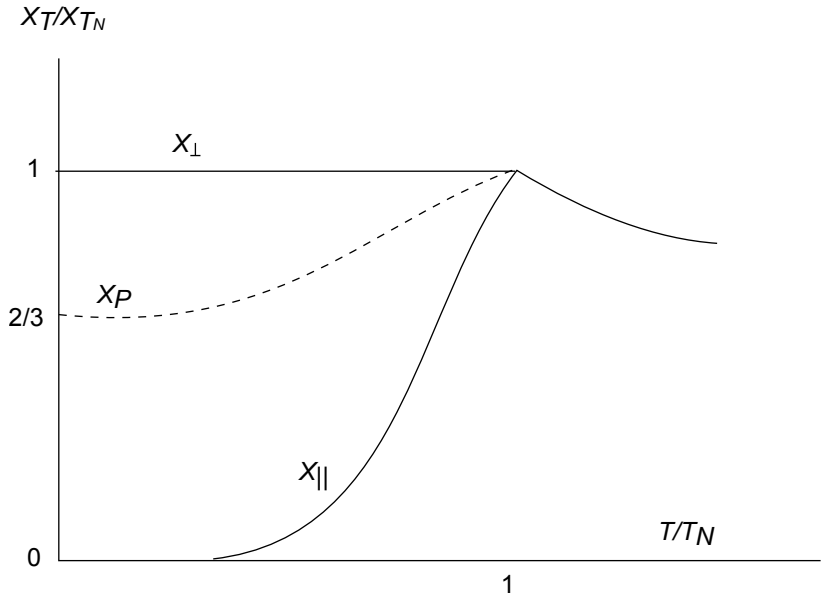

Figure 3: Theoretical temperature dependence of the susceptibility of an antiferromagnet (see [22]). The variable $\chi_\perp$ is a susceptibility when measured by applying the magnetic field $\boldsymbol{H}$ along the direction perpendicular to the antiferromagnetic vector $\boldsymbol{L}$ and the variable $\chi_\parallel$ is the magnetic susceptibility measured by applying magnetic field along the antiferromagnetic vector $\boldsymbol{L}$. χ_p, on the other hand, is the susceptibility of the polycrystalline material for which it is impossible to measure $\chi_\perp$ and $\chi_\parallel$ separately. Thus χ_p has values between those of $\chi_\perp$ and $\chi_\parallel$. According to simplified calculation, the susceptibility of a powder χ_p at temperature near 0K is $3/2$ times smaller than its value near T_N.

if we express the relation (6.1.2) in the linear form as

$$\frac{1}{\chi} = \frac{T+\theta}{C} \tag{6.1.3}$$

we can use the linear least square method again and find parameters θ and C by fitting the curve to the computed values of T and χ.

At the Neel temperature T_N, the susceptibility of powder material reaches to its maximum value and below the Neel temperature susceptibility starts to decrease in value. However, theory and experiment with single crystals show that if the field is applied along the direction which is perpendicular to the antiferromagnetic vector the susceptibility is more or less constant below the Neel temperature. This is demonstrated in Figure 3. Only after transition point the susceptibility start decreasing according to the Curie-Weiss law.

We would like to start this section by first modeling and reproducing the properties of generic antiferromagnetic materials described above. Then we will introduce magnetoelectric effect and see how the properties of the antiferromagnet is modified by this phenomena.

We should note that in the whole subsequent discussion of the theory we implicitly assume that the material under considerations is not a conductor of electricity. This will allow us to consider magnetic moments as localized in lattice sites. This is not a very severe restriction for our purpose since most of the antiferromagnets are insulators. With this assumption we can model the interaction between magnetic moments $\boldsymbol{m}_i$ and $\boldsymbol{m}_j$ by the energy

$$W_{AFM} = -\sum_{ij} A_{ij}(\boldsymbol{m}_i \cdot \boldsymbol{m}_j) \tag{6.1.4}$$

When we introduced an effective Hamiltonian for the $BiFeO_3$ system, we presented energy of this form as part of magnetic energy, see e.g. (5.1.20), and (5.1.21). If the parameters A_{ij} of the energy (6.1.4) are mostly negative[22] ($A_{ij} < 0$), the minimum of the energy (6.1.4) is when moments i and j are pointing along opposite directions. Therefore, this type of energy will favor antiparallel alignment of the corresponding magnetic moments. Our task is then to model magnetic system with such energy and see if the structure we obtain in this way is really G-type antiferromagnet with all of its characteristic behavior. In particular we are interested in the evolution of the susceptibility with temperature. The results of the simulation are given in Figures 4, 5, 6, and 7. Figure 4 is just an illustration of the procedure for calculating the magnetic susceptibility. For that purpose we apply a magnetic field H along the $[111]$ direction (which is perpendicular to the antiferromagnetic vector) and compute the components and magnitude of the magnetization M. To compute the perpendicular component of the susceptibility we fit M vs. H curve with a straight line

$$f(x) = a + bx$$

[22]This term "mostly negative" can be made mathematically strict by introducing the notion of negative definite matrices. However, it is not necessary for such a simple case.

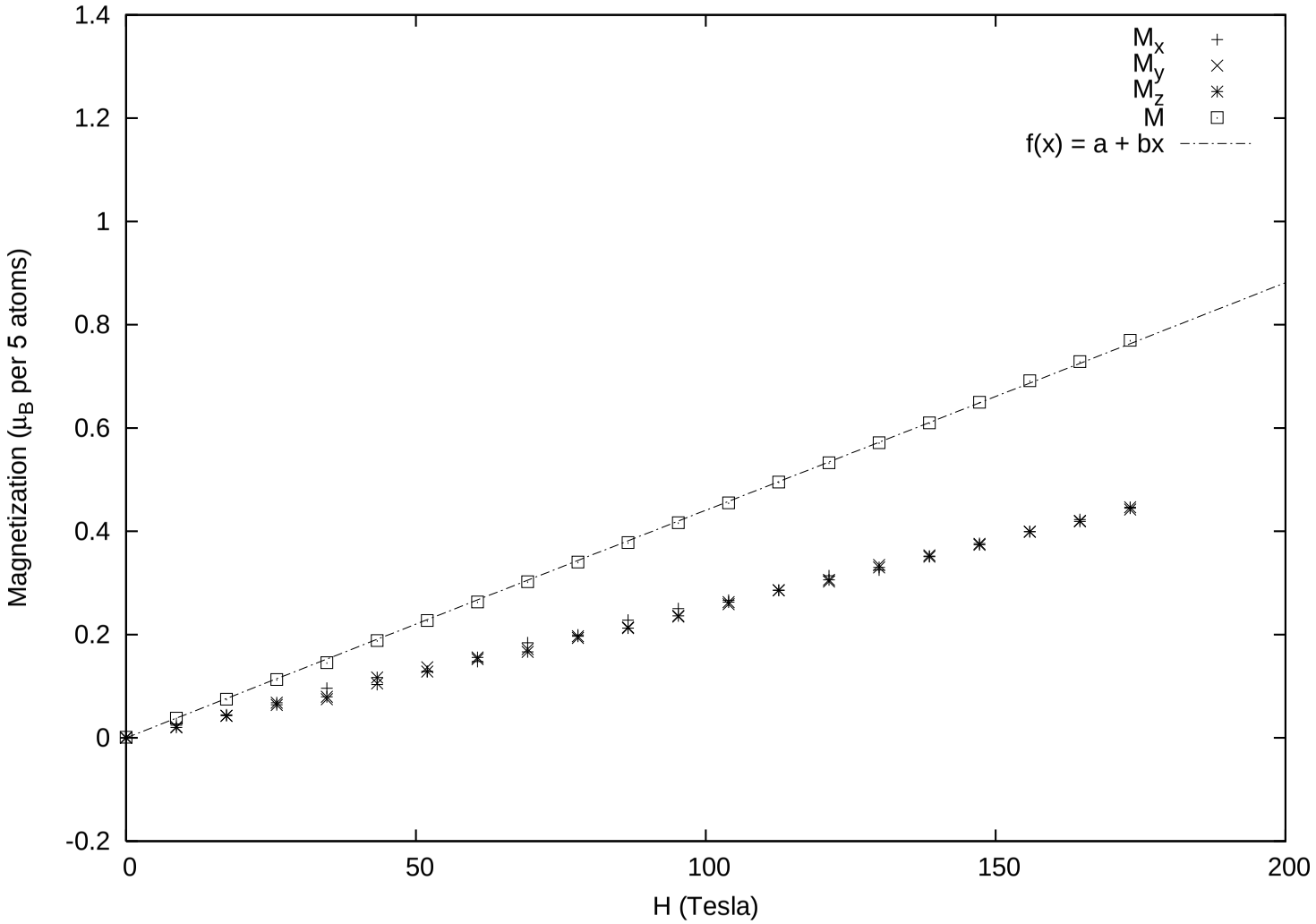

Figure 4: The components and magnitude of the magnetization $\boldsymbol{M}$ vs. magnetic field $\boldsymbol{H}$ applied along the [111] direction.

Then a will give the value of intrinsic magnetization M_0 and b will give us the value of the magnetic susceptibility. As we can see from Figure 4, the intrinsic magnetization M_0 for this particular system is zero, i.e. the M vs H curve can be fitted with the b parameter alone. In a similar manner we can compute the magnetic susceptibility for all different temperatures and get the dependence of the susceptibility on the temperature T. This dependence is given by the curve in Figure 7 labeled as "$\gamma_{me} = 0$". The curves labeled as "$\gamma_{me} = 0$" of Figures 5 and 6 shows the evolution of the antiferromagnetic vector $\boldsymbol{L}$ and of the local modes $\boldsymbol{u}$ with temperature, respectively. As one can see from these figures, the behavior of our system is the same as for the pure antiferromagnetic and ferroelectric systems. The next step in our investigation is to include magnetoelectric effect and see the change in properties of the system due to this effect.

To include such magnetoelectic effect, we again have to introduce the energy term that models that effect. We use the following expression for such energy

$$W_{ME} = -\sum_{i,j,\alpha,\gamma,\nu,\delta} \gamma_{me} \cdot E_{ij,ref} m_{i,\alpha} m_{j,\gamma} u_{i,\nu} u_{i,\delta} \tag{6.1.5}$$

This energy was also previously discussed with γ_{me} as a part of the total energy expression, see e.g. (5.1.22). In the expression (6.1.5), the vector $\boldsymbol{u}_i$ represents the local mode of the system. The value of the electric dipole moment at site i is proportional to $\boldsymbol{u}_i$. Thus, for our purpose, we can use $\boldsymbol{u}$ instead of $\boldsymbol{p}$, and it is also convenient to introduce the quantity E_{ij} defined as

$$E_{ij} = \gamma_{me} E_{ij,ref}$$

where $E_{ij,ref}$ is the value of the $E_{ij,\alpha\gamma\nu\delta}$ coefficients obtained from first-principle calculations. In that case, γ_{me} quantifies the strength of the magnetoelectric (ME) couplings: $\gamma_{me} = 0$ means no ME coupling, while $\gamma_{me} = 1$ is associated with the ME coupling computed by first-principle calculations.

An increase in γ_{me} in the equation (6.1.5) therefore characterizes the strength of magnetoelectric effect. The curves labeled as "$\gamma_{me} = 0.125$" in Figures 5, 6, 7 show the the results for our

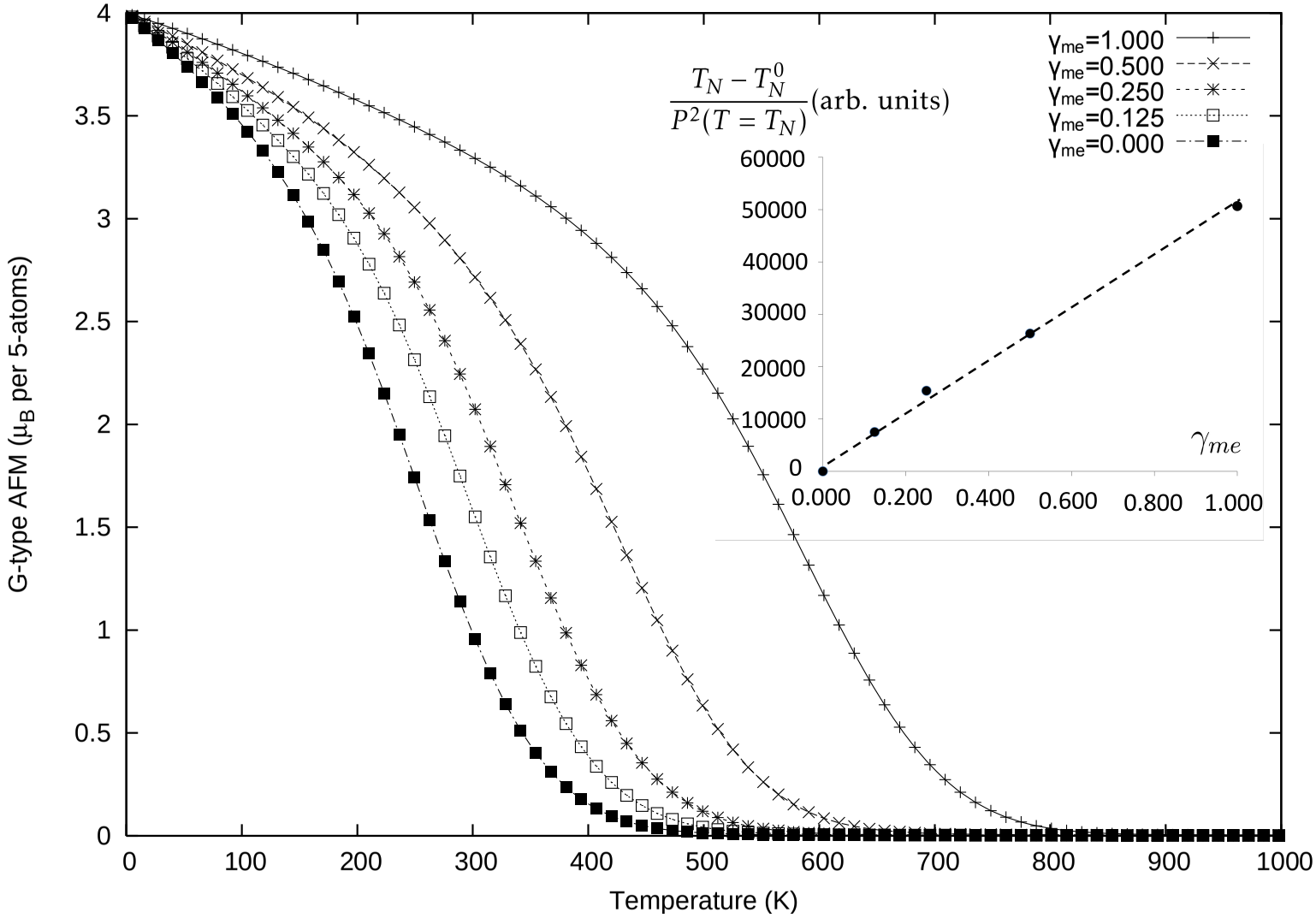

Figure 5: Temperature dependency of the magnitude of the G-type antiferromagnetic vector, for five different selected sets of the ME-related E_{ij} parameters of Eq. (6.1.5). The numerical data are shown by symbols while the lines are guide for the eyes (see appendix).

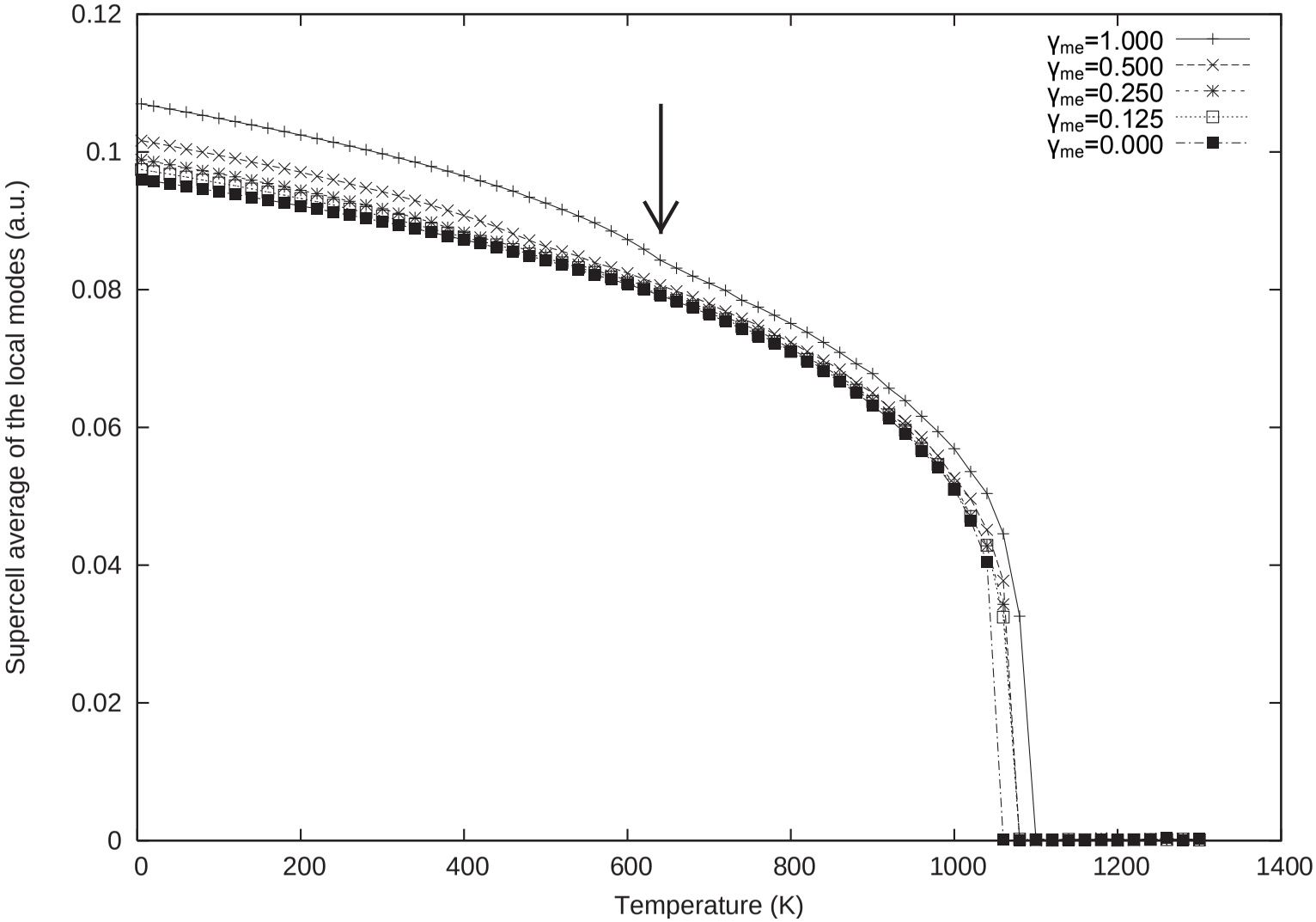

Figure 6: Same as Figure 5, but for the magnitude of the supercell average of the local modes. The arrow shows the location of the Neel temperature when $\gamma_{me} = 1$ (see appendix).

computations when $\gamma_{me} = 0.125$. The rest of the curves correspond to the successively doubling of γ_{me}. When $\gamma_{me} = 1$, the curves will correspond to the E_{ij} that are obtained from first principles.

From those figures, one can see that the properties of our system have been modified by introduction of the magnetoelectric effect. First of all, the antiferromagnetic transition temperature increases by increasing the value of γ_{me}. This also means, of course, that the magnitude of the antiferromagnetic vector is larger for large values of γ_{me} for temperatures below 250K. Exception for this is the point $T = 0$K, where all the curves give the same value for antiferromagnetic degrees of freedom. The magnitude of the $\boldsymbol{L}$ vector at 0K corresponds to the state in which all the magnetic moments are perfectly aligned (in parallel and antiparallel fashion) along the same axis. In this state, the $\boldsymbol{L}$ vector takes its maximum possible value, i.e. it saturates. The saturation value of the $\boldsymbol{L}$ vector is $4\mu_B$, as consistent with first-principle calculations [37].

As it is the case with antiferromagnetic order parameter, the evolution of the ferroelectric order parameter, i.e. local modes (and consequently polarization) with temperature also experiences modifications (Figure 6). First of all, as a result of magnetoelectric effect, the ferroelectric transition temperature T_C increases: larger γ_{me} corresponds to larger T_C. However, the increase in T_C is much smaller in comparison with the increase in the antiferromagnetic transition temperature T_N. This is not surprising. Since Neel temperature for our system is smaller than the ferroelectric transition temperature, the effect of the magnetic order on T_C is is not as strong. Another consequence of the magnetoelectic effect is in the magnitude of the local mode $\boldsymbol{u}$: larger value of γ_{me} leads to larger value of $\boldsymbol{u}$ at a given temperature (below the Curie temperature). Also the disappearance of the magnetic order parameter at T_N leads to a fast decrease in the magnitude of $\boldsymbol{u}$ when increasing temperature near T_N, when γ_{me} is large.

The magnetoelectric effect also introduces modifications on the curve representing the evolution of the magnetic susceptibility with temperature. For example, the value of the magnetic susceptibility decreases as γ_{me} increases for given values of temperatures below T_N. Also, the ferroelectric transition temperature has an effect on the χ vs. T curve in the form of a jump near $T_C \sim 1070$K. This jump decreases as γ_{me} becomes smaller and disappears for $\gamma_{me} = 0$. For $\gamma_{me} = 0$,

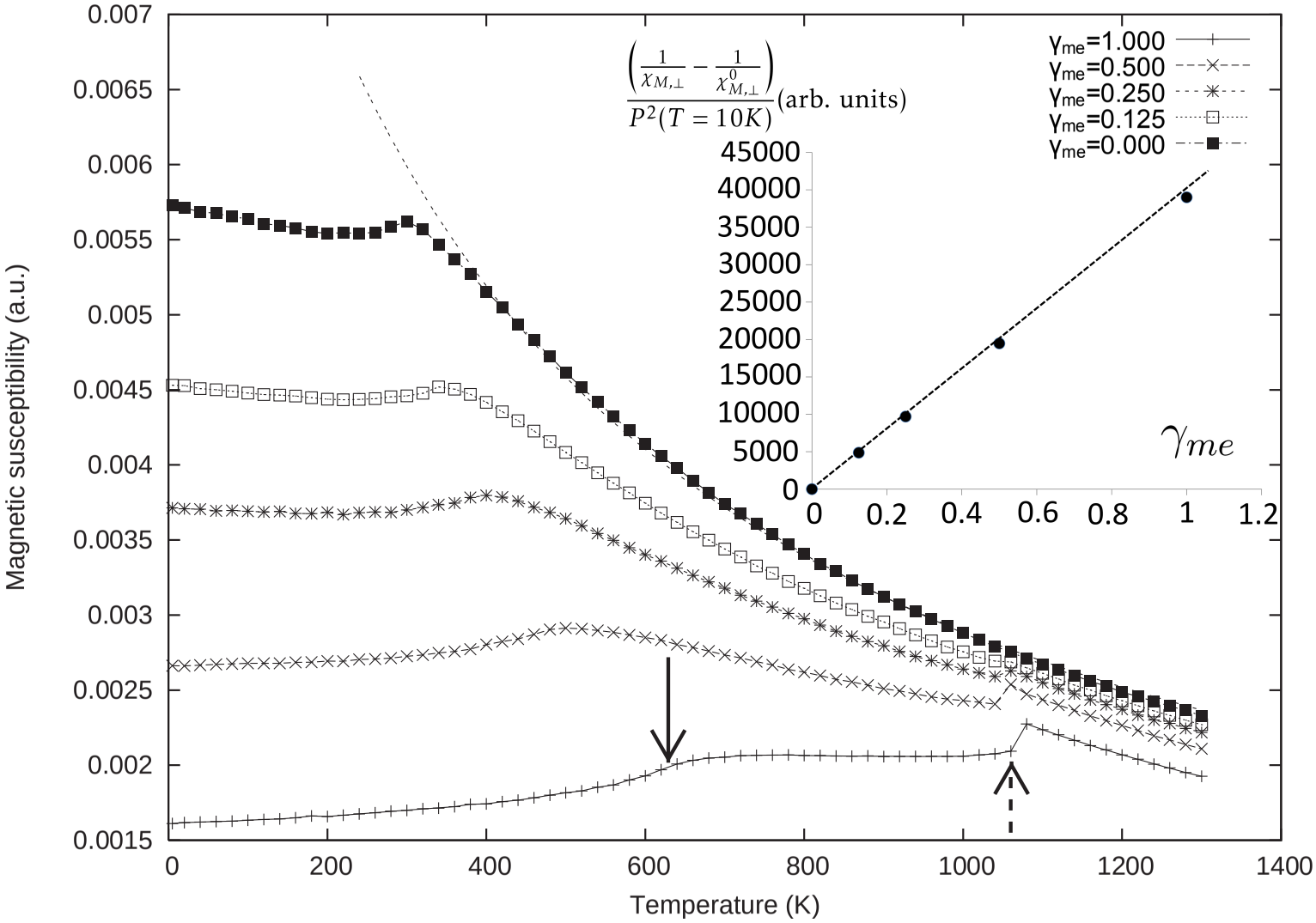

Figure 7: Same as Figures 5 and 6, but for the perpendicular component of the magnetic susceptibility. The dashed line shows the fit of the magnetic susceptibility by $\frac{const}{T+T_N^0}$, in case of *no* ME coupling. The solid and dashed arrows depict the values of T_N and T_C, respectively, when $\gamma_{me} = 1$. The inset demonstrated one of the relations which is direct consequence of the equation (6.1.18) (see appendix).

the region above T_N behaves according to the Curie-Weiss law.[23] This is shown by a dashed line in Figure 7. Below T_N, the value of the magnetic susceptibility is more or less constant when $\gamma_{me} = 0$. However, this typical behavior for antiferromagnets below and above T_N is modified by the introduction of the magnetoelectric effect. For example, for $\gamma_{me} = 1$, below T_N, χ is increasing monotonically with temperature and above T_N it is more or less constant between the Neel temperature T_N and the Curie temperature T_C. Also, as we already mentioned, we observe a jump in χ vs T curves, at T_C, for $\gamma_{me} \neq 0$, which is obviously a deviation from a Curie-Weiss behavior.

Our task now is to understand the character of this modifications qualitatively and quantitatively. We can approach this problem from two different sites. One of them is based on microscopic theory (in analogous to Weiss molecular field theory for ferromagnets). Another is macroscopic theory (based on Landau type expansion of free energy). Here we will describe microscopic theory and give excerpt from the article that describes macroscopic theory (see the appendix).

When there is no interaction between the magnetic moments (i.e., when the parameter A_{ij} is zero) magnetic moments point along arbitrary directions making the overall magnetization of the system under consideration zero. Under external field $\boldsymbol{H}$ however, magnetic moments start to align themselves along the field. The alignment will not be perfect at finite temperature. Higher the magnitude of the field the stronger the alignment and consequently larger the magnetization. On the other hand if we increase the temperature more misaligned the system of magnetic moments becomes. Thus the competition between disorder introduced by temperature and alignment because of external field will determine the final magnetisation of the system.

The phenomena described above can be reproduced by adding the energy term of the form

$$\text{ENERGY} = -\sum_i m_i H \tag{6.1.6}$$

[23] By Curie-Weiss law we mean the following relation that describes the evolution of the magnetic susceptibility χ with temperature above the transition temperature T_N:

$$\chi = \frac{C}{T + T_N}$$

where C is a constant that depends on the system under considerations.

which will tend to align the moments along the field $\boldsymbol{H}$. One can show analytically that indeed the magnetization of such system will be proportional to the external field for a small values of $\boldsymbol{H}$

$$\text{ENERGY} = -\sum_i m_i H \quad \Longrightarrow \quad M = \chi_m H \tag{6.1.7}$$

and the proportionality coefficient χ_m of the expression (6.1.7) is called magnetic susceptibility. The susceptibility of such system will depend on temperature according to the law

$$\boxed{\chi_m = \frac{C}{T}} \tag{6.1.8}$$

This law is called Curie law and the coefficient C is known as the Curie constant. Systems that obey such a law are pure paramagnets. These are systems with non zero magnetic moments but those moments do not interact with each other.

Let's now introduce the interaction between magnetic dipole moments. We can do so by expressing the total magnetic energy of the system as

$$\text{ENERGY} = -\sum_i m_i H - \sum_{i,j} m_i (A_{ij} m_j) \tag{6.1.9}$$

We can rewrite the energy expression (6.1.9) as follows

$$\begin{aligned} \text{ENERGY} &= -\sum_{i,j} m_i \underbrace{[H + A_{ij} m_j]}_{H_t} \\ &= -\sum_i m_i H_t \end{aligned} \tag{6.1.10}$$

where H_t can be considered as a total (effective) magnetic field experienced by individual magnetic moments. This allows us to treat the system as a paramagnetic system under the effective magnetic field. We also can introduce effective or theoretical susceptibility as follows

$$\chi = \frac{M}{H + AM} = \frac{C}{T} \tag{6.1.11}$$

We can solve the above equation for M:

$$M = \frac{CH}{T - CA} \tag{6.1.12}$$

The measured or "true" susceptibility defined as

$$\chi_m = \frac{M}{H}$$

will have the following dependency on temperature:

$$\boxed{\chi_m = \frac{C}{T - (CA)}} \tag{6.1.13}$$

Depending on the sign of the parameter A, the intersection of the $1/\chi_m$ vs. T curve with the T axis can be in the positive values of the temperature or in the negative values of the temperature. If that intersection happens in the positive values of T, the system is ferromagnetic and if the intersection is in the negative values of T the system is antiferromagnetic. Thus the dependence of the $1/\chi_m$ on temperature T gives us the information about the microscopic structure of the system. It tells us about the sign of the coefficient A, and as we discussed above, the sign of A will determine the magnetic structure of the system.

Theory of the ferromagnetic or antiferromagnetic *multiferroics* can be developed along the same lines. First we replace the true magnetic field with an effective magnetic field as follows

$$\begin{aligned} \text{ENERGY} &= -\sum_i m_i H - \sum_{i,j} m_i (A m_j) - \sum_{i,j} m_i (E m_j P_j^2) \\ &= -\sum_{i,j} m_i \underbrace{[H + (A + E P_j^2) m_j]}_{H_t} \\ &= -\sum_i m_i H_t \end{aligned} \tag{6.1.14}$$

Then the expression for the measured magnetic susceptibility will take the following analytic form

$$\boxed{\chi_m = \frac{C}{T-(CA+CEP^2)}} \tag{6.1.15}$$

As it is usually the case, it is much better to work with the inverse quantity $1/\chi_m$ which has thus the following form

$$\boxed{\frac{1}{\chi} = T/C - A - EP^2} \tag{6.1.16}$$

Below the Neel temperature, *pure* antiferromagnetic has almost constant value for its susceptibility

$$\frac{1}{\chi} = A \tag{6.1.17}$$

This is not hard to understand if one imagines that magnetic moments of antiferromagnets in the absence of any anisotropy will align themselves to be perpendicular to external magnetic field to minimize the energy and then start canting along the field. The angle of that canting will be proportional to the external magnetic field and for small fields one can approximately assume that the proportionality coefficient is constant. This proportionality constant is also proportional to the magnetic susceptibility.

For antiferromagnetic multiferroic, this expression can be generalized to be

$$\frac{1}{\chi} = A + EP^2 \tag{6.1.18}$$

The study based on the ideas outlined above are also described in the article titled: "Magnetoelectric signature in the magnetic properties of antiferromagnetic multiferroics: atomistic simulations and phenomenology" (not yet published). The article discusses different approach in deriving the expression for χ. It also draws many conclusions related to magnetic susceptibility of the multiferroic antiferromagnets. The excerpt from that article is given in the Appendix.

6.2 $BiFeO_3$ under electric field

There have been conducted series of outstanding experimental works [5, 6, 8] to study properties of $BiFeO_3$ under external electric field $\boldsymbol{E}$. As a result of such remarkable experimental studies it was demonstrated that the direction of the magnetic order parameter can be switched by applying an electric field. Electric field is directly coupled to the polarization $\boldsymbol{P}$ through the energy term in the form $\boldsymbol{E} \cdot \boldsymbol{P}$. As we know, in $BiFeO_3$ the polarization is coupled with its magnetic degrees of freedom. Thus through a change in the direction of the polarization it was shown that it is also possible to change the direction of antiferromagnetic vector. Coupling between magnetic and electric degrees freedom can be strong.

The coupling between electric and magnetic degrees of freedom is believed to be in a such character that the magnetic order parameter desires to remain perpendicular to the polarization. However, our simulations showed that the more general rule is that the magnetic order parameter $\boldsymbol{L}$ remains perpendicular to the axis about which the oxygen octahedra tilt.[24] Also we conducted simulations with higher values of electric fields and with electric field applied along different directions. Such study gave us more detailed information about the phenomena of switching the direction of the $\boldsymbol{L}$ vector. For example, complete paths through different phases were mapped and described in a detail. Since, all of those result were described in a published article [27], we will not repeat them here, but refer the reader to that article for details.

6.3 Weak ferromagnetism in $BiFeO_3$

Along with antiferromagnetism, BFO also possesses weak ferromagnetism. To understand the nature of such ordering, we have conducted first principle calculations. As a result of that, we concluded that the tilting of the oxygen octahedra is responsible for weak ferromagnetism to occur. Once produced, this magnetization $\boldsymbol{M}$ points along the direction that is perpendicular to both he $\boldsymbol{L}$ and $\boldsymbol{\omega}$ vectors. This findings suggest what kind of energy term should be included in the effective

[24] These results were published in the article: "Electric-field-induced paths in multiferroic $BiFeO_3$ from atomistic simulations," S. Lisenkov, D. Rahmedov and L. Bellaiche, Physical Review Letters 103, 047204 (2009).

Hamiltonian to get ferromagnetism in BFO.

To model and investigate ferromagnetism in BFO, we thus included an energy term of the following form

$$E_{\text{FERRO}} = \sum_{ij} K_{ij}(\boldsymbol{\omega}_i - \boldsymbol{\omega}_j) \cdot (\boldsymbol{m}_i \times \boldsymbol{m}_j) \tag{6.3.1}$$

This is a specific part of the energy associated with magnetic degrees of freedom and we mentioned it before (equation (5.1.25)).

Using this energy term in the MC simulation gives the resulting vectorial order parameters at $T = 10\text{K}$:

$$\begin{aligned}
\boldsymbol{M}(\mu_B) = &\quad 0.011 \cdot \hat{\boldsymbol{x}} + 0.011 \cdot \hat{\boldsymbol{y}} - 0.022 \cdot \hat{\boldsymbol{z}} && (6.3.2)\\
\boldsymbol{L}(\mu_B) = &\quad 2.723 \cdot \hat{\boldsymbol{x}} - 2.877 \cdot \hat{\boldsymbol{y}} + 0.045 \cdot \hat{\boldsymbol{z}} && (6.3.3)\\
\boldsymbol{P}(\text{C/m}^2) = &\quad 0.404 \cdot \hat{\boldsymbol{x}} + 0.404 \cdot \hat{\boldsymbol{y}} + 0.404 \cdot \hat{\boldsymbol{z}} && (6.3.4)\\
\boldsymbol{\omega}(\text{radians}) = &\quad -0.137 \cdot \hat{\boldsymbol{x}} - 0.137 \cdot \hat{\boldsymbol{y}} - 0.137 \cdot \hat{\boldsymbol{z}} && (6.3.5)
\end{aligned}$$

One can see from this result that $\boldsymbol{M} \perp \boldsymbol{L} \perp \boldsymbol{\omega}$, as consistent with Eq. (6.3.1).

Such addition to the model gave us a more accurate picture about BFO systems. For example, Figure 8 shows the temperature dependency of the antiferromagnetic moment and ferromagnetic moment obtained as a result of our simulation. These results were published in an article [31].

We also find that inclusion of the weak feromagnetism leads to important consequences in relation to the magnetoelectric effect in BFO. It was believed that BFO in its "pure" antiferromagnetic state always show linear magnetoelectric effect. However, according to our computations the inclusion of energy term that leads to weak feromagnetism is essential for such effect to occur. Without such energy term, the BFO system can be characterized just by quadratic magnetoelectric effect. As it is shown in Figure 9 for non-zero K_{ij} parameter of Eq. (6.3.1) the value of the polarization under magnetic field experiences a faster decrease as the value of the magnetic field increases. The

resulting curve can be fitted when $K_{ij} \neq 0$ by

$$P = P_0 + \alpha H + \beta H^2 \tag{6.3.6}$$

where $P_0 = 0.698\text{C/m}^2$, $\beta = 1.75 \pm 0.01 \times 10^{-8}\text{C/T m}^2$, and $\alpha = -8.1 \pm 0.3 \times 10^{-7}\text{C/T}^2\text{m}^2$. And of course larger value of K_{ij} parameter of Eq. (6.3.1) leads to larger values of the linear magnetoelectric coefficient, α. Note that β is the quadratic magnetoelectric effect.

7 Magnetic cycloid of $BiFeO_3$

So far we mentioned that $BiFeO_3$ has antiferromagnetic vector $\boldsymbol{L}$, electric polarization $\boldsymbol{P}$, weak magnetization $\boldsymbol{M}$, and antiferrodistortive vector $\boldsymbol{\omega}$. However, $BiFeO_3$ bulks possess one more degree of freedom related to its magnetic structure. Pages below will describe this important property of this material.

Geometry of magnetic cycloid of $BiFeO_3$ bulk. In its ground state, the magnetic structure of $BiFeO_3$ can be characterised by a magnetic cycloid propagating along one of the $\boldsymbol{k}_1 = [0\bar{1}1]$, $\boldsymbol{k}_2 = [\bar{1}01]$, and $\boldsymbol{k}_3 = [\bar{1}10]$ directions. All of those directions are equivalent and can be observed experimentally. Usually crystal is divided into magnetic domains each of which has different cycloids characterised by different propagation vector $\boldsymbol{k}$. The plane on which the magnetic moments rotate is called the cycloidal plane. In $BiFeO_3$, the cycloidal plane coincides with the plane made by the direction along which polarisation $\boldsymbol{P}$ points, i.e. the $[111]$ direction, and the corresponding propagation vector $\boldsymbol{k}$. In the rest of this work, unless otherwise specified, we will consider only the case where $\boldsymbol{k} = \boldsymbol{k}_1 = [0\bar{1}1]$. All the subsequent results can be easily extended for the case of $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$.

Cycloidal energy. Now we would like to modify the energy of the system to incorporate the cycloidal degree of freedom. To model the magnetic cycloid, we propose the following energy

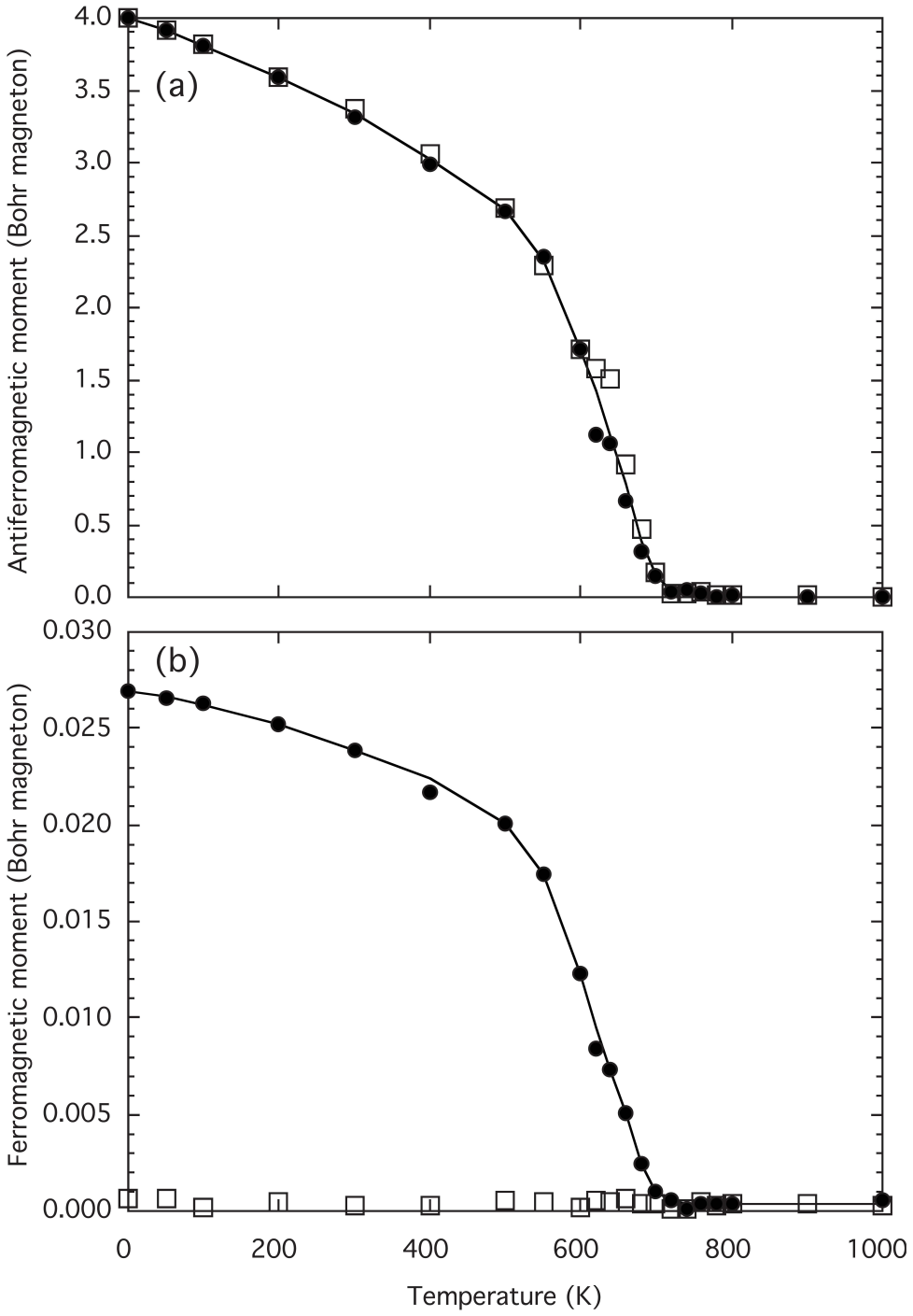

Figure 8: Magnitude of the AFM vector [Panel a] and of the magnetization [Panel b] as a function of temperature in a BFO thick film. The filled (respectively, open) symbols correspond to simulations in which the K_{ij} parameter of Eq. (6.3.1) has been turned on (respectively, off). Lines are guides for the eyes (see [31]).

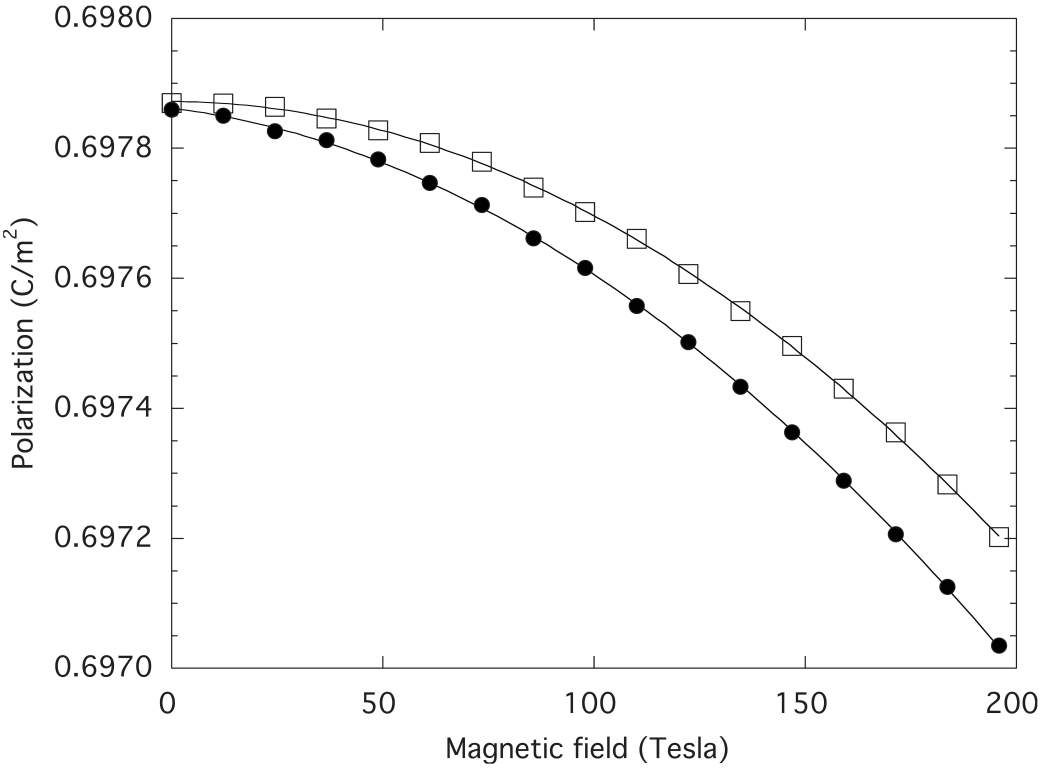

Figure 9: Evolution of the polarization when a magnetic field $\boldsymbol{H}$ is applied along the $[11\bar{2}]$ direction at $T = 20$ K. The filled circles correspond to the case in which the K_{ij} parameter of Eq. (6.3.1) has zero value, while the open squares correspond to the state in which no weak ferromagnetism exists (see [31]).

expression

$$W_p = -\sum_{ij} C_{ij}(\boldsymbol{p}_i \times \boldsymbol{e}_{ij}) \cdot (\boldsymbol{m}_i \times \boldsymbol{m}_j) \tag{7.0.7}$$

In the expression (7.0.7) the index i runs over all the Fe sites while the index j runs over the second-nearest neighbors of the Fe site i, and $\boldsymbol{e}_{ij}$ is a vector along the direction joining the i and j sites. Figure 10(a) shows the resulting magnetic structure of the system after adding the energy term (7.0.7). As one can see, our model reproduces the known geometry of the magnetic cycloid very well.

Energy parameter C_{ij} and the length of the magnetic cycloid. The length of the cycloid can be tuned by changing the the C_{ij} parameter of the W_p energy term. Smaller values of C_{ij} parameter will result in a longer cycloid while larger C_{ij} values will make the cycloid shorter. In the simulation that we will perform however, the size of the cycloid stays fixed and will coincide with the supercell size used in the simulation which is $18 \times 18 \times 18$. In our simulation, this will be reflected by the fact that we have minimum and maximum values for C_{ij} between which the cycloid is stable. The true magnetic cycloid of BFO is much longer than the one that is used in our simulation. Thus the value for C_{ij} coefficient that we use in our simulation might be larger than for the true cycloid.

7.1 Spin Density Wave

More careful analysis of our results showed us that the magnetic cyclod of BFO does not lie in a single plane but deviates from this plane a little bit. This deviation also has harmonic character. All of these are clearly demonstrated in Figure 10(b). Figure 11 shows angular representation of the in-plane and out-of-pane modulations of the magnetic dipoles.

Almost simultaneously with our work, an experimental study [41] on BFO was published and confirmed this out-of-plane spin density wave. Our results are in agreement with this experimental work. The ability of our model to capture such small effect gives us confidence about the accuracy of our model.

7.2 Anharmonicity

There have been experimental studies on anharmonicity of the cycloid of $BiFeO_3$ [33, 34, 35, 36, 38, 39, 40]. Harmonic cycloid means that sine and cosine functions with single wave vector k can accurately describe the cycloid. For such cycloid, the end of the magnetic dipole vector $\boldsymbol{m}$ traces a circular path as we move from one Fe site to the next along the propagation direction of the cycloid:

$$m_x = m_0 \cdot \sin(ki) \tag{7.2.1}$$

$$m_y = m_0 \cdot \cos(ki) \tag{7.2.2}$$

where i are positions of lattice sites along the cycloidal propagation direction. This is a picture of a perfect harmonic cycloid and any deviations from such cycloid can be called anharmonicity. There are many ways of measuring such anharmonicity depending on its nature. We took the following approach. The perfect cycloid will have only one wave vector k, thus in reciprocal space it will be represented by a single point. Our simulation indeed shows that the intensity of a single point in the reciprocal space is dominant. The ratio of the intensity of that point with the intensity of the point with the next highest intensity can be used as a measure of the anharmonicity of the cycloid.

Thus, we suggest that maybe instead of trigonometric functions it may be more accurate to describe the cycloid of $BiFeO_3$ using Jacobi elliptic functions. We do not need to go into long discussion about the properties of Jacobi elliptic functions here. It is sufficient to say that these functions have a parameter denoted m, the value of which varies between 0 and 1 ($0 \leq m \leq 1$). When $m = 0$, the Jacobi elliptic functions become trigonometric function and the cycloid becomes harmonic. We studied these possibilities and calculated the value of the m parameter. The main lesson that we learned from this study was that the anharmonicity of the cycloid is very small (i.e., m is very small) and probably can be ignored for most purposes. To make this point quantitatively, let's look at numerical results obtained from our simulations. When we look at Fourier transformation of the magnetic structure of BFO, 99.12% of the total intensity (square of the Fourier

components) is concentrated in the point corresponding to the cycloid. The point with the next highest intensity has 0.09% of the total intensity. If we had cycloid described by Jacoby function with the m parameter about 0.4, we would have the same distribution of the intensity in the Fourier transformed space. Thus, we can conclude that the value of the m parameter is about 0.4. This result is in agreement with the most recent measurements described in the references [36, 38].

We also conducted simulations to study the behavior of the magnetic cycloid under external magnetic field. The direction of the applied field was chosen to be $[\bar{2}11]$. We considered two different cases: one case with the energy term that gives weak ferromagnetism and another case without this energy term. In the section 5.1 devoted to effective Hamiltonian, Equation (5.1.25) is the term that is responsible for the existence of weak ferromagnetism in this system. By setting the K coefficient of that energy term to zero and by setting it to its first-principle calculated value, we can turn on or off the weak ferromagnetism in the antiferromagnetic state of BFO. For these cases, the evolution of the magnetism of BFO is illustrated in Panel (a) of Figure 12. If we interpolate the curve corresponding to $K \neq 0$, we find that it intersects the y-axis (corresponding to the spontaneous magnetization) at the point $0.027\mu_B$, i.e. at the point corresponding to the magnetization of the antiferromagnetic BFO with weak ferromagnetism. On the other hand, the curve for which $K = 0$ shows that, without the energy term (5.1.25), the system under external magnetic field transforms to the purely antiferromagnetic system without weak ferromagnetism. As a result, we conclude that the transition from cycloidal to the antiferromagnetic state does not require the existence of a weak ferromagnetism. These results are in agreement with experiment [45]. The panel (b) of Figure 12 shows the evolution of the intensity of the points in the reciprocal space corresponding to the antiferromagnetic state and to the cycloidal state. This Panel shows how the cycloid is continuously disappearing and how the structure of BFO is gradually transforming to the antiferromagnetic structure.

Magnetic cycloid of $BiFeO_3$ under external electric field. In this section, we study the magnetic cycloid of $BiFeO_3$ under an external electric field $\boldsymbol{E}$ applied along different cubic directions.

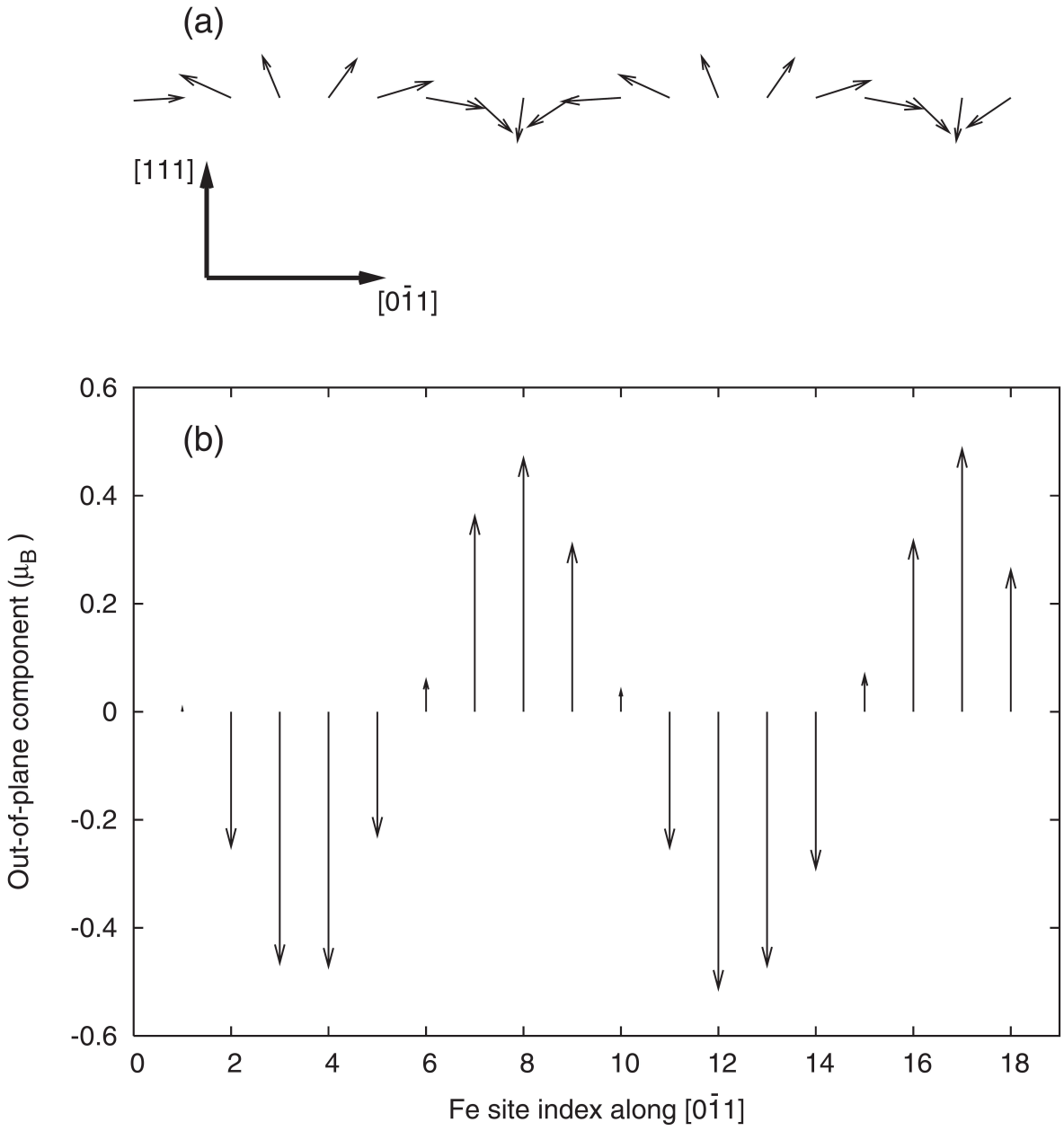

Figure 10: Snapshot of the magnetic dipoles' configuration at 0.2 K when the C_{ij} parameter of Eq. (7.0.7) is chosen to be 5×10^{-6} Hartree/Bohrμ_B^2. Panel (a) shows the projections of the magnetic dipoles into the $(\bar{2}11)$ cycloidal plane for a line of dipoles centered along the same $[0\bar{1}1]$ direction. Panel (b) displays the out-of-plane components of these dipoles (see [30]).

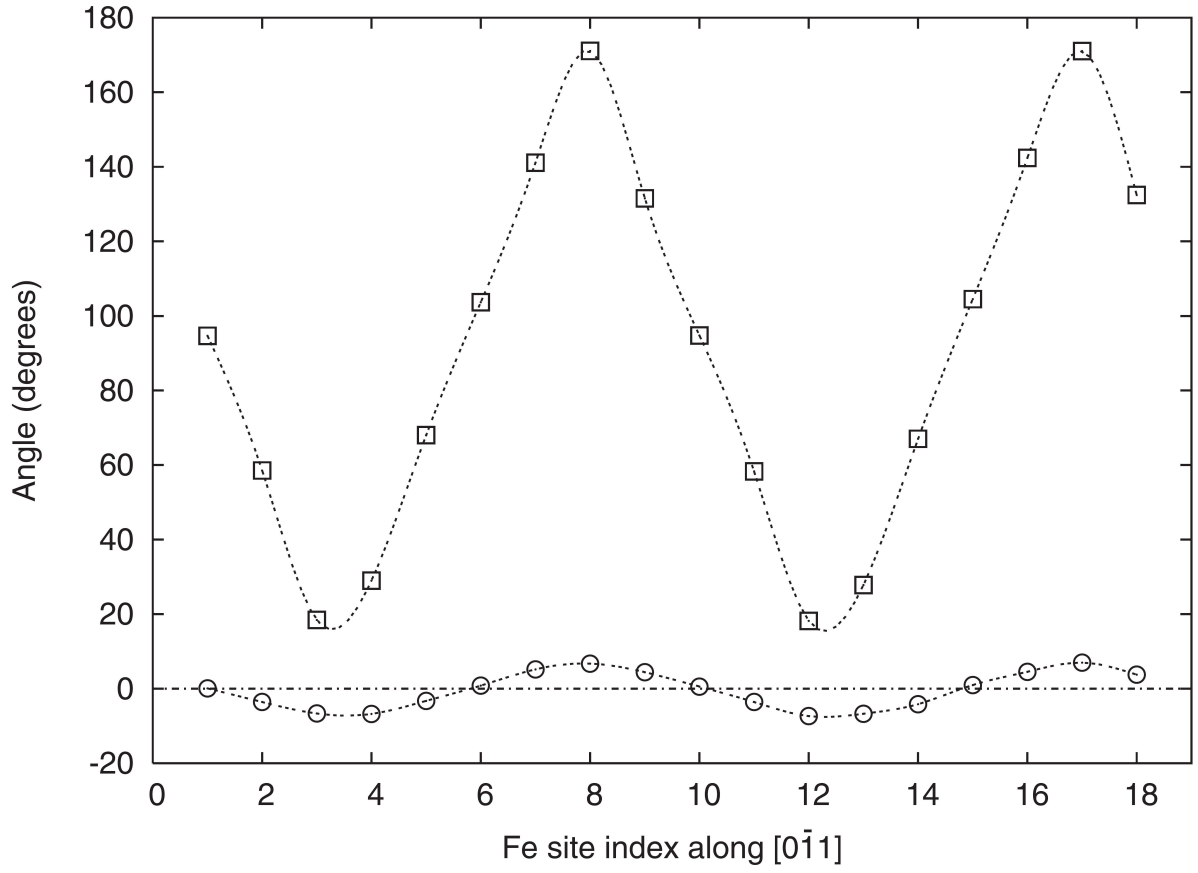

Figure 11: Angle made by the magnetic dipoles with the polarization (open squares) and angle made by these dipoles out of the cycloidal plane (open circles), for a line of dipoles centered along the same $[0\bar{1}1]$ direction. C_{ij} is chosen to be 5×10^{-6} Hartree/Bohrμ_B^2 and the temperature is 0.2 K. The horizontal dashed line corresponds to zero angles (see [30]).

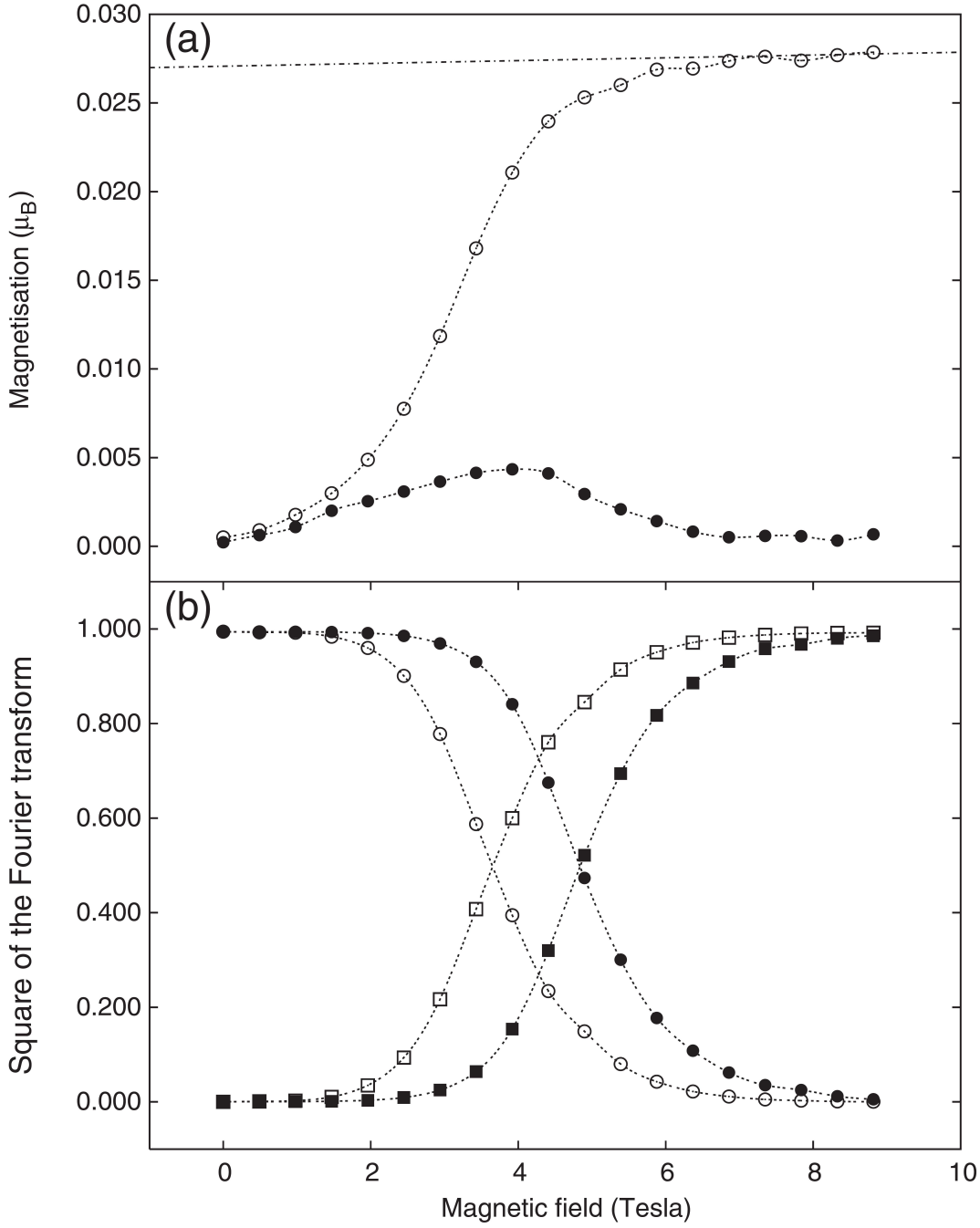

Figure 12: Magnitude of the magnetization (Panel a) and of the square of the Fourier transforms of the local magnetic dipoles' configuration (Panel b) at the k-point corresponding to the cycloidal state (circles) and the k-point corresponding to the antiferromagnetic state (squares), as a function of the magnitude of the magnetic field applied along the $[\bar{2}11]$ direction when the K_{ij} parameter of Eq. (6.3.1) is switched on (open symbols) or off (filled symbols). The C_{ij} parameter is chosen to be 4×10^6 Hartree/Bohrμ_B^2 and the temperature is 5 K. The dashed line in Panel (a) shows the interpolation down to 0 Tesla of the M-versus-H straight line existing above 7 Tesla when K_{ij} is switched on (see [30]).

We observe the destruction of the magnetic cycloid for high enough values of the electric fields, and we give energetic insight to explain this phenomena. Before the destruction, the cycloid undergoes different transformations such as rotation of the cycloidal plane and change in its propagation directions. Also, the unusual magnetic structure of $BiFeO_3$ gives rise to an unique type of population inversion that leads to realization of negative temperature that is observed for the first time in this work.

Electric polarization of $BiFeO_3$ under external electric field. In its ground state, $BiFeO_3$ has a polarization pointing along the $[111]$ direction. Polarisation of $BiFeO_3$ is coupled to the external electric field $\boldsymbol{E}$ through the $W_1 = -\boldsymbol{E} \cdot \boldsymbol{P}$ energy term. If we apply external electric field $\boldsymbol{E}$ along one of the cube diagonal which is different than the initial $[111]$ direction, the polarisation $\boldsymbol{P}$ will change its direction to point along the direction of applied electric field. This phenomena is investigated in our previous work. For this to happen, the magnitude of the electric field must be larger than some critical value.

Energy term favorable to antiferromagnetic state. The magnetic degrees of freedom of $BiFeO_3$ are coupled to its electric polarisation. This coupling can be classified into two kinds. The first kind is characterized by the energy that has the form

$$W_{p^2} = \sum_{ij} E_{ij} \, (\boldsymbol{p}_i \cdot \boldsymbol{p}_j)(\boldsymbol{m}_i \cdot \boldsymbol{m}_j) \tag{7.2.3}$$

Equation (7.2.3) is also discussed when we talked about the effective Hamiltonian of the system in Section 5.1 (see, e.g., Equation (5.1.22)). We denote this energy term as W_{p^2} to indicate that it depends on square of polarisation. This term alone will result in antiferromagnetic state and this state has been studied extensively in our previous work.

Energy term favorable to cycloidal state. The second kind of coupling is characterised by the energy term that has the form

$$W_p = -\sum_{ij} C_{ij}\,(\boldsymbol{p}_i \times \boldsymbol{e}_{ij}) \cdot (\boldsymbol{m}_i \times \boldsymbol{m}_j) \tag{7.2.4}$$

This is a term of the total energy and we mentioned it in the section 5.1 (see, e.g., Equation (5.1.25)). We denote the energy term of Equation (7.2.4) as W_p since its dependency on polarisation is linear. This second kind of interaction leads to the cycloidal magnetic structure. The final magnetic structure will be determined as a result of a competition between these two kind of couplings.

The question that we are asking in this work is what will happen to the magnetic cycloid when an external electric field $\boldsymbol{E}$ is applied to the system. We will start with the case when $\boldsymbol{E}$ is applied along the polarization, i.e. along the $[111]$ direction. One might expect that the cycloid should not experience any change since in this case the propagation direction $\boldsymbol{k}$ and the normal of the cycloidal plane remains perpendicular to the polarization $\boldsymbol{P}$. Also the expression for the energy term W_p shows that an increase in $\boldsymbol{P}$ will lower this energy term which is favorable for the cycloidal state. However, as we pointed out above, the final magnetic structure is determined as a result of a competition between W_{p^2} and W_p. Important thing to notice is that W_{p^2} depend quadratically on $\boldsymbol{P}$ while the dependence of W_p on $\boldsymbol{P}$ is linear. This suggests that, for sufficiently large values of electric field $\boldsymbol{E}$, the energy W_{p^2} will be lower than W_p. Our simulation indeed shows that for every value of the C parameter, there is a critical field beyond which the W_{p^2} term is dominant. This is demonstrated in Figure 13 for two values of the cycloidal C parameter.

In Figure 13, we plot the difference in total energy between the cycloidal state and antiferromagnetic state. Initially, the cycloidal state has a lower energy (energy difference is negative). However, after some critical value of $\boldsymbol{E}$, the antiferromaganetic state has lower energy (energy difference is positive). One also can see that the larger value of C will demand stronger field for the cycloid to be destroyed. This reasoning should be applicable for arbitrary direction of the external

electric field, but as we will see, when $\boldsymbol{E}$ is applied along other directions, other phenomena can also occur.

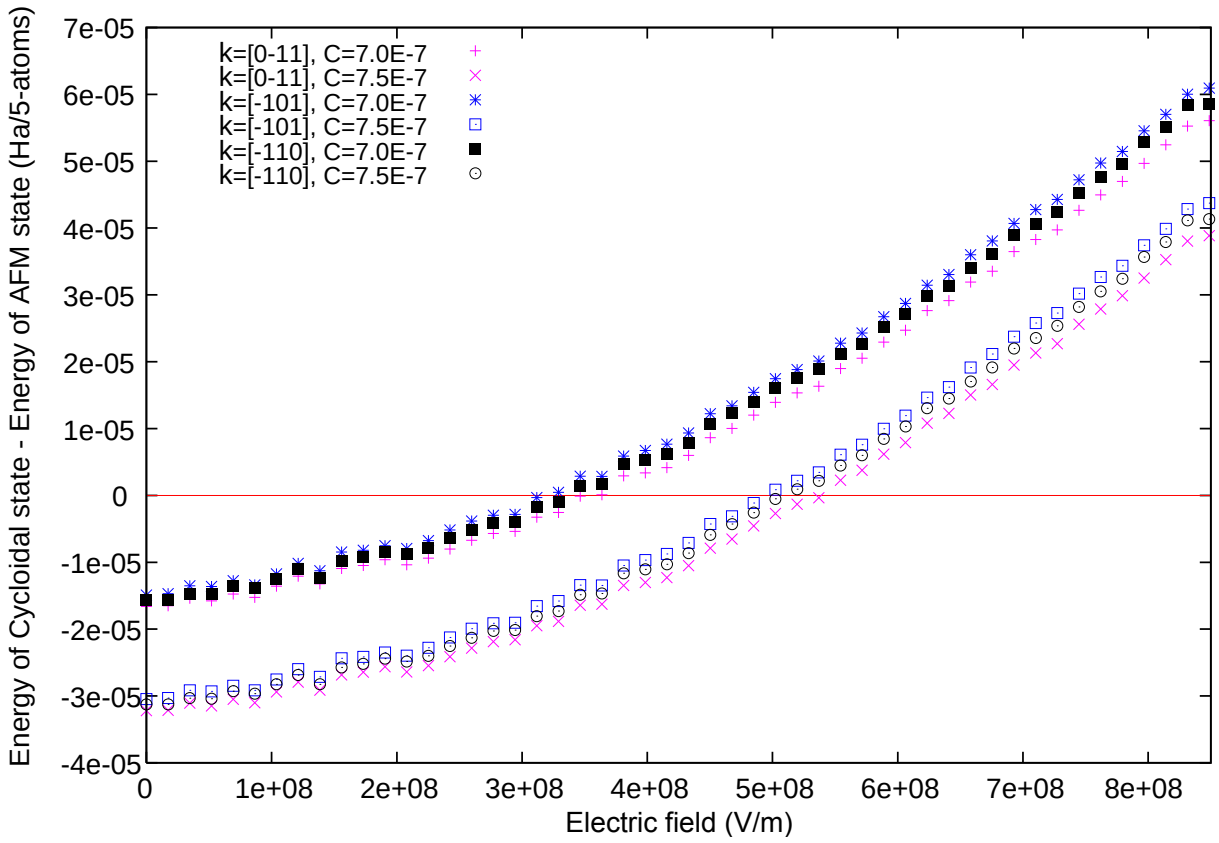

Figure 13: Difference in energy between the cycloidal and antiferromagnetic states when the electric field is applied along the [111] direction.

Let us now consider the case when $\boldsymbol{E}$ is applied along the $[\bar{1}11]$ direction. The magnitude of the applied electric field increases from 0 to 10^9 V·m^{-1}. Again, after certain critical value of electric field $\boldsymbol{E}$, the initial polarization will point along the $[\bar{1}11]$ direction. Now the cycloidal plane does not coincide with the plane made by the initial polarisation $\boldsymbol{P}$ and propagation direction $\boldsymbol{k}$. Our simulation shows that, in this case, the cycloidal plane will rotate to coincide with the plane spanned by the new polarization $\boldsymbol{P}$ direction and the $\boldsymbol{k}$ vector. This is demonstrated in Figure 14. One thing to note here is that the number of MC steps for the cycloidal plane to rotate to a new state is relatively long in comparison with the flipping of the polarization $\boldsymbol{P}$ to a new direction. Thus, the system for a long time remains in a state with high energy. This is an example of population inversion like phenomena in magnetic systems.

This phenomena of inversion in multiferroic BFO is more dramatically illustrated when the

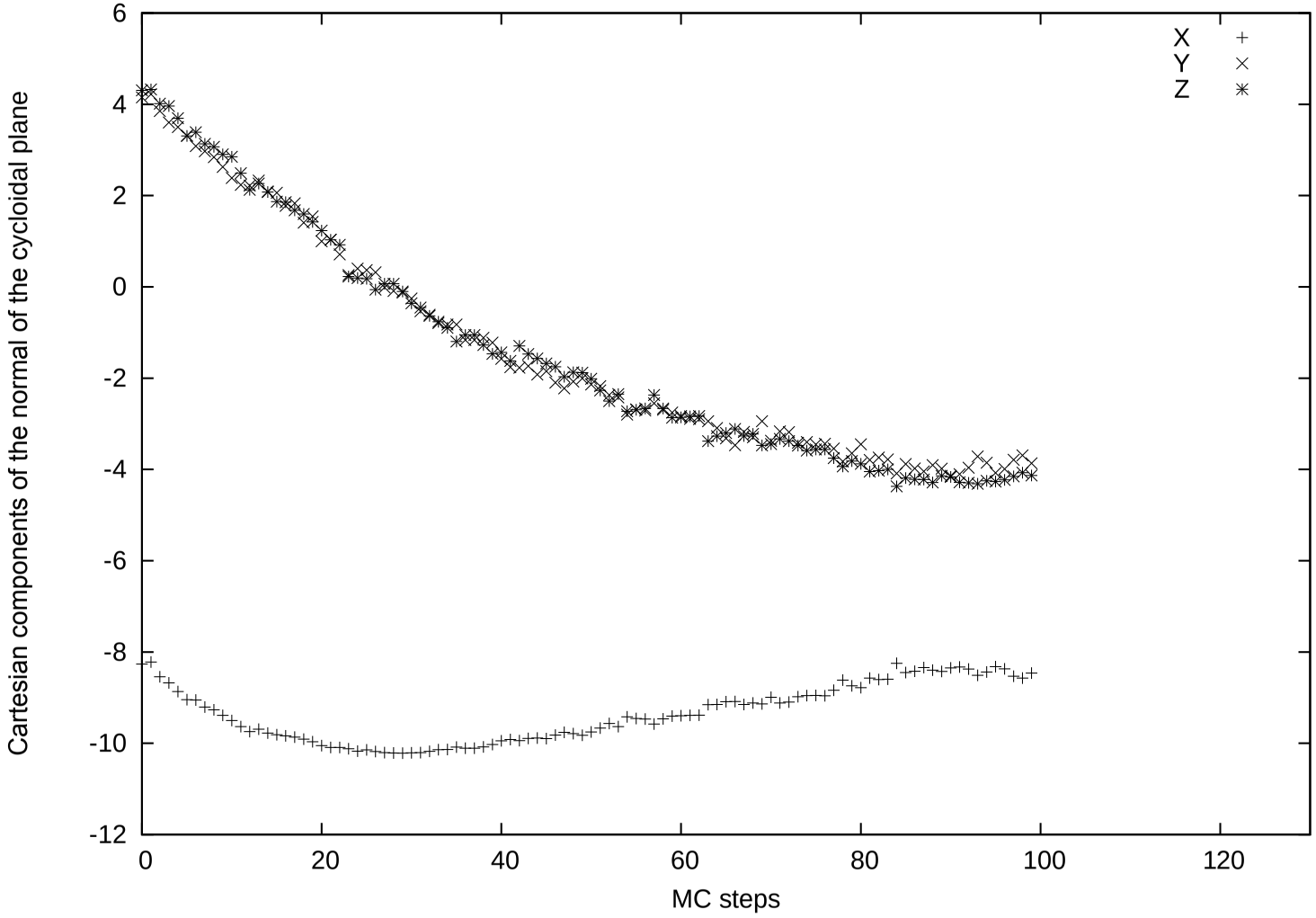

Figure 14: Change of the normal of the cycloidal plane vs. Monte Carlo steps. Cycloidal plane is defined by the direction of its normal (computed as $\boldsymbol{m}_i \times \boldsymbol{m}_j$ where i and j refer to the neighboring sites along the propagation direction of the cycloid). Electric field is applied along the $[\bar{1}11]$ direction.

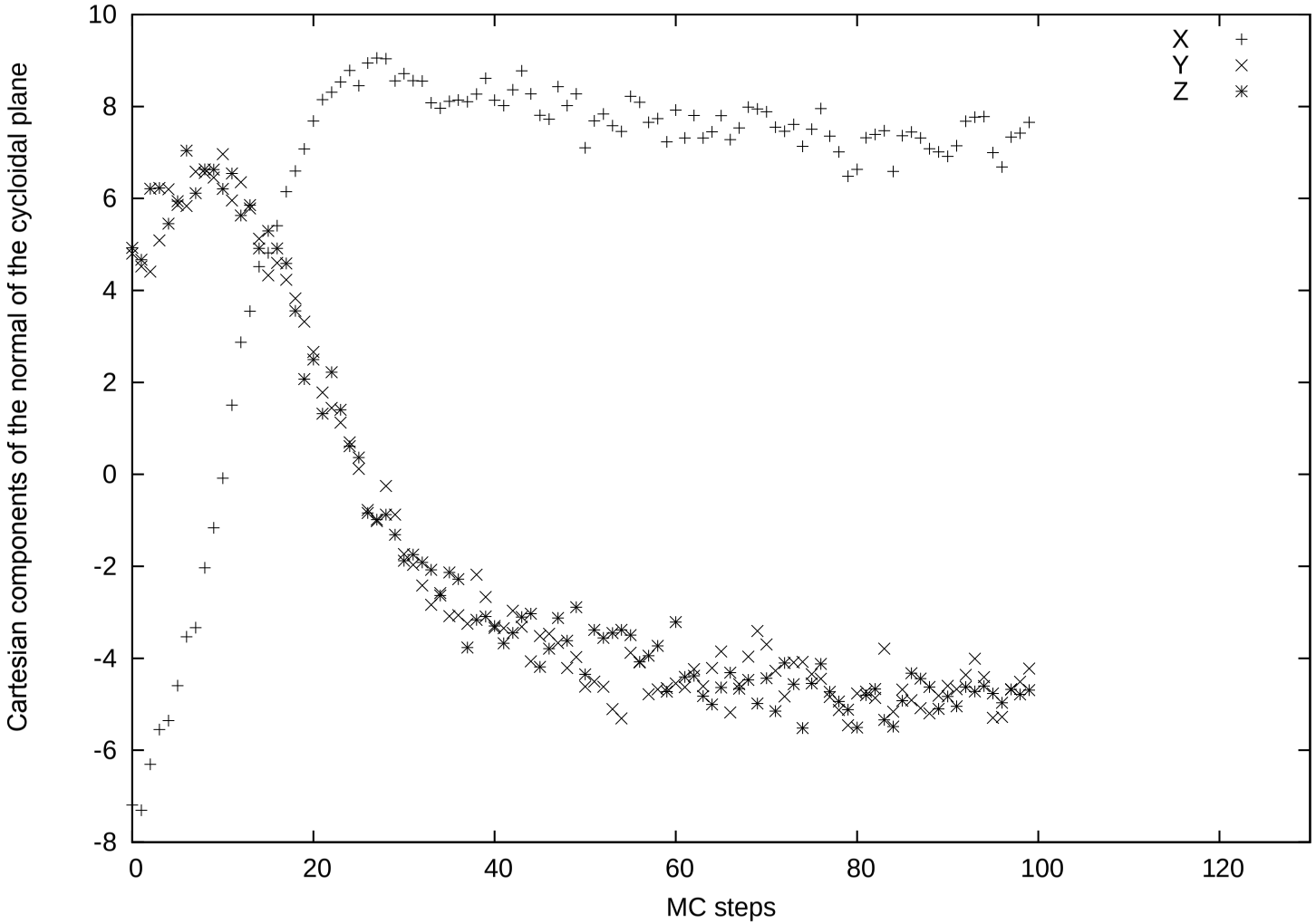

Figure 15: Change of the normal of the cycloidal plane vs. Monte Carlo steps. Electric field is applied along $[\bar{1}\bar{1}\bar{1}]$.

field is applied along the $[\bar{1}\bar{1}\bar{1}]$ direction. The magnitude of the applied electric field increases from 0 to 10^9 V·m^{-1}. Again, the polarization $\boldsymbol{P}$ will flip to point along the new $[\bar{1}\bar{1}\bar{1}]$ direction. One might think that the cycloidal state should not experience any transformation. In this case, however, the energy W_p does not remain the same, but rather changes its sign. Our simulation shows that the cycloidal plane will experience rotation along the $\boldsymbol{k}$ direction by 180 degrees. This is illustrated in Figure 15. Thus, the final state looks the same as the initial state, however the flipping of the polarization $\boldsymbol{P}$ will result in observable physical phenomena, i.e. turning of the cycloidal plane. Again, the magnetic degrees of freedom transforms much slower than electric the degrees of freedom, thus resulting in population inversion.

Another interesting question that we should address is the possibility of changing the propagation direction of cycloid by applying electric field. To achieve that, we should flip polarisation to the direction that is not perpendicular to the propagation direction of the cycloid. An example of such direction for our cycloid is $[1\bar{1}1]$. Again, the polarisation flip can be observed easily. Our simulation shows that, in this case, the cycloid will not remain the same. However, the new propagation direction is not perpendicular to $\boldsymbol{P}$, i.e. is not along $[\bar{1}01]$ or $[011]$ directions. On the other hand, if we heat the system and cool it down, again the new cycloid will be perpendicular to the polarisation (bi.e. can be $[\bar{1}01]$ or $[011]$ direction). Why can't we observe the change in $\boldsymbol{k}$ at a given temperature? One possible answer to this question may be the fact that the size of the supercell that we used in our simulation is too small or the energetic barrier to overcome is too large.

From an experimental point of view, the easiest directions for applying electric field are $[001]$ and $[00\bar{1}]$. Therefore, we also performed simulations for those cases. The result is shown in Figure 16 for different values of C. This results show that the antiferromagnetic state is always preferable to the cycloidal state for large enough value of magnetic field and that the critical field at which the transition from cycloidal to AFM state occurs increases as the C coefficient increases in the magnitude. This is in agreement with our proposed explanation for the destruction of the magnetic cycloid in favor of antiferromagnetic state, since our reasoning is independent on direction of the applied field.

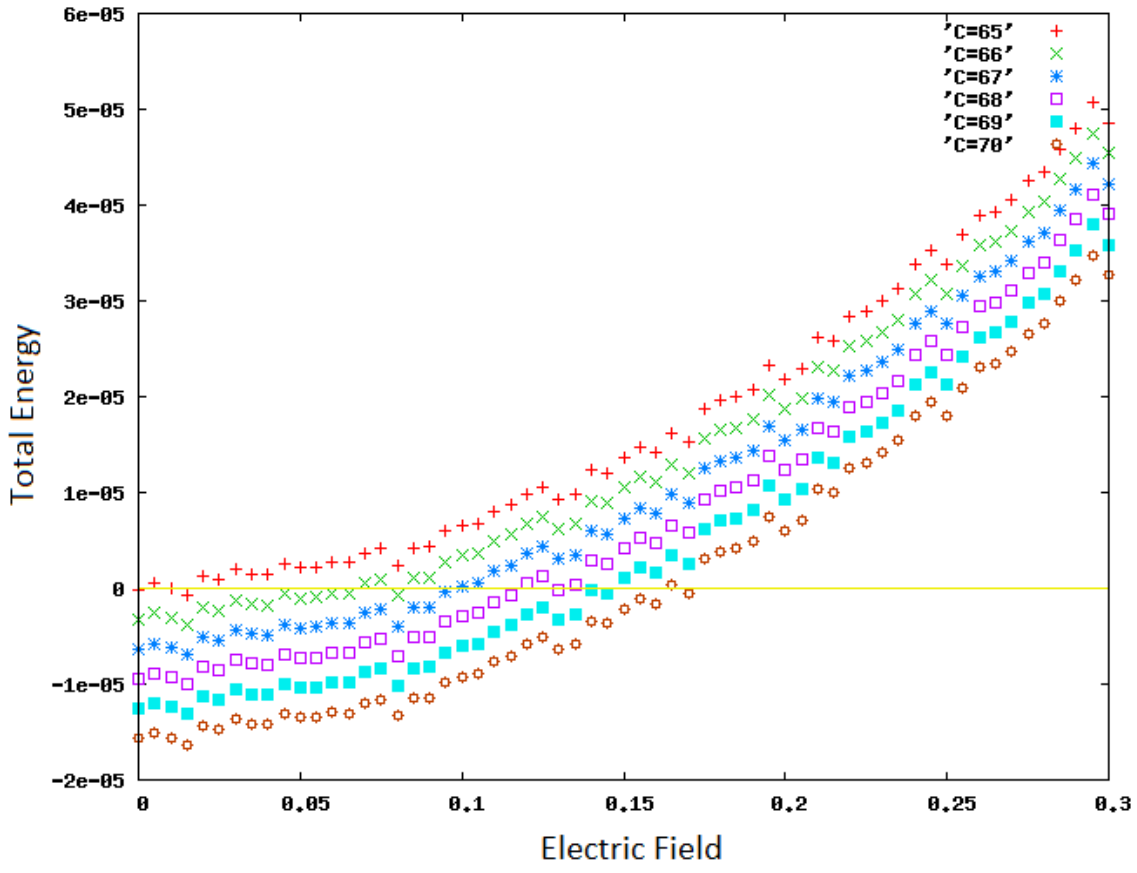

Figure 16: Difference in total energy (in atomic units) between cycloidal state and antiferromagnetic state when different C parameters are used vs. the electrical field (in units of 10^9 V m^{-1}). Electric field is applied along the [001] direction. Different lines corresponds to different values of the C parameters. For example, the curve labeled as $C = 65$ corresponds to $C = 65 \times 10^{-8}$ a.u.

In conclusion, we performed MC simulation to study the magnetic cycloid of $BiFeO_3$ under electric field applied along different directions. Energy considerations shows that under large electric fields, the cycloid will be destroyed in favor of an antiferromagnetic state. We also observed the rotation of the cycloidal plane. When it comes to the change in the direction of propagation vector, however, it is hard to achieve it with the computational resources available to us. One thing is certain – cycloid will not remain along the same propagation direction, but we can't be sure if it will go to the new direction which is perpendicular to $\boldsymbol{P}$ at a given temperature. However, by heating and cooling down the system, the correct propagation direction can be found.

7.3 Magnetic cycloid of BFO under an epitaxial strain

Experiments with BFO indicated that under strain a different type of cycloid, so called type-2 cycloid, could be obtained. This cycloid has a propagation direction along $[110]$, which is thus not perpendicular to the direction along which polarization points. These results are described in an paper [59]. To see if our model reproduces this effect, we conducted simulations in which we applied an epitaxial strain in the $x-y$ plane.

Our first step was getting the state with configuration of magnetic dipole moments corresponding to the type-2 cycloid. This can be achieved in two different ways. The first way consists in changing the coordinate system in the configuration file "by hand". This will give us an artificial way of getting the type-2 cycloid. The second way is by changing the direction of polarization under external field. As we discussed it already, the magnetic state will respond to this change by changing its configuration. After getting the new configuration for magnetic dipoles by this way, we can use it with old configurations for the rest of the degrees of freedom i.e. local modes and antiferrodistrortive vectors. In our study, we got the type-2 cycloid using both ways.

It is difficult to get energy diagram for different magnetic configurations using Monte Carlo method because Monte Carlo methods lead to one final state and to compare the energies of two different states (one of which is, of course, unstable) for the strain, we have to use some tricks. Our tricks are based on the fact that the speed by which magnetic structure riches its equilibrium states

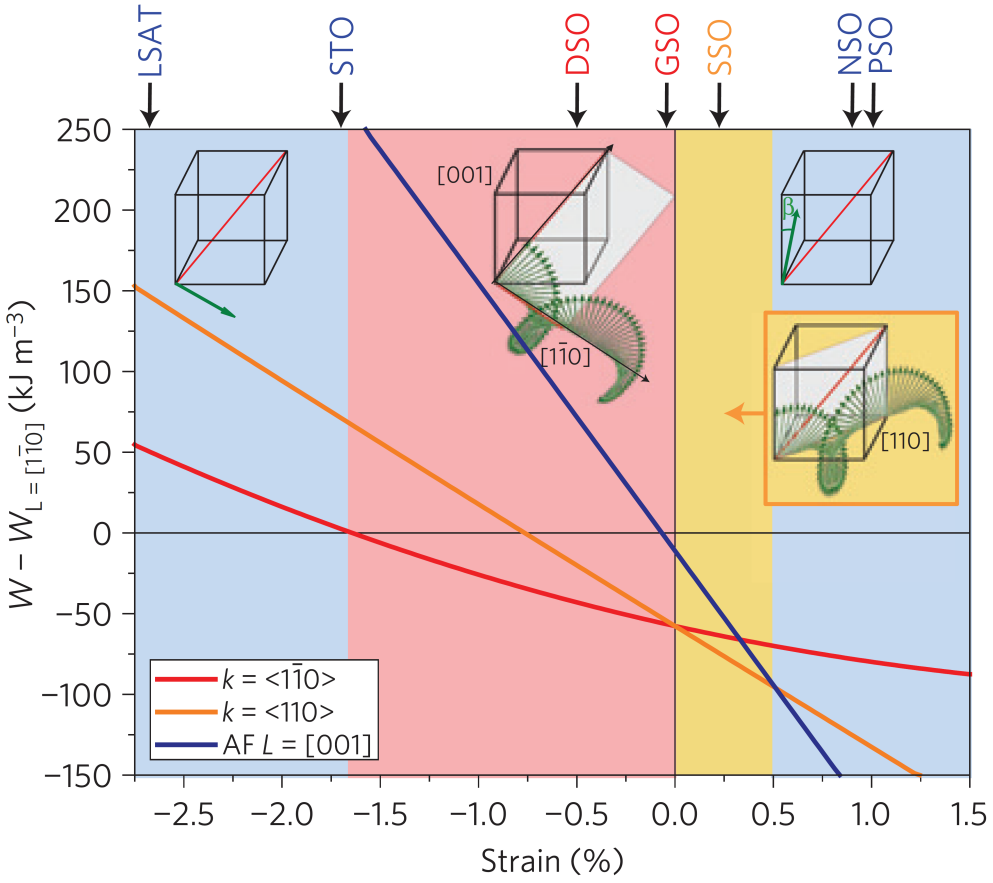

Figure 17: Energy diagram of BFO films under an epitaxial (x,y) strain [59]. To apply such strain experimentally, BFO films can be grown on the following substrates: LSAT – $(LaAlO_3)_{0.3}(Sr_2AlTaO_6)_{0.7}$, STO – $SrTiO_3$, DSO – $DyScO_3$, GSO – $GdScO_3$, SSO – $SmScO_3$, NSO – $NdScO_3$, and PSO – $PrScO_3$. The figure shows three curves corresponding to the energy of the three magnetic states with respect to the antiferromagnetic state with the $\boldsymbol{L}$ vector pointing along the $[1\bar{1}0]$ direction. The curve labeled as $k =< 1\bar{1}0 >$ represents the magnetic structure with type 1 cycloid, i.e. cycloid with propagation direction along the $[1\bar{1}0]$ direction. The curve labeled as $k =< 110 >$ represents the energy of the type 2 cycloid, i.e. cycloid with propagation direction along the $[110]$ direction. The curve labeled as $AF\ L = [001]$ represents the energy of the antiferromagnetic state with the $\boldsymbol{L}$ vector pointing along the $[001]$ direction.

is very slow (more than 100 times) in comparison with other degrees of freedom. Thus we conduct MC simulation with smaller number of steps. These number of steps were just enough for the other degrees of freedom to get their ground state while the magnetic configuration experienced almost no change. This type of computations indeed shows that it is possible to obtain type-2 cycloid under strain. Again we have to point out that, in our model, the value of the C parameter that leads to the cycloid is not computed from first principles and results that we obtain have to be considered more qualitatively than quantitatively.

Figure 17 shows the energy diagram for the BFO films under strain obtained from Landau-Ginzburg theory.[25] The vertical axis in this figure represents the difference between the energy of a given state and the energy of the antiferromagnetic state with $\boldsymbol{L}$ being along the $[1\bar{1}0]$ direction. Since the result from the effective Hamiltonian method is consistent with this diagram (results not shown here), we would like to describe some features of Figure 17. First of all, for large enough value of the strain (compressive or tensile) the state of the system is always antiferromagnetic. However, under tensile strain the antiferromagnetic vector is pointing along the $[001]$ direction (i.e., this state has lower energy than the state for which $\boldsymbol{L}$ points along the $[1\bar{1}0]$ direction). On the other hand, when the strain becomes compressive, an opposite effect is observed, i.e. the $\boldsymbol{L}$ vector is now pointing along the $[1\bar{1}0]$ direction. Furthermore, Figure 17 also reveals that the type-1 cycloid (which is the cycloid present in $BiFeO_3$ bulk) occurs in the epitaxial film, when the strain ranges between $\simeq 0$ and $\simeq -1.5\%$ (compressive strain). This Figure also shows that there is a small region in the energy diagram where the type-2 cycloid is energetically stable. This region is narrow and corresponds to small tensile strain. All of these are in agreement with our computations (not shown here).

[25]The studies based on Landau-Ginzburg theory were done by A. P. Pyatakov and A. K. Zvezdin [59].

Part III

Conclusions and prospects

8 Conclusions

Multiferroics are very complex systems. The existence of many degrees of freedom and their mutual interactions makes these materials out of reach of (or very challenging for) many theoretical methods. In the course of this work, we developed first-principle-based approach and we successfully applied this method to study multiferroic $BiFeO_3$. This framework starts from first-principle based calculations. The results of these calculations are used to construct an effective Hamiltonian. This Hamiltonian constitutes energetic description of the interactions between different degrees of freedom. The effective Hamiltonian is then used in Monte Carlo methods to obtain finite-temperature properties of our system. As a result of application of our method to $BiFeO_3$, we reproduced many experimentally observed properties of BFO systems. Some of those properties are

1. Ferroelectric state with a polarization $\boldsymbol{P}$ pointing along the $[111]$ direction, and the Curie temperature $T_C \simeq 1100$K;
2. Antiferromagnetic state $\boldsymbol{L}$ vector perpendicular to the $\boldsymbol{P}$ and with the Neel temperatures $T_N \simeq 640$K;
3. Tilting of the oxygen octahedra about the $[111]$ direction described by antiferrodistortive vector $\boldsymbol{\omega}$;
4. Week ferromagnetism with the magnetization vector $\boldsymbol{M}$ perpendicular to both $\boldsymbol{L}$ and $\boldsymbol{P}$ vectors;
5. Magnetic cycloid with a correct propagation direction and cycloidal plane;
6. Anharmonicity of the magnetic cycloid;
7. Out-of-plane spin-density waves;

8. Posibilities of type-2 cycloid under epitaxial strain.

Magnetic energy of the BFO can be divided into three types, each of which have different physics behind them. The first type tries to make BFO purely antiferromagnetic, the second is responsible for its weak ferromagnetism and the third gives to this material its cycloidal state.

As a result of our investigations for the microscopic origin of the weak ferromagnetism in BFO, we concluded that the interaction between the tilting of the oxygen octahedra and magnetic moments are responsible for this magnetization. The importance of the weak ferromagnetism for the linear magnetoelectric coefficient to occur was also discovered. In addition to that, we computed the evolution of the $\boldsymbol{L}$ and $\boldsymbol{M}$ vectors with temperature and showed that both of them disappear at the same (Neel) temperature.

The response of the magnetic cycloid to the flipping of the direction of electric polarisation was also studied. These results indicate that if the propagation vector $\boldsymbol{k}$ of the cycloid is perpendicular to the new direction of the polarization $\boldsymbol{P}$, then the cycloidal plane will respond by rotating, in order to conserve its perpendicularity to both of those $\boldsymbol{k}$ and $\boldsymbol{P}$ vectors. In addition to that, because of the coupling between chirality of the magnetic cycloid and polarization, flipping of the polarization vector $\boldsymbol{P}$ results in rotation of the cycloidal plane in some unexpected ways.

Our simulation of the magnetic cycloid of BFO shows that under high enough values of magnetic field, electric field, and epitaxial strain, the cycloid is always destroyed in favor of the antiferromagnetic state with weak ferromagnetism.

In addition to the simulations, in the course of this work, we also conducted analytic developments to describe the behavior of magnetic susceptibility of antiferromagnets and how this behavior is modified in multiferroic antiferromagnets. The analytical results obtained by this way were compared with the simulations. As a result of this comparison, we confirmed the consistency between two methods and better understand novel phenomena related to magnetic susceptibility (e.g., deviation from Curie-Weiss behavior).

The value of the method developed in this work is not limited in its predictive power. For example, the knowledge about the energy term that is responsible for particular phenomena clari-

fies the physics behind that phenomena. Such knowledge also gives ideas about how to suppress or/and amplify certain effects. In addition to that, future experiments can be based on the energetic insights obtained as a result of our method.

9 Future prospects

When it comes to the future research prospects about materials like BFO, the list of possible studies would be very long. The physics of BFO is very rich and many important questions about this material still remain unanswered. However, in this section, we will limit ourselves only to the effects that can be studied with the tools that we have already developed during our study and described in previous pages.

Magnetic susceptibility

We started the result section of this dissertation by developing theory for the magnetic susceptibility for the multiferroic BFO. One can extend this study of the magnetic susceptibilty in many ways. For example, in our system, we did not have magnetic anisotropy which limited our study only to the perpendicular component of the magnetic susceptibility. Magnetic anisotropy can be studied by modifying magnetic energy of the system, which would lead to such anisotropy. Such anisotropy would also allow us to compute the parallel component of the magnetic susceptibility. The next possible step could be an attempt for obtaining magnetic susceptibility from fluctuations of the system (e.g., using correlation functions involving magnetization).

Some toy models to consider

The energy terms of BFO can be easily modified to give us ferromagnetic multiferroic. This would involve changing signs of some energy terms. By changing values of certain coefficients, the magnetic and ferroelectric transition temperature can be modified. This can give us a system with a ferrolectric transition temperature smaller than the magnetic transition temperature. The classification of all of these models and exploring phenomena associated with each of them could

be interesting theoretical projects. Exploring such toy models can result in very interesting effects. For example, it is possible to make a system with an antiferromagnetic structure that transforms into a ferromagnetic structure under the external electric field.

Generalization of the cycloidal energy

In reproducing the magnetic cycloid of BFO bulk, we were guided by simplicity. In other words, we included an energy term that is as simple as possible and that would reproduce experimentally observed features of the cycloid. There are many possible avenues for generalizing this energy term and those generalizations can lead to some interesting new effects. For example, one can add the first nearest neighbors into picture and see their impact to the cycloidal state.

First-principle value of the C

As we made it clear in the manuscript, the C parameter of the cycloidal energy was not obtained from first-principle calculations. Although we made some estimations about the value of C, the possibility of calculating this energy parameter using quantum mechanical methods still remains an open question.

Strain on different planes

In our study, strain was applied only on a particular plane (namely in a (001) plane) and it was not studied using Monte Carlo method. Thus, it would be interesting to explore magnetic cycloids of BFO under strain applied on different planes using the Monte Carlo method. Depending on the plane and value of the strain, the magnetic cycloid of BFO can respond in unexpected ways to such external influences.

Magnetic susceptibility of the cycloidal state

As we did with the antiferromagnetic state of BFO, one can develop theory for the magnetic susceptibility of the cycloidal state of BFO. In the case of the antiferromagnetic state, we got interesting relations between different degrees of freedom and susceptibility. It would be interesting to get

analogous expression for the cycloidal state.

Magnetoelectic coefficients

Finally, it is time to introduce and study magnetoelectic coefficients of BFO. These coefficients are very small in magnitude and it is challenging to obtain them experimentally. However, we still can study them theoretically and obtain some interesting relations between them and other properties of BFO. In particular, the temperature dependence of those coefficients could be studied and the relation between these coefficients and other properties of BFO like magnetic susceptibility, electric polarization, and antiferromagnetic vector can be explored.

Dynamics of cycloid

Monte Carlo methods are limited to the study of the equilibrium states and it can not be used when it comes to *dynamics* of the system. For example, the Monte Carlo method can show that the plane of the magnetic cycloid rotates when applying specific electric fields but it can not tell us the exact path of rotation and the switching time. Therefore, our scheme can be improved if we couple it with molecular dynamics method. However, since Monte Carlo method is computationally much more efficient in comparison with molecular dynamics method, it can not be replaced by these methods completely, i.e. Monte-Carlo method should still be efficiently used for the study of specific static-like properties.

References

[1] G.A. Smolenskii and I.E. Chupis. Sov. Phys. Usp. 25, 475 (1982).

[2] W. Eerenstein, N.D. Mathur, and J.F. Scott. Nature (London) 442, 759 (2006).

[3] N.A. Spaldin and M. Fiebig. Science 309, 391 (2005).

[4] J. B. Neaton, C. Ederer, U. V. Waghmare, N. A. Spaldin, and K. M. Rabe. Phys. Rev. B 71, 014113 (2005).

[5] T. Zhao, A. Scholl, F. Zavaliche, K. Lee, M. Barry, A. Doran, M. P. Cruz, Y. H. Chu, C. Ederer, N. A. Spaldin, R. R. Das, D. M. Kim, S. H. Baek, C. B. Eom and R. Ramesh. Nature Mater. 5, 823 (2006).

[6] D. Lebeugle, D. Colson, A. Forget, M. Viret, A. M. Bataille, and A. Gukasov. Phys. Rev. Lett. 100, 227602 (2008).

[7] I. M. Gelfand, *Lectures on Linear Algebra.* Dover Publications (September 1, 1989).

[8] Seongsu Lee, Taekjib Choi, W. Ratcliff, II, R. Erwin, S-W. Cheong, and V. Kiryukhin. Phys. Rev. Lett. 100, 227602 (2008).

[9] H. Bea, M. Bibes, A. Barthelemy, K. Bouzehouane, E. Jacquet, A. Khodan, J.-P. Contour, S. Fusil, F. Wyczisk, A. Forget, D. Lebeugle, D. Colson, and M. Viret. Appl. Phys. Lett. 87, 072508 (2005).

[10] H. Bea, M. Bibes, S. Petit, J. Kreisel, and A. Barthelemy. Philos. Mag. Lett. 87, 165 (2007).

[11] R. Mazumder, P. Sujatha Devi, Dipten Bhattacharya, P. Choudhury, A. Sen, and M. Raja. Appl. Phys. Lett. 91, 062510 (2007).

[12] R. Haumont, J. Kreisel, P. Bouvier, and F. Hippert. Phys. Rev. B 73, 132101 (2006).

[13] C. Ederer and N.A. Spaldin. Phys. Rev. B 71, 060401(R) (2005).

[14] Igor A. Kornev, S. Lisenkov, R. Haumont, B. Dkhil, and L. Bellaiche. Phys. Rev. Lett. 99, 227602 (2007).

[15] S. Lisenkov, I.A. Kornev, and L. Bellaiche. Phys. Rev. B 79, 012101 (2009).

[16] I. Sosnowska, T.P. Neumaier, and E. Steichele. J. Phys. C 15, 4835 (1982).

[17] A. Garcia and D. Vanderbilt. Appl. Phys. Lett. 72, 2981 (1998).

[18] L. Bellaiche, A. Garcia, and D. Vanderbilt. Phys. Rev. B 64, 060103(R) (2001).

[19] Igor A. Kornev, L. Bellaiche, P.-E. Janolin, B. Dkhil, and E. Suard. Phys. Rev. Lett. 97, 157601 (2006).

[20] V. Ranjan, S. Bin-Omran, and L. Bellaiche. Phys. Rev. B 71, 195302 (2005).

[21] D. Vanderbilt and W. Zhong. Ferroelectrics 206, 181 (1998).

[22] B. D. Cullity, and C. D. Graham. *Introduction to magnetic materials (Second Edition).* Published by John Wiley & Sons, Inc. 2009.

[23] L. D. Landau, and E.M. Lifshitz *Quantum Mechanics Non-Relativistic Theory, Third Edition: Volume 3* Published by Butterworth-Heinemann 1981.

[24] L. D. Landau, and E.M. Lifshitz *Electrodynamics of Continuous Media, Second Edition: Volume 8* Published by Butterworth-Heinemann 1984.

[25] J. Wang, J. B. Neaton, H. Zheng, V. Nagarajan, S. B. Ogale, B. Liu, D. Viehland, V. Vaithyanathan, D. G. Schlom, U. V. Waghmare, N. A. Spaldin, K. M. Rabe, M. Wuttig, R. Ramesh, *Science* **299**, 1719 (2003).

[26] S. Lee et al.,*Phys. Rev. B* **78**, 100101(R) (2008).

[27] S. Lisenkov, D. Rahmedov, and L. Bellaiche, *Phys. Rev. Lett.* **103**, 047204 (2009)

[28] Ira N. Levine *Quantum Chemistry.* Prentice Hall 1991.

[29] Parr, R. G. Yang, W. *Density-Functional Theory of Atoms and Molecules.* Oxford University Press 1989.

[30] D. Rahmedov, Dawei Wang, Jorge Iniguez and L. Bellaiche, PRL 109, 037207 (2012).

[31] D. Albrecht, S. Lisenkov, W. Ren, D. Rahmedov, I.A. Kornev, and L. Bellaiche, Phys. Rev. B 81, 140401(R) (2010).

[32] I. Sosnowska, M. Loewenhaupt, W. I. F. David, and R. M. Ibberson, Physica (Amsterdam) 180B181B, 117 (1992).

[33] A.V. Zalesskii, A. A. Frolov, A. K. Zvezdin, A. A. Gippius, E. N. Morozova, D. F. Khozeevc, A. S. Bush, and V. S. Pokatilov, Sov. Phys. JETP 95, 101 (2002).

[34] A.V. Zalesskii, A. K. Zvezdin, A. A. Frolov, and A. A. Bush, JETP Lett. 71, 465 (2000).

[35] A. A. Bush, A. A. Gippius, A.V. Zalesskii, and E. N. Morozova, JETP Lett. 78, 389 (2003).

[36] M. Ramazanoglu, W. Ratcliff II, Y. J. Choi, S. Lee, S.-W. Cheong, and V. Kiryukhin, Phys. Rev. B 83, 174434 (2011).

[37] J. B. Neaton, C. Ederer, U. V. Waghmare, N. A. Spaldin, and K. M. Rabe, Phys. Rev. B 71, 014113 (2005).

[38] I. Sosnowska and R. Przenioslo, Phys. Rev. B 84, 144404 (2011).

[39] R. Przenioslo, A. Palewicz, M. Regulski, I. Sosnowska, R. M. Ibberson, and K. S. Knight, J. Phys. Condens. Matter 18, 2069 (2006).

[40] V. S. Pokatilov and A. S. Sigov, J. Exp. Theor. Phys. 110, 440 (2010).

[41] M. Ramazanoglu, M. Laver, W. Ratcliff II, S. M. Watson, W. C. Chen, A. Jackson, K. Kothapalli, S. Lee, S.-W. Cheong, and V. Kiryukhin, Phys. Rev. Lett. 107, 207206 (2011).

[42] I. Sosnowska and A. L. Zvezdin, J. Magn. Magn. Mater. 140144, 167 (1995).

[43] R. de Sousa and J. E. Moore, Appl. Phys. Lett. 92, 022514 (2008).

[44] J. Jeong et al., Phys. Rev. Lett. 108, 077202 (2012).

[45] D.Wardecki, R. Przenioso, I. Sosnowska, Y. Skourski, and M. Loewenhaupt, J. Phys. Soc. Jpn. 77, 103709 (2008).

[46] W. Zhong, D. Vanderbilt, and K. M. Rabe, Phys. Rev. Lett. 73, 1861 (1994); Phys. Rev. B 52, 6301 (1995).

[47] I. Kornev, L. Bellaiche, P.-E. Janolin, B. Dkhil, and E. Suard, Phys. Rev. Lett. 97, 157601 (2006).

[48] P. Fischer, M. Polomska, I. Sosnowska, and M. Szymanski, J. Phys. C 13, 1931 (1980).

[49] S. Lisenkov, I.A. Kornev, and L. Bellaiche, Phys. Rev. B 79, 012101 (2009); Phys. Rev. B 79, 219902(E) (2009).

[50] S.V. Kiselev, R. P. Ozerov, and G. S. Zhdanov, Sov. Phys. Dokl. 7, 742 (1963); G. A. Smolenskii, V. Isupov, A. Agranovskaya, and N. Krainik, Sov. Phys. Solid State 2, 2651 (1961).

[51] J. R. Teague, R. Gerson, and W. J. James, Solid State Commun. 8, 1073 (1970).

[52] A. M. George, J. Iniguez, and L. Bellaiche, Phys. Rev. B 65, 180301(R) (2002).

[53] H. Katsura, N. Nagaosa, and A.V. Balatsky, Phys. Rev. Lett. 95, 057205 (2005).

[54] A. M. Kadomtseva, A.K. Zvezdin, Y. F. Popov, A. P. Pyatakov, and G. P. Vorobev, JETP Lett. 79, 571 (2004).

[55] G. Catalan and J. F. Scott, Adv. Mater. 21, 2463 (2009).

[56] D. Talbayev, S. A. Trugman, S. Lee, H. T. Yi, S.-W. Cheong, and A. J. Taylor, Phys. Rev. B 83, 094403 (2011).

[57] M. Cazayous, Y. Gallais, A. Sacuto, R. de Sousa, D. Lebeugle, and D. Colson, Phys. Rev. Lett. 101, 037601 (2008).

[58] P. Rovillain, R. de Sousa, Y. Gallais, A. Sacuto, M. A. Measson, D. Colson, A. Forget, M. Bibes, A. Barthelemy, and M. Cazayous, Nature Mater. 9, 975 (2010).

[59] D. Sando, A. Agbelele, D. Rahmedov, J. Liu, P. Rovillain, C. Toulouse, I. C. Infante, A. P. Pyatakov, S. Fusil, E. Jacquet, C. Carretero, C. Deranlot, S. Lisenkov, D. Wang, J-M. Le Breton, M. Cazayous, A. Sacuto, J. Juraszek, A. K. Zvezdin, L. Bellaiche, B. Dkhil, A. Barthelemy & M. Bibes. Nature Materials 12, 641646 (2013).

Appendices

In the appendix A I will give copyright information from two publishers: Nature Publishing Group (NPG) and American Physical Society (APS). In the appendix B I will present an excerpt from the article (not yet published) that gives an alternative derivation for the relations that were described in the section 6.1.

A Copyright information

A.1 Copyright policies of the Nature Publishing Group (NPG)

As an author I contacted the NPG to get the permission for using the figure 17 from the article:
Title: Crafting the magnonic and spintronic response of $BiFeO_3$ films by epitaxial strain
Author: D. Sando, A. Agbelele, D. Rahmedov, J. Liu, P. Rovillain, C. Toulouse, I. C. Infante, A. P. Pyatakov, S. Fusil, E. Jacquet, C. Carretero, C. Deranlot, S. Lisenkov, D. Wang, J-M. Le Breton, M. Cazayous, A. Sacuto, J. Juraszek, A. K. Zvezdin, L. Bellaiche, B. Dkhil, A. Barthlmy, M. Bibes
Publication: Nature Materials
Publisher: Nature Publishing Group
Date: Apr 28, 2013

I was granted the permission and below I am including their copyright policies for authors:

Title:	Crafting the magnonic and spintronic response of $BiFeO_3$ films by epitaxial strain
Author:	D. Sando, A. Agbelele, D. Rahmedov, J. Liu, P. Rovillain, C. Toulouse, I. C. Infante, A. P. Pyatakov, S. Fusil, E. Jacquet, C. Carrétéro, C. Deranlot, S. Lisenkov, D. Wang, J-M. Le Breton, M. Cazayous, A. Sacuto, J. Juraszek, A. K. Zvezdin, L. Bellaiche, B. Dkhil, A. Barthélémy, M. Bibes
Publication:	Nature Materials
Publisher:	Nature Publishing Group
Date:	Apr 28, 2013

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select an alternative Requestor Type to obtain a quick price or to place an order.

Ownership of copyright in the article remains with the Authors, and provided that, when reproducing the Contribution or extracts from it, the Authors acknowledge first and reference publication in the Journal, the Authors retain the following non-exclusive rights:

a) To reproduce the Contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).

b) They and any academic institution where they work at the time may reproduce the Contribution for the purpose of course teaching.

c) To reuse figures or tables created by them and contained in the Contribution in other works created by them.

d) To post a copy of the Contribution as accepted for publication after peer review (in Word or Text format) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the Journal article on NPG's web site (eg through the DOI).

NPG encourages the self-archiving of the accepted version of your manuscript in your funding agency's or institution's repository, six months after publication. This policy complements the recently announced policies of the US National Institutes of Health, Wellcome Trust and other research funding bodies around the world. NPG recognises the efforts of funding bodies to increase access to the research they fund, and we strongly encourage authors to participate in such efforts.

Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read NPG's online author reuse guidelines.

For full paper portion: Authors of original research papers published by NPG are encouraged to submit the author's version of the accepted, peer-reviewed manuscript to their relevant funding

body's archive, for release six months after publication. In addition, authors are encouraged to archive their version of the manuscript in their institution's repositories (as well as their personal Web sites), also six months after original publication.

v2.0

BACK CLOSE WINDOW

A.2 Copyright policies of the American Physical Society (APS)

As an author I used the figures from the articles that I published in the journals of the American Physical Society (APS) and below I am including the copyright policies of the American Physical Society (APS) as shown in their website relevant to for authors.

Does transferring copyright affect my patent rights?
No. Copyright is separate from any patent rights, and the APS transfer agreement specifically states that patent rights are not affected. However, you should be aware that submitting a manuscript to a journal without first taking steps to protect your patent rights (e.g., filing for a patent) could endanger those rights. Consult your patent attorney.

As the author of an APS-published article, may I post my article or a portion of my article on my own website?
Yes, the author or the author's employer may use all or part of the APS published article, including the APS-prepared version (e.g., the PDF from the online journal) without revision or modification, on the author's or employer's website as long as a fee is not charged. If a fee is charged, then APS permission must be sought. In all cases, the appropriate bibliographic citation and notice of the APS copyright must be included.

What happens if the author has posted an APS-published article on a free access e-print server or on the authors' or institutions' web page and subsequently a fee is imposed for access to those sites?
When a fee is imposed, the author must either obtain permission from APS or withdraw the article from the e-print server or Institutional Repository.

As the author of an APS-published article, may I post my article or a portion of my article on an e-print server?
The author has the right to post and update the article on a free-access e-print server using files prepared and formatted by the author. Any such posting made or updated after acceptance of the article for publication by APS shall include a link to the online abstract in the APS journal or to the entry page of the journal. In all cases, the appropriate bibliographic citation and notice of the APS copyright must be included. If the author wishes to use the APS-prepared version (e.g., the PDF from the online journal) on an e-print server other than authors' or employer's website, then APS permission must be sought. Similarly, if the author wishes to post the article (any version) on an e-print server that charges a fee for use, APS permission must be sought.

As the author of an APS-published article, can I post my article or a portion of my article on a web resource like wikipedia or quantiki?
Sites like wikipedia and quantiki are strict about permissions and require that authors hold copyright to articles that they post there. In order to allow authors to comply with this requirement, APS permits authors to hold copyright to a "derived work" based on an article published in an APS journal as long as the work contains at least 10% new material not covered by APS's copyright and does not contain more than 50% of the text (including equations) of the original article.

As the author of an APS-published article, will I hold copyright to a "derived work", as described above, even if the original article was published prior to 1 October 2008?
Yes. The APS will extend this author right to all papers published in APS journals.

As the author (or the author's employer) of an APS-published article, may I use copies of part or all of my article in the classroom?
Yes, the author or his/her employer may use all or part of the APS-prepared version for educational purposes without requesting permission from the APS as long as the appropriate bibliographic citation is included.

As the author of an APS-published article, may I use figures, tables, graphs, etc. in future publications?
Yes, as the author you have the right to use figures, tables, graphs, etc. in subsequent publications using files prepared and formatted by you or the APS-prepared versions. The appropriate bibliographic citation must be included.

As the author of an APS-published article, may I include my article or a portion of my article in my thesis or dissertation?
Yes, the author has the right to use the article or a portion of the article in a thesis or dissertation without requesting permission from APS, provided the bibliographic citation and the APS copyright credit line are given on the appropriate pages.

As the author of an APS-published article, may I give permission to a colleague or third party to republish all or part of the article in a print publication?
Yes, as the author you can grant permission to third parties to republish print versions of the article provided the APS-prepared version (e.g., the PDF from the online journal, or a copy of the article from the print journal) is not used for this purpose, the article is not published in another journal, and the third party does not charge a fee. The appropriate bibliographic citation and notice of the APS copyright must be included.

As the author of an APS-published article, may I give permission to a colleague or third party to republish all or part of the article in an online journal, book, database compilation, etc.?
Authors should direct the third party request to APS.

As the author of an APS-published article, may I provide a PDF of my paper to a colleague or third party?
The author is permitted to provide, for research purposes and as long as a fee is not charged, a PDF copy of his/her article using either the APS-prepared version or the author prepared version.

As a third party (not an author), may I republish an article or portion of an article published by APS?
Yes, APS will grant permission to republish articles or portions of articles (e.g., tables, graphs, excerpts) published by APS. Depending on the reuse and medium APS has the right to grant permission subject to APS terms and conditions and a fee may be assessed.

As a third party, may I use articles published by APS for lecture and classroom purposes?
Yes, you may use photocopied articles published by APS for lecture and classroom purposes for a single semester without asking permission from APS. However, if the article becomes part of your course material beyond one semester, you must obtain permission from APS. Also, there is no limitation on the use of APS articles using links to the material accessible through institutional subscriptions.

How do I request permission to republish APS-copyrighted material?
To request permission to republish APS-copyrighted material, please provide the following information:

1. Title of journal
2. Title of article
3. Name of author

4. Volume number, page number (or article identifier), year
5. Indicate if you are requesting to republish in print, online, CD-ROM, and/or other format
6. Indicate if you wish to republish all or portion of article; if a portion describe the specific material, e.g., figure numbers, excerpt
7. Indicate how the material will be used, e.g., in a book, journal, proceeding, thesis, etc.
8. Indicate the title of the article/thesis/chapter etc., and the name of the publication in which your work will appear
9. Indicate the name of the publisher
10. Indicate whether or not a fee will be charged for the publication

*To prevent clerical error, please include all requests in a single email or letter.

All permission requests must be in writing (email is acceptable). Blanket permissions are not granted. Please note all requests are subject to APS terms and conditions and a fee may be assessed.

Please allow 5-7 business days for us to respond to a permission request provided all the above information is provided at the time of the request.

Send all permission requests to:

Associate Publisher
American Physical Society
One Physics Ellipse
College Park, MD 20740
Email: assocpub@aps.org

If your questions have not been addressed and you need further assistance, please call: 301-209-3283.

How do I provide a proper bibliographic citation and notice of the APS copyright?
Provide the following information in this order:

Authors names, journal title, volume number, page number (or article identifier), year of publication. "Copyright (year) by the American Physical Society."

Further information
For further information about copyright in general, please refer to the Library of Congress FAQ at: http://www.copyright.gov/help/faq/

Journals published by the American Physical Society can be found at http://publish.aps.org/

FAQ Version: October 1, 2008

B Magnetoelectric signature in the magnetic properties of antiferromagnetic multiferroics: atomistic simulations and phenomenology

D. Rahmedov, S. Prosandeev, J.Iniguez, L. Bellaiche. PRB 88, 224405 (2013) "Copyright (2014) by the American Physical Society."

B.1 Numerical method

Our "numerical toy model" originates from the effective Hamiltonian of $BiFeO_3$ (BFO) developed in Refs. [2, 3]. More precisely, the total energy, E_{tot}, is written as:

$$E_{tot} = E_{struct}\left(\{\mathbf{u}_i\},\{\eta_i\},\{\omega_i\}\right) + E_{Mag}(\{\mathbf{m}_i\},\{\mathbf{u}_i\}) \quad , \tag{B.1.1}$$

where the first term, E_{struct}, gathers energies associated with structural degrees of freedom and their mutual interactions. These structural degrees of freedom are (i) the $\mathbf{u}_i$ local soft mode in unit cell i, which is directly proportional to the electrical dipole centered on that site; (ii) the η_i strain in unit cell i, that contains homogeneous and inhomogeneous parts [9]; and (iii) the $\{\omega_i\}$ vector that characterizes the oxygen octahedral tilting in unit cell i. The analytic form of E_{struct} is given in Ref. [10]. The second term of Eq. (B.1.1), E_{Mag}, gathers the interactions between the $\mathbf{m}_i$ magnetic moments of Fe ions at different cells i, *and* the interactions between these magnetic moments and the local soft modes. The analytical expression of E_{Mag} is the one provided in Ref. [3] with the exception that the interactions between magnetic moments and both the strains and oxygen octahedral tiltings are neglected here (which slightly simplifies our atomistic tool with respect to the real effective Hamiltnian of BFO, in order to determine the sole effects of ME couplings on physical properties). E_{Mag} thus reads:

$$E_{\mathrm{Mag}}(\{\mathbf{m}_i\},\{\mathbf{u}_i\}) = \sum_{i,j,\alpha,\gamma} D_{ij,\alpha\gamma} m_{i,\alpha} m_{j,\gamma} + \sum_{i,j,\alpha,\gamma,\nu,\delta} E_{ij,\alpha\gamma\nu\delta} m_{i,\alpha} m_{j,\gamma} u_{i,\nu} u_{i,\delta} \tag{B.1.2}$$

where α, γ, ν and δ denote Cartesian components. The sums over i run over all the Fe-sites while the sums over j run over the first, second and third nearest neighbors of the Fe-site i. The first term of E_{Mag} represents the exchange interactions between magnetic moments at sites i and j. The second term represents how the local soft-mode affect these magnetic exchange interactions, which is precisely the (bi-quadratic) ME coupling that we are going to study here. Note that this latter coupling is allowed by symmetry in *any* multiferroic material, which implies that our results should be of general nature [11]. All the parameters entering the analytical expression of E_{tot} (including the D_{ij} coefficients) are extracted from first-principles calculations on $BiFeO_3$ [12], with the sole exception of the E_{ij} parameters that are allowed to vary from their first-principles values – in order to precisely assert the effect of the ME coupling on electric and magnetic properties.

Monte-Carlo (MC) simulations using E_{tot} for the internal energy are performed on a $14 \times 14 \times 14$ supercell, with the local magnetic moments $\{\mathbf{m}_i\}$ having a fixed magnitude of $4\mu_B$ (as consistent with first-principles computations of BFO [1]). Practically, the system is cooled down to very low temperature (of the order of the Kelvin) under an electric field that is applied along the pseudo-cubic [111] direction, in order to reach a R3c ground state. Then, this field is removed and the temperature is increased until reaching 1300 K. The numerical results shown here corresponds to this increase-in-temperature path under no field.

B.2 Results

B.2.1 Atomistic Simulations

In the present study, three different physical quantities are determined : (1) the G-type antiferromagnetic vector, which is defined as $\mathbf{L} = \frac{1}{N}\sum_i \mathbf{m}_i(-1)^{n_x(i)+n_y(i)+n_z(i)}$ where the sum runs over all the Fe sites and where $n_x(i)$, $n_y(i)$ and $n_z(i)$ are integers locating the cell i (more precisely, the Fe-site i is centered at $[n_x(i)\mathbf{x} + n_y(i)\mathbf{y} + n_z(i)\mathbf{z}]a$, where a is the 5-atom lattice constant and where $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ are unit vectors along the Cartesian axes); (2) the $\mathbf{u}$ supercell average of the $\{\mathbf{u}_i\}$ local soft modes, which is directly proportional to the spontaneous polarization; and (3) the perpendicular component of the magnetic susceptibility, $\chi_{M,\perp}$, which is computed as the linear slope of the func-

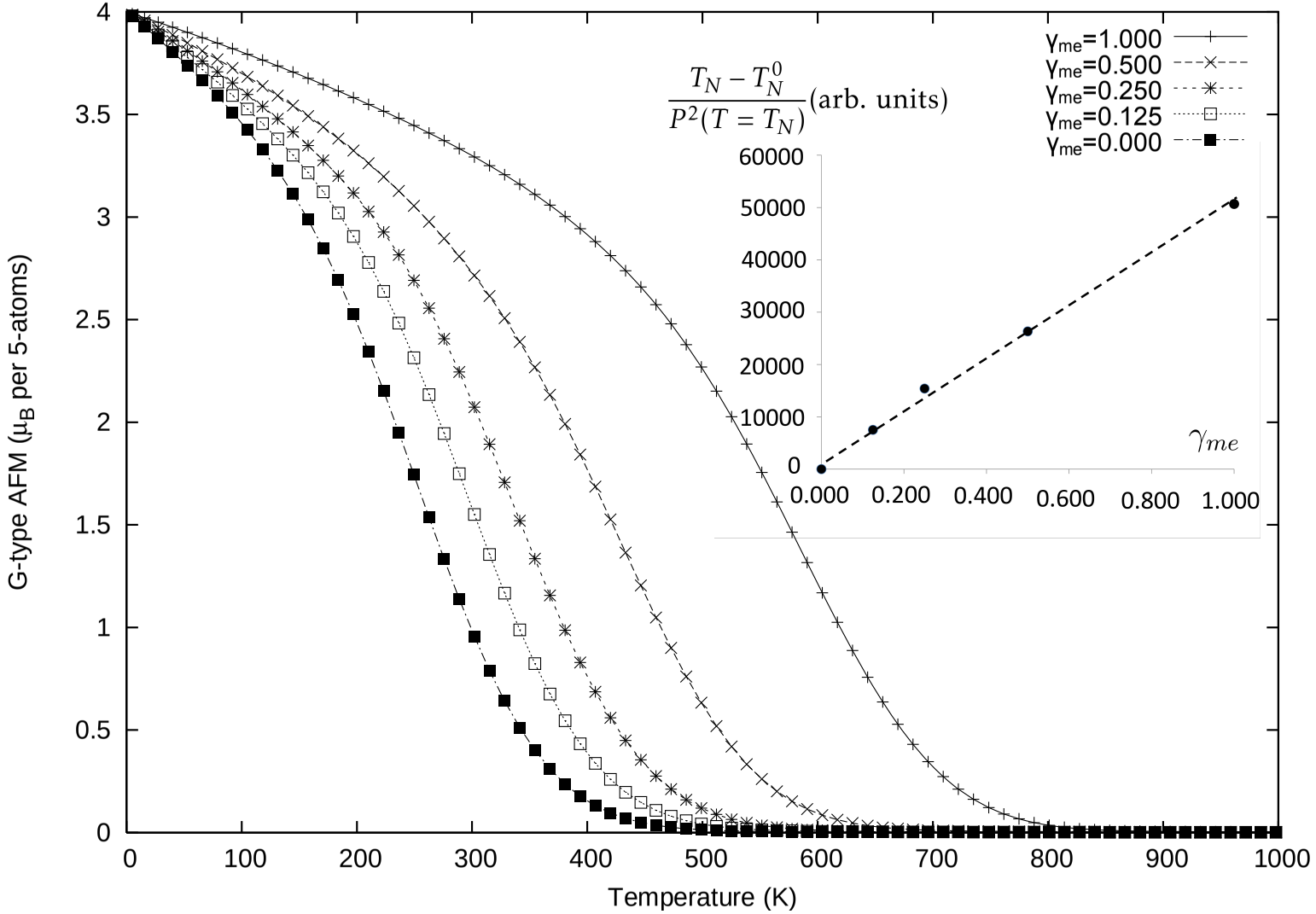

Figure 18: Temperature dependency of the magnitude of the G-type antiferromagnetic vector, for five different selected sets of the ME-related E_{ij} parameters of Eq. (B.1.2). The numerical data are shown by symbols while the lines are guide for the eyes. The inset shows $\frac{T_N - T_N^0}{P^2(T=T_N)}$ as a function of γ_{me}, where $P^2(T = T_N)$ is the square of the polarization at the Neel temperature and where T_N^0 is the "bare" Neel temperature (i.e., in the case of no ME coupling, that is corresponding to γ_{me}=0). This inset confirms the validity of the phenomenological Equation (B.2.7).

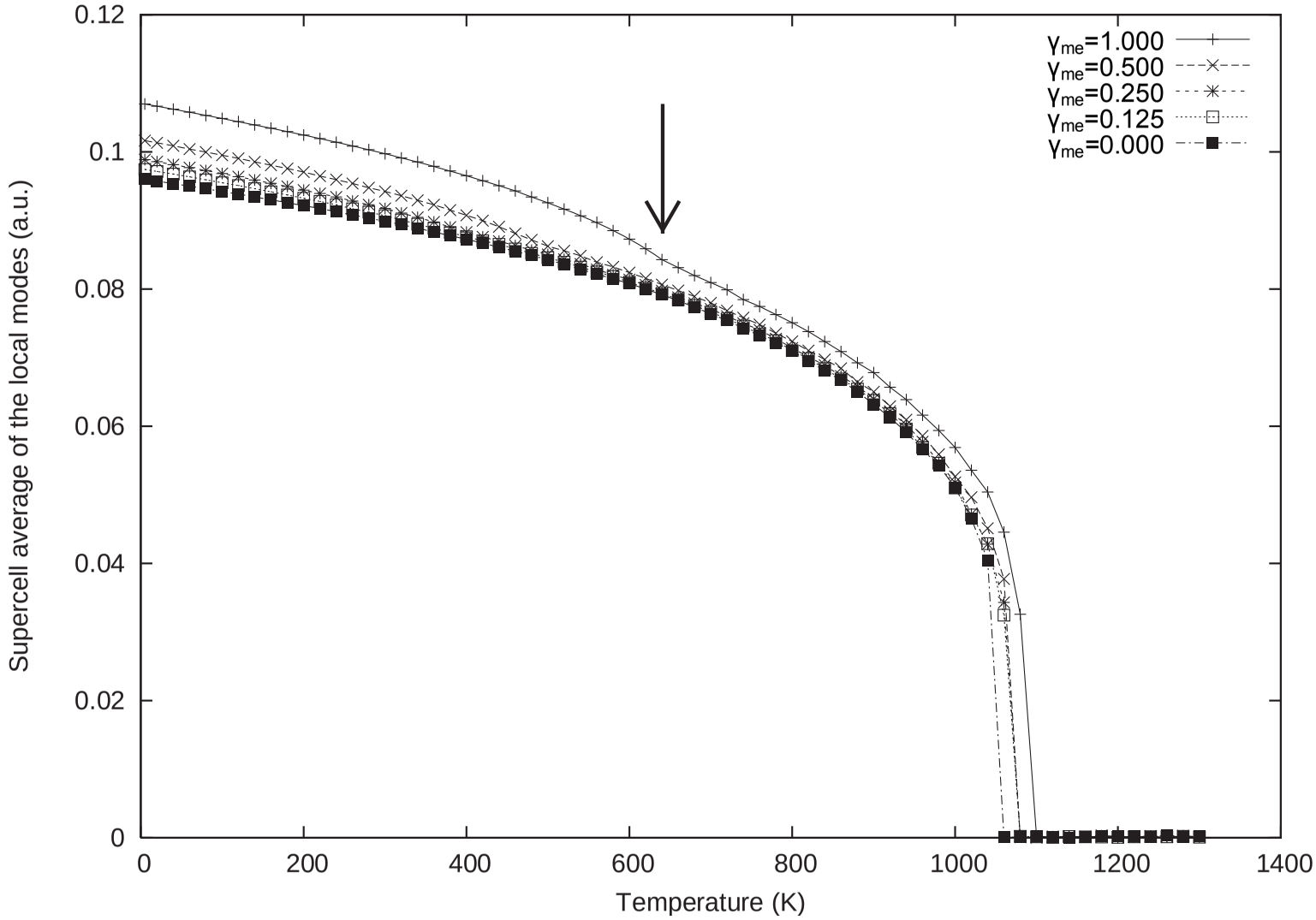

Figure 19: Same as Figure 18, but for the magnitude of the supercell average of the local modes. The arrow shows the location of the Neel temperature when $E_{ij} = E_{ij,ref}$ (or, equivalently, γ_{me}=1).

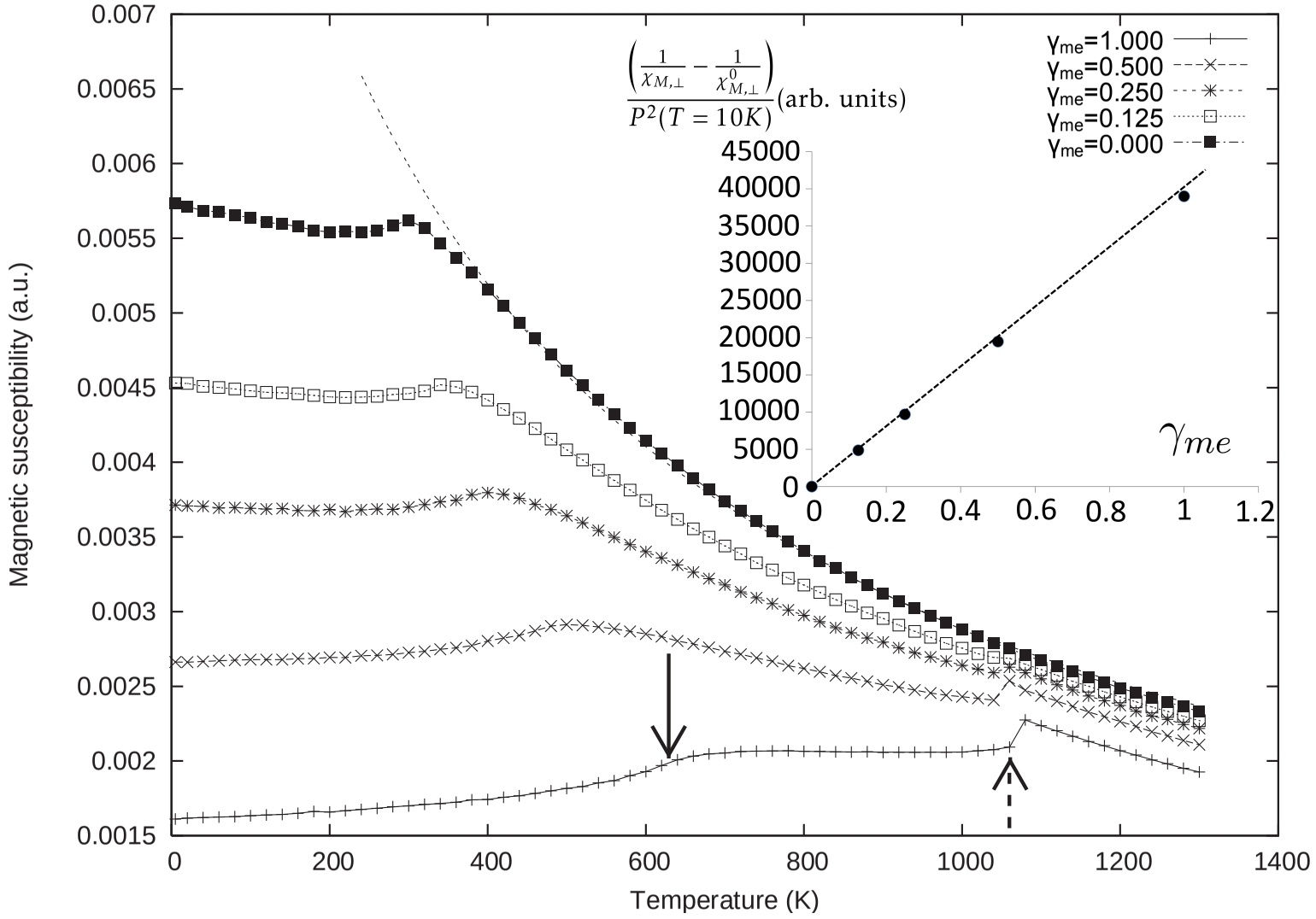

Figure 20: Same as Figures 18 and 19, but for the perpendicular component of the magnetic susceptibility. The dashed line shows the fit of the magnetic susceptibility by $\frac{b_2}{T+T_N^0}$, in case of *no* ME coupling. The solid and dashed arrows depict the values of T_N and T_C, respectively, when $E_{ij} = E_{ij,ref}$ (or, equivalently, γ_{me}=1). The inset displays $(\frac{1}{\chi_{M,\perp}} - \frac{1}{\chi^0_{M,\perp}})/P^2(T = 10K)$ as a function of γ_{me} (see text), where $P^2(T = 10K)$ is the square of the polarization at 10K. This inset demonstrates that the phenomenological Equation (B.2.13) is obeyed.

tion representing the dependence of the magnetization (which is simply the supercell average of the $\{\mathbf{m}_i\}$'s) on an applied magnetic field. This latter field is oriented along the $[111]$ pseudo-cubic direction, that is along the polarization direction while being perpendicular to the G-type AFM vector of BFO [3]. The applied magnetic field is allowed to have a magnitude ranging between zero and 100 Tesla, in order to precisely compute $\chi_{M,\perp}$. Figures 18, 19 and 20 show the magnitude of $\mathbf{L}$, the magnitude of $\mathbf{u}$ and $\chi_{M,\perp}$, respectively, as a function of temperature for five different sets of E_{ij} coefficients. If we denote the first-principles values of the E_{ij} coefficients in BFO as $E_{ij,ref}$, then these five sets correspond to E_{ij} being equal to $\gamma_{me}E_{ij,ref}$, with γ_{me} =1, 0.5, 0.25, 0.125 and 0.0, respectively (as characteristic of progressively weaker magneto-electric couplings until they fully vanish).

Figure 18 reveals that the Neel temperature, T_N, (which is taken as the temperature at which the magnitude of $\mathbf{L}$ possesses an inflection point) strongly increases when the magneto-electric coefficients grow in strength: for $\gamma_{me} = 0$, it is about 275K (this temperature will be denoted as T_N^0 in the following) while it significantly increases up to 635K when γ_{me}=1. Another effect that is visible in Fig. 18 is the enhancement of the AFM vector resulting from the increase of γ_{me} for any temperature below T_N^0= 275K (at the sole exception of 0K for which the antiferromagnetic vectors are all equal to $4\mu_B$ in magnitude, as consistent with quantum mechanics).

Furthermore, Figure 19 shows that the Curie temperature, T_C, which is the temperature at which the polarization suddenly jumps from a vanishing to non-zero value (via a first-order transition), is less sensitive to the ME couplings: T_C varies from 1050K to 1090K when the γ_{me} coefficient changes from 0 to 1. However, Fig. 19 also demonstrates that the polarization-versus-temperature dependence changes its curvature around T_N. This effect is more pronounced for stronger γ_{me}, therefore indicating a significant effect of ME couplings on some *electric* properties.

In addition to Fig. 18, another consequence of ME couplings on magnetic properties can be clearly seen from Figure 20. As a matter of fact, in the case of $\gamma_{me} = 0$, the magnetic susceptibility adopts the "normal" behavior inherent to antiferromagnetic systems [22], that is a kind of a plateau for temperatures lower than T_N^0 and then a monotonic decrease (that we numerically found to be

inversely proportional to $T+T_N^0$, as consistent with Refs [22, 24]) when heating the systems above T_N^0. In contrast, switching on the ME couplings has three dramatic effects: (1) the value of the magnetic susceptibility decreases as the ME couplings increase in strength for any temperature $T \leq T_N^0$, with the plateau occurring when $\gamma_{me} = 0$ being even replaced by a slightly increasing function when increasing the temperature up to T_N^0 for the largest studied γ_{me} parameters; (2) the magnetic susceptibility is not anymore inversely proportional to $T+T_N$ when heating the system from T_N to T_C. In fact, $\chi_{M,\perp}$ is found to be nearly independent of temperature for the strongest γ_{me} coefficients; and (3) a sudden jump of $\chi_{M,\perp}$ is clearly seen at the Curie temperature.

B.2.2 Phenomenology

To reveal the origins of all these effects and better understand them, let us develop a phenomenology for which the free energy, $\mathcal{F}$, is given by:

$$\begin{aligned}\mathcal{F} = \mathcal{F}_0 + \tfrac{A_2}{2}P^2 + \tfrac{A_4}{4}P^4 + \tfrac{A_6}{6}P^6 + \tfrac{B_2}{2}M^2 + \tfrac{C_2}{2}L^2 + \tfrac{C_4}{4}L^4 \\ + \tfrac{\beta_{PM}}{2}P^2M^2 + + \tfrac{\beta_{PL}}{2}P^2L^2 + \tfrac{\beta_{LM}}{2}L^2M^2 - MH \end{aligned} \quad \text{(B.2.1)}$$

where M and L are the magnetization and G-type antiferromagnetic moment, respectively, while P and H are the electrical polarization and applied magnetic field, respectively. Note that, in the case of the simulations described above, the polarization, induced magnetization and magnetic field are all along the pseudo-cubic [111] direction while the AFM vector is along a direction being perpendicular to [111]. As a result, we are only concerned about the magnitude, rather than direction, of the physical quantities appearing in Eq.(B.2.1).

As consistent with phenomenologies of antiferromagnets undergoing a second-order magnetic transition and of ferroelectrics undergoing a first-order structural transition, the C_4 and A_6 coefficients are both positive and constant, while the A_4 parameter is also constant but is negative [24].

On the other hand, the C_2, B_2 and A_2 coefficients are *temperature-dependent* [24]:

$$\begin{aligned} C_2 &= c_2(T - T_N^0) \\ B_2 &= b_2(T + T_N^0) \\ A_2 &= a_2(T - T_C^0) \end{aligned} \tag{B.2.2}$$

where c_2, b_2 and a_2 are positive constants, and with T_C^0 being related to the "bare" Curie temperature, i.e. corresponding to the case of *no* ME coupling [13]. Notice the difference in sign in front of T_N^0 between the first and second line of Eq.(B.2.2).

Moreover, the β_{PL} coefficient of Eq. (B.2.1) is considered here to be a *negative* constant, since Figs. 18 and 19 show that increasing γ_{me} results in an enhancement of both L and P (at a fixed temperature below T_N^0). This enhancement also implies that the electric-dipole-mediated exchange parameters appearing in the second energy term of Eq. (B.1.2) disfavor ferromagnetism even more when γ_{me} increases in magnitude. As a result, the β_{PM} coefficient is positive, and enhancing the strength of the E_{ij} coefficients of Eq. (B.1.2) (or equivalently, γ_{me}) increases the magnitude of both the β_{PL} and β_{PM} parameters of Eq. (B.2.1). Finally, the β_{LM} parameter is a *positive* constant, as characteristic of a competition between the magnetization and AFM vector.

If we take into account that, *for temperatures below* T_C, $A_6(\beta_{PL}L^2 + A_2)$ is negative and that the magnetization always vanishes in the studied AFM system under no field, the minimization of Eq. (B.2.1) with respect to P gives:

$$P^2 = \frac{-A_4 + \sqrt{A_4^2 - 4A_6(\beta_{PL}L^2 + A_2)}}{2A_6} \quad for\ T \leq T_C \tag{B.2.3}$$

Inserting Eq. (B.2.2) into Eq. (B.2.3), and distinguishing the temperature ranges below and

above the Neel temperature (at which the AFM vector vanishes) thus give:

$$P^2 = \frac{-A_4+\sqrt{A_4^2+4A_6a_2(T_C^0-T)}}{2A_6} \quad for\ T_N \leq T \leq T_C \tag{B.2.4}$$
$$P^2 = \frac{-A_4+\sqrt{A_4^2+4A_6a_2(T_C^0-T)-4A_6\beta_{PL}L^2}}{2A_6} \quad for\ T \leq T_N \leq T_C$$

The fact that the first and second lines of Eq. (B.2.4) differ by the presence of $-4A_6\beta_{PL}L^2$ under the square-root (which is a quantity that is always positive) successfully explains the upward change of slope in the polarization-versus-temperature curves of Fig. 19 below T_N, when the ME effect is turned on (that is when the β_{PL} coefficient does not vanish), since L is also temperature dependent – as shown in Fig. 18. Our Landau-type phenomenological model can therefore reproduce and explain some striking features revealed by the atomistic simulations.

Furthermore, minimizing Eq. (B.2.1) with respect to L in our AFM multiferroic (for which M is zero for any temperature, when no field is applied) gives:

$$L^2 = \frac{-(C_2+\beta_{PL}P^2)}{C_4} \quad for\ T \leq T_N \leq T_C \tag{B.2.5}$$

Inserting Eq. (B.2.2) into this latter equality then yields

$$L^2 = \frac{c_2(T_N^0-T)}{C_4} - \frac{(\beta_{PL}P^2)}{C_4} \quad for\ T \leq T_N \leq T_C \tag{B.2.6}$$

Since β_{PL} is negative while C_4 is positive, the second term in the right-hand side of Eq. (B.2.6) explains another significant result of the atomistic simulations, namely why increasing γ_{me} enhances the magnitude of the AFM vector L for any finite temperature below T_N (see Fig. 18).

Moreover, setting Eq. (B.2.6) to zero provides the "renormalized" Neel temperature, that is the Neel temperature that takes into account ME effects:

$$T_N = T_N^0 - \frac{(\beta_{PL}P^2)}{c_2} \tag{B.2.7}$$

This equation indicates that the difference between T_N and T_N^0 should be proportional to both γ_{me} and the square of the polarization (computed at the Neel temperature). As shown in the inset of Fig. 18, such proportionality is well satisfied by the results of the atomistic simulations. Equation (B.2.7), that is derived from a Landau-type-model, thus provides a successful explanation of some key features of Fig. 18 (that arises from atomistic calculations), namely (1) why the Neel temperature is larger than the "bare" Neel temperature, T_N^0, when ME effects are switched on; and (2) why T_N increases when increasing γ_{me}. It would therefore be a mistake to determine the bare magnetic exchange parameters from the experimental Neel temperature of multiferroics, especially if these latter exhibit strong ME parameters.

Let us now try to explain and deeply understand some striking results shown in Fig. 20. For that, one first has to minimize Eq. (B.2.1) with respect to M:

$$M = \frac{H}{B_2 + \beta_{LM} L^2 + \beta_{PM} P^2} \tag{B.2.8}$$

Taking the derivative of this latter equality with respect to H then gives:

$$\chi_{M,\perp} = \frac{1}{B_2 + \beta_{LM} L^2 + \beta_{PM} P^2} \tag{B.2.9}$$

This latter equation can be separated into three different equalities, depending on the range of temperatures for which the AFM vector and/or polarization vanish or not:

$$\begin{aligned} \chi_{M,\perp} &= \tfrac{1}{B_2} \quad for\ T_N \leq T_C \leq T \\ \chi_{M,\perp} &= \tfrac{1}{B_2 + \beta_{PM} P^2} \quad for\ T_N \leq T \leq T_C \\ \chi_{M,\perp} &= \tfrac{1}{B_2 + \beta_{LM} L^2 + \beta_{PM} P^2} \quad for\ T \leq T_N \leq T_C \end{aligned} \tag{B.2.10}$$

The first and second lines of Eq. (B.2.10) indicate that the inverse of the magnetic susceptibility should exhibit a sudden change of $\beta_{PM} P^2$ at the ferroelectric phase transition, when the polarization appears via a first-order transition. Such feature therefore explains the significant increase of χ_M

numerically found when increasing the temperature through the Curie temperature, for the largest studied ME coefficients (see Fig. 20) [16].

Moreover, inserting the second line of Eq. (B.2.2) and the first line of Eq. (B.2.4) into the second line of Eq. (B.2.10) gives:

$$\chi_{M,\perp} = \frac{1}{b_2(T+T_N^0) + \beta_{PM}\{\frac{-A_4+\sqrt{A_4^2+4A_6a_2(T_C^0-T)}}{2A_6}\}} \quad for\ T_N \leq T \leq T_C \tag{B.2.11}$$

This latter equation indicates that the magnetic susceptibility follows the "usual" $\frac{1}{b_2(T+T_N^0)}$ behavior when there is no ME coupling. On the other hand, switching the β_{PM} coefficient (by making γ_{me} non-null) leads to a violation of such traditional law, and can result in unusual behavior. For instance, let us assume, for simplicity, that $4A_6a_2(T_C^0 - T)$ is much smaller than A_4^2. Then using a Taylor expansion of the square root results in the rewriting of Eq. (B.2.11) as:

$$\chi_{M,\perp} = \frac{1}{(b_2 + \frac{\beta_{PM}a_2}{A_4})T + (b_2T_N^0 - \frac{\beta_{PM}A_4}{A_6} - \frac{\beta_{PM}a_2T_C^0}{A_4})} \quad for\ T_N \leq T \leq T_C \tag{B.2.12}$$

In that case, an exact cancellation of b_2 and $\frac{\beta_{PM}a_2}{A_4}$ (recall that b_2 is positive while $\frac{\beta_{PM}a_2}{A_4}$ is negative) would render the magnetic susceptibility independent of the temperature when this latter ranges between T_N and T_C, which is nearly the case for the largest investigated γ_{me} as shown by the numerical data of Fig. 20!

Finally, the insertions of Eq. (B.2.6) and of the second line of Eq. (B.2.2) into the third line of Eq. (B.2.10) result in:

$$\chi_{M,\perp} = \frac{1}{(b_2 - \frac{\beta_{LM}c_2}{C_4})T + (b_2 + \frac{\beta_{LM}c_2}{C_4})T_N^0 + (\beta_{PM} - \beta_{LM}\frac{\beta_{PL}}{C_4})P^2} \quad for\ T \leq T_N \leq T_C \tag{B.2.13}$$

if we assume that $b_2 = \frac{\beta_{LM}c_2}{C_4}$ then the magnetic susceptibility becomes independent of temperature for $T \leq T_N$ *when there is no ME coupling* (i.e., when β_{PL}=β_{PM}=0), as nearly consistent with Fig. 20. Moreover, the last term of the denominator of Eq. (B.2.13) is positive (since β_{PL} is

negative while β_{PM}, β_{LM} and C_4 are all positive) and involves the square of the polarization. Since this latter decreases when increasing the temperature up to the Neel temperature (see Fig. 19), the phenomenological Equation (B.2.13) naturally explains aother numerical result, namely why $\chi_{M,\perp}$ does not exhibit anymore a plateau and, in fact, increases when heating the system to T_N for the largest considered γ_{me} coefficient.

Note that Eq. (B.2.13) also tells us that $\frac{1}{\chi_{M,\perp}} - \frac{1}{\chi^0_{M,\perp}}$, with $\chi^0_{M,\perp}$ being the magnetic susceptibility in case of *no* ME coupling, is equal to $(\beta_{PM} - \beta_{LM}\frac{\beta_{PL}}{C_4})P^2$ and thus should be directly proportional to both the γ_{me} parameter and the square of the polarization. The inset of Fig. 20 reveals that such proportionality indeed holds for the results of the atomistic simulations, which further asserts the validity of the phenomenological model developed here.

B.3 Conclusions

In summary, we have demonstrated, via the use of an atomistic scheme, that magnetic properties can be strongly affected by the ME coupling in an antiferromagnet multiferroic. This includes several strong deviations of the perpendicular component of the magnetic susceptibility from the universal behavior seen in "pure" antiferromagnets. Let us also emphasize that our phenomenological model (that allowed to reproduce and understand novel key features of magnetic properties of our model AFM multiferroic) can be easily extended to AFM multiferroics for which the Neel temperature is larger than the ferroelectric Curie temperature, as well as to system exhibiting a second-order paraelectric-to-ferroelectric phase transition or even to ferromagnet ferroelectrics (unlike the case we studied here). In all these situations, $\chi_{M,\perp}$ will likely exhibit anomalous features that should be reproduced and understood by these phenomenologies. We thus hope that the present work is of broad interest and deepens our current knowledge of multiferroics and antiferromagnets.

References

[1] J.B. Neaton *et al*, *Phys. Rev. B* **71**, 014113 (2005).

[2] I. Kornev *et al*, *Phys. Rev. Lett.* **99**, 227602 (2007).

[3] D. Albrecht, S. Lisenkov, Wei Ren, D. Rahmedov, Igor A. Kornev and L. Bellaiche, *Phys. Rev. B* **81**, 140401(R) (2010).

[4] J. Wang *et al*, *Science* **299**, 1719 (2003).

[5] J.C. Wojdel and J. Iniguez, *Phys. Rev. Lett.* **103**, 267205 (2009).

[6] W. Eerenstein, N. D. Mathur , J. F. Scott, *Nature* **442**, 759 (2006).

[7] N.A. Spaldin,and M. Fiebig, *Science* **309**, 391 (2005).

[8] C. Zhong, J. Fang and Q. Jiang, *J. Phys. C* **16**, 9059 (2004).

[9] W. Zhong, D. Vanderbilt and K.M. Rabe, *Phys. Rev. Lett.* **73**, 1861 (1994); *Phys. Rev. B* **52**, 6301 (1995).

[10] I. Kornev, L. Bellaiche, P.-E. Janolin, B. Dkhil and E. Suard, *Phys. Rev. Lett.* **97**, 157601 (2006).

[11] In fact, the predictions of this manuscript are technically valid for the (numerous) multiferroics for which the paraelectric-to-ferroelectric transition temperature is higher than the magnetic Neel temperature. This is because our toy model originates from the effective Hamiltonian of BFO, which is a material for which such hierarchy in temperature exists. A brief discussion is provided at the end of this manuscript to address other cases.

[12] V.I. Anisimov, F. Aryasetiawan and A.I. Lichtenstein, *J. Phys: Condens. Matter* **9**, 767 (1997).

[13] T_C^0 is not exactly the bare Curie temperature because the paraelectric-to-ferroelectric transition is of first order. In that case, the bare Curie temperature is equal to $T_C^0 + \frac{3A_4^2}{16a_2A_6}$ [14]. Note also that the paraelectric-to-ferroelectric transition predicted by the presently used effective Hamiltonian of BFO is rather special, namely it is a trigger-type transition that is driven by the collaborative coupling between the polarization and the tilting of oxygen octahedra [15]. As a result, the free energy is likely more complicated than the one provided in Eq. (B.2.1). However, we decided to use this latter for the sake of simplicity and because it can easily reproduce and explain several ferroelectric and magnetic anomalies for temperatures below the Curie temperature.

[14] See, e.g., $http://yclept.ucdavis.edu/course/240C/Notes/Landau/LandauPhaseTrans.pdf$

[15] I. Kornev and L. Bellaiche, *Phys. Rev. B* **79**, 10015(R) (2009).

[16] Note that plugging the second line of Eq. (B.2.2) into the first line of Eq. (B.2.10) results in a magnetic susceptibility that is given by $\frac{1}{b_2(T+T_N^0)}$ for temperature above T_C. This implies that such magnetic susceptibility should be independent of the ME coupling. Figure 20 indicates that this independency is not really obeyed, therefore pointing to an inadequacy of the phenomenology introduced here for temperature at or above the Curie point. This inadequacy

is further emphasized when realizing that the phenomenology also predicts that the Curie point should be insensitive to the γ_{me} parameter according to the first line of Eq. (B.2.4), which is not exactly the case (see Fig. 19). We numerically found that these deviations from the numerical data are in fact caused by *local* effects, namely the local electric dipoles still interact with the local magnetic moments even at or above the bare Curie temperature.