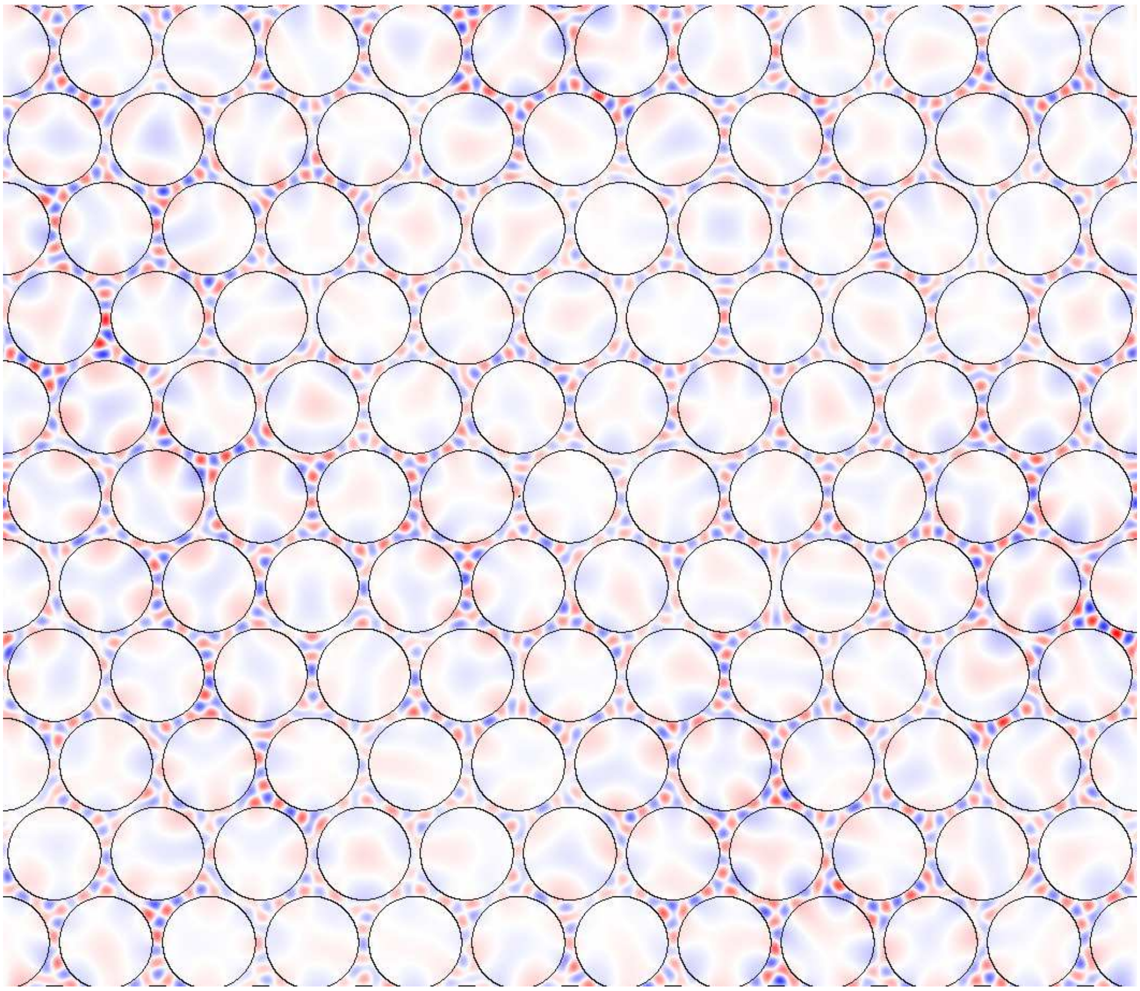


The previous page shows a rather delocalized Wannier function constructed by a direct lattice Fourier transform of the original Bloch functions as obtained from band structure computation. The very last page shows the corresponding symmetry compliant and maximally localized Wannier function created by means of the methods decribed in this thesis.

Matthias Hubertus Schillinger

Maximally localized photonic Wannier functions for the highly efficient description of integrated Photonic Crystal circuits

Maximally localized photonic Wannier functions for the highly efficient description of integrated Photonic Crystal circuits

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der Fakultät für Physik der Universität
Karlsruhe (TH)

genehmigte

DISSERTATION

von

Dipl. Phys. Matthias Hubertus Schillinger
aus München

Tag der mündlichen Prüfung: 01.12.2006
Referent: Prof. Dr. K. Busch
Korreferent: Prof. Dr. P. Wölfle

Contents

List of Figures

1 Overview

Over the next decades, one of the main challenges that the computing and communication industries will have to face, is to deliver more data faster. Consumers will demand to download full-length movies, not just photos or music files, and, of course, in shorter and shorter times. In addition, the number of Internet users is steadily increasing. To meet these future demands, the bandwidth of todays, especially on-chip, communication channels has to be increased.

One potential solution is photonics. Photonics offers much greater bandwidth than traditional copper networks and can carry multiple signals simultaneously without interference. Optical fibers have already started to replace electrical wires even in short distance communication such as local area networks and other computer interconnects during the last several years.

However, the realization of a completely seamless, all-optical network, where the communication and, potentially, the entire information processing would be carried out by laser light rather than electricity, requires novel optical materials capable of confining and micro-moulding the flow of light. While the surrounding air acts for electrical devices as an insulator, accomplishing that the electrical charge carriers are confined to their circuit board tracks even around sharp edges, electromagnetic pulses can easily escape into free space, when their light paths, e.g. conventional optical fibers, are strongly bent or distorted.

Photonic Crystals (PCs) represent a novel class of optical materials which can act as an insulator for light if their composition is properly engineered. In these periodic dielectric structures, a combination of microscopic scattering resonances from individual elements of the periodic array and Bragg scattering from the underlying lattice leads, in analogy to the appearance of electronic band structures in crystalline solids, to the formation of a so-called Photonic Band Structure (PBS). This flexibility in material composition, lattice periodicity, and symmetry allows one to tailor the photonic dispersion relations to almost any application need. Suitably engineered, these PCs exhibit frequency ranges over which the propagation of light is forbidden regardless of direction [1, 2, 3]. For frequencies inside these complete Photonic Band Gaps (PBGs) the PC acts as an completely insulating material for light, i.e. as a perfect mirror, and thus, enables for instance, the reduction of light leakage out of sharply bent optical fibers [4, 5].

Using now such a PC with a complete PBG as a background material, the breaking of the perfect periodicity of the structure by deliberately creating defect structures inside the PC, yields narrow prescribed wave-guiding channels that confine electromagnetic radiation to a region of the order of only several wavelengths. By this, PCs allow the creation of micro-circuit-like wave-guiding structures inside a perfect optical insulator[1].

These prospects have triggered enormous experimental activities aimed at the fabrication of two-dimensional (2D) and three-dimensional (3D) PC structures for telecommunication applications with PBGs in the near infrared frequency range. Considering that the first Bragg resonance occurs when the lattice constant equals half the wavelength of light, fabrication of PCs with band gaps in the near infrared regime, requires substantial technological efforts. For 2D PCs, advanced planar micro-structuring techniques borrowed from semiconductor technology can greatly simplify the fabrication process and high-quality PCs with embedded defects and waveguides have been fabricated in various material systems such as semiconductors [6, 7, 8, 9, 10, 11], polymers [12, 13], and glasses [14, 15]. Radiation losses [16, 17, 18], occurring in 2D PCs along the third dimension, can be avoided altogether if light is guided within the complete PBG of a 3D PC. Therefore, substantial efforts towards the manufacturing of suitable 3D PCs have been made. These structures include layer-by-layer structures [19, 20], inverse opals [21, 22, 23] as well as the fabrication of templates via laser holography [24, 25] and two-photon polymerization [26, 27, 28, 29, 30]. By double inversion of such polymer templates, 3D silicon-based Photonic Crystals with complete PBGs have recently been realized [31].

Given this tremendous flexibility in the fabrication of PCs, it is clear that any experimental exploration, or technological exploitation, of PCs has to be accompanied by a quantitative theoretical analysis. Detailed and accurate numerical simulations of PC structures allow the interpretation of experimental data and help to extract relevant parameters.

As alluded to earlier, substantial efforts have to be made to fabricate PCs featuring a complete PBG. The production of those containing embedded defect structures is even more sophisticated and, therefore, an expensive endeavor. Due to this, the creation of integrated wave-guiding structures in such PCs is hardly advisable without having already a detailed design blueprint, obtained by means of accurate theoretical simulations, that meets the envisaged functionality.

However, as mentioned above, the design space of free parameters is tremendous, and the best design found is only the best solution out of a limited set of combinations analyzable in the available time. Through this, the demand for highly efficient numerical tools enabling the accurate theoretical investigation of defect structures embedded into PCs becomes apparent.

To date, theoretical investigations of defect structures in PCs have almost exclusively been carried out by employing Finite Difference Time Domain (FDTD) discretizations of Maxwell's equations [32, 33]. Further methods being used are Galerkin-type methods [34] utilizing Hermite-Gauss functions as expansion basis [35, 36, 37], multiple multipole methods based on cylindrical [38, 39] or vector spherical [40] harmonics, and plane wave supercell calculations. For a detailed overview of these methods please see Ref. [41].

---

[1] PCs with complete PBGs are often regarded as the optical analogue of electronic semiconductors where a controlled "doping" allows one to tailor optical properties and specific functionalities.

All these methods disregard any information about the underlying PC. However, this information is readily obtainable by standard photonic band structure computation [41, 42], which provides dispersion relations and associated eigenmodes, so-called Bloch functions, for infinitely extended, perfectly periodic PCs. The Bloch functions themselves are less suited as an expansion basis for the theoretical description of defect structures with locally confined defect mode profiles, as they are, in principle infinitely, extended.

However, an ideal basis is stated by the so-called Wannier functions (WFs) [43], borrowed from solid state theory. WFs are obtained from Bloch functions via a Lattice Fourier Transform (LFT), i.e. a (continuous) unitary transformation. These functions contain all information about the periodic structure of the underlying PC and can be chosen to be, at least in principle, exponentially localized [44, 45]. Unfortunately, the construction of maximally localized Wannier functions (MLWFs) is problematic due to their non-uniqueness in definition with respect to additional phase factors and entanglement of the underlying Bloch functions (see Chap. 4) and, therefore, their practical importance in computational electronic structure theory and computational electrodynamics has been fairly minimal.

The situation changed after the publication of a novel method for calculating MLWFs for an isolated band complex in the context of crystalline solids by Marzari and Vanderbilt [46]. An extension to this method for the treatment of entangled groups of bands was later introduced by Souza et al. [47].

The aim of this thesis was to transfer these methods to periodically structured dielectric materials, i.e. Photonic Crystals, for the construction of maximally localized photonic Wannier functions [48] and to show their applicability and efficiency as an expansion basis for defect calculations in PCs.

The thesis is structured as follows. In Chap. 2 the fundamental concepts for the theoretical description of PCs are briefly reviewed and the utilized model systems are introduced. The accuracy and efficiency of the Wannier function approach for the theoretical description of defect structures embedded in Photonic Crystals is demonstrated in Chap. 3, where novel designs for basic functional elements, such as waveguide bends, intersections and couplers are presented. In addition, a rough comparison of the computational performance compared to FDTD calculations is given (Sec. 3.5). In Chap. 4, first the history of research on maximally localized Wannier functions is briefly reviewed and the complications arising in their construction are explained. Subsequently, the method of Marzari and Vanderbilt will be presented and applied to a model system. The extension of Souza et al. is introduced in Chap. 5, followed by the application of this method to a sample PC. One of the major differences and complications in applying these methods to PCs is, that the quanta of light, the photons, are bosons. Whereas it is usually sufficient in the context of crystalline solids to construct maximally localized Wannier functions only in a limited energy window around the Fermi energy, the description of defect structures embedded into PCs demands in general way more MLWFs, depending on the material structure of the PC and the strength of the introduced defect. These complications have been overcome by means of a newly developed bottom-up approach to the construction of arbitrary many, symmetry compliant and maximally localized Wannier functions, which is based on the methods mentioned above. The details of this new approach and its successful application to the construction of the first 38 maximally localized Wannier functions for a triangular lattice model system are presented in Chap. 6. A summary and future prospects are concluding this thesis.

The manuscript is also intended to serve as a reference to the computer software package that has been developed during my dissertation. Therefore, and since certain minor differences in the derivation of the underlying methods of electronic structure theory [46, 47] have been obtained, the calculations are presented in a rather detailed fashion. However, in order to keep the structure of this thesis clear, certain technical details such as longer calculations and proofs have been sourced out into several appendices.

2 Basic principles

In the following chapter the fundamental equations and concepts for the theoretical description of Photonic Crystals are derived, starting from Maxwell's equations up to the definition of photonic Wannier functions. Thereby, a general formulation will be used addressing the full vectorial problem in all three spatial dimensions at first. After introducing 1D, 2D, and 3D Photonic Crystals, the derived equations and definitions will be specialized for 2D systems in Sec. 2.3, as the construction of maximally localized Wannier functions is demonstrated for 2D PCs in the subsequent chapters. The model systems for this demonstration will be described in Sec. 2.3.2. In addition, some fundamental concepts of group theory are provided for these model systems in Sec. 2.3.3.

2.1 Wave propagation in dielectric media

Wave propagation in dielectric media is governed by the source- and current-free Maxwell's equations[1]

$$\nabla\,\mathbf{D}(\mathbf{r},t) = 0\;, \qquad \nabla\times\mathbf{E}(\mathbf{r},t) = -\frac{\partial\mathbf{B}(\mathbf{r},t)}{\partial t}\;, \tag{2.1}$$

$$\nabla\,\mathbf{B}(\mathbf{r},t) = 0\;, \qquad \nabla\times\mathbf{H}(\mathbf{r},t) = \frac{\partial\mathbf{D}(\mathbf{r},t)}{\partial t}\;, \tag{2.2}$$

which connect the electric field **E**, the electric displacement **D**, the magnetic field **H**, and the magnetic flux density **B**. Using that all known natural materials do not show a magnetic response at optical frequencies[2] and considering only field intensities where the used dielectrics respond linearly to external electric fields, one can identify

$$\mathbf{D}(\mathbf{r},t) = \varepsilon_0\varepsilon(\mathbf{r})\mathbf{E}(\mathbf{r},t)\;, \qquad \mathbf{B}(\mathbf{r},t) = \mu_0\mathbf{H}(\mathbf{r},t). \tag{2.3}$$

---

[1] In SI units.

[2] Nonzero magnetic response usually occurs only in the context of artificially created metamaterials [49]. In a certain subclass of those, meanwhile established as so-called left-handed materials, a negative magnetic permeability μ in combination with a negative electric permittivity ε results in a negative index of refraction [50, 51] accompanied with several highly interesting phenomena [52, 53, 54, 55].

Here, $\varepsilon(\mathbf{r})$ may also be tensorial, describing anisotropic media[3]. By combining now the curl equations in (2.1) and (2.2), and replacing the electric displacement $\mathbf{D}$ and magnetic field $\mathbf{B}$ by means of the material equations (2.3) yields a wave equation for the electric (2.4) as well as for the magnetic field (2.5),

$$\eta(\mathbf{r})\,\nabla\times\nabla\times\mathbf{E}+\frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}=0\,, \tag{2.4}$$

$$\nabla\times\eta(\mathbf{r})\,\nabla\times\mathbf{H}+\frac{1}{c^2}\frac{\partial^2\mathbf{H}}{\partial t^2}=0\,, \tag{2.5}$$

where the vacuum speed of light $c = 1/\sqrt{\varepsilon_0\mu_0}$ and the inverse dielectric function $\eta(\mathbf{r}) = \varepsilon^{-1}(\mathbf{r})$ were introduced. Since the main interest lies on stationary states we implicitly transfer above equations to Fourier space by assuming time harmonic dependencies of the fields according to

$$\mathbf{E}(\mathbf{r},t)=\mathbf{E}(\mathbf{r},\omega)\;e^{i\omega t}\,, \qquad\qquad \mathbf{H}(\mathbf{r},t)=\mathbf{H}(\mathbf{r},\omega)\;e^{i\omega t}\;\;. \tag{2.6}$$

Using these dependencies Eqs. (2.4) and (2.5) read in the frequency domain

$$\eta(\mathbf{r})\,\nabla\times\nabla\times\mathbf{E}(\mathbf{r},\omega)-\frac{\omega^2}{c^2}\,\mathbf{E}(\mathbf{r},\omega)=0\,, \tag{2.7}$$

$$\nabla\times\eta(\mathbf{r})\,\nabla\times\mathbf{H}(\mathbf{r},\omega)-\frac{\omega^2}{c^2}\,\mathbf{H}(\mathbf{r},\omega)=0\,. \tag{2.8}$$

Above equations might be recast into standard eigenvalue form

$$\mathcal{L}^{(\mathbf{E})}\,\mathbf{E}(\mathbf{r},\omega)=\frac{\omega^2}{c^2}\mathbf{E}(\mathbf{r},\omega)\,, \tag{2.9}$$

$$\mathcal{L}^{(\mathbf{H})}\,\mathbf{H}(\mathbf{r},\omega)=\frac{\omega^2}{c^2}\mathbf{H}(\mathbf{r},\omega)\,, \tag{2.10}$$

by introducing differential operators

$$\mathcal{L}^{(\mathbf{E})} := \eta(\mathbf{r})\,\nabla\times\nabla\times\;\;, \tag{2.11}$$

$$\mathcal{L}^{(\mathbf{H})} := \nabla\times\eta(\mathbf{r})\,\nabla\times\;\;\;. \tag{2.12}$$

Eqs. (2.7)–(2.10) are the starting point for the calculation of photonic band structures (Sec. 2.2.3) as well as for the computation of frequencies and field distributions of defect structures (Chap. 3) embedded into Photonic Crystals. Therefore no special restrictions, like for instance periodicity, have been imposed on the dielectric function $\varepsilon(\mathbf{r})$ so far. The special case of a periodically varying $\varepsilon(\mathbf{r})$ constituting a perfect Photonic Crystal will be discussed subsequently.

---

[3] In the case of Photonic Crystals two cases can be distinguished here: The first is that the materials constituting the PC are anisotropic themselves which leads to inherently tunable band structures [56]. Secondly, the case may be treated that the underlying PC consists of isotropic materials but the embedded defect structures contain anisotropic elements [57].

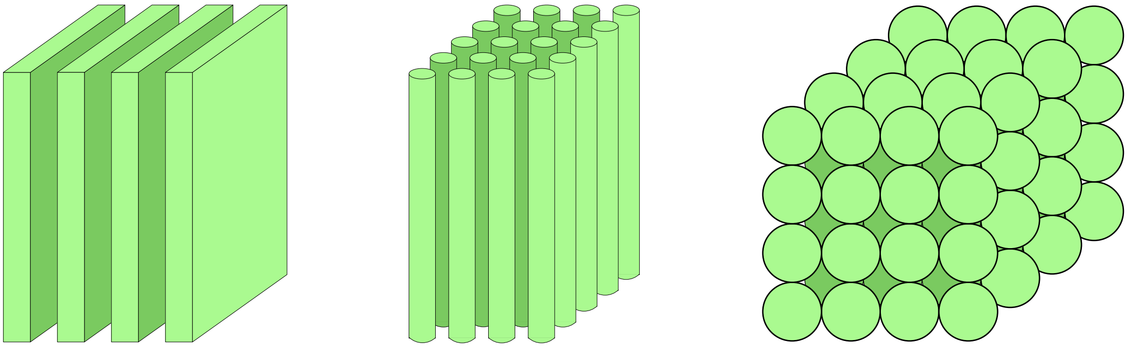

(a) 1D (b) 2D (c) 3D

Figure 2.1: *Schematics of one-, two-, and three-dimensional Photonic Crystals.*

2.2 Wave propagation in Photonic Crystals

In this section the basics of wave propagation in PCs are briefly reviewed. Starting from the definition of PCs and the direct lattice (Sec. 2.2.1), over the concept of the reciprocal lattice (Sec. 2.2.2) up to terms like the photonic band structure and photonic Bloch functions (Sec. 2.2.3) will be considered. After browsing through these fundamental concepts the step towards the introduction of photonic Wannier functions is performed (Sec. 2.2.4), followed by some remarks on the discretization of Fourier space.

2.2.1 Definition of Photonic Crystals and the direct lattice

PCs are periodically structured dielectric materials and can therefore be described mathematically by a periodically varying dielectric function

$$\varepsilon_{\mathrm{p}}(\mathbf{r}) = \varepsilon_{\mathrm{p}}(\mathbf{r} + \mathbf{R}) \ , \tag{2.13}$$

where the periodicity is meant with respect to the set

$$\mathcal{R} := \{\mathbf{R} \mid \mathbf{R} = \sum_{i=1}^{d} m_i \mathbf{a}_i; \ m_i \in \mathbb{Z}\} \tag{2.14}$$

of lattice vectors $\mathbf{R}$, generated by the primitive translations $\mathbf{a}_i$, $i = 1 \dots d$, of the underlying PC lattice. The set $\mathcal{R}$ is often referred to as the direct lattice. Its counterpart in Fourier space, the reciprocal lattice, will be introduced in Sec. 2.2.2.

Depending on the dimensionality d of the lattice one can distinguish three- (3D), two- (2D) and one-dimensional (1D) Photonic Crystals. For illustration, some sample schematics of those are depicted in Fig. 2.1. The focus of this thesis is set on 2D PCs, however, the subsequent derivations are performed in the full vectorial 3D formulation and will later in Sec. 2.3 be specialized for the case of 2D PCs.

(a) Square lattice. (b) Triangular lattice.

Figure 2.2: *Schematics of sample 2D PCs: arrangement of cylindrical holes on* ***(a)*** *a square and* ***(b)*** *a triangular lattice. The primitive translations* $\mathbf{a}_1$ *and* $\mathbf{a}_2$ *of each lattice are indicated as black arrows. The corresponding Wigner-Seitz Cells are enframed by black lines.*

Since the dielectric function is periodic, the spatial region of computation can be limited to a Wigner-Seitz Cell (WSC)[4], which is defined as the set of all spatial vectors $\mathbf{r}$ which are closer to the origin than to any other lattice vector $\mathbf{R}$. The WSCs for two sample 2D systems are depicted in Fig. 2.2.

In the following, the case of a perfectly periodic PC is considered where

$$\varepsilon(\mathbf{r}) \equiv \varepsilon_{\mathrm{p}}(\mathbf{r}) = \varepsilon_{\mathrm{p}}(\mathbf{r}+\mathbf{R}) \; . \tag{2.15}$$

The partial differential equations (2.7) and (2.8) have to be solved with respect to periodic boundary conditions for the corresponding field intensities. This is synonymous to the fact that for a periodic dielectric function $\varepsilon_{\mathrm{p}}(\mathbf{r})$ the operators $\mathcal{L}^{(\mathbf{E})}$ and $\mathcal{L}^{(\mathbf{H})}$ commute with the discrete translation operator $T_{\mathbf{R}}$, defined by its action on a general function $\boldsymbol{\psi}(\mathbf{r})$ as

$$T_{\mathbf{R}}\boldsymbol{\psi}(\mathbf{r}) = \boldsymbol{\psi}(\mathbf{r}+\mathbf{R}) \; , \tag{2.16}$$

for all lattice vectors $\mathbf{R} \in \mathcal{R}$ defined in Eq. (2.14). The eigenfunctions of $T_{\mathbf{R}}$ are given by plane waves $\boldsymbol{\psi}_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}$ obeying the eigenvalue equation

$$T_{\mathbf{R}}\, e^{i\mathbf{kr}} = e^{i\mathbf{kR}}\, e^{i\mathbf{kr}} \; , \tag{2.17}$$

with eigenvalues $e^{i\mathbf{kR}}$ and being associated with a real wave vector $\mathbf{k}$. They might be utilized as an expansion basis for the solutions of eigenvalue equations (2.9) and (2.10) since it is well known from operator theory that commutating operators can be diagonalized within the same basis.

2.2.2 Reciprocal lattice

The special set of wave vectors $\mathbf{k} = \mathbf{G}$ belonging to eigenvalues $e^{i\mathbf{GR}} = 1$ constitutes the so-called reciprocal lattice [58]. In analogy to the direct lattice (2.14), it is defined

---

[4] A restriction to a WSC is in general preferable, but its numerical realization quite cumbersome for slanted bases such as the triangular lattice considered in Sec. 2.3.2. For those systems, usually ordinary elementary cells are used for computations, i.e. rhomboids instead of hexagons for the above example.

as

$$\mathcal{G} = \{\mathbf{G} \mid \mathbf{G} = \sum_{i=1}^{d} n_i \mathbf{u}_i;\ n_i \in \mathbb{Z}\} \tag{2.18}$$

where the reciprocal lattice vectors (RLVs) $\mathbf{G}$ are generated by the primitive translations $\mathbf{u}_i$, $i = 1 \dots d$. Using column vector notation latter ones are defined by

$$\big[\mathbf{u}_1 \dots \mathbf{u}_d\big]^T = 2\pi \big[\mathbf{a}_1 \dots \mathbf{a}_d\big]^{-1} \tag{2.19}$$

and, therefore, obey

$$\mathbf{u}_i\, \mathbf{a}_j = 2\pi \delta_{ij}\ , \tag{2.20}$$

where δ_{ij} is the Kronecker delta. The Wigner-Seitz Cell of the reciprocal lattice is called the first Brillouin Zone (BZ). Besides exhibiting the same symmetry as the reciprocal lattice itself this definition of an elementary cell for the reciprocal lattice has the advantage, that wave vectors $\mathbf{k}$ belonging to the BZ boundary obey the Laue condition

$$\mathbf{k}\mathbf{G} = \frac{1}{2} G^2\ . \tag{2.21}$$

Hence, Bragg scattering leads to the development of standing waves for these values of $\mathbf{k}$, observable in the band structure as states with zero group velocity perpendicular to the BZ boundary.

2.2.3 Bloch functions and band structure

For partial differential equations, such as (2.9)–(2.10), Bloch's theorem[5] states that their solutions $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$ can be written as a lattice periodic function

$$\mathbf{u}_{n\mathbf{k}}(\mathbf{r}) = \mathbf{u}_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) \tag{2.22}$$

times a plane wave $e^{i\mathbf{k}\mathbf{r}}$:

$$\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r}) = \mathbf{u}_{n\mathbf{k}}(\mathbf{r})\, e^{i\mathbf{k}\mathbf{r}}\ . \tag{2.23}$$

These functions are called Bloch functions and are labeled with a band index n, a wave vector $\mathbf{k}$ and belong to the eigenfrequencies $\omega_{n\mathbf{k}}$ of

$$\mathcal{L}_{\mathrm{p}}^{(\mathbf{E}/\mathbf{H})}\, \boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r}) = \frac{{\omega_{n\mathbf{k}}}^2}{c^2} \boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r}) \tag{2.24}$$

where $\mathcal{L}_{\mathrm{p}}^{(\mathbf{E}/\mathbf{H})}$ denotes the operators $\mathcal{L}^{(\mathbf{E}/\mathbf{H})}$ defined in (2.11)–(2.12) for the special case $\eta(\mathbf{r}) \equiv \eta_{\mathrm{p}}(\mathbf{r})$. The newly occurring band index n originates in the possibility to restrict the wave vectors $\mathbf{k}$ to lie within the first Brillouin Zone. Here, and also in the remaining parts of this thesis, the same general labeling $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$ for the solutions

---

[5] For a rigorous proof please see [58].

of $\mathcal{L}_{\mathrm{p}}^{(\mathbf{E})}$ and $\mathcal{L}_{\mathrm{p}}^{(\mathbf{H})}$ will be used for ease of notation later on. Hence, it might be advantageous for the reader to remember that

$$\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r}) := \begin{cases} \mathbf{E}_{n\mathbf{k}}(\mathbf{r}) := \mathbf{E}(\mathbf{r}, \omega_{n\mathbf{k}}) & \text{for } \mathcal{L}_{\mathrm{p}}^{(\mathbf{E})} \,, \\ \mathbf{H}_{n\mathbf{k}}(\mathbf{r}) := \mathbf{H}(\mathbf{r}, \omega_{n\mathbf{k}}) & \text{for } \mathcal{L}_{\mathrm{p}}^{(\mathbf{H})} \,. \end{cases} \tag{2.25}$$

The name $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$ will be used for properties which are common to both, the electric and the magnetic field. Nevertheless, it should be emphasized that the same physical situation is described by both, the magnetic and electric field, of course, and therefore either formulation can be chosen to fully describe the envisaged problem. In practice, though, the magnetic field formulation is usually preferred for 3D PCs [42] since $\operatorname{div}\mathbf{H} = 0$ eases numerical calculations significantly and, if desired, the corresponding electric field $\mathbf{E}(\mathbf{r})$ can always be obtained[6] by means of Maxwell's equations (2.1)–(2.2). For 1D PCs there is no advantage for either polarization as $\operatorname{div}\mathbf{E} = 0$ is fulfilled there as well and for 2D systems the polarization of the considered electromagnetic wave influences the choice of a convenient formulation as will be clarified in Sec. 2.3.

As mentioned above, the Bloch functions are eigenfunctions of the operators $\mathcal{L}_{\mathrm{p}}^{(\mathbf{E}/\mathbf{H})}$ belonging to their associated eigenfrequencies $\omega_{n\mathbf{k}}$. Considering these frequencies as continuous functions of $\mathbf{k}$ yields the dispersion relation for a certain band n:

$$\omega_n(\mathbf{k}) := \omega_{n\mathbf{k}} \,. \tag{2.26}$$

The entirety of those dispersion curves is called the Photonic Band Structure (PBS). In Fig. 2.4 on p. 17 such PBSs are depicted for two example systems of 2D Photonic Crystals which are going to be introduced in Sec. 2.3. It can be seen from these figures, that the dispersion relations (2.26) deviate strongly from the linear dispersion in homogeneous media. Especially at the BZ boundaries the dispersion becomes flat accompanied by the development of standing waves originating from Bragg resonances, i.e. from constructive interference of electromagnetic waves scattered back from all single scatterers located at sites $\mathbf{R}$ of the entire lattice $\mathcal{R}$. For certain well-chosen dielectric functions, the Bragg scattering can open frequency ranges, where no linear wave propagation through the crystal is possible[7]. In these frequency regions the PC acts as a perfect mirror for incoming electromagnetic waves. In addition, this so-called PBG enables the confinement of waves to a width of the order of the lattice constant, a, and by this gives rise to the development of wave-guiding structures embedded into the Photonic Crystal, which will be discussed in Sec. 3.3.

The Bloch functions $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$ constitute an orthonormal set of basis functions and, by this, obey the orthonormality relation

$$\langle \boldsymbol{\psi}_{m\mathbf{k}} | \boldsymbol{\psi}_{n\mathbf{k}'} \rangle = \delta_{mn}\, \delta(\mathbf{k} - \mathbf{k}') \,, \tag{2.27}$$

---

[6] Please be aware of the fact that this mapping is not unique due to the gauge freedom of adding arbitrary gradient fields.

[7] However, for sufficiently high intensities the propagation of non-linear electromagnetic waves is possible, giving rise to the occurrence of solitary waves and interesting novel interaction effects between those [59, 60, 61].

where Dirac's delta distribution and bra-ket notation for scalar products between different Bloch functions have been used. Though the operator $\mathcal{L}^{(\mathbf{E})}$ is not hermitian with respect to the standard scalar product, this drawback can be eliminated by choosing a different definition of scalar products. Thus, it is convenient to define

$$\langle \boldsymbol{\psi}_{m\mathbf{k}} | \boldsymbol{\psi}_{n\mathbf{k}'} \rangle = \begin{cases} \int_{\mathbb{R}^d} d^d\mathbf{r}\; \varepsilon(\mathbf{r})\, \boldsymbol{\psi}^*_{m\mathbf{k}}(\mathbf{r})\, \boldsymbol{\psi}_{n\mathbf{k}'}(\mathbf{r}) & \text{for } \mathcal{L}^{(\mathbf{E})}\,, \\ \int_{\mathbb{R}^d} d^d\mathbf{r}\; \qquad \boldsymbol{\psi}^*_{m\mathbf{k}}(\mathbf{r})\, \boldsymbol{\psi}_{n\mathbf{k}'}(\mathbf{r}) & \text{for } \mathcal{L}^{(\mathbf{H})}\,, \end{cases} \tag{2.28}$$

where the integration is performed over all space. To obtain numerical solutions for these Bloch functions $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$ and the corresponding eigenvalues $\omega_{n\mathbf{k}}$, a straightforward way is to expand all the periodic functions, $u_{n\mathbf{k}}(\mathbf{r})$ and $\varepsilon_{\mathrm{p}}(\mathbf{r})$, into Fourier series over the reciprocal lattice $\mathcal{G}$, thereby transforming the differential equation (2.24) into an infinite matrix eigenvalue problem, which may be suitably truncated and solved numerically. Details of this plane wave method (PWM) for isotropic systems can be found, for instance, in Ref. [42] and for anisotropic systems in Ref. [56]. Iterative plane wave based methods are available as well [62, 63].

While the PWM provides a straightforward approach to computing the band structure of PCs, it also exhibits a number of shortcomings, such as slow convergence associated with the truncation of Fourier series in the presence of discontinuous changes in the dielectric constant. Therefore, an efficient real space approach to computing photonic band structures has been developed recently [64]. Within this approach, the wave equation is discretized in real space through finite differences or more advanced finite element techniques leading to sparse matrix problems. The solution of the resulting algebraic problems are obtained by employing multigrid methods, which guarantee an efficient solution by taking full advantage of the smoothness of the photonic Bloch functions.

Besides the determination of dispersion relations and associated Bloch functions, other physical quantities of interest, such as the density of states, the local density of states [42] and the group velocity [65, 64], can be directly obtained from these band structure calculations.

It should also be emphasized at this point that the Bloch functions are the fundamental building blocks for the construction of Wannier functions, and, by this, also for the accurate computation of frequencies and mode profiles of defect structures embedded into the Photonic Crystal. Due to this fundamental role of the Bloch functions, their numerical quality with respect to convergence, orthonormality, and behavior under symmetry operations of the direct lattice, is of paramount importance.

2.2.4 Photonic Wannier functions

Photonic WFs are formally defined as Lattice Fourier Transform (LFT) of Bloch functions via

$$\mathbf{W}_{n\mathbf{R}}(\mathbf{r}) := \frac{1}{\mathrm{V}_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k}\; e^{i\mathbf{k}\mathbf{R}}\, \boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})\,, \tag{2.29}$$

where the integration is performed over the entire first Brillouin Zone and $\mathrm{V_{BZ}}$ denotes the volume of the latter. It is worth mentioning here that the integration is performed with respect to the wave vector index $\mathbf{k}$ and not with respect to the position vector $\mathbf{r}$. In the remaining part of the thesis notations like $\mathbf{W}^{(\mathbf{E})}_{n\mathbf{R}}(\mathbf{r})$ and $\mathbf{W}^{(\mathbf{H})}_{n\mathbf{R}}(\mathbf{r})$ will be used to distinguish between WFs for the electric and the magnetic field, respectively. In contrast to the infinitely extended Bloch functions, the WFs are supposed to be a set of localized functions, labeled with a lattice vector $\mathbf{R}$, indicating that the center of the function is located inside a WSC around the lattice site $\mathbf{R}$, and a band index n, inherited from the corresponding set of Bloch functions.

In practice, however, achieving really localized WFs is not an easy task. Already the non-uniqueness of Bloch functions with respect to a global phase $e^{i\phi(\mathbf{k})}$ changes the localization properties tremendously. When dealing with the case of entangled bands, a proper disentangling of those bands (a more detailed look on this topic follows in Chap. 4) can only be achieved by allowing linear combinations of Bloch functions (BFs) at each wave vector $\mathbf{k}$, facilitating the demixing of such entangled bands. The problems of non-unique phase choices and the demixing of bands can be simultaneously addressed by introducing so-called generalized Bloch functions, $\widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r})$, which are connected to the original BFs for each wave vector $\mathbf{k}$ by mutually different unitary matrices $U^{(\mathbf{k})}_{mn}$ according to

$$\widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r}) = \sum_m U^{(\mathbf{k})}_{mn}\, \boldsymbol{\psi}_{m\mathbf{k}}(\mathbf{r}) \ , \tag{2.30}$$

where the summation runs over all entangled bands m. The resulting functions $\widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r})$ are disentangled in a certain sense, which will be illuminated in Sec. 4.1. In analogy to Eq. (2.29) a corresponding set of generalized Wannier functions can be defined by

$$\begin{aligned}\mathbf{W}_{n\mathbf{R}}(\mathbf{r}) :&= \frac{1}{\mathrm{V_{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k}\ e^{i\mathbf{k}\mathbf{R}}\, \widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r}) &(2.31)\\ &= \frac{1}{\mathrm{V_{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k}\ e^{i\mathbf{k}\mathbf{R}} \sum_m U^{(\mathbf{k})}_{mn}\, \boldsymbol{\psi}_{m\mathbf{k}}(\mathbf{r}) \ , &(2.32)\end{aligned}$$

where the index n can no longer be associated with a certain band due to the mixing of bands by means of the matrices $U^{(\mathbf{k})}$. It is rather an integer label for the generalized WFs and BFs. As a matter of principle, these functions should intrinsically be named $\widetilde{\mathbf{W}}_{n\mathbf{R}}(\mathbf{r})$ as they are constructed by a LFT of generalized BFs, $\widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r})$. However, as this entire thesis is concerned with the localization of generalized Wannier functions only, the tilde is dropped for notational convenience. As the underlying BFs do, these functions build a complete set of orthogonal functions, too, obeying

$$\langle \mathbf{W}_{n\mathbf{R}} | \mathbf{W}_{m\mathbf{R}'} \rangle = \delta_{nm}\, \delta_{\mathbf{R}\mathbf{R}'} \ , \tag{2.33}$$

where the scalar product is defined as indicated in Eq. (2.28). In addition, the Bloch theorem (2.23) translates into a discrete translational invariance condition for the

Wannier functions,

$$\mathbf{W}_{n\mathbf{R}}(\mathbf{r}) = \mathbf{W}_{n\mathbf{0}}(\mathbf{r}-\mathbf{R}) \; . \tag{2.34}$$

Due to this translational invariance it is sufficient to construct Wannier functions only for the WSC around the origin, as the WFs at other lattice sites $\mathbf{R}$ can be obtained by just shifting the functions $\mathbf{W}_{n\mathbf{0}}(\mathbf{r})$ by the considered lattice vector $\mathbf{R}$.

2.2.5 Discrete Fourier space and finite domains

For the computation of Wannier functions it is obvious that the integration occurring in the LFT (2.31) cannot be performed analytically. In fact, the integration has to be approximated by a discrete sum over a finite set of wave vectors, $\mathbf{k}$. As suggested in [46], a uniformly spaced Monkhorst-Pack mesh [66] is used for this purpose. Based on this mesh, the representation of integrations changes according to

$$\frac{1}{\mathrm{V}_{\mathrm{BZ}}} \int d^d\mathbf{k} \;\longrightarrow \frac{1}{N_{\mathrm{kp}}} \sum_{\mathbf{k}} \; , \tag{2.35}$$

where N_{kp} denotes the total number of wave vectors $\mathbf{k}$ contained in the mesh. After performing this step towards a discrete set of wave vectors the definition of WFs reads

$$\mathbf{W}_{n\mathbf{R}}(\mathbf{r}) := \frac{1}{N_{\mathrm{kp}}} \sum_{\mathbf{k}} e^{i\mathbf{k}\mathbf{R}} \, \widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r}) \; . \tag{2.36}$$

As mentioned in Sec. 2.2.4 the knowledge of the Wannier functions at the central WSC is sufficient, due to the translational invariance condition (2.34). These functions adopt in discrete Fourier space the rather simple form

$$\mathbf{W}_{n\mathbf{0}}(\mathbf{r}) := \frac{1}{N_{\mathrm{kp}}} \sum_{\mathbf{k}} \widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r}) \; . \tag{2.37}$$

By this expression it might become apparent that the Wannier function $\mathbf{W}_{n\mathbf{0}}(\mathbf{r})$ can be regarded as the first Brillouin Zone average of Bloch functions $\widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r})$ for the band n. In addition, by discretizing the BZ a finite and periodically repeating computational domain on which the WFs are defined is inherently introduced. While being defined formerly on all space $\mathbb{R}^d$ the WFs are now only defined on a domain of N_{cells} unit cells. It is noteworthy that the total number of cells N_{cells} coincides with the total number of wave vectors N_{kp}. This fact might be comprehensible by two different arguments. The first is that the WFs have to fulfill Born-von Karman boundary conditions which determine the proper set of wave vectors $\mathbf{k}$ depending on the size of the supercell considered. The second and probably easier understandable argument is a dimensional one. The dimension of the functional space made up of the generalized BFs for a single band n is N_{kp}. On the other hand the set of lattice sites $\mathbf{R}$ labeling the WFs is limited to N_{cells}. Since the dimension of vector spaces should not be affected by a LFT, N_{kp} and N_{cells} have to coincide.

The introduction of a discrete sampling of the BZ is also responsible for the fact that the translational invariance (2.34) is only fulfilled approximately. Therefore, for correct numerical calculations it has to be assured that $N_{\mathrm{cells}} = N_{\mathrm{kp}}$ is chosen sufficiently large, when using (2.34) for the construction of WFs located at lattice sites $\mathbf{R} \neq \mathbf{0}$.

Due to the strong localization of MLWFs it is mostly sufficient for numerical calculations to limit the WFs to a smaller computational domain with $N_{\mathrm{cd}} < N_{\mathrm{cells}}$ elementary cells. This limitation is dependent on the localization of the WF itself and has to be determined empirically by employing certain accuracy limits for orthogonality, normalization, and overlap matrix elements.

2.3 Two-dimensional Photonic Crystals

This thesis — although being applicable to 3D Photonic Crystals as well — is focused on the generation of maximally localized Wannier functions for 2D Photonic Crystals. For this reason the equations obtained in the previous sections will be specialized for the case of two-dimensional systems. Later on, two example systems will be introduced: a square and a triangular lattice of cylindrical air pores etched into silicon [6]. The square lattice system will be utilized for the demonstration of defect calculations and for determining maximally localized WFs for an isolated group of bands addressed in Chaps. 3 and 4, respectively. The triangular lattice of air pores can be considered as one of the most difficult cases for the calculation of maximally localized Wannier functions[8]. It will later be used for the demonstration of the construction of MLWF for higher order, wildly entangled bands.

2.3.1 Decoupling of polarizations

In the following, for two-dimensional systems the dielectric function $\varepsilon(\mathbf{r})$ is assumed to vary in the xy-plane while being homogeneous along the z-direction. This leads to restrictions

$$\frac{\partial \varepsilon(\mathbf{r})}{\partial z} = 0 \qquad \frac{\partial \mathbf{E}(\mathbf{r})}{\partial z} = 0 \qquad \frac{\partial \mathbf{H}(\mathbf{r})}{\partial z} = 0 \, , \tag{2.38}$$

for the dielectric function $\varepsilon(\mathbf{r})$ and the electric and magnetic fields. Substituting these into the wave equations (2.7) and (2.8) yields one differential equation for the E_z-component of the electric field

$$\eta(\mathbf{r}) \left[\frac{\partial^2}{\partial x} + \frac{\partial^2}{\partial y} \right] E_z(\mathbf{r}) + \frac{\omega^2}{c^2} E_z(\mathbf{r}) = 0 \tag{2.39}$$

which is decoupled from the other components, and two coupled differential equations for the remaining field components, $E_x(\mathbf{r})$ and $E_y(\mathbf{r})$. The same is valid for

---

[8] Reasons for this will be given in Sec. 5.3.

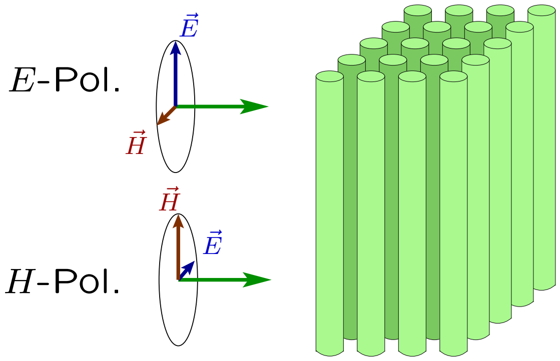


Figure 2.3: *Decoupling of polarizations for 2D systems for a sample PC made of a square lattice of rods in air. The dielectric function is constant along the vertical axis (z-axis). E-Polarization: the E-field is oriented parallel to the z-axis. H-Polarization: the H-field is oriented parallel to the z-axis.*

the H-field. One obtains a decoupled wave equation for the H_z-component

$$\left[\frac{\partial}{\partial x}\eta(\mathbf{r})\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\eta(\mathbf{r})\frac{\partial}{\partial y}\right]H_z(\mathbf{r})+\frac{\omega^2}{c^2}H_z(\mathbf{r})=0 \tag{2.40}$$

and two coupled equations for the $H_x(\mathbf{r})$ and $H_y(\mathbf{r})$ components. Due to this decoupling one can consider two different polarizations depending on whether the E-field (E-Polarization) or the H-field (H-polarization) is parallel to the z-axis illustrated in Fig. 2.3. In both cases the field distributions are conveniently describable by scalar wave equations for the z-components of the respective fields. Arbitrary polarizations can be handled by taking appropriate linear combinations. For a perfectly periodic and isotropic 2D PC, $\varepsilon(\mathbf{r})\equiv\varepsilon_{\mathrm{p}}(\mathbf{r})$, this decoupling leads to two scalar eigenvalue equations

$$\eta_{\mathrm{p}}(\mathbf{r})\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right]E_{n\mathbf{k}}(\mathbf{r})=-\frac{\omega_{n\mathbf{k}}^2}{c^2}E_{n\mathbf{k}}(\mathbf{r})\;, \tag{2.41}$$

$$\left[\frac{\partial}{\partial x}\eta_{\mathrm{p}}(\mathbf{r})\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\eta_{\mathrm{p}}(\mathbf{r})\frac{\partial}{\partial y}\right]H_{n\mathbf{k}}(\mathbf{r})=-\frac{\omega_{n\mathbf{k}}^2}{c^2}H_{n\mathbf{k}}(\mathbf{r})\;, \tag{2.42}$$

where the Bloch functions $E_{n\mathbf{k}}(\mathbf{r})$ and $H_{n\mathbf{k}}(\mathbf{r})$ represent the z-component of the electric and the magnetic field, respectively. Based on this, scalar 2D Wannier functions might be defined as

$$W_{n\mathbf{R}}^{(E)}(\mathbf{r}):=\frac{1}{\mathrm{V_{BZ}}}\int\limits_{\mathrm{BZ}}d^2\mathbf{k}\;e^{i\mathbf{kR}}\,E_{n\mathbf{k}}(\mathbf{r})\;, \tag{2.43}$$

$$W_{n\mathbf{R}}^{(H)}(\mathbf{r}):=\frac{1}{\mathrm{V_{BZ}}}\int\limits_{\mathrm{BZ}}d^2\mathbf{k}\;e^{i\mathbf{kR}}\,H_{n\mathbf{k}}(\mathbf{r})\;. \tag{2.44}$$

The lattice sites $\mathbf{R}$ and wave vectors $\mathbf{k}$ can be considered as two-dimensional vectors lying within the xy-plane of the three-dimensional space and the set of vectors $\mathbf{R}$ belongs to a 2D lattice.

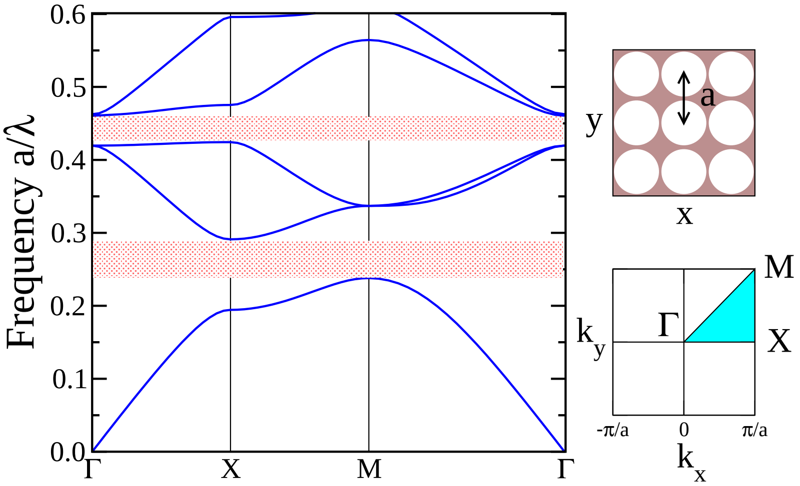
2.3.2 Model systems

For two-dimensional periodic arrangements of cylindrical objects such as holes in a dielectric, there exist five distinct so-called Bravais lattices [see Ref. 67, p. 6]: the square, rectangular, centered rectangular, hexagonal (triangular), and oblique lattice. The most commonly used systems for experimentally realized 2D PCs are arrangements of pores or rods on a square or triangular [6] lattice, which have already been illustrated in Figs. 2.2(a) and 2.2(b), respectively.

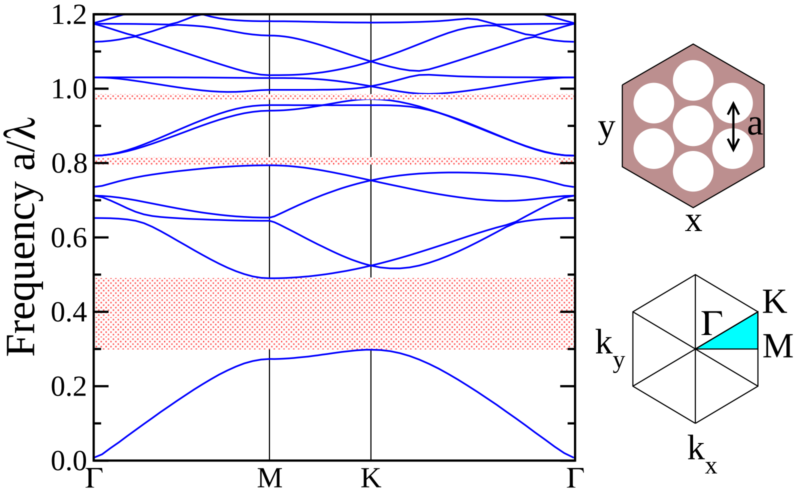
In this thesis two model systems of different Bravais lattice type and for different polarizations will be considered:

The first is a square lattice (with lattice constant a) of air pores (ε_{c}=1) of radius r=0.475a etched in silicon (ε_{b}=12), which will be utilized for the demonstration of defect calculations in a Wannier function basis (Chap. 3) and for the construction of MLWFs for an isolated group of bands (Chap. 4). This system exhibits for E-polarized light two band gaps, with the fundamental band gap extending from a/λ=0.238 to a/λ=0.291 (20% of the mid-gap frequency).

The second system is a triangular lattice of air pores of radius $r = 0.45a$ etched in silicon [6]. This PC exhibits for H-polarized radiation a large fundamental band gap between a/λ=0.298 and a/λ=0.490 (48.7% of the mid-gap frequency). Also two smaller higher order PBGs are visible and another one, not visible here, occurs between bands 38 and 39, and will serve as a naturally given upper bound for the number of bands for which MLWF will be constructed. At first this model system will be utilized for the demonstration of the construction of maximally localized Wannier functions for non-isolated groups of bands in Chap. 5. Later in Chap. 6 the first 38 MLWFs for this system will be constructed by means of the newly developed bottom-up approach [68]. The corresponding band structures, direct lattices, and first Brillouin Zones for both systems are depicted in Fig. 2.4.



(a) Square lattice of air pores with $r = 0.475a$. This PC exhibits for E-polarized light two band gaps, with the fundamental band gap extending from $a/\lambda = 0.238$ to $a/\lambda = 0.291$ (20% of the mid-gap frequency).



(b) Triangular lattice of air pores with $r/a = 0.45$. This PC exhibits a large fundamental band gap between a/λ=0.298 and a/λ=0.490 (48.7% of the mid-gap frequency). Also two smaller higher order PBGs are visible and another one, not visible here, occurs between bands 38 and 39.

Figure 2.4: *Photonic band structures for two 2D macroporous silicon PCs with different lattice types and polarizations. The frequencies are given in dimensionless units of lattice constant a over the vacuum wavelength λ which corresponds to the calculated frequency. Band gaps are shaded in red. The upper insets show the underlying real space lattice (with lattice constant a) of air pores with radius r and dielectric constant* $\varepsilon_c = 1$ *in silicon* ($\varepsilon_b = 12$) *for each system, whereas the lower ones depict the first Brillouin Zone of the corresponding reciprocal lattice with their irreducible wedges highlighted.*

2.3.3 Symmetries of 2D Bravais lattices

As symmetry is in general a mighty concept in nature and particularly with regard to maximally localized Wannier functions, this section is devoted to a brief group theoretical description of the symmetries of 2D Photonic Crystals.

2.3.3.1 Square lattice

In addition to the translational symmetry by a lattice vector $\mathbf{R}$, (2.13), 2D Photonic Crystals have other spatial symmetries, too. For instance, the considered square lattice model system is invariant under mirror reflections σ'_v, σ''_v, σ'_d, and σ''_d, which symbolize mirror reflections at mirror axes, pointing from the origin to the nearest neighbors (σ'_v, σ''_v) and second nearest neighbors (σ'_d, σ''_d). It is also invariant under rotations of 90°, 180°, and 270°. These symmetry operations are denoted by C_4, $C_2(\equiv C_4^2)$, and $C_4^{-1}(\equiv C_4^3)$, respectively, where C_n in general designates rotations by $2\pi/n$. Together with the identity operation, I, which keeps the structure as it is, the above mentioned symmetry operations constitute the C_{4v} point group[9]:

$$C_{4v} = \left\{ I, C_4, C_4^{-1}, C_2, \sigma'_v, \sigma''_v, \sigma'_d, \sigma''_d \right\} \tag{2.45}$$

A symmetry operation R_1 is called conjugate to a different operation R_2, if the relation

$$R_1 = R R_2 R^{-1} \tag{2.46}$$

holds for some $R \in C_{4v}$. In the case of a square lattice, (σ'_v, σ''_v), (σ'_d, σ''_d), and (C_4, C_4^{-1}) are pairs of conjugate operations as

$$\sigma''_v = C_4 \sigma'_v C_4^{-1} \ , \tag{2.47}$$

$$\sigma''_d = C_4 \sigma'_d C_4^{-1} \ , \tag{2.48}$$

$$C_4^{-1} = \sigma'_v C_4 {\sigma'_v}^{-1} \ . \tag{2.49}$$

Due to a certain arbitrariness in choosing the underlying coordinate system[10], conjugate operations might be considered as substantially equal. Hence, the group C_{4v} might be written as well as

$$C_{4v} = \left\{ I, 2C_4, C_2, 2\sigma_v, 2\sigma_d \right\} \ , \tag{2.50}$$

where, for instance, $2\sigma_d$ designates two operations of type σ_d.

A representation of a general group $\mathcal{G}$ on an n-dimensional linear vector space L is a realization of the group as a set of linear operators acting on L. I.e. each element $R \in \mathcal{G}$ is associated with an operator $D(R)$ acting on L as

$$D(R)\, f_i(\mathbf{r}) = \sum_{j=1}^{n} D_{ji}(R) f_j(\mathbf{r}) \ , \tag{2.51}$$

---

[9] The notation used here is known as Schönflies notation.

[10] E.g. σ'_v becomes σ''_v when rotating the coordinate system by 90°.

where the functions $f_i(\mathbf{r})$ constitute a basis of L. The dimension of such a representation is given by the rank of its representation matrices $D_{ji}(R)$. While $D(R)$ is the representation of one single element $R \in \mathcal{G}$, the set

$$D := \big\{ D(R) \mid R \in \mathcal{G} \big\} \ , \tag{2.52}$$

is called the representation of the group $\mathcal{G}$. A representation is called reducible if at least two true subspaces $L^{(\alpha)}$ of L can be found so that the subspaces $L^{(\alpha)}$ themselves are invariant under all $D(R) \in D$. If it is not possible to find such a set of subspaces $L^{(\alpha)}$, i.e. L is the only vector space which is invariant under all $D(R)$, the representation is called irreducible. For the case of a reducible representation (RREP) the vector space L is separable into direct sums of invariant subspaces $L^{(\alpha)}$ as

$$L = L^{(1)} \oplus L^{(2)} \oplus \cdots \ , \tag{2.53}$$

and therefore all $D(R)$ can, by means of equivalence transformations, be brought into block diagonal form

$$D(R) = D^{(1)}(R) \oplus D^{(2)}(R) \oplus \cdots \ , \tag{2.54}$$

where the $D^{(\alpha)}(R)$ are the irreducible representations (IREPs) of group elements R leaving the corresponding vector space $L^{(\alpha)}$ invariant. As it might happen that a RREP contains several non-equivalent IREPs one may write

$$D(R) = \bigoplus_{\alpha} m_\alpha D^{(\alpha)}(R) \quad \text{with } m_\alpha \in \mathbb{N} \ . \tag{2.55}$$

The trace of an IREP is called its character $\chi^{(\alpha)}(R)$ with respect to R, i.e.

$$\chi^{(\alpha)}(R) := \sum_{i=1}^{n} D_{ii}^{(\alpha)}(R), \tag{2.56}$$

and if $\widetilde{f}_\alpha(\mathbf{r})$, connected to the $f_i(\mathbf{r})$ by unitary transformations, is a simultaneous eigenfunction to $D^{(\alpha)}(R)$ for all $R \in \mathcal{G}$, Eq. (2.51) simplifies to[11]

$$D^{(\alpha)}(R)\, \widetilde{f}_\alpha(\mathbf{r}) = \chi^{(\alpha)}(R)\, \widetilde{f}_\alpha(\mathbf{r}) \ , \tag{2.57}$$

and, hence, the function might be associated with the IREP $D^{(\alpha)}$ of the group $\mathcal{G}$,

$$D^{(\alpha)} = \big\{ D^{(\alpha)}(R) \mid R \in \mathcal{G} \big\} \ , \tag{2.58}$$

also uniquely characterized by the set of characters $\{\chi^{(\alpha)}(R) \mid R \in \mathcal{G}\}$. The matrix-like arrangement of $\chi_{\alpha R} := \chi^{(\alpha)}(R)$ is called a character table. Such are depicted in the left panels of Figs. 2.5 and 2.7.

---

[11] For one-dimensional IREPs.

C_{4v}	I	$2C_4$	C_2	$2\sigma_v$	$2\sigma_d$
A_1	1	1	1	1	1
A_2	1	1	1	-1	-1
B_1	1	-1	1	1	-1
B_2	1	-1	1	-1	1
E	2	0	-2	0	0

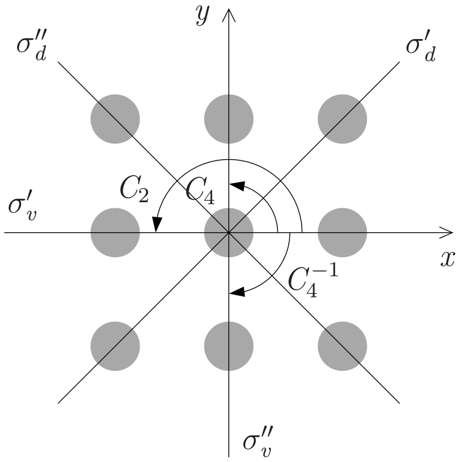


Figure 2.5: *Character table (left panel) and corresponding symmetry operations (right panel) for the C_{4v} point group of the square lattice.*

In general, assuming that the 2D PC is invariant under any symmetry operation R that belongs to the point group $\mathcal{M}$ of the lattice, the differential operator $\mathcal{L}_{\mathrm{p}}^{(\mathbf{E}/\mathbf{H})}$ defined on p. 9 commutes with all such elements of $\mathcal{M}$. Group theory tells now [69, 70, 71] that any eigenfunction, $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$, is then associated with an IREP of the $\mathbf{k}$-group $\mathcal{M}_{\mathbf{k}}$, which is the subgroup of $\mathcal{M}$ that keeps the wave vector $\mathbf{k}$, labeling the eigenmodes $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$, invariant[12]. Hence, a classification of the eigenmodes according to the IREPs of $\mathcal{M}_{\mathbf{k}}$ is possible [72].

Of special interest (see Sec. 4.3.3) for this thesis are only the BFs at the Γ-point of the reciprocal lattice, i.e. at $\mathbf{k} = \mathbf{0}$. It might readily be seen that in general all elements of the point group $\mathcal{M}$ leave the Γ-point invariant and thus

$$\mathcal{M}_\Gamma \equiv \mathcal{M} \, . \tag{2.59}$$

In the case of the square lattice $\mathcal{M}_\Gamma \equiv C_{4v}$ holds, and hence the Bloch functions at the Γ-point can be characterized according to the irreducible representations of the C_{4v} point group. This group has four one-dimensional IREPs named A_1, A_2, B_1, B_2 and one two-dimensional, designated with E, where two-dimensional implies that the corresponding modes are doubly degenerate. The spatial symmetry of these IREPs is described by their characters $\chi^\alpha(R)$ which are given for the C_{4v} point group in the left panel of Fig. 2.5. A sample set of Γ-point BFs for the square lattice model system is depicted in Fig. 2.8, where the first 24 BFs are classified according to IREPs of C_{4v}. Later on, it will be referred to such functions as A_1-modes or functions of symmetry type A_1, depending on their transformation properties under C_{4v}.

---

[12] Up to a reciprocal lattice vector $\mathbf{G}$.

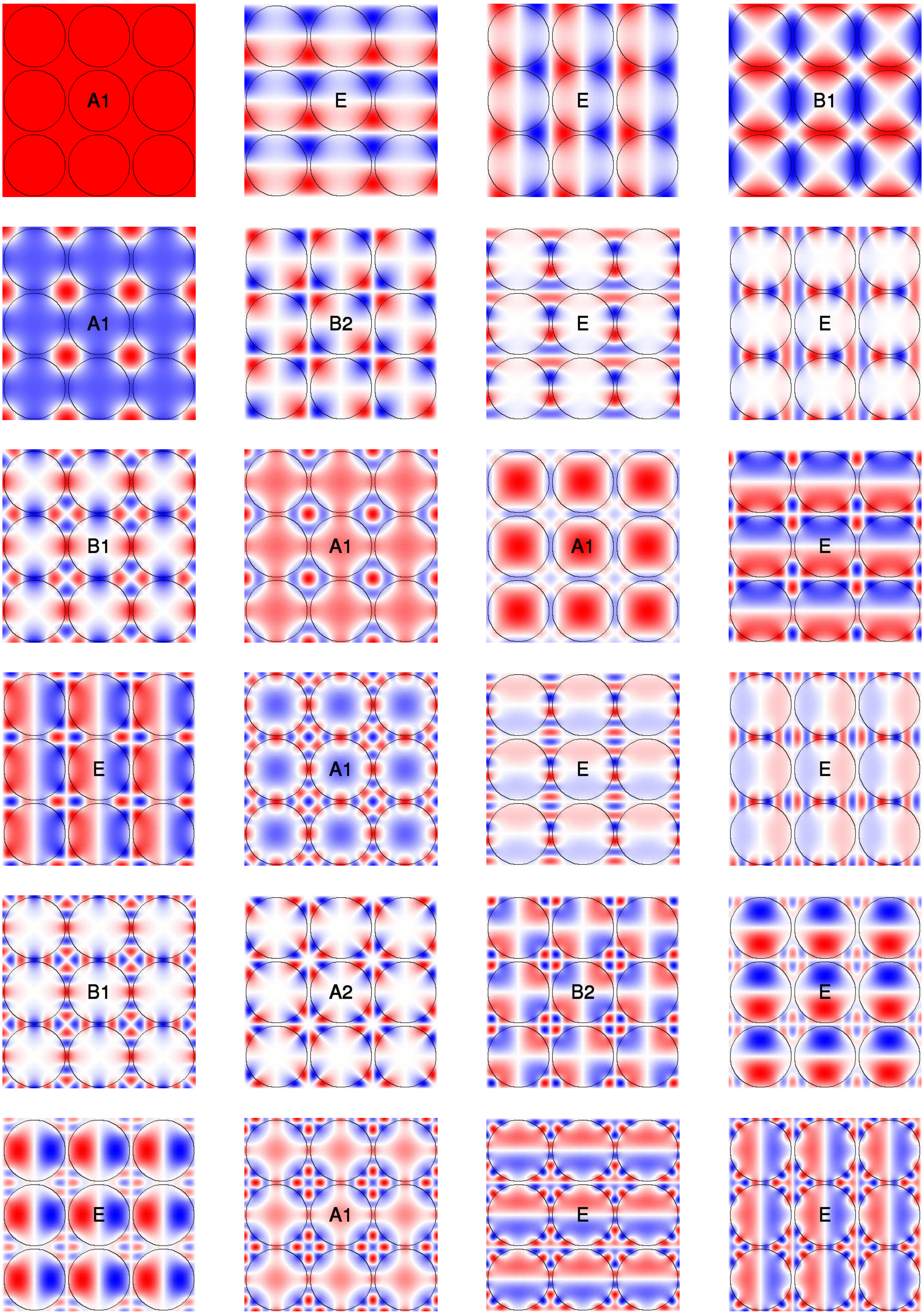


Figure 2.6: *The first 24 Bloch functions at the* Γ*-point classified with respect to the irreducible representations of the* C_{4v} *point group of the underlying square lattice.*

C_{6v}	I	$2C_6$	$2C_3$	C_2	$3\sigma_x$	$3\sigma_y$
A_1	1	1	1	1	1	1
A_2	1	1	1	1	-1	-1
B_1	1	-1	1	-1	-1	1
B_2	1	-1	1	-1	1	-1
E_1	2	1	-1	-2	0	0
E_2	2	-1	-1	2	0	0

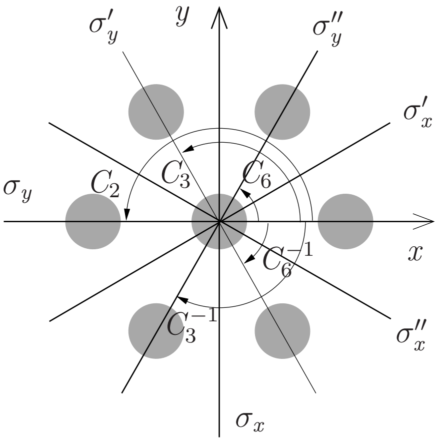


Figure 2.7: *Character table (left panel) and corresponding symmetry operations (right panel) for the C_{6v} point group of the triangular lattice.*

2.3.3.2 Triangular lattice

In analogy to the square lattice the right panel of Fig. 2.7 shows the symmetry operations for the arrangement of cylindrical objects on a triangular lattice. This structure is invariant under rotations of multiples of 60°. Those rotations are denoted by C_6, $C_3(\equiv C_6^2)$, $C_2(\equiv C_6^3)$, $C_3^{-1}(\equiv C_6^4)$, and $C_6^{-1}(\equiv C_6^5)$. The structure is also invariant under σ_x or σ_y mirror reflections and their equivalent mirror reflections denoted by σ'_x, σ''_x and σ'_y, σ''_y, respectively. Together with the identity operation I, these symmetry operations constitute the C_{6v} point group:

$$C_{6v} = \left\{I, C_6, C_6^{-1}, C_3, C_3^{-1}, C_2, \sigma_x, \sigma'_x, \sigma''_x, \sigma_y, \sigma'_y, \sigma''_y\right\} , \tag{2.60}$$

or merging conjugate operations

$$C_{6v} = \left\{I, 2C_6, 2C_3, C_2, 3\sigma_x, 3\sigma_y\right\} . \tag{2.61}$$

This point group has six irreducible representations named $A_1, A_2, B_1, B_2, E_1, E_2$ whose characters are depicted in the left panel of Fig. 2.7. In addition, Fig. 2.8 shows the first 38 BFs of the considered triangular lattice model system classified according to these IREPs.

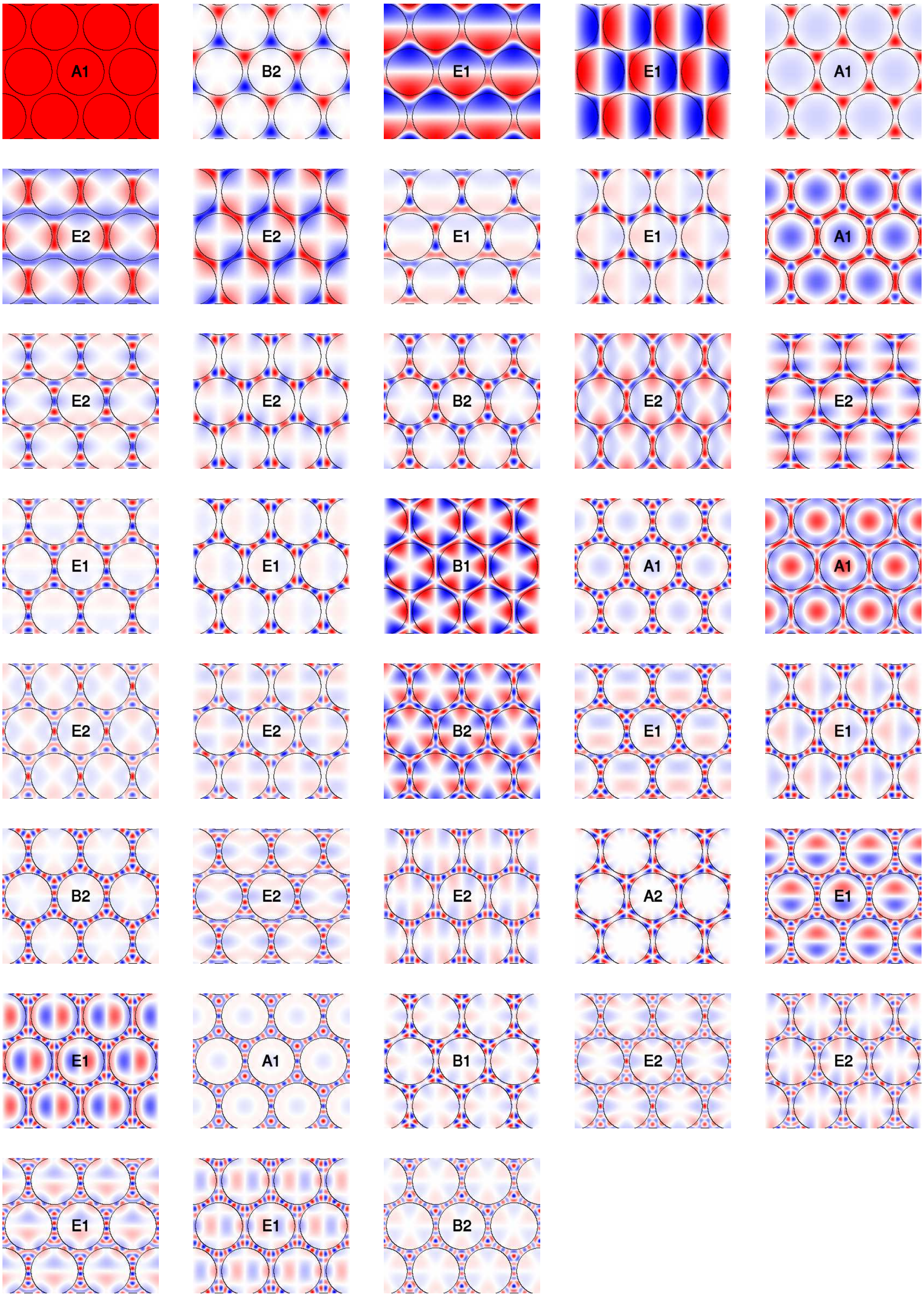


Figure 2.8: *The first 38 Bloch functions at the Γ-point classified with respect to the irreducible representations of the C_{6v} point group of the underlying triangular lattice.*

3 Defect computations in a Wannier function basis

Light paths for frequencies inside a PBG can be created by deliberately embedding defects into a PC. For instance, if a single pore in the square lattice model system, specified in 2.3.2, is modified or not edged at all, an optical micro-cavity is formed leading to a localized mode of light with a frequency lying inside the PBG. A chain of such point defects can act as a linear waveguiding channel and facilitates the construction of very sharp waveguide bends [73](for a PC of dielectric rods in air). Combinations of these basic elements can provide ultra-small beam splitters, Mach-Zehnder interferometers, and functional micro-optical elements such as ultra-compact mode separators [74]. Properly designed, a set of such functional elements may allow one to realize a PC circuitry capable of guiding electromagnetic radiation in a way similar to the guidance of electrons in electronic micro-circuits as alluded already in the introduction. In this section the usage and efficiency of the Wannier function approach for the description of the optical properties of PC circuits is illustrated by presenting basic functional elements, like straight waveguides (Sec. 3.3), highly transmissive waveguide bends and low-crosstalk waveguide crossings (Sec. 3.4), for the model PC specified in 2.3.2. Since this thesis is focused on the construction of maximally localized Wannier functions the demonstration of the efficiency of the Wannier function approach is limited to the the case of E-polarized light and only the basic equations for Wannier function based lattice models of defect structures embedded into PCs are derived. Applications to H-polarized radiation in combination with anisotropic defect structures is discussed in Ref. [57] and a detailed description of the entire approach has been published in Ref. [48].

3.1 Wannier function based lattice models

The description of defect structures embedded in PCs starts with the corresponding wave equation (2.39)

$$\nabla^2 E(\mathbf{r}) + \left(\frac{\omega}{c}\right)^2 \left\{\varepsilon_{\mathrm{p}}(\mathbf{r}) + \delta\varepsilon(\mathbf{r})\right\} E(\mathbf{r}) = 0 \; . \tag{3.1}$$

Here, the dielectric function $\varepsilon(\mathbf{r})$ was decomposed into a periodic part, $\varepsilon_{\mathrm{p}}(\mathbf{r})$, and the contribution $\delta\varepsilon(\mathbf{r})$ that describes the defect structures. Within the Wannier function approach, the electromagnetic field is expanded into WFs according to

$$E(\mathbf{r}) = \sum_{n,\mathbf{R}} E_{n\mathbf{R}} \, W_{n\mathbf{R}}(\mathbf{r}) \; , \tag{3.2}$$

with unknown amplitudes $E_{n\mathbf{R}}$. The number of WFs and lattice sites $\mathbf{R}$ is limited to N_{WF} and N_R, respectively. Inserting this expansion into the wave equation (3.1) leads to the basic equation for lattice models of defect structures embedded in PCs

$$\sum_{n',\mathbf{R}'} \left\{ \delta_{nn'}\delta_{\mathbf{RR}'} + D^{nn'}_{\mathbf{RR}'} \right\} E_{n'\mathbf{R}'} = \left(\frac{c}{\omega}\right)^2 \sum_{n',\mathbf{R}'} A^{nn'}_{\mathbf{RR}'} E_{n'\mathbf{R}'} \ . \tag{3.3}$$

The matrix $A^{nn'}_{\mathbf{RR}'}$ depends only on the dispersion relation and mixing matrices $U^{(\mathbf{k})}_{mn}$ of the WFs and is defined through

$$A^{nn'}_{\mathbf{RR}'} = -\int_{\mathbb{R}^2} d^2\mathbf{r} \ \ W^*_{n\mathbf{R}}(\mathbf{r})\, \nabla^2\, W_{n'\mathbf{R}'}(\mathbf{r}) \tag{3.4}$$

$$= \frac{1}{V_{\mathrm{BZ}}} \int_{\mathrm{BZ}} d^2\mathbf{k} \ \ e^{i\mathbf{k}(\mathbf{R}-\mathbf{R}')} \sum_m U^{(\mathbf{k})\dagger}_{nm} \left(\frac{\omega_{m\mathbf{k}}}{c}\right)^2 U^{(\mathbf{k})}_{mn'} \ . \tag{3.5}$$

Due to the smoothness of the photonic dispersion relation $\omega_{n\mathbf{k}}$ with respect to the wave vector $\mathbf{k}$, the exponential factor in Eq. (3.5) leads to a very rapid decay of the magnitude of matrix elements with increasing separation $|\mathbf{R}-\mathbf{R}'|$ between lattice sites, effectively making the matrix $A^{nn'}_{\mathbf{RR}'}$ sparse. Furthermore, it may be shown that the matrix $A^{nn'}_{\mathbf{RR}'}$ is symmetric and positive definite. Similarly, once the Wannier functions of the underlying PC are determined, the matrix $D^{nn'}_{\mathbf{RR}'}$ depends solely on the overlap of these functions, mediated by the defect structure:

$$D^{nn'}_{\mathbf{RR}'} = \int_{\mathbb{R}^2} d^2\mathbf{r} \ \ W^*_{n\mathbf{R}}(\mathbf{r})\, \delta\varepsilon(\mathbf{r})\ W_{n'\mathbf{R}'}(\mathbf{r}) \ . \tag{3.6}$$

As a consequence of the localization properties of both the Wannier functions and the defect dielectric function, the hermitian matrix $D^{nn'}_{\mathbf{RR}'}$, too, is sparse. Moreover, since MLWFs can be chosen to be real [46], both matrices, $A^{nn'}_{\mathbf{RR}'}$ and $D^{nn'}_{\mathbf{RR}'}$ are real and symmetric.

Depending on the nature of the defect structure, one is interested in (i) frequencies of localized defect modes, (ii) dispersion relations of straight waveguides, and (iii) transmission and reflection spectra through waveguide bends and other, more complex functional elements. In the following, each of these cases is discussed consecutively.

3.2 Single defects

At first, simple cavities created by changing the dielectric constant $\varepsilon_{\mathrm{def}}$ of a single pore as shown in the inset of Fig. 3.2 are considered. Such a change could be created by infiltrating a pore with e.g. a polymer or a liquid crystal [75]. The deviation from the perfect PC is then described by

$$\delta\varepsilon(\mathbf{r}) = (\varepsilon_{\mathrm{def}} - \varepsilon_{\mathrm{c}})\, \Theta(\mathbf{r} - \mathbf{R}_{\mathrm{def}}) \ , \tag{3.7}$$

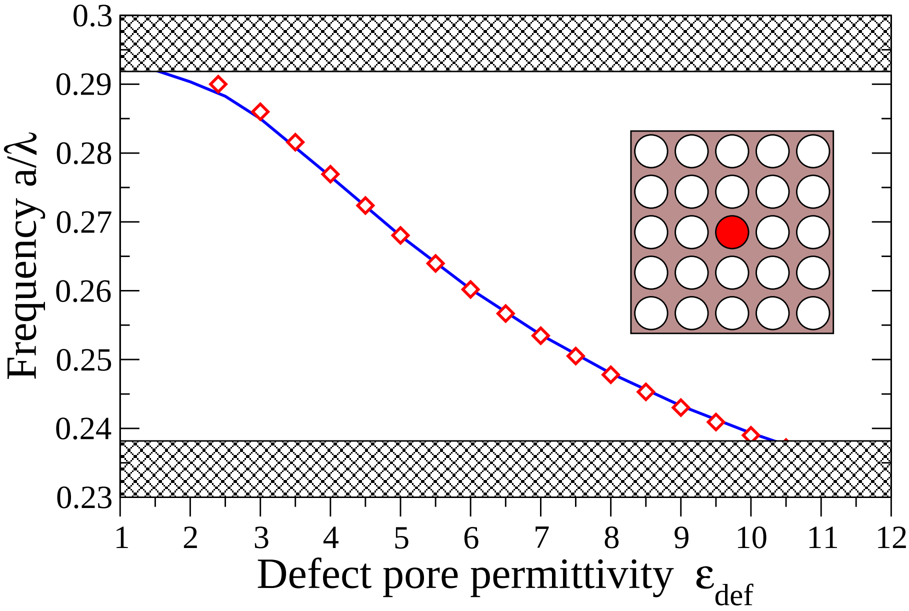


Figure 3.1: *Frequencies of the localized defect modes associated with a single pore infiltrated with a material with dielectric constant* ε_{def}. *Values indicated with diamonds are calculated with the Wannier function approach by directly solving Eq. (3.3) as a generalized eigenvalue problem for the cavity frequencies that lie within the photonic band gap. The solid lines are calculated correspondingly using plane-wave-based supercell calculations [63]. The parameters and corresponding band structure of the underlying model system are given in Fig. 2.4(a). The results of the Wannier function approach are in perfect agreement with numerically exact results of supercell calculations.*

where $\Theta(\mathbf{r})$ defines the area of the defect pore and ε_{c} is the dielectric constant of the air voids ($\varepsilon_{\mathrm{c}} = 1$). For a single infiltrated pore $\Theta(\mathbf{r}) = 1$ inside the pore at $\mathbf{R}_{\mathrm{def}} = \mathbf{0}$ and $\Theta(\mathbf{r}) = 0$ outside. After calculating the corresponding matrix elements, Eq. (3.3) is solved as a generalized eigenvalue problem for the cavity frequencies that lie within the photonic band gap. In Fig. 3.2 the dependence of the eigenfrequencies of the monopole-like cavity modes on the dielectric permittivity of the infiltrated material $\varepsilon_{\mathrm{def}}$ for the model system specified in Fig. 2.4 is depicted. It shows how a monopole-like defect mode emerges from the upper band edge and traverses through the Photonic Band Gap as $\varepsilon_{\mathrm{def}}$ increases. Fig. 2.4 also confirms that results obtained with the Wannier function approach are in complete agreement with plane-wave-based supercell calculations [63].

The symmetry properties of the cavity modes, which are shown in Figs. 3.2(b), clearly correlate with the symmetry of the underlying Wannier functions. This suggests that the convergence properties of the Wannier function approach depends on the nature and symmetry properties of the cavity modes under consideration. To determine the WFs which provide the largest contribution to the defect mode, it is helpful to define a measure V_n of the strength of the contributions from the

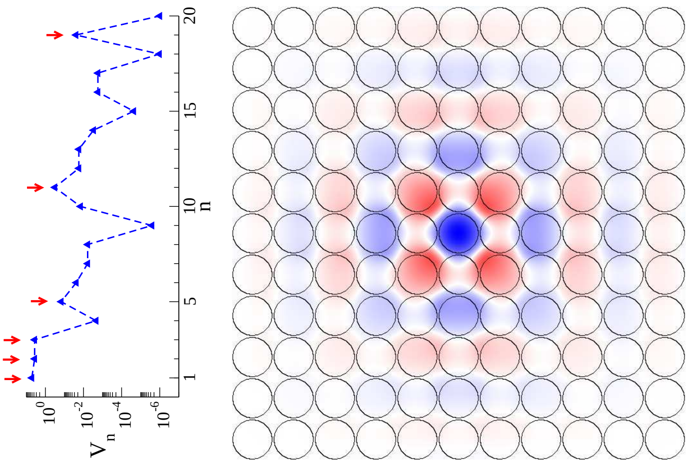


Figure 3.2: *Electric field distribution for the monopole-like cavity mode arising from filling a single pore with a material of refractive index* $\varepsilon_{def} = 2.40$. *The contributions* $V_n = \sum_R |E_{n\mathbf{R}}|^2$ *to this cavity mode of each single Wannier function, labeled with generalized index* n*, are displayed in the left panel. The WFs with highest contribution are marked with red arrows.*

individual WF associated with band n via

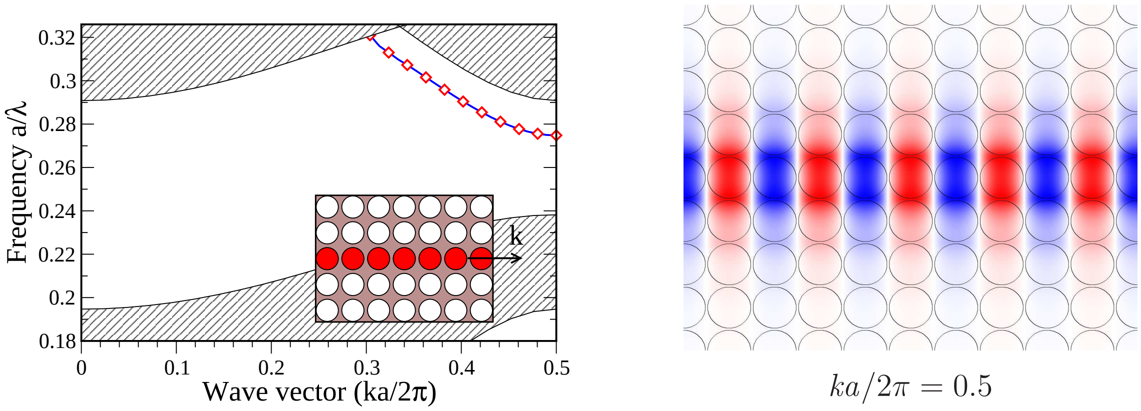
$$V_n = \sum_{\mathbf{R}} |E_{n\mathbf{R}}|^2 \,. \tag{3.8}$$

In the left panel of Fig. 3.2 the dependence of the parameter V_n on the band index n is depicted for the cavity mode shown in the right panel. According to their contributions a subset of $N_{\mathrm{WF}} = 6$ of former $N_{\mathrm{WF}} = 20$ WFs is chosen for the subsequent calculations without worsening the reached accuracy. Thus, the calculation of defect modes might be used to “gauge” the WF approach to numerically converged supercell calculations to determine a minimal set of WFs that yields the same accuracy of results. Though being already a highly efficient approach, the picking of suitable WFs improves the efficiency even more. For comparison: while plane wave based supercell calculations require the solution of an eigenvalue problem with matrix dimensions of the order of several thousands, the WF approach reaches results of the same accuracy by just solving a generalized eigenvalue problem (3.3) with matrix dimensions of the order of $(N_{\mathrm{WF}} \times N_{\mathrm{R}}) \times (N_{\mathrm{WF}} \times N_{\mathrm{R}}) = 130 \times 130$.

3.3 Dispersion relations of waveguides

Arguably the most important types of defect clusters in PCs are one or several adjacent straight rows of defects. Properly designed, such defect rows form a PC waveguide which allows the efficient guiding of light for frequencies within a photonic band gap [73, 76]. Due to the one-dimensional periodicity of such a waveguide, its guided modes obey the 1D Bloch-Floquet theorem

$$E(\mathbf{r} + \mathbf{s}_{\mathrm{w}}) = e^{\mathbf{i}\mathbf{k}(\omega)\mathbf{s}_{\mathrm{w}}}\, E(\mathbf{r}) \tag{3.9}$$



$ka/2\pi = 0.5$

Figure 3.3: *Dispersion relation (left) and exemplary mode profile (right) for the propagating guided modes of a waveguiding structure created by infiltration of a row of pores with materials specified in Fig. 3.2. This waveguide is mono-moded throughout the complete passing band ranging from* $a/\lambda = 0.275$ *to* $a/\lambda = 0.291$*, where the upper limit of the passing band is set by the upper edge of the band gap. The results obtained with the Wannier function approach, using* $N_{WF} = 6$ *Wannier functions are indicated with diamonds and are in complete agreement with the results of plane-wave based supercell computations shown as solid line. The parameters and corresponding band structures of the underlying model systems are given in Fig. 2.4.*

and thus can be labeled by a complex and frequency dependent wave vector, $\mathbf{k}(\omega)$, parallel to the waveguide director, $\mathbf{s}_{\mathrm{w}}=w_1\mathbf{a}_1+w_2\mathbf{a}_2$. Here, $\mathbf{a}_1$ and $\mathbf{a}_2$ are the primitive lattice vectors of the PC, and the integers w_1 and w_2 define the direction of the waveguide (for instance the so-called W1-waveguide depicted in the inset of Fig. 3.3 can be described through w_1=1 and w_2=0 where $\mathbf{a}_1$ is pointing along the horizontal axis). In Fig. 3.3 the dispersion relation of the propagating guided modes (i.e. modes with $\mathrm{Im}\,\mathbf{k}(\omega) = 0$) for a waveguide embedded (as indicated in the corresponding inset) into the model PC described in Fig. 2.4 is presented. The results obtained with the Wannier function approach are compared to calculations carried out by employing the supercell technique [63, 42] and show excellent agreement.

At this point it should be emphasized that, in contrast to the supercell technique, the Wannier function approach, features also the calculation of dispersion relations for evanescent guided modes, which obey Eq. (3.9) for complex wave vectors $\mathbf{k}(\omega)$ with non-zero imaginary part. Clearly, these modes are largely irrelevant in perfectly periodic straight waveguides. However, they do play an important role as soon as the perfect periodicity of the waveguide is broken either through imperfections due to fabrication tolerances, or through the deliberate creation of deviations from periodicity such as bends or coupled cavity-waveguide systems. In such cases, these evanescent guided modes give rise to light localization effects and determine the non-trivial transmission and reflection properties of PC circuits. Furthermore, they allow for the calculation of the complex transmission and reflection amplitudes of multi-port devices which will be presented subsequently (see also Ref. [77]). For details on the computation of dispersion relations of waveguides by means of the Wannier function approach please see Ref. [48].

3.4 Functional elements for Photonic Crystal circuits

Having established the basic properties of localized defect modes and waveguides based on the infiltration of pores with low-index materials, a set of functional elements for the creation of tunable PC circuits that combine several attractive advantages is presented:

1. The PC circuits are based on 2D PCs consisting of air pores in high-refractive-index dielectrics and, therefore, can easily be fabricated.

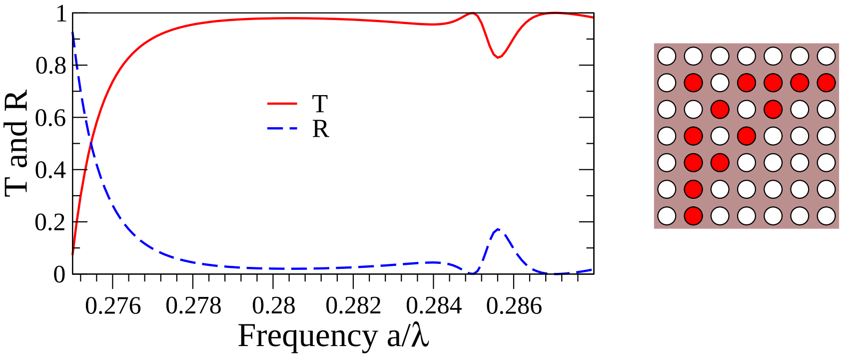
2. They exploit the non-degenerate, monopole-like defect modes for E-polarized light that result from infiltrating individual pores with appropriate low or moderate refractive index materials such as liquid crystals and/or polymers (see Fig. 3.2). By construction, this leads to essentially mono-moded PC waveguides. Furthermore, a peculiar symmetry of the square lattice cavity modes may be exploited to obtain designs for broad-band non-reflecting waveguide bends and beam splitters as well as broad-band low-crosstalk waveguide intersections operating in the same frequency range.

3. Owing to the tunability of the infiltrated materials the resulting circuits will be tunable.

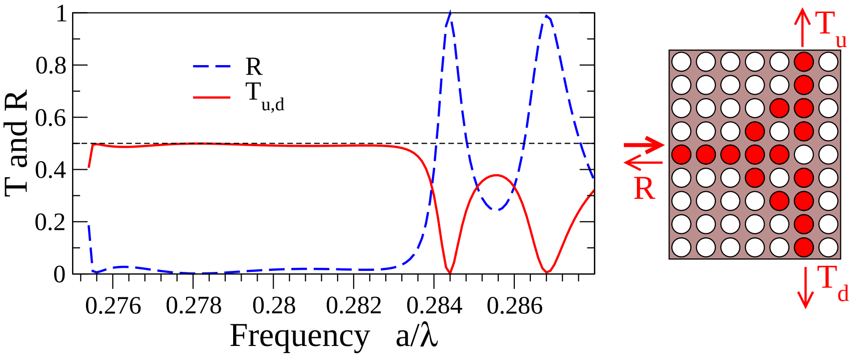
The results for the transmission characteristics are obtained within the Wannier function approach by solving Eq. (3.3) in the basis of Wannier functions as a system of linear equations with the frequency as a given parameter. To this end, certain expansion coefficients $E_{n\mathbf{R}}$ are assigned values to specify an incoming waveguide mode, either propagating or evanescent. Of significant importance is the fact that, in contrast to most standard methods, the Wannier function approach permits the accurate and efficient calculation of the complete scattering matrices of PC devices which consist of the complex scattering amplitudes connecting all, i.e., both propagating and evanescent, incoming with all outgoing channels and, therefore, contain all the phase relations and localization properties between these channels. As a consequence, it provides the possibility to model large-scale PC devices exploiting all the advantages of a guided-mode scattering matrix technique which is described in detail in Ref. [48] and Ref. [77], and is conceptually similar to the well-known cascaded impedance matrices of microwave theory [78]. To some extent[1] this technique might also be associated with the concept of Jones matrices [79, 80].

---

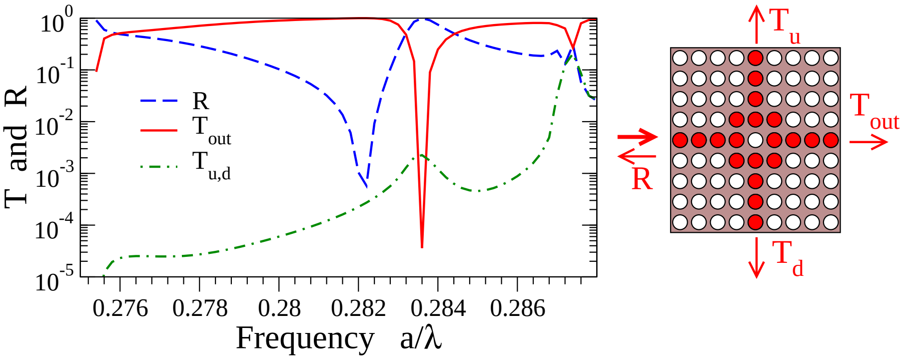
[1] Jones matrices represent only the change of the polarization of a light beam when traversing a certain optical device and the corresponding calculus employs standard matrix-matrix multiplication rules in contrast to those of scattering matrices.



(a) 90° waveguide bend.

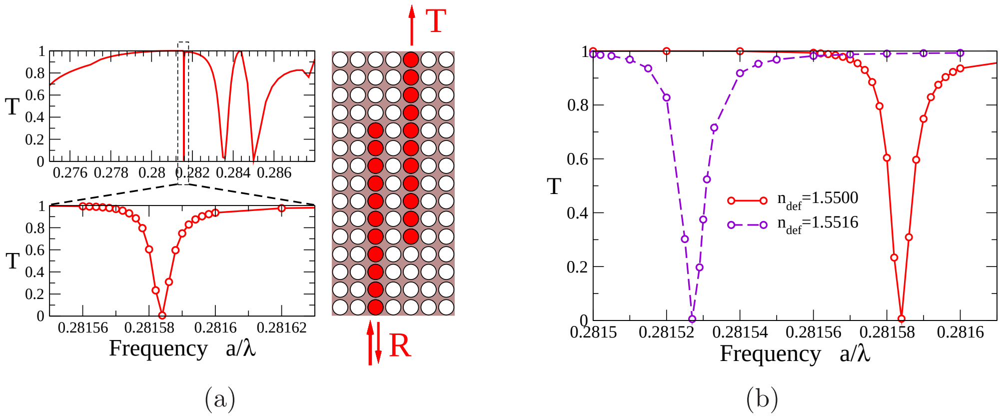


(b) 90° beam splitter.



(c) Waveguide intersection.

Figure 3.4: *Designs and transmission characteristics of different devices created by single pore infiltration with materials of dielectric permittivity* $\varepsilon_{def} = 2.40$. *All these devices exhibit high transmittance values over a common and broad frequency range and might therefore be utilized as basic building blocks for large scale integrated PC circuits.*



(a) (b)

Figure 3.5: ***(a)*** *Design and transmission spectrum for a folded directional coupler embedded into a square lattice of air pores.* ***(b)*** *Resonance dip of the transmittance for different values of refractive index n of the infiltrated material. The parameters and corresponding band structures of the underlying model systems are given in Fig. 2.4.*

In Fig. 3.4(a) a design for a waveguide bend with an angle of 90° is shown. This bend design exhibits excellent broad-band transmission characteristics. Based on this bend, it was possible to design a 90° beam splitter, depicted in Fig. 3.4(b), with minimal reflectance over the same broad frequency range. Fig. 3.4(c) shows a waveguide intersection created for the model system, where it was possible to almost completely eliminate parasitic cross-talk between the waveguides without having to take recourse to high-Q resonances as suggested in Ref. [81]. As a result, the achieved waveguide intersection operates over a reasonable broad range of frequencies. In addition to the here presented basic functional elements for a square lattice system, analogous calculations for a triangular lattice of air pores can be found in [82].

It should be emphasized that the designs discussed above are created out of physical intuition based on the symmetry of the underlying localized single cavity mode, which leads to several devices with nearly optimal transmission properties that all work in a common frequency band.

Further improvements of the above designs may be obtained by applying recently developed methods for PC-based device design optimization that utilize the full numerical efficiency of the WF approach [74] together with the possibilities for fine tuning of device properties provided by the single pore infiltration technique through infiltrating different pores with materials of different refractive index.

Finally, to demonstrate the tunability provided by the single pore infiltration technique a Folded Directional Coupler (FDC) [83] is presented. A schematic of the coupler is depicted in Fig. 3.5(a). This FDC exhibits a sharp resonance with a quality factor $Q > 40,000$. Fig. 3.5(b) shows the change of the resonance frequency with respect to a change $\Delta n = 1.6 \cdot 10^{-3}$ of the refractive index. The resulting tunability may be expressed through the shift $\Delta\lambda/\lambda$ of the resonance wavelength. The obtained numerical results can be cast in form of a linear tuning in Δn: $\Delta\lambda/\lambda \approx 0.13\Delta n$. This tunability is solely based on the tunability of the infiltrated material. For instance, the birefringence of nematic liquid crystals can lead to values for Δn as large as 0.2, corresponding to rather substantial tuning ranges of the resonance frequency of the FDC.

3.5 Design optimization

The aim of the development of any functional element for integrated PC circuits is certainly to find a device design with optimal characteristics for a certain and predefined functional purpose. However, due to the vast number of design possibilities[2] finding a design that meets the required specifications is, using standard simulation tools like FDTD, a long and tedious enterprise, if manageable at all.

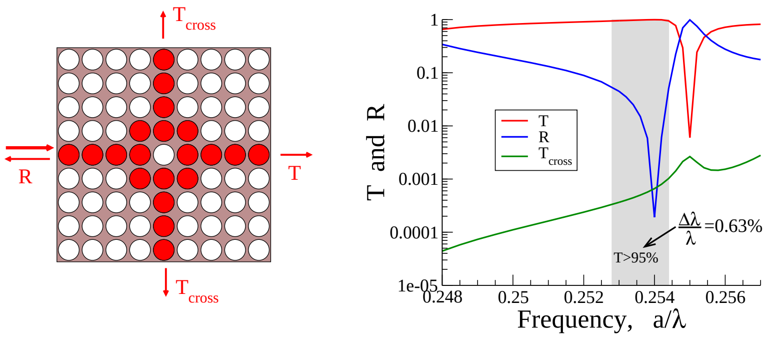
Recent developments utilize the efficiency of the WF approach together with a numerical small rank update technique [74] to provide a highly efficient tool for developing optimal designs meeting predefined performance criteria. By using this technique, it was possible to further improve the design of the waveguide intersection depicted in Fig. 3.4(c) for achieving a high bandwidth (2% of the center frequency) as well as low crosstalk (-40dB) and reflectance (-30dB) [84]. Starting from an initial design, the optimization using small rank adjustments was carried out by allowing each pore of a domain, consisting of 7 by 7 unit cells around the crossing, to be either empty or filled with a low-index material. Due to symmetry, there have been a total of 1024 possible designs (air pore filling patterns). Using a Solaris workstation (400MHz UltraSparc II, 4GB RAM), this extremely efficient technique required roughly 1h to check the transmission properties for all possible designs of such a waveguide crossing, i.e., only 3.5 s are required to check a single pattern. The results for the best design are shown in Fig. 3.6(b). Compared to the initial design, Fig. 3.6(a), the frequency bandwidth of the operational window for which the transmittance exceeds 95% has been more than tripled.

Finally, the achieved optimal design could even be further optimized by allowing the filled pores to be infiltrated with different materials. However, for continuous variations of the refractive indexes between 1.5 and 1.75 the design space is too large to search brute force through all possible configurations. Instead, a recently developed Wannier basis gradients sensitivity analysis [85] was employed, yielding the fine-tuned design depicted in Fig. 3.6(c) which features a transmittance of nearly unity by simultaneously no reduction in the operation bandwidth.

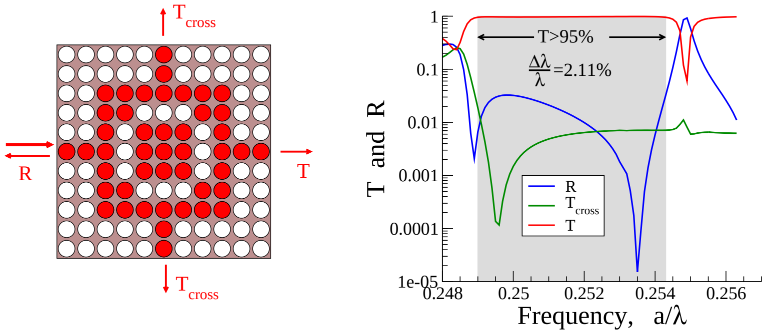
In conclusion, recently developed Wannier based methods like the small-rank update technique [86], can help to speed up already highly efficient WF based numerical calculations even more. For comparison of efficiency it should be mentioned that calculations using the ”bare” Wannier function approach, i.e., without the small rank adjustment method, checking all of the 1024 possible designs would have required about 500h on the same hardware. Using standard FDTD analysis would have required about 10 000h.

---

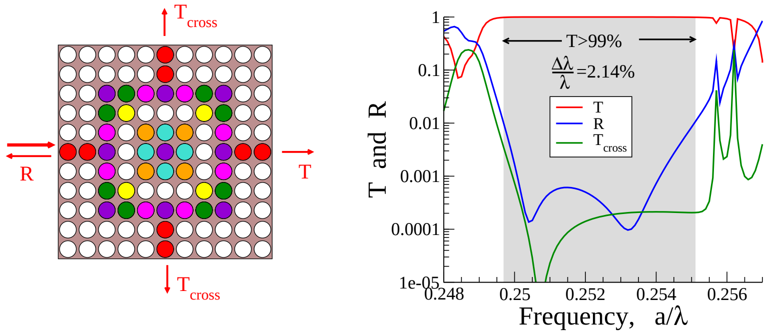
[2] For the example considered here: which pores should be infiltrated and with which material.



(a) Design and transmission characteristics for the initial waveguide intersection.



(b) Design and transmission characteristics for the coarsely optimized intersection.



(c) Design and transmission characteristics for the fine-tuned waveguide intersection.

Figure 3.6: *Transmission, reflection, and crosstalk spectra for three different waveguide crossing designs:* ***(a)*** *the original design;* ***(b)*** *the best design that can be obtained by infilling air pores in 7 by 7 unit cells around the crossing with a low-index material (n = 1.7);* ***(c)*** *the optimized design obtained by refining design (b) through fine-tuning of the refractive indexes of the infiltrated material. The different indexes are color encoded over the range from 1.5 to 1.75: n = 1.75 (violet), 1.74 (green), 1.68 (turquoise), 1.58 (orange), 1.56 (magenta), and 1.5 (yellow). Please note that the optimization procedure was performed for a conceptually equivalent square lattice PC with a pore radius of* $r=0.45a$ *instead of the model system with* $r=0.475a$ *introduced in Sec. 2.3.2. The change of the underlying PC was performed to ease its accessibility to an experimental realization.*

With this last demonstration of the efficiency and accuracy of the Wannier function approach this chapter might be closed with the remark that the entire approach depends crucially on the existence of a set of symmetric and maximally localized Wannier functions, whose construction by means of the newly developed bottom-up approach is illustrated in the following chapters. The set of WFs used for the above described optimization of a waveguide intersection is exhibited in Fig. 3.7 and might contribute to the reader's motivation to learn more about their construction.

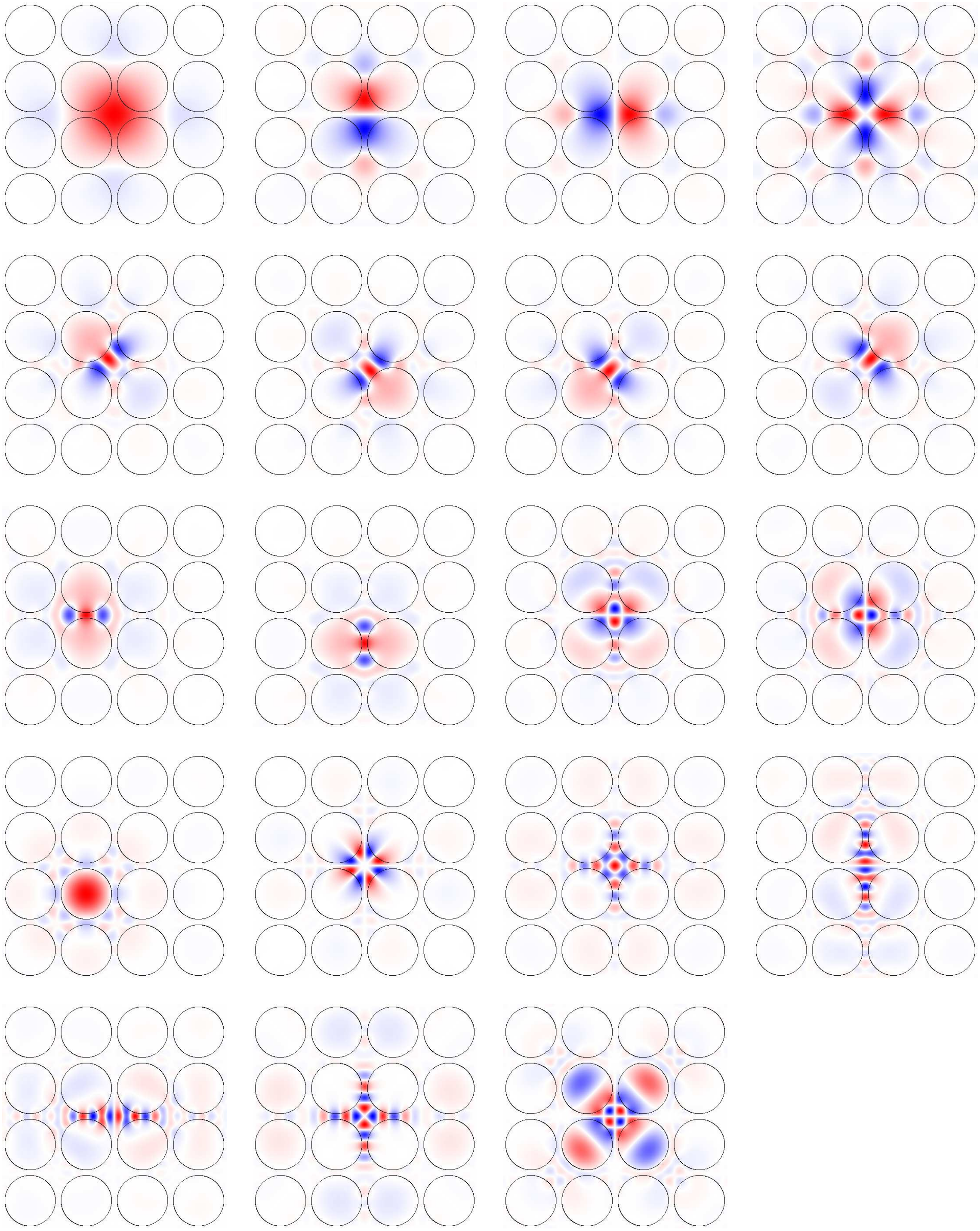


Figure 3.7: *The first 19 Wannier functions which were used for the optimization of the waveguide intersection displayed in Fig. 3.6(c). It is worth mentioning that only the 13th WF has its center inside an air pore. All other functions are centered around positions inside the background material.*

4 Maximally localized Wannier functions for an isolated group of bands

Wannier functions are named after their inventor, Gregory Hugh Wannier, who has introduced this set of localized functions for studying electronic excitations in electronic insulators already in 1937 [43]. However, they have seen only little practical use in computational electronic structure theory until the end of the 20th century. This is mainly caused by their non-uniqueness in definition: even in the case of a single isolated band, it is well known that the corresponding Wannier function is not unique, due to a phase indeterminacy of the underlying Bloch functions. However, for this case, the conditions for obtaining a maximally localized, exponentially decaying WF are known [87, 88]. For the case of several non-isolated bands, an even stronger indeterminacy is given due to the mixing of bands, leading to the definition (2.31) of generalized WFs. It was shown by Des Cloiseaux [89] that achieving well localized and symmetric WFs simultaneously is only possible for a set of such generalized WFs. A closer look on this fact will be taken in Sec. 4.1. For this more complicated case, a numerical method [46], was published in 1997 by N. Marzari and D. Vanderbilt, allowing to obtain maximally localized WFs for an isolated group of bands, i.e. for a group of bands which are crossing themselves but are separated to higher and lower bands by band gaps. In 2002, Souza *et al.* provided in [47] an additional algorithm for disentangling a non-isolated group of bands from higher- and lower-lying bands. This algorithm is designed to be applied as a preprocessing step to the method of Marzari and Vanderbilt.

In the context of research on the optical properties of defect structures in Photonic Crystals, Wannier functions have first been mentioned in Ref. [90] by K. M. Leung in 1993. In 2002 a tight-binding Wannier function method was suggested, where localized approximate WFs are obtained by means of the standard Slater-Koster formalism [91, 92, 93]. The idea of obtaining maximally localized photonic Wannier function according to the method of Marzari and Vanderbilt, and applying those to defect structures embedded in a PC for E-polarized light was made public around 2002 [94, 95]. The construction of some of the lowest lying MLWFs with the above mentioned methods for the H-field was presented in Ref. [96] by Whittaker and Croucher. The complete Wannier function approach, including the calculation of defect modes, dispersion relations of linear waveguides, and transmission characteristics of multiport devices was introduced by Busch *et al.* in Ref. [48].

The aim of this thesis was to transfer the methods described in [46] and [47] to periodically structured dielectric materials, i.e. Photonic Crystals, and to check their applicability for the construction of maximally localized photonic Wannier functions. This transfer process started with the analytical recalculation of [46] and its subsequent implementation in computer codes. Therefore, this chapter is contributed to the construction of maximally localized WFs for the case of an isolated group of bands.

After a brief description of the complications arising from the mixing of bands in Sec. 4.1, the algorithm of Marzari and Vanderbilt is explained in detail. Longer calculations have been sourced-out into several appendices for better readability of the manuscript. Special emphasis should be devoted to Sec. 4.3.3.2 where a new set of initial trial functions is proposed, tremendously easing the calculation of maximally localized WFs. Its usefulness is proven subsequently for the construction of MLWFs for the example system specified in Fig. 2.4(a).

The more difficult case of non-isolated groups of bands will be discussed in Chap. 5. A major complication arising from the fact that in the case of PCs, Wannier functions have to be constructed for bosonic "particles" (photons) is discussed in Chap. 6. In addition, a novel bottom-up approach for the construction of arbitrarily many symmetric and well localized WFs is presented there.

4.1 Entanglement of bands

As already stated in Sec. 2.2.4, Wannier functions are defined as a Lattice Fourier Transform of Bloch functions $\psi_{n\mathbf{k}}(\mathbf{r})$:

$$W_{n\mathbf{R}}(\mathbf{r}) := \frac{1}{\mathrm{V_{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k}\; e^{i\mathbf{k}\mathbf{R}}\, \psi_{n\mathbf{k}}(\mathbf{r}) \,, \tag{4.1}$$

where the involved fields are scalar quantities representing the z-component of either the electric or the magnetic field, as explained in Sec. 2.3, since the main focus of this thesis lies on 2D systems. However, the generalization to 3D is straightforward by just considering these fields as vectorial with bold symbols $\mathbf{W}_{n\mathbf{R}}(\mathbf{r})$, $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$, and $\mathbf{u}_{n\mathbf{k}}(\mathbf{r})$. For this purpose the notation is kept d-dimensional.

Though the definition and, therefore, the construction recipe looks somewhat straightforward, the construction of a set of maximally localized WFs is in practice a rather complicated task [97]. The reasons for this will be pointed out in the following.

As mentioned earlier, the definition of Bloch functions is not unique. There is an indeterminacy with respect to a phase factor $\phi_n(\mathbf{k})$

$$\psi_{n\mathbf{k}}(\mathbf{r}) \rightarrow \psi_{n\mathbf{k}}(\mathbf{r})\, e^{i\phi_n(\mathbf{k})} \,, \tag{4.2}$$

which might be different for every band n and wave vector $\mathbf{k}$. This means that if a Bloch function is multiplied with such a phase factor, the resulting function is still a valid Bloch function . On a first sight such a non-uniqueness does not seem to cause greater difficulties. However, as Eq. (4.1) is not invariant under phase transformations (4.2), one consequently obtains a different WF for each distinct set of phases $\{\phi_n(\mathbf{k})\}$. In addition, all of these Wannier functions have completely different localization properties. But among all those combinations of phases there is one[1] set which leads to a Wannier function that is maximally localized in a certain sense. The practical problem which has to be solved is to find exactly this combination of phases that leads to a MLWF.

Though there exists an analytic way of finding this combination in 1D systems [87, 88], a corresponding method does not exist for higher-dimensional systems[2]. For such systems, there are even further complications arising from the occurrence of crossing and repulsive[3] bands. In Fig. 4.1 this complication is illustrated. It might be obvious that the mode structure or mode profile is not following the band structure anymore. Especially the case of repulsive bands states the most complex problem, as in this case the mode structure is transfered continuously from one mode to the other over the region in **k**-space where the dispersion curves have non-zero and opposite signed curvatures. In contrast to this, crossing bands are only degenerate at certain **k**-points and can be disentangled at all other positions by just interchanging the band indices if required.

For demixing repulsive bands one has to increase the degree of freedom further. This is realized by introducing unitary matrices $U_{mn}^{(\mathbf{k})}$ obeying the unitarity condition

$$U^{(\mathbf{k})^\dagger} U^{(\mathbf{k})} = \mathbb{1} \tag{4.3}$$

which allow linear combination of Bloch functions $\boldsymbol{\psi}_{n\mathbf{k}}(\mathbf{r})$ belonging to the same wave vector **k** according to

$$\widetilde{\psi}_{n\mathbf{k}}(\mathbf{r}) := \sum_{m=1}^{N} U_{mn}^{(\mathbf{k})} \, \psi_{m\mathbf{k}}(\mathbf{r}) \; , \tag{4.4}$$

where the functions $\widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r})$ are called generalized Bloch functions and N is the number of bands that build up a composite group.

For illuminating the meaning of these generalized BFs the Bloch functions displayed in Fig. 4.1 are demixed[4] at different wave vectors **k** as follows: taking the bands at the X-point as reference, i.e.

$$\mathbf{k} = \frac{2\pi}{a}(0.5, 0.5) : \qquad \begin{pmatrix} \widetilde{\psi}_5(\mathbf{r}) \\ \widetilde{\psi}_6(\mathbf{r}) \\ \widetilde{\psi}_7(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \psi_5(\mathbf{r}) \\ \psi_6(\mathbf{r}) \\ \psi_7(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} \psi_5(\mathbf{r}) \\ \psi_6(\mathbf{r}) \\ \psi_7(\mathbf{r}) \end{pmatrix} ,$$

---

[1] Up to permutations of band indices and for requiring that the WF should be centered in the WSC around the origin.

[2] Reasons for this are given in [46].

[3] Sometimes also called avoided crossings.

[4] Not considering the problem of a proper phase choice $e^{i\phi_{n\mathbf{k}}}$ for each function.

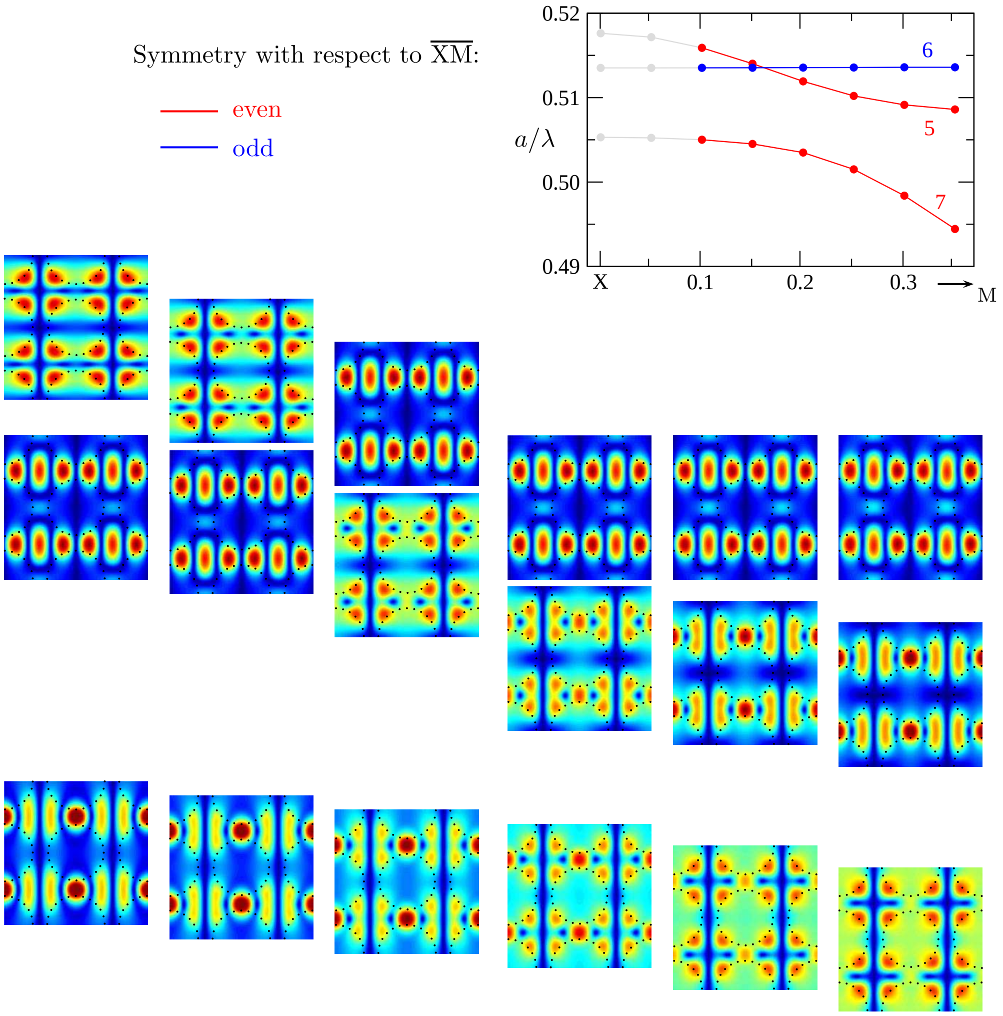


Figure 4.1: *Absolute values of Bloch functions* $|\psi_{n\mathbf{k}}(\mathbf{r})|$ *of bands* $5-7$ *of a sample square lattice PC plotted along the* $\overline{XM}$ *direction, i.e.* $\mathbf{k} = (0.5, k_y)$ *with* $k_y = 0 \ldots 0.5$ *in units of* $2\pi/a$*. The arrangement of the function plots is chosen to reflect the corresponding dispersion relation* $\omega_n(\mathbf{k})$ *for each band, depicted in the right upper inset as a function of* $\mathbf{k} = (0.5, k_y)$*. Starting at the* X*-point, i.e.* $k_y = 0$*, first the sixth and seventh band cross each other at* $k_y \approx 0.15$*, afterwards the former band seven and band five are repelling each other over the region* $k_y \approx 0.2 - 0.3$ *with a continuous transfer of the mode structure between these modes. Close to the* M*-point band five carries now the mode structure of the former seventh band, the sixth band corresponds to the former fifth band, and band seven shows the mode structure of the former band six. For more details on crossing and repulsive bands see Ref. [98].*

the bands close to the M-point in Fig. 4.1 could be demixed by the unitary transformation[5]

$$\mathbf{k} = \frac{2\pi}{a}(0.5, 0.5) : \qquad \begin{pmatrix} \widetilde{\psi}_5(\mathbf{r}) \\ \widetilde{\psi}_6(\mathbf{r}) \\ \widetilde{\psi}_7(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_5(\mathbf{r}) \\ \psi_6(\mathbf{r}) \\ \psi_7(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} \psi_7(\mathbf{r}) \\ \psi_5(\mathbf{r}) \\ \psi_6(\mathbf{r}) \end{pmatrix} .$$

For a wave vector between these k-points the unitary transformation would have the form

$$\begin{pmatrix} \widetilde{\psi}_5(\mathbf{r}) \\ \widetilde{\psi}_6(\mathbf{r}) \\ \widetilde{\psi}_7(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} c & 0 & \sqrt{1-c^2} \\ -\sqrt{1-c^2} & 0 & c \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_5(\mathbf{r}) \\ \psi_6(\mathbf{r}) \\ \psi_7(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} c\,\psi_5(\mathbf{r}) + \sqrt{1-c^2}\,\psi_7(\mathbf{r}) \\ -\sqrt{1-c^2}\,\psi_5(\mathbf{r}) + c\,\psi_7(\mathbf{r}) \\ \psi_6(\mathbf{r}) \end{pmatrix} ,$$

with an in general complex number c. In this way it could be reached that BFs having similar mode profiles share the same band index n. This means, for the generalized BFs the mode and band labeling coincide. When applying now the LFT (2.29) on these functions

$$\mathbf{W}_{n\mathbf{R}}(\mathbf{r}) = \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k}\; e^{-i\mathbf{k}\mathbf{R}}\; \widetilde{\boldsymbol{\psi}}_{n\mathbf{k}}(\mathbf{r}) \tag{4.5}$$

$$= \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k}\; e^{-i\mathbf{k}\mathbf{R}} \sum_{m=1}^{N} U_{mn}^{(\mathbf{k})}\, \boldsymbol{\psi}_{m\mathbf{k}}(\mathbf{r}) , \tag{4.6}$$

the resulting Wannier functions, $W_{n\mathbf{R}}(\mathbf{r})$, are inherently symmetric and with a proper phase choice highly localized. For simplicity the same notation is used here as for the Wannier functions obtained by the direct lattice Fourier transform defined in (2.29).

Concluding, for higher-dimensional systems not a set of phases has to be determined numerically, but a combination of unitary matrices $\{U_{mn}^{(\mathbf{k})}\}$ that yields a set of maximally localized Wannier functions for a group of N composite bands simultaneously. How this matrices can be calculated will be explained in detail in the following sections. Since the main goal is to achieve MLWFs, one first has to look for an appropriate measure of localization which will be derived subsequently.

4.2 Measure of localization

In order to obtain a set of maximally localized Wannier functions a reliable measure of localization is required. A natural choice for a single Wannier function with band

---

[5] Eq. 4.4 is for the following examples denoted as a standard matrix-vector product, where the summation is carried out over the second dimension of the matrix. Therefore, to get matrices $U_{mn}^{(\mathbf{k})}$ complying Eq. 4.4 the transposed of the displayed matrices have to be taken.

index n and centered inside the central WSC, i.e. $\mathbf{R} = \mathbf{0}$, might be its moment of second order or variance

$$\sigma_n^2 = \langle r^2 \rangle_n - \mathbf{r}_n^2 \tag{4.7}$$

with

$$\mathbf{r}_n := \langle \mathbf{r} \rangle_n := \big\langle W_{n\mathbf{0}} \big| \mathbf{r} \big| W_{n\mathbf{0}} \big\rangle \qquad \text{and} \qquad \langle r^2 \rangle_n := \big\langle W_{n\mathbf{0}} \big| r^2 \big| W_{n\mathbf{0}} \big\rangle, \tag{4.8}$$

where the scalar products are defined as indicated in (2.28). As is well known from statistics, the variance measures the width of a probability distribution around its average value. Here, σ_n^2 measures the spread of the Wannier function $W_{n\mathbf{0}}(\mathbf{r})$ around its center $\mathbf{r}_n$. A smaller spread σ_n^2 indicates better localization. Hence, a maximally localized function is the one having the smallest spread among all considered functions.

As pointed out earlier, for higher-dimensional systems one has to deal with a group of N bands which have to be disentangled. For such a situation a reasonable measure of localization is given by the functional

$$\Omega = \sum_{n=1}^{N} \sigma_n^2 = \sum_{n=1}^{N} \Big[\langle r^2 \rangle_n - \mathbf{r}_n^2 \Big] , \tag{4.9}$$

which is just the sum of the spreads of all N Wannier functions considered. Recalling that the generalized Wannier functions depend on the unitary matrices $U_{mn}^{(\mathbf{k})}$ it's evident that the spread functional is a function of these matrices, too. To emphasize this, one may write

$$\Omega := \Omega\big[\{ U_{mn}^{(\mathbf{k})} \}\big] . \tag{4.10}$$

Based on this notation a mathematical criterion for the determination of the set of unitary matrices that lead to maximally localized Wannier functions is readily available. It is exactly the set which solves the minimum problem

$$\Omega\big[\{ U_{mn}^{(\mathbf{k})} \}\big] = \text{Min.} \tag{4.11}$$

The solution of this multidimensional non-linear minimization problem has to be calculated numerically and the corresponding procedure will be outlined subsequently.

4.3 Minimization of the spread functional

As pointed out already earlier, the conditions for reaching the global minimum of (4.11) can be given analytically only for certain systems[6] like an isolated band in 1D [87]. In general Eq. (4.11) has to be solved numerically. For such a numerical treatment one has to derive a discrete formulation of (4.9) that can be minimized by means of computer algorithms.

---

[6] Please see Sec. IV.C and Sec. V.A in Ref. [46] for details.

4.3.1 Discrete formulation

The easiest way of which one could think of to solve (4.11) is just to start with a certain set of unitary matrices $U^{(\mathbf{k})}_{mn}$, construct the corresponding generalized Bloch functions and from these the generalized Wannier functions. Afterwards calculate the value of the spread functional and compare with results of other functions constructed already. This procedure could be continued until the global minimum is found. However, the LFT (4.5) is quite time consuming and an algorithm based on this procedure would be somewhat inefficient. On the other hand, the Wannier functions are connected to the Bloch functions via (4.5). Therefore, an expression for the spread functional (4.9) in terms of Bloch functions $\psi_{n\mathbf{k}}(\mathbf{r})$ or their periodic parts $u_{n\mathbf{k}}(\mathbf{r})$ might be derivable.

Indeed, as shown by Blount in Ref. [99], matrix elements of the position operator between Wannier functions take the form

$$\langle \mathbf{R}m|\mathbf{r}|\mathbf{0}n\rangle = \frac{i}{V_{\mathrm{BZ}}}\int\limits_{\mathrm{BZ}} d^d\mathbf{k}\; e^{i\mathbf{kR}}\, \langle u_{m\mathbf{k}}|\nabla_{\mathbf{k}}|u_{n\mathbf{k}}\rangle \;, \tag{4.12}$$

$$\langle \mathbf{R}m|r^2|\mathbf{0}n\rangle = -\frac{1}{V_{\mathrm{BZ}}}\int\limits_{\mathrm{BZ}} d^d\mathbf{k}\; e^{i\mathbf{kR}}\, \langle u_{m\mathbf{k}}|\nabla^2_{\mathbf{k}}|u_{n\mathbf{k}}\rangle \;, \tag{4.13}$$

where the operator $\nabla_{\mathbf{k}}$ is understood to act to the right only, i.e. only on the ket. Using the above, one can reexpress the matrix elements occurring in the spread functional by

$$\mathbf{r}_n = \frac{i}{V_{\mathrm{BZ}}}\int\limits_{\mathrm{BZ}} d^d\mathbf{k}\; \langle u_{n\mathbf{k}}|\nabla_{\mathbf{k}}|u_{n\mathbf{k}}\rangle \;, \tag{4.14}$$

$$\langle r^2\rangle_n = \frac{1}{V_{\mathrm{BZ}}}\int\limits_{\mathrm{BZ}} d^d\mathbf{k}\; \big||\nabla_{\mathbf{k}} u_{n\mathbf{k}}\rangle\big|^2 \;. \tag{4.15}$$

Eq. (4.15) follows from Eq. (4.13) after integration by parts and utilizing the geometric properties of the first Brillouin zone. A detailed derivation is given in Appendix A.1. In practice, the integration over the Brillouin zone has to be approximated by a finite sum over a set of N_{kp} discrete $\mathbf{k}$-points

$$\frac{1}{V_{\mathrm{BZ}}}\int\limits_{\mathrm{BZ}} \longrightarrow \frac{1}{N_{kp}}\sum_{\mathbf{k}} \tag{4.16}$$

where a uniformly spaced Monkhorst-Pack mesh [66] is used for the discretization of the Brillouin zone. For this kind of BZ sampling the LFT of Bloch functions reads

$$W_{n\mathbf{R}}(\mathbf{r}) = \frac{1}{N_{\mathrm{kp}}}\sum_{\mathbf{k}} e^{i\mathbf{kR}}\, \psi_{n\mathbf{k}}(\mathbf{r}) \;. \tag{4.17}$$

Using the translational properties of WFs, it is sufficient to construct the WFs at the center cell $\mathbf{R} = \mathbf{0}$

$$W_{n\mathbf{0}}(\mathbf{r}) = \frac{1}{N_{\mathrm{kp}}} \sum_{\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}) \tag{4.18}$$

and from these the ones at other lattice sites via Eq. (2.34). In addition the gradient $|\nabla_{\mathbf{k}} u_{n\mathbf{k}}\rangle$ has to be expressed by finite differences as well. For this purpose the discretization scheme

$$\nabla_{\mathbf{k}} f(\mathbf{k}) = \sum_{\mathbf{b}} w_b \mathbf{b} \left[f(\mathbf{k}+\mathbf{b}) - f(\mathbf{k}) \right] \tag{4.19}$$

$$\left| \nabla_{\mathbf{k}} f(\mathbf{k}) \right|^2 = \sum_{\mathbf{b}} w_b \left[f(\mathbf{k}+\mathbf{b}) - f(\mathbf{k}) \right]^2 \tag{4.20}$$

is used, where the vector $\mathbf{b}$ points from $\mathbf{k}$ to its nearest neighbors and w_b is a properly chosen weight factor obeying $\sum_{\mathbf{b}} w_b b_\alpha \, b_\beta = \delta_{\alpha\beta}$ and depending only on $b = |\mathbf{b}|$. Using the above substitutions the discretized versions of Eqs. (4.14) and (4.15) read

$$\mathbf{r}_n = \frac{i}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \mathbf{b} \Big[\langle u_{n\mathbf{k}} | u_{n\mathbf{k}+\mathbf{b}} \rangle - 1 \Big] \tag{4.21}$$

$$\langle r^2 \rangle_n = \frac{1}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \Big[2 - 2\,\mathrm{Re} \langle u_{n\mathbf{k}} | u_{n\mathbf{k}+\mathbf{b}} \rangle \Big] \; . \tag{4.22}$$

In the continuum limit ($N_{kp} \to \infty, \; b \to 0$) these expressions reduce to Eq. (4.14) and Eq. (4.15), respectively. However, using these the spread functional Ω is not invariant under the special phase transformation

$$u_{n\mathbf{k}}(\mathbf{r}) \to u_{n\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{R}} \; , \quad \text{for all } \mathbf{k}, \tag{4.23}$$

which should lead to a translation of the Wannier center $\mathbf{r}_n \to \mathbf{r}_n + \mathbf{R}$ only, not affecting the extent of the function and, therefore, leaving the spread unchanged. Since this invariance condition is crucial for the development of a stable algorithm for the spread minimization —as the spread of a function should remain the same irrespective of the WSC it is located in— the discretized expressions (4.21) and (4.22) are ill-suited.

Due to this fact, other finite difference expressions have to be found that guarantee invariance under (4.23) and coincide with expressions (4.14) and (4.15) to first orders in the mesh spacing b.

For convenience one might define an element of the matrix of overlap integrals as

$$M_{mn}^{(\mathbf{k},\mathbf{b})} := \langle u_{m\mathbf{k}} | u_{n\mathbf{k}+\mathbf{b}} \rangle \tag{4.24}$$

and expand the diagonal elements up to second order in b

$$M_{nn}^{(\mathbf{k},\mathbf{b})} = 1 + ixb + \frac{1}{2} y b^2 + \mathcal{O}\left(b^3\right) \tag{4.25}$$

where x and y, following [46], are supposed to be real numbers. However, it was possible to prove this fact for x only[7]. Interestingly, the algorithm works irrespective of y being real or not, as only its real part enters the equations. Adopting the above abbreviation for the scalar products, Eqs. (4.21) and (4.22) take the form:

$$\begin{aligned} \mathbf{r}_n &= \frac{i}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \mathbf{b} \left[M_{nn}^{(\mathbf{k},\mathbf{b})} - 1 \right] && (4.26) \\ \langle r^2 \rangle_n &= \frac{1}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \left[2 - 2\,\mathrm{Re} M_{nn}^{(\mathbf{k},\mathbf{b})} \right] . && (4.27) \end{aligned}$$

Using (4.25) it is obvious that the expression in brackets in Eq. (4.26) can be expanded as[8]

$$M_{nn}^{(\mathbf{k},\mathbf{b})} - 1 = ixb + \mathcal{O}\left(b^2\right) . \tag{4.28}$$

One might remember here that for a complex number z, $\ln(z) \approx z - 1$ around $z = 1$ and, therefore,

$$\ln M_{nn}^{(\mathbf{k},\mathbf{b})} \approx M_{nn}^{(\mathbf{k},\mathbf{b})} - 1 . \tag{4.29}$$

Combining the above with Eq. (4.28) yields

$$\ln M_{nn}^{(\mathbf{k},\mathbf{b})} = ixb + \mathcal{O}\left(b^2\right) . \tag{4.30}$$

One can identify

$$xb = \mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} , \tag{4.31}$$

which leads directly to

$$i \; \mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} = ixb + \mathcal{O}\left(b^2\right) . \tag{4.32}$$

This expression shows exactly the same scaling behavior for small b as Eq. (4.28). Hence, Eq. (4.26) might be recast to

$$\mathbf{r}_n = -\frac{1}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \mathbf{b} \left[\mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} \right] , \tag{4.33}$$

where it is worth noting that the Wannier centers $\mathbf{r}_n$ are directly related to the complex phases $\arg(M_{nn}^{(\mathbf{k},\mathbf{b})}) = \mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})}$ of the overlap matrix elements $M_{nn}^{(\mathbf{k},\mathbf{b})}$.

Similar to Eq. (4.30), the expression in brackets in Eq. (4.27) can be expanded as

$$2 - 2\,\mathrm{Re}\, M_{nn}^{(\mathbf{k},\mathbf{b})} = -\,\mathrm{Re}[y]\, b^2 + \mathcal{O}\left(b^3\right) . \tag{4.34}$$

---

[7] The proof and additional comments are given in Appendix A.2. In principle, x and y should be labeled with a band index n. However, for compliance with the original notation in [46] the index is dropped.

[8] The spread functional Ω contains terms of $\mathbf{r}_n^2$, so it's sufficient to consider terms up to first order in b only.

To find an expression that scales in the same way, the expansion

$$\begin{aligned}\left|M_{nn}^{(\mathbf{k},\mathbf{b})}\right|^2 &= M_{nn}^{(\mathbf{k},\mathbf{b})*} M_{nn}^{(\mathbf{k},\mathbf{b})} \\ &= (1 - ixb + \frac{1}{2} y^* b^2 + \cdots)(1 + ixb + \frac{1}{2} y b^2 + \cdots) \\ &= 1 + \mathrm{Re}[y]\, b^2 + x^2 b^2 + \mathcal{O}\left(b^3\right) \qquad (4.35)\end{aligned}$$

is examined. Using

$$x^2 b^2 = \left[\mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})}\right]^2$$

one might identify

$$1 - \left|M_{nn}^{(\mathbf{k},\mathbf{b})}\right|^2 + \left[\mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})}\right]^2 = -\,\mathrm{Re}[y]\, b^2 + \mathcal{O}\left(b^3\right) \ , \qquad (4.36)$$

which shows the same scaling behavior as (4.34). Thus the second moments might be reexpressed through

$$\left\langle r^2 \right\rangle_n = \frac{1}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \left\{ 1 - \left|M_{nn}^{(\mathbf{k},\mathbf{b})}\right|^2 + \left[\mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})}\right]^2 \right\} . \qquad (4.37)$$

Testing the behavior of Eq. (4.33) and (4.37) under the transformation (4.23) yields[9]

$$\begin{aligned}\mathbf{r}_n &\xrightarrow{(4.23)} \mathbf{r}_n + \mathbf{R} \qquad (4.38) \\ \left\langle r^2 \right\rangle_n &\xrightarrow{(4.23)} \left\langle r^2 \right\rangle_n + 2\mathbf{r}_n \mathbf{R} + R^2 \ , \qquad (4.39)\end{aligned}$$

which implies

$$\Omega \xrightarrow{(4.23)} \Omega \ . \qquad (4.40)$$

Hence, by carrying out the replacements

$$\begin{aligned}M_{nn}^{(\mathbf{k},\mathbf{b})} - 1 &\xrightarrow{\mathcal{O}\left(b^2\right)} i \ \mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} \ , \qquad (4.41a) \\ 2 - 2\,\mathrm{Re}\, M_{nn}^{(\mathbf{k},\mathbf{b})} &\xrightarrow{\mathcal{O}\left(b^3\right)} 1 - \left|M_{nn}^{(\mathbf{k},\mathbf{b})}\right|^2 + \left[\mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})}\right]^2 , \qquad (4.41b)\end{aligned}$$

a formulation of the spread functional Ω has been found that is invariant under the transformation (4.23). At this point one should emphasize, however, that Ω is invariant under (4.23), but, of course, not under a general unitary transformation (4.4). This feature will be used in the following section to minimize the spread functional numerically.

---

[9] For details please see Appendix A.3.

4.3.2 Steepest descent minimization of the spread functional

For solving an unconstrained highly non-linear optimization problem like Eq. (4.11), several distinct methods exist [100]. One of those is the Steepest Descent (SD) method, also known as gradient descent, which will later be used to minimize Eq. (4.9). To give a short overview about available methods, some of the most prominent algorithms for solving unconstrained non-linear minimization problems are listed subsequently:

- **Steepest Descent**
  Requires the computation of the gradient, either analytically or numerically. Converges to the minimum as long as the initial guess is sufficiently close to it.
- **Conjugate Gradients**
  Converges in general much faster than the SD, especially for narrow valley structures. However, for non-linear optimization problems not only the gradient, but also the Hessian matrix have to be computed.
- **Simulated Annealing**
  This method is mostly applied to problems where an analytical expression for the gradient of the functional is not derivable [101].

The first two methods might be combined with line searches along the gradient directions to find an optimal step width. Since an analytical expression for the gradient of Eq. (4.9) is derivable, as will be shown later in Sec. 4.3.2.2, the gradient methods are favorable. The advantage of non-linear Conjugate Gradients over SD is that the convergence speed is higher. However, the computational efforts which have to be made are higher, too, as higher order derivatives have to be taken into account. In some cases the analytical derivation of the Hessian matrix might be cumbersome. If at the same time the computational expenses for a single iteration step are rather low, a larger number of iteration steps might easily be tolerated in favor of prevention from additional analytical acrobatics. As this situation is given for the problem considered, SD is the method of choice.

In the next section the method of Steepest Descent will be described in a more detailed fashion and illustrated by means of a simple scalar function as an example.

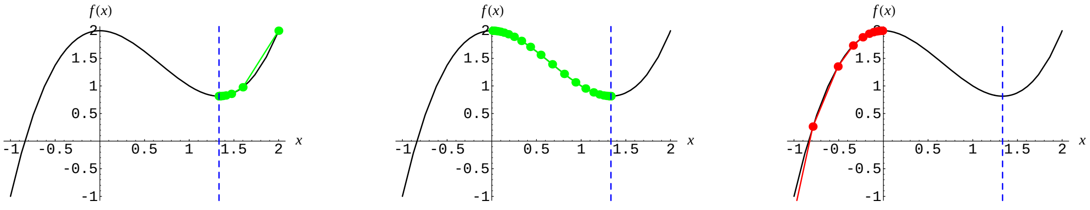
4.3.2.1 Method of Steepest Descent

As already stated, the method of Steepest Descent is an algorithm for finding the nearest local minimum of a function which presupposes that the gradient of the function can be computed. The SD algorithm starts at a point $\mathbf{x}_0$ and, as many times as needed, moves from $\mathbf{x}_i$ to $\mathbf{x}_{i+1}$ by minimizing along the line extending from $\mathbf{x}_i$ in the direction of $-\nabla f(\mathbf{x}_i)$, called the local downhill gradient.

When applied to a one-dimensional function, the method takes the form of iterating

$$x_{i+1} = x_i - \alpha f'(x_i) \,, \tag{4.42}$$

starting at an initial point x_0 for some small $\alpha > 0$ until a fixed point is reached. As an illustration of this procedure Fig. 4.2 shows the graph of a sample third



(a) $x_0 = 2$ (b) $x_0 = 0.01$ (c) $x_0 = -0.01$

Figure 4.2: *Illustration of the gradient descent for the simple example of a third order polynomial* $f(x) = x^3 - 2x^2 + 2$*. The graph of* $f(x)$ *is the continuous black line. The red and green dots indicate function values* $f(x_i)$*, where the* x_i *are obtained through several iteration steps of Eq.* (4.42)*. The finite step width was chosen to be* $\alpha = 0.1$*. Figs. 4.2(a)–4.2(c) show the iteration procedure for different initial values* x_0*. While the iteration procedures shown in Fig. 4.2(a) and 4.2(b) successfully converge to the local minimum at* $x_{min} = \frac{4}{3}$ *the gradient descent in 4.2(c) fails, since the initial value* $x = -0.01$ *is not lying in the basin of attraction of the local minimum.*

order polynomial, $f(x)$, as well as the function values $f(x_i)$ (filled circles), where the points x_i are obtained by iterating Eq. (4.42), but for different initial values x_0. It emphasizes that the choice of good initial values is crucial for the successful application of this method. In addition, the number of iteration steps required, depends significantly on the choice of initial values. The closer x_0 is to the minimum, the less iteration steps have to be performed. In Sec. 4.3.3 the construction of good initial guesses for the minimization of Eq. (4.9) will be discussed and a novel set of initial trial functions will be introduced that achieves for certain systems[10] that the initial set of unitary matrices $\{U_{mn}^{(\mathbf{k})(0)}\}$ already coincides with the solution $\{U_{mn}^{(\mathbf{k})(\mathrm{min})}\}$ of (4.11). But first the derivation of the gradient of the spread functional is carried out in the next section.

4.3.2.2 Gradient of the spread functional

As mentioned in the section before, a Steepest Descent approach to solving the functional (4.9) prerequisites an analytical expression for the gradient. To obtain the desired expression one first reexpresses the unitary matrices as

$$U^{(\mathbf{k})} = e^{W^{(\mathbf{k})}} , \tag{4.43}$$

with anti-hermitian matrices ${W^{(\mathbf{k})}}^{\dagger} = -W^{(\mathbf{k})}$. The gradient will be determined with respect to this newly introduced matrices $W^{(\mathbf{k})}$. The change $d\Omega$ of the spread

---

[10] It works especially well for E-polarization.

functional Ω induced by infinitesimal small changes $dW^{(\mathbf{k})}_{mn}$ is given by

$$d\Omega = \sum_{\mathbf{k}} \sum_{m,n} \frac{d\Omega}{dW^{(\mathbf{k})}_{mn}} \, dW^{(\mathbf{k})}_{mn} \, . \tag{4.44}$$

Using the definition

$$\frac{d\Omega}{dW^{(\mathbf{k})}_{mn}} =: \left(\frac{d\Omega}{dW^{(\mathbf{k})}} \right)^{*}_{mn} = G^{(\mathbf{k})*}_{mn} \tag{4.45}$$

which differs from the one chosen in [46], the change of the spread functional can be conveniently expressed as

$$\begin{aligned} d\Omega &= \sum_{\mathbf{k}} \sum_{m,n} G^{(\mathbf{k})*}_{mn} \, dW^{(\mathbf{k})}_{mn} \\ &= \sum_{\mathbf{k}} \mathrm{Tr}\Big(G^{(\mathbf{k})\dagger} \, dW^{(\mathbf{k})} \Big) \, . \end{aligned} \tag{4.46}$$

To get an idea how an expression for the gradient might be derived, a simple scalar function

$$f : \mathbf{r} \rightarrow f(\mathbf{r}) \tag{4.47}$$

is considered as an example at first. Calculating the first order change of the function $f(\mathbf{r})$ with respect to an infinitesimal change in space $\mathbf{r} \rightarrow \mathbf{r} + d\mathbf{r}$

$$f(\mathbf{r} + d\mathbf{r}) = f(\mathbf{r}) + \nabla f(\mathbf{r}) \, d\mathbf{r} + \mathcal{O}\left(d\mathbf{r}^2 \right) \tag{4.48}$$

the gradient can be determined by relating first orders in $d\mathbf{r}$. As an illustration the example $f : \mathbf{r} \rightarrow \mathbf{r}^2$ is picked. Applying a small change in space leads to

$$f(\mathbf{r} + d\mathbf{r}) = (\mathbf{r} + d\mathbf{r})^2 = \mathbf{r}^2 + 2\mathbf{r} \, d\mathbf{r} + \mathcal{O}\left(d\mathbf{r}^2 \right) \tag{4.49}$$

and the gradient can be read off as $\nabla f(\mathbf{r}) = 2\,\mathbf{r}$. The derivation of an expression for the gradient of the spread functional Ω is now performed in complete analogy to the above example.

After identifying the first order change of the unitary matrices $U^{(\mathbf{k})}_{mn}$ to an infinitesimal small change $W^{(\mathbf{k})}_{mn} \rightarrow W^{(\mathbf{k})}_{mn} + dW^{(\mathbf{k})}_{mn}$ as[11]

$$dU^{(\mathbf{k})}_{mn} = \delta_{mn} + dW^{(\mathbf{k})}_{mn} \tag{4.50}$$

the change of the periodic part, $u_{n\mathbf{k}}(\mathbf{r})$ is given by

$$u_{n\mathbf{k}}(\mathbf{r}) \rightarrow u_{n\mathbf{k}}(\mathbf{r}) + \sum_{m} dW^{(\mathbf{k})}_{mn} u_{m\mathbf{k}}(\mathbf{r}) \, . \tag{4.51}$$

---

[11] Using linearization of the exponential.

After inserting (4.51) into the definition of the spread functional (4.9), a longer calculation, carried out in Appendix A.5, leads to the desired first order change of the spread

$$d\Omega = \sum_{\mathbf{k}} \mathrm{Tr}\left\{ \frac{4}{N_{\mathrm{kp}}} \sum_{\mathbf{b}} w_b \Big(\mathcal{A}\left[R^{(\mathbf{k},\mathbf{b})}\right] - \mathcal{S}\left[T^{(\mathbf{k},\mathbf{b})}\right] \Big)\, dW^{(\mathbf{k})} \right\} , \tag{4.52}$$

where the following definitions have been made

$$R_{mn}^{(\mathbf{k},\mathbf{b})} := M_{mn}^{(\mathbf{k},\mathbf{b})}\, M_{nn}^{(\mathbf{k},\mathbf{b})^*} \tag{4.53}$$

$$T_{mn}^{(\mathbf{k},\mathbf{b})} := \frac{M_{mn}^{(\mathbf{k},\mathbf{b})}}{M_{nn}^{(\mathbf{k},\mathbf{b})}}\, q_n^{(\mathbf{k},\mathbf{b})} \tag{4.54}$$

$$q_n^{(\mathbf{k},\mathbf{b})} := \mathrm{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} + \mathbf{b}\,\mathbf{r}_n , \tag{4.55}$$

and abbreviations for anti-symmetric and symmetric combinations of matrices have been introduced according to

$$\mathcal{A}\,[B] := \frac{1}{2}\big\{ B - B^\dagger \big\} , \tag{4.56}$$

$$\mathcal{S}\,[B] := \frac{1}{2i}\big\{ B + B^\dagger \big\} . \tag{4.57}$$

By comparison with Eq. (4.46) an expression for the gradient is finally given by

$$G^{(\mathbf{k})} = -\frac{4}{N_{\mathrm{kp}}} \sum_{\mathbf{b}} w_b \Big(\mathcal{A}\left[R^{(\mathbf{k},\mathbf{b})}\right] - \mathcal{S}\left[T^{(\mathbf{k},\mathbf{b})}\right] \Big) , \tag{4.58}$$

which differs from the one given in [46], apart from the sign, by a factor of N_{kp}^{-1}, which leads later to a different definition of the finite step width α for a single descent step.

4.3.2.3 Single descent step and optimal step width

Since an expression for the gradient is now available, a single step of the gradient descent is demonstrated in the following. Later, in 4.3.2.4, the complete iterative procedure will be described.

By introducing an infinitesimal change along the gradient direction, i.e.

$$dW^{(\mathbf{k})} = \epsilon G^{(\mathbf{k})} , \tag{4.59}$$

where ϵ is a positive infinitesimal, one realizes that the first order change $d\Omega$ is positive in this case due to

$$\begin{aligned} d\Omega &= \sum_{\mathbf{k}} \mathrm{Tr}\left[G^{(\mathbf{k})^\dagger}\, \epsilon\, G^{(\mathbf{k})} \right] \\ &= \epsilon \sum_{\mathbf{k}} \left\| G^{(\mathbf{k})} \right\|^2 , \end{aligned} \tag{4.60}$$

and, therefore, by using the definition (4.45) the gradient $G^{(\mathbf{k})}$ points in the direction of the strongest increase of Ω. Here, the latter equation was derived by inserting (4.59) into (4.46) and $\|\cdot\|$ denotes the Frobenius norm. In practice, a finite step

$$\Delta W^{(\mathbf{k})} = -\alpha \frac{N_{\mathrm{kp}}}{4w} \, G^{(\mathbf{k})} \; , \tag{4.61}$$

with $w = \sum_{\mathbf{b}} w_b$ against the gradient direction is performed to achieve a decrease of Ω. The normalization[12] was chosen such that α can generally be chosen to lie within the interval $(0, 1]$. Stable convergence of the algorithm was always obtained for $\alpha = 0.5$. If the number of k-points, N_{kp}, is large enough, values closer to 1 can speed up convergence slightly.

The question which arises at this point is how a finite change $\Delta W^{(\mathbf{k})}$ affects the unitary matrix $U^{(\mathbf{k})}$. Since $\Delta W^{(\mathbf{k})}$ and $W^{(\mathbf{k})}$ commute , i.e.

$$\left[W^{(\mathbf{k})}, \Delta W^{(\mathbf{k})} \right] = 0 \; , \tag{4.62}$$

the usage of the well-known Campbell-Baker-Hausdorff formula yields that the matrix $U^{(\mathbf{k})}$ transforms under $W^{(\mathbf{k})} \rightarrow W^{(\mathbf{k})} + \Delta W^{(\mathbf{k})}$ like

$$U^{(\mathbf{k})} \rightarrow e^{W^{(\mathbf{k})} + \Delta W^{(\mathbf{k})}} = e^{W^{(\mathbf{k})}} e^{\Delta W^{(\mathbf{k})}} = U^{(\mathbf{k})} \, \Delta U^{(\mathbf{k})} \; , \tag{4.63}$$

and, therefore, the change of the unitary matrix $\Delta U^{(\mathbf{k})}$ is just given by

$$\Delta U^{(\mathbf{k})} = \exp(\Delta W^{(\mathbf{k})}) \; . \tag{4.64}$$

The above equation could have been obtained equally well through the limiting procedure

$$\Delta U^{(\mathbf{k})} = \lim_{n \rightarrow \infty} \left(\mathbb{1} + \frac{\Delta W^{(\mathbf{k})}}{n} \right)^n = \exp(\Delta W^{(\mathbf{k})}) \; . \tag{4.65}$$

Either way, one finally arrives at the following update equations

$$U^{(\mathbf{k})} \mapsto U^{(\mathbf{k})} \exp(\Delta W^{(\mathbf{k})}) \; , \tag{4.66}$$

$$M^{(\mathbf{k},\mathbf{b})} \mapsto U^{(\mathbf{k})\dagger} M^{(\mathbf{k},\mathbf{b})^{(0)}} U^{(\mathbf{k}+\mathbf{b})} \; , \tag{4.67}$$

for performing a single descent step against the gradient direction $G^{(\mathbf{k})}$. For updating the inner product matrices $M_{mn}^{(\mathbf{k},\mathbf{b})}$ the "rotated" matrices $U^{(\mathbf{k})}$ from Eq. (4.66) are used.

---

[12] With the above definition of the gradient (4.58) and the chosen step width $\epsilon = -\alpha \frac{N_{\mathrm{kp}}}{4w}$ the first order change given in Eq. (4.60) scales as $d\Omega \propto \frac{1}{N_{\mathrm{kp}}} \sum_{\mathbf{k}}$ whereas the same equation would yield $d\Omega \propto \sum_{\mathbf{k}}$ using the corresponding expressions given in [46]. Latter behavior would imply that $|d\Omega|$ increases steadily with increasing number of k-points, N_{kp}, and diverges, therefore, in the continuum limit.

4.3.2.4 Iterative procedure

After clarifying how a single descent step, that leads to a decrease in Ω, can be performed, the complete iterative procedure of the gradient descent is sketched subsequently.

1. At the beginning, i.e. at iteration step $i = 0$, all the matrix elements

$$M_{mn}^{(\mathbf{k},\mathbf{b})\,(0)} = \langle u_{m\mathbf{k}}^{(0)} | u_{n\mathbf{k}+\mathbf{b}}^{(0)} \rangle \tag{4.68}$$

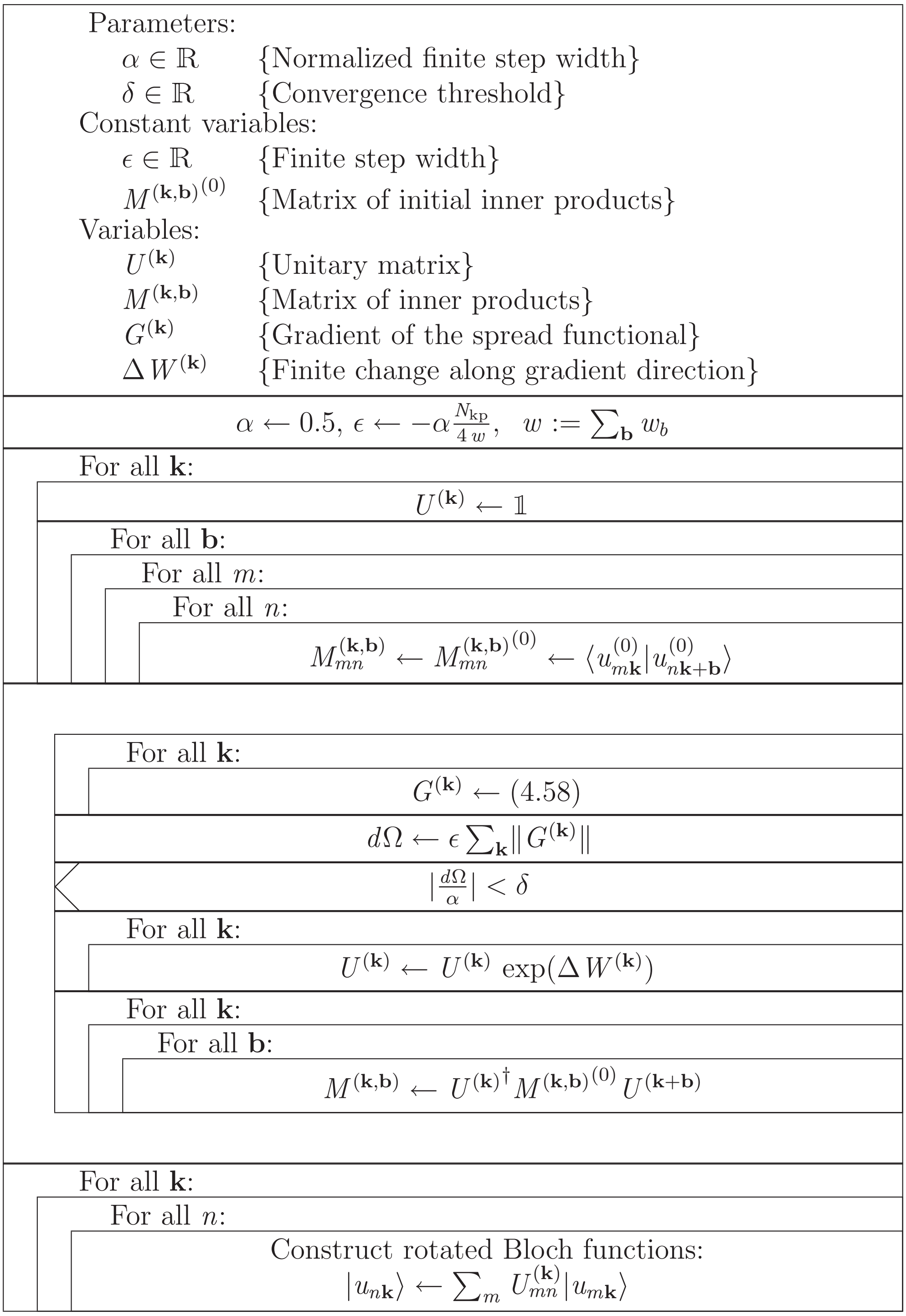
   have to be calculated numerically. The periodic functions $u_{n\mathbf{k}}^{(0)}(\mathbf{r})$ are either those constructed directly through a band structure program or those that have already passed a certain preprocessing step explained in 4.3.3 later on.

2. Subsequently the gradients $G^{(\mathbf{k})}$ have to be calculated according to Eq. (4.58) and from these the matrices $\Delta W^{(\mathbf{k})}$, specified in Eq. (4.61), have to be constructed. If the absolute change $|d\Omega|$ of the spread functional, calculated according to Eq. (4.60), is smaller than a reasonably chosen threshold δ, the minimum is reached and the unitary matrices $U^{(\mathbf{k})}$ are used to construct the new set of Bloch functions via (4.4).

3. If the gradient didn't under-run the given threshold the matrices $\Delta W^{(\mathbf{k})}$ are used to proceed to the next iteration step by performing the updates (4.66) and (4.67). Afterwards the algorithm continues with step 2.

Steps 2 and 3 are executed as long as $|d\Omega| > \delta$. After $|d\Omega|$ has under-run this threshold a set of generalized Bloch functions is constructed according to (4.4) using the obtained set of unitary matrices, $U_{mn}^{(\mathbf{k})(\mathrm{min})}$, which solves the minimum problem (4.11). Afterwards the corresponding Wannier functions are constructed by means of the Lattice Fourier Transform (4.18). In Fig. 4.3 the complete algorithm is depicted as a Nassi-Shneiderman [102] chart.

Steepest Descent Algorithm



```
Parameters:
    $\alpha \in \mathbb{R}$    {Normalized finite step width}
    $\delta \in \mathbb{R}$    {Convergence threshold}
Constant variables:
    $\epsilon \in \mathbb{R}$    {Finite step width}
    $M^{(\mathbf{k},\mathbf{b})^{(0)}}$    {Matrix of initial inner products}
Variables:
    $U^{(\mathbf{k})}$    {Unitary matrix}
    $M^{(\mathbf{k},\mathbf{b})}$    {Matrix of inner products}
    $G^{(\mathbf{k})}$    {Gradient of the spread functional}
    $\Delta W^{(\mathbf{k})}$    {Finite change along gradient direction}

$\alpha \leftarrow 0.5,\ \epsilon \leftarrow -\alpha \frac{N_{\mathrm{kp}}}{4\,w},\quad w := \sum_{\mathbf{b}} w_b$
For all k:
    $U^{(\mathbf{k})} \leftarrow \mathbb{1}$
    For all b:
        For all m:
            For all n:
                $M_{mn}^{(\mathbf{k},\mathbf{b})} \leftarrow M_{mn}^{(\mathbf{k},\mathbf{b})^{(0)}} \leftarrow \langle u_{m\mathbf{k}}^{(0)} | u_{n\mathbf{k}+\mathbf{b}}^{(0)} \rangle$

    For all k:
        $G^{(\mathbf{k})} \leftarrow (4.58)$
    $d\Omega \leftarrow \epsilon \sum_{\mathbf{k}} \| G^{(\mathbf{k})} \|$
    $|\frac{d\Omega}{\alpha}| < \delta$
    For all k:
        $U^{(\mathbf{k})} \leftarrow U^{(\mathbf{k})} \exp(\Delta W^{(\mathbf{k})})$
    For all k:
        For all b:
            $M^{(\mathbf{k},\mathbf{b})} \leftarrow U^{(\mathbf{k})\dagger} M^{(\mathbf{k},\mathbf{b})^{(0)}} U^{(\mathbf{k}+\mathbf{b})}$

For all k:
    For all n:
        Construct rotated Bloch functions:
        $|u_{n\mathbf{k}}\rangle \leftarrow \sum_m U_{mn}^{(\mathbf{k})} |u_{m\mathbf{k}}\rangle$
```

Figure 4.3: *Nassi-Shneiderman [102] chart of the steepest descent algorithm.*

4.3.3 Trial function based preprocessing

As alluded earlier a good choice of the initial values for a Steepest Descent minimization is of paramount importance. They govern whether a minimum can be reached at all and influence in addition the number of iterations steps which are required to converge.

To construct now exactly such a set of good initial values, $U_{mn}^{(\mathbf{k})(0)}$, for the unitary matrices the following procedure is used. First a set of trial functions $T_n(\mathbf{r})$, chosen to be a rough initial guess to the Wannier functions, is projected onto all considered BFs at each wave vector $\mathbf{k}$,

$$|\phi_{n\mathbf{k}}\rangle = \sum_m |\psi_{m\mathbf{k}}\rangle\langle\psi_{m\mathbf{k}}|T_n\rangle = \sum_m A_{mn}^{(\mathbf{k})}\,|\psi_{m\mathbf{k}}\rangle \tag{4.69}$$

As this functions are not orthogonal a symmetric orthonormalization is performed to obtain a set of orthonormal functions

$$|\psi_{n\mathbf{k}}^{(0)}\rangle = \sum_m \big(S^{-1/2}\big)_{mn}|\phi_{m\mathbf{k}}\rangle\ , \tag{4.70}$$

where the overlap matrix is defined as $S_{mn} = \langle\phi_{m\mathbf{k}}|\phi_{n\mathbf{k}}\rangle$ and its inverse square root is calculated by transforming S into diagonal form, taking the inverse square roots of the diagonal elements and transform back afterwards. The resulting functions $\psi_{n\mathbf{k}}^{(0)}(\mathbf{r})$ are labelled with iteration index $i = 0$ because they are the initial functions for the spread minimization algorithm described earlier in 4.3.2 and are connected to the initial functions $\psi_{n\mathbf{k}}(\mathbf{r})$ via

$$|\psi_{n\mathbf{k}}^{(0)}\rangle = \sum_m \big(AS^{-1/2}\big)_{mn}\,|\psi_{m\mathbf{k}}\rangle =: \sum_m U_{mn}^{(\mathbf{k})(0)}\,|\psi_{m\mathbf{k}}\rangle\ . \tag{4.71}$$

Criteria for “good” trial functions are:

1. They should be real valued functions, as maximally localized Wannier functions can be chosen to be purely real for time reversal symmetric systems [45]. This choice also ensures that $\psi_{n-\mathbf{k}}^{(0)}(\mathbf{r}) = \psi_{n\mathbf{k}}^{(0)^*}(\mathbf{r})$.

2. They should reflect the symmetry of the underlying lattice, i.e. they should transform under symmetry operations according to an irreducible or reducible representation of the point group of the direct lattice.

3. They should be localized around a pronounced center.

Among several possibilities for suitable trial functions $T_n(\mathbf{r})$ the ones which have been used during the synthesis of this thesis are presented in the following sections, together with brief explanations for which systems which functions are best suited.

4.3.3.1 Random Gaussians

A set of functions which is obtainable quite easily is a set of randomly arranged Gaussian functions. Admittedly, these functions indeed state a really crude initial guess on the Wannier functions. However, for certain systems they are already sufficiently close to the envisaged solution and the minimization converges. It is important that their centers are sufficiently far away from high-symmetry lines of the underlying lattice to prevent the matrices $A_{mn}^{(\mathbf{k})}$ from becoming singular for certain values of $\mathbf{k}$.

Their disadvantage is that it usually takes more iterations to converge to the global minimum compared to symmetrically chosen trial functions suggested in Sec. 4.3.3.2 and sometimes no convergence is obtained. If one is interested in hybridized WFs for square lattice systems, random Gaussians are a good choice, but in general they are used rarely.

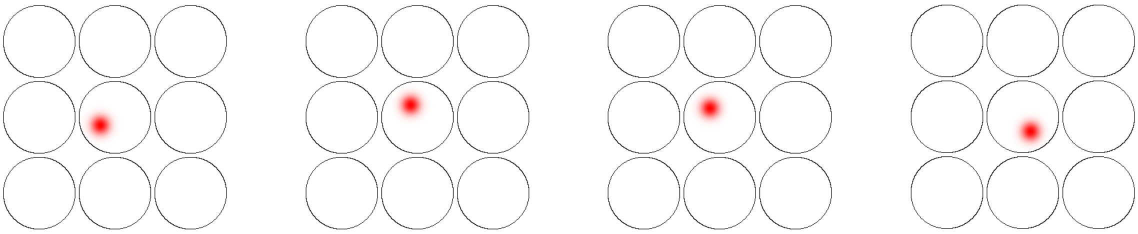


Figure 4.4: *Randomly positioned Gaussians as trial functions for the square lattice model system described in Fig. 2.4(a).*

4.3.3.2 Localized Γ-point Bloch functions

The set of initial trial functions presented here arose from the question: what are the Wannier functions for a 1×1 k-point sampling of the first Brillouin Zone? In this case the only wave vector which is contained in the Monkhorst-Pack mesh is the Γ-point, i.e. $\mathbf{k} = \mathbf{0}$. Since the point group of this wave vector coincides with the point group of the direct lattice, the corresponding BFs can be associated with an irreducible representations of this group, and hence, reflect perfectly the symmetry of the underlying lattice. In addition, the BFs at the Γ-point can be chosen purely real. One might, therefore, conjecture that the BFs at the Γ-point should have a certain affinity to the MLWFs for finer k-space meshes. And indeed —at least for the case of E-polarized light— the similarity is striking, giving rise to the assumption that these functions might be ideally suited as trial function. The only catch is that being periodic functions they are, of course, not localized. But luckily, it proves to be a successful strategy to artificially "localize" these functions by enforcing function values of zero outside a WSC around well-chosen centers. In this regard it turned out to be useful to use two different sets of these artificially localized functions.

The first set consists of all functions localized around the origin, $\mathbf{r} = \mathbf{0}$. All functions constructed in this manner are transforming according to an IREP of the

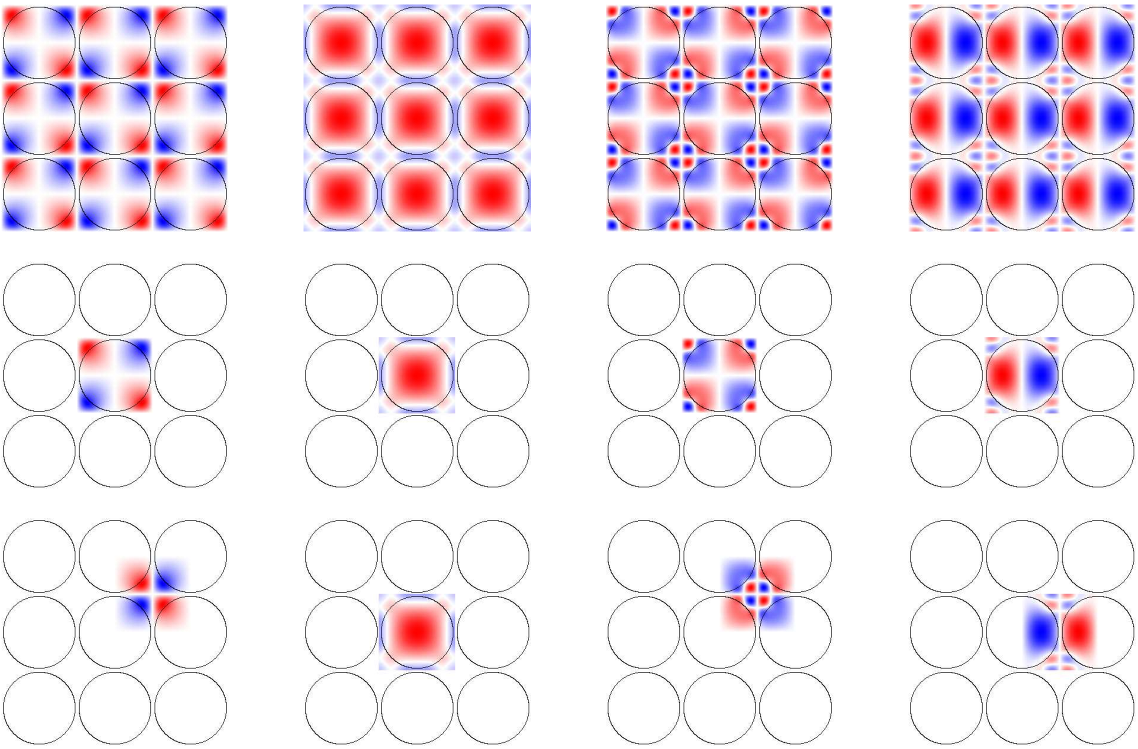


Figure 4.5: *Artificially localized* Γ*-point Bloch functions created for the square lattice model system described in Fig. 2.4(a). The upper row shows the underlying Bloch function at the* Γ*-point,* $\psi_{n\mathbf{0}}(\mathbf{r})$*. The functions in the mid row are created by "cutting out" a WSC around the origin. For the functions in the bottom row the WSC was clipped around high-symmetry points where the underlying BF has pronounced intensities, yielding the smallest spread of all possible artificially localized functions having high-symmetry points as centers.*

point group $\mathcal{M}$ of the underlying lattice. One might, therefore, refer to this set as the full-symmetric set of trial functions. This set is well-suited for handling the H-polarized case and states also a perfect set of trial functions for the subspace optimization procedure, described later in Secs. 5.1.2 and 5.1.6. A sample set for the considered square lattice model system is exhibited in the mid row of in Fig. 4.5.

The second set which is especially well-suited for the construction of E-polarized MLWFs features functions centered around those high-symmetry points of the direct lattice, where the spread $\langle T_n|r^2|T_n\rangle - \langle T_n|\mathbf{r}|T_n\rangle^2$ of these functions is minimal. Due to this fact, one might refer to it as the minimum-spread set. Such a set is depicted in the bottom row of Fig. 4.5. In the case of E-polarized light, preprocessing with this set yields in most cases instantly the global minimum of (4.11). It might also be used to create well-localized and symmetric WFs stating a local minimum of a large group of bands.

At this point it should be mentioned that there is —irrespective of a general and wide-spread tendency of people to use smooth functions for such a purpose— absolutely no need to localize the BF with anything else then a step function. In this way, the resulting function can be normalized properly and the area of integration can be restricted to a single WSC which speeds up the entire preprocessing significantly. Most important is that a true WSC is cut-off, so that the clipped area is symmetric with respect to its center.

4.4 Application to the square lattice model system

After the complete algorithm for obtaining MLWFs was described in detail, this section is devoted to its application to the square lattice model system described in Sec. 2.3.2. This system features two lower-lying band gaps (see Fig. 2.4(a)): the first occurs between bands one and two and the second between bands three and four. Therefore, the bands 1–3 build two isolated groups of bands as defined in the introduction to this chapter. The first group contains only the first band, as a single isolated band, and the second group consists of bands two and three. Hence, the method of Marzari and Vanderbilt is applicable to both of them. In Fig. 4.6(a) the obtained MLWFs[13] are displayed. The maximally localized Wannier function of the first band is an A_1-type[14] function centered at the M-point of the direct lattice, i.e. $\mathbf{r} = (0.5, 0.5)a$, whose point group coincides with the point group, C_{4v}, of the underlying lattice. The WFs for band two and three belong to the E-representation and build, therefore, a basis of a two-dimensional irreducible representation of C_{4v}. In correspondence to atomic orbitals, these functions could be compared to s- and p-orbitals.

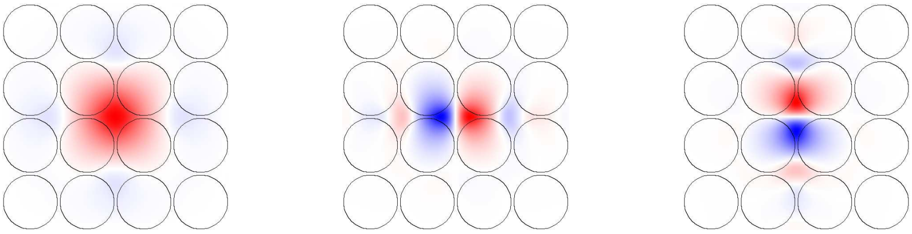
Equally well, one could consider all three bands as a composite group which is itself isolated to higher bands through a band gap. When minimizing the spread functional of all three functions together a new global minimum occurs yielding a set of hybridized functions depicted in Fig. 4.6(b). Again, in analogy to the atomic case, these functions might be considered to represent three sp^2 hybrid orbitals [103]. However, in contrast to “true” sp^2-hybrids they are not 120° rotational images of each other, as the underlying square lattice enforces that these modes transform according to a representation of C_{4v} and not to a representation of a point group with at least 3-fold rotational symmetry.

In conclusion, this example shows that the choice of which bands are grouped together leads to different global minima of the spread functional (4.9) and, consequently, to different sets of MLWFs. Besides being perceived from a personal point of view as less symmetric, the sp^2-hybrids have some drawbacks with respect to numerics. For instance, their centers can be determined less accurately compared to those being high-symmetry points, and thus, their transformation under symmetry operations is inferior as well. This might lead, for instance, to significant asymmetries in the transmittance of beam-splitting devices.

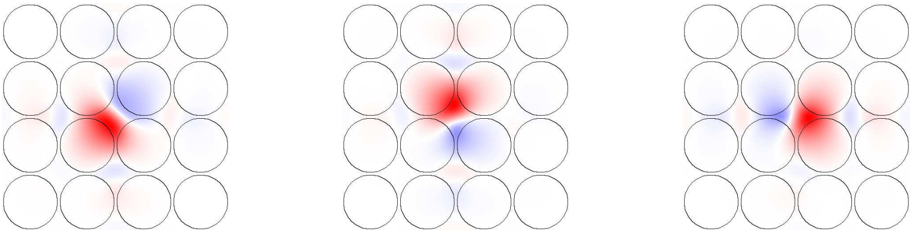
---

[13] The usage of minimum-spread trial functions presented in Sec. 4.3.3.2 yields directly the global minimum of the spread functional.

[14] See Sec. 2.3.3 for classification of functions according to irreducible representations of the point group of the lattice.



(a) Maximally localized Wannier functions for considering the band range 1–3 as build up from two isolated groups of bands: the first group consists of only the first band and the second group consists of bands two and three. The resulting functions are of symmetry type A_1 and E, corresponding to atomic s- and p-orbitals.



(b) Maximally localized Wannier functions for considering the band range 1–3 as one composite group of bands. In this case, there exists a new global minimum of the spread functional for this band complex yielding a sort of sp^2-hybridized functions. However, the square lattice prevents the functions from having real 120° rotational symmetry, as they have to transform according to a representation of C_{4v} and not with respect to a point group with at least 3-fold rotational symmetry.

Figure 4.6: *Set of maximally localized Wannier functions for the first three bands of the considered square lattice model system exemplifying the freedom of choice in the grouping of bands. The parameters of the model system and corresponding band structure are given in Fig. 2.4(a) on p. 17.*

A second fact that should be mentioned here, is, that the MLWFs tend to be localized in the material with higher refractive index. This fact is inherent already to the underlying Bloch functions and one might, therefore, distinguish between eigenmodes of the PC, having significant intensities in the material with higher refractive index (material-modes), and eigenmodes that have pronounced intensity inside the air pores (air-modes). In general, it is observable that there are more material-modes than air-modes occurring in a PC in the low frequency regime. In addition, the Wannier functions transform under symmetry operations of the point group $\mathcal{M}$ of the underlying lattice according to either reducible or irreducible representations of $\mathcal{M}$ with respect to symmetry centers whose point group coincides with $\mathcal{M}$. For the square lattice the M-point is such a center, and thus, MLWF belonging to irreducible representations of C_{4v} centered at this high-symmetry point can exist. However, knowing that in a triangular lattice such a center, besides the origin, does not exist, one might argue that MLWF corresponding to material-modes and transforming according to an IREP of C_{6v} might not occur.

Finally, one might also perceive that in general only a few isolated groups of bands occur in PCs and with the method of Marzari and Vanderbilt [46] alone, only a small set of WF can be constructed, which are in general not complete enough to describe embedded defect structures sufficiently well. However, if it is possible to artificially isolate a certain group of bands from higher ones by means of another algorithm which is, for instance, implemented as an additional preprocessing step, the algorithm of Marzari and Vanderbilt would be applicable again.

5 Maximally localized Wannier functions for entangled bands

As can been seen by inspecting the complete band structure of the triangular lattice model system which is depicted in Fig. 5.1 up to band 40, isolated groups of bands occur rather seldomly. Instead, one is in general faced with a huge bulk of wildly crossing bands, which also might be called to be entangled.

As indicated at the end of the last chapter, it would be desirable to have a possibility to isolate or disentangle a certain group of N bands from higher bands, with whom they are entangled. Such an algorithm was developed in the context of solid state physics by Souza *et al.* [47] and is intended to be a preprocessing step to the spread minimization algorithm of Marzari and Vanderbilt [46].

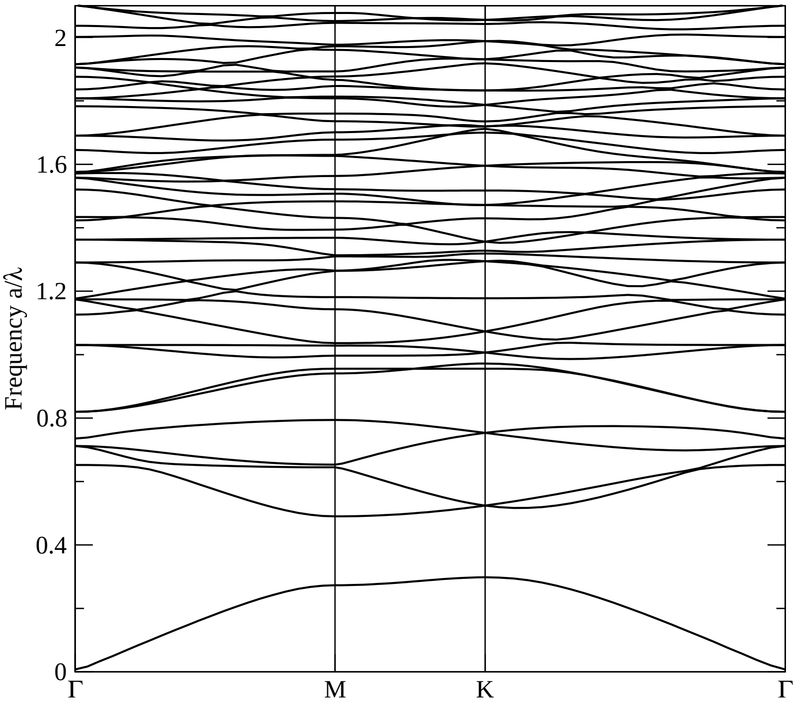


Figure 5.1: *Complete band structure up to band 40 for the triangular lattice model system, introduced in Sec. 2.3.2. The band structure is dominated by regions of wildly crossing bands. Isolated groups of bands occur only rarely.*

5.1 Artificial creation of an isolated group of bands

The algorithm of Souza *et al.* utilizes that the spread functional (4.9) is separable into three parts[1],

$$\Omega = \Omega_{\mathrm{I}} + \Omega_{\mathrm{OD}} + \Omega_{\mathrm{D}} \ . \tag{5.1}$$

Here, Ω_{I} , Ω_{OD}, and Ω_{D} are called the invariant, off-diagonal, and diagonal part of the spread functional, respectively, and are defined in discretized form as

$$\Omega_{\mathrm{I}} = \frac{1}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \Big(N - \sum_{m,n} \big| M_{mn}^{(\mathbf{k},\mathbf{b})} \big|^2 \Big) \ , \tag{5.2}$$

$$\Omega_{\mathrm{OD}} = \frac{1}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \sum_{m \neq n} \big| M_{mn}^{(\mathbf{k},\mathbf{b})} \big|^2 \ , \tag{5.3}$$

$$\Omega_{\mathrm{D}} = \frac{1}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \sum_{n} \Big(-\operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} - \mathbf{b}\mathbf{r}_n \Big)^2 \ . \tag{5.4}$$

As its name already indicates, Ω_{I} is invariant under all unitary transformations $U_{mn}^{(\mathbf{k})}$ between Bloch functions, and hence, states a lower bound for the spread functional Ω.

5.1.1 Invariant part of the spread functional and its interpretation

To examine the meaning of Ω_I, it is re-expressed as

$$\Omega_I \quad = \quad \frac{1}{N_{kp}} \sum_{\mathbf{k}} \omega_I(\mathbf{k}) \ , \tag{5.5}$$

where the functional $\omega_I(\mathbf{k})$ is defined as

$$\omega_I(\mathbf{k}) = \sum_{\mathbf{b}} w_b \, T_{\mathbf{k},\mathbf{b}} \ , \tag{5.6}$$

with

$$T_{\mathbf{k},\mathbf{b}} = N - \sum_{m,n}^{N} \big| M_{mn}^{(\mathbf{k},\mathbf{b})} \big|^2 = \operatorname{Tr} \Big[\hat{P}_{\mathbf{k}} \, \hat{Q}_{\mathbf{k}+\mathbf{b}} \Big] \ . \tag{5.7}$$

$T_{\mathbf{k},\mathbf{b}}$ is called the “spillage” between neighboring spaces $\mathcal{S}(\mathbf{k})$ and $\mathcal{S}(\mathbf{k}+\mathbf{b})$, measuring the degree of mismatch between them, vanishing when they are identical. $\hat{P}_{\mathbf{k}} = \sum_{n=1}^{N} |u_{n\mathbf{k}}\rangle\langle u_{n\mathbf{k}}|$ is the projector onto the space $\mathcal{S}(\mathbf{k})$ spanned by the N periodic parts, $u_{n\mathbf{k}}(\mathbf{r})$, and $\hat{Q}_{\mathbf{k}} = 1 - \hat{P}_{\mathbf{k}}$. Therefore, $\omega_I(\mathbf{k})$ might be interpreted as “combined spillage” over all neighboring subspaces $\mathcal{S}(\mathbf{k}+\mathbf{b})$, vanishing when they are all identical. This means that the smaller $\omega_I(\mathbf{k})$ the better the agreement of $\mathcal{S}(\mathbf{k})$ with all its neighboring subspaces $\mathcal{S}(\mathbf{k}+\mathbf{b})$. Concluding, it can readily be seen that Ω_I can be interpreted as a measure of “global smoothness of connection” between the subspaces $\mathcal{S}(\mathbf{k})$ throughout the first Brillouin Zone.

---

[1] See Appendix A.4.

5.1.2 Subspace optimization

That a certain group of N bands is entangled with higher bands, means, that the corresponding BFs, building up the vector space $\mathcal{S}(\mathbf{k})$, carry mode informations or contain mode structures of higher bands (see Fig. 4.1 on p. 40), which could again be eliminated by means of unitary transformations, $U_{mn}^{(\mathbf{k})}$. However, the dimension of the vector space is predefined by the number, N, of WFs that should be constructed, and thus, the incompatible mode informations can not leave this subspace, as the unitary matrices $U_{mn}^{(\mathbf{k})}$ can only mix the BFs already contained in $\mathcal{S}(\mathbf{k})$. If one wants to get rid of such incongruous mode informations, one has to provide a possibility for them to leave the vector space $\mathcal{S}(\mathbf{k})$.

This is realized by defining a frequency window, i.e. a cutoff frequency ω_c, such that the number of Bloch functions, belonging to frequencies inside this window, is in general larger than the number of Wannier functions, N, to be constructed[2]. This means that the frequency window defines an $N_\mathbf{k}$-dimensional space $\mathcal{F}(\mathbf{k})$ with $N_\mathbf{k} \geq N$ at each k-point. This situation is illustrated in Fig. 5.2. The cutoff frequency ω_c has to be chosen in such a way that all bands, presumably carrying mode informations of bands which are element of $\mathcal{S}(\mathbf{k})$, are contained in $\mathcal{F}(\mathbf{k})$ at all k-points where entanglement occurs.

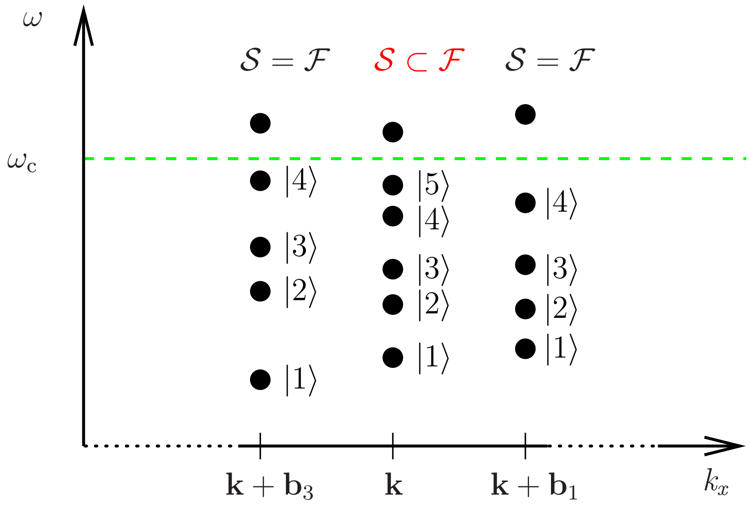


Figure 5.2: *A sample dispersion relation along the k_x-axis through the Brillouin zone. At each k-point the cutoff frequency ω_c defines an N_k dimensional vector space $\mathcal{F}(\mathbf{k})$ spanned by all states that belong to eigenfrequencies $\omega_{n\mathbf{k}} \leq \omega_c$. In this example, it is intended to construct $N = 4$ Wannier functions. The cutoff frequency selects $N_k = 5$ states at k-point* $\mathbf{k}$ *and, therefore, an optimal subset $\mathcal{S}(\mathbf{k})$ with dimension $N = 4$ has to be determined at this wave vector. Since $\mathcal{S} = \mathcal{F}$ holds for the neighboring mesh nodes, no subspace optimization has to be performed there.*

The group of N bands can now be isolated from higher ones by finding at each

---

[2] At some k-points the bands might not be entangled, and thus no additional bands have to be taken into account there.

k-point an optimal N-dimensional subspace $\mathcal{S}(\mathbf{k}) \subseteq \mathcal{F}(\mathbf{k})$ that leads to the best "global smoothness of connection". This means, at each $\mathbf{k}$, a subspace $\mathcal{S}(\mathbf{k})$ of $\mathcal{F}(\mathbf{k})$ has to be determined that fits best to its neighboring spaces $\mathcal{S}(\mathbf{k}+\mathbf{b})$. Clearly, this is the set of subspaces, $\{\mathcal{S}(\mathbf{k})\}$, that minimizes Ω_I.

According to Eq. (5.5), Ω_I is minimal when all $\omega_I(\mathbf{k})$ are as small as possible. Therefore, the problem of determining an optimal $\mathcal{S}(\mathbf{k})$ reduces to minimizing the functional $\omega_I(\mathbf{k})$ at each k-point with the additional constraint that the functions spanning $\mathcal{S}(\mathbf{k})$ are still orthonormal. As $\omega_I(\mathbf{k})$ couples functions at different wave vectors $\mathbf{k}$, the minimization has to be carried out self-consistently throughout the first Brillouin Zone. This means, at each iteration step i one has to traverse through all k-points of the Monkhorst-Pack mesh and find for each of them N orthonormal functions, $u_{n\mathbf{k}}^{(i)}(\mathbf{r})$, defining the subspace $\mathcal{S}^{(i)}(\mathbf{k}) \subseteq \mathcal{F}(\mathbf{k})$ such that the "combined spillage",

$$\omega_I^{(i)}(\mathbf{k}) = \sum_{\mathbf{b}} w_b \, T_{\mathbf{k},\mathbf{b}}^{(i),(i-1)} = \sum_{\mathbf{b}} w_b \;\mathrm{Tr}\left[\hat{P}_{\mathbf{k}}^{(i)} \hat{Q}_{\mathbf{k}+\mathbf{b}}^{(i-1)}\right] \,, \tag{5.8}$$

over all neighboring subspaces $\mathcal{S}^{(i-1)}(\mathbf{k}+\mathbf{b})$ of the previous iteration step, $i-1$, is as small as possible. Convergence is reached when $\mathcal{S}^{(i)}(\mathbf{k}) = \mathcal{S}^{(i-1)}(\mathbf{k})$ at all mesh nodes $\mathbf{k}$.

5.1.3 Minimization of the combined spillage

The minimization of the combined spillage, $\omega_I(\mathbf{k})$, has to be carried out with respect to a subset of N functions out of a superset of $N_{\mathbf{k}}$ functions. Using Lagrange multipliers, $\Lambda_{nm,\mathbf{k}}^{(i)}$, to enforce orthonormality, the Lagrange function at the ith iteration step reads

$$L^{(i)}(\mathbf{k}) := \omega_I^{(i)}(\mathbf{k}) + \sum_{m=1}^{N} \sum_{n=1}^{N} \Lambda_{nm,\mathbf{k}}^{(i)} \left[\langle u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}}^{(i)} \rangle - \delta_{mn}\right] , \tag{5.9}$$

which is stationary when the first order variation $\delta L^{(i)}(\mathbf{k})$ vanishes. The variation, $\delta\omega^{(i)}(\mathbf{k})$, of the first term is given by

$$\begin{aligned}
\omega^{(i)}(\mathbf{k}) + \delta\omega^{(i)}(\mathbf{k}) &= Nw - \sum_{m=1}^{N}\sum_{n=1}^{N}\sum_{\mathbf{b}} w_b \langle u_{m\mathbf{k}}^{(i)} + \delta u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}+\mathbf{b}}^{(i-1)}\rangle\langle u_{n\mathbf{k}+\mathbf{b}}^{(i-1)} | u_{m\mathbf{k}}^{(i)} + \delta u_{m\mathbf{k}}^{(i)}\rangle \\
&= Nw - \sum_{m=1}^{N}\sum_{n=1}^{N}\sum_{\mathbf{b}} w_b \Bigg\{ \langle u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}+\mathbf{b}}^{(i-1)}\rangle\langle u_{n\mathbf{k}+\mathbf{b}}^{(i-1)} | u_{m\mathbf{k}}^{(i)}\rangle \\
&\qquad + \langle u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}+\mathbf{b}}^{(i-1)}\rangle\langle u_{n\mathbf{k}+\mathbf{b}}^{(i-1)} | \delta u_{m\mathbf{k}}^{(i)}\rangle \\
&\qquad + \langle \delta u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}+\mathbf{b}}^{(i-1)}\rangle\langle u_{n\mathbf{k}+\mathbf{b}}^{(i-1)} | u_{m\mathbf{k}}^{(i)}\rangle \\
&\qquad + O\left((\delta u_{m\mathbf{k}}^{(i)})^2\right) \Bigg\} \; . \qquad (5.10)
\end{aligned}$$

The variation of the second term which might be defined as

$$W(\mathbf{k}) := \sum_{m=1}^{N}\sum_{n=1}^{N} \Lambda_{nm,\mathbf{k}}^{(i)} \left[\langle u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}}^{(i)}\rangle - \delta_{mn}\right] \qquad (5.11)$$

is determined by

$$\begin{aligned}
W^{(i)}(\mathbf{k}) + \delta W^{(i)}(\mathbf{k}) &= \sum_{m=1}^{N}\sum_{n=1}^{N} \Lambda_{nm,\mathbf{k}}^{(i)} \left[\langle u_{m\mathbf{k}}^{(i)} + \delta u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}}^{(i)} + \delta u_{n\mathbf{k}}^{(i)}\rangle - \delta_{mn}\right] \\
&= \sum_{m=1}^{N}\sum_{n=1}^{N} \Lambda_{nm,\mathbf{k}}^{(i)} \Bigg\{ \left[\langle u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}}^{(i)}\rangle - \delta_{mn}\right] + \\
&\qquad \left[\langle u_{m\mathbf{k}}^{(i)} | \delta u_{n\mathbf{k}}^{(i)}\rangle + \langle \delta u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}}^{(i)}\rangle + O(\delta^2 u)\right] \Bigg\} . \qquad (5.12)
\end{aligned}$$

From Eqs. (5.10) and (5.12) it may readily be seen that

$$\begin{aligned}
\frac{\delta\omega_I^{(i)}(\mathbf{k})}{\delta u_{m\mathbf{k}}^{(i)\,*}} &= -\sum_{\mathbf{b}} w_b \sum_{n=1}^{N} |u_{n\mathbf{k}+\mathbf{b}}^{(i-1)}\rangle\langle u_{n\mathbf{k}+\mathbf{b}}^{(i-1)} | u_{m\mathbf{k}}^{(i)}\rangle \\
&= -\sum_{\mathbf{b}} w_b \hat{P}_{\mathbf{k}+\mathbf{b}}^{(i-1)} |u_{m\mathbf{k}}^{(i)}\rangle \qquad (5.13)
\end{aligned}$$

where $\hat{P}^{(i-1)}_{\mathbf{k}+\mathbf{b}}$ is the projector onto neighboring subspaces $\mathcal{S}^{(i-1)}(\mathbf{k}+\mathbf{b})$. Similarly, the functional derivative of $W^{(i)}(\mathbf{k})$ is

$$\frac{\delta\, W^{(i)}(\mathbf{k})}{\delta\, u^{(i)\,*}_{m\mathbf{k}}} = \sum_{n=1}^{N} \Lambda^{(i)}_{nm,\mathbf{k}} |u^{(i)}_{n\mathbf{k}}\rangle. \tag{5.14}$$

The stationarity of $L^{(i)}(\mathbf{k})$ leads to two stationary conditions:

$$\frac{\delta\omega^{(i)}_{I}(\mathbf{k})}{\delta\, u^{(i)\,*}_{m\mathbf{k}}} + \sum_{n=1}^{N} \Lambda^{(i)}_{nm,\mathbf{k}} \frac{\delta}{\delta\, u^{(i)\,*}_{m\mathbf{k}}} \left[\langle u^{(i)}_{m\mathbf{k}} | u^{(i)}_{n\mathbf{k}} \rangle - \delta_{mn} \right] = 0 \tag{5.15}$$

and its complex conjugate version, where the matrices of Lagrangian multipliers $\Lambda^{(i)}_{nm,\mathbf{k}}$ have dimensions $N \times N$. Inserting now the just derived functional derivatives into the stationary condition (5.15) yields

$$\underbrace{\left[\sum_{\mathbf{b}} w_b \hat{P}^{(i-1)}_{\mathbf{k}+\mathbf{b}} \right]}_{=:\hat{Z}^{(i-1)}_{\mathbf{k}}} |u^{(i)}_{m\mathbf{k}}\rangle - \sum_{n=1}^{N} \Lambda^{(i)}_{nm,\mathbf{k}} |u^{(i)}_{n\mathbf{k}}\rangle = 0\ . \tag{5.16}$$

By defining an operator $\hat{Z}^{(i-1)}_{\mathbf{k}}$ as indicated above, this equation adopts the rather simple form

$$\hat{Z}^{(i-1)}_{\mathbf{k}} |u^{(i)}_{m\mathbf{k}}\rangle = \sum_{n=1}^{N} \Lambda^{(i)}_{nm,\mathbf{k}} |u^{(i)}_{n\mathbf{k}}\rangle. \tag{5.17}$$

Choosing now the $|u^{(i)}_{m\mathbf{k}}\rangle$ to be eigenvectors of $\hat{Z}^{(i-1)}_{\mathbf{k}}$ yields the eigenvalue equation

$$\hat{Z}^{(i-1)}_{\mathbf{k}} |u^{(i)}_{m\mathbf{k}}\rangle = \lambda^{(i)}_{m\mathbf{k}} |u^{(i)}_{m\mathbf{k}}\rangle. \tag{5.18}$$

Until now, it is not obvious, in which manner this eigenvalue equation might influence the minimization of the functional $\omega^{(i)}_{\mathrm{I}}(\mathbf{k})$. It becomes apparent when $\omega^{(i)}_{\mathrm{I}}(\mathbf{k})$

is reformulated in terms of the newly introduced operator $\hat{Z}_{\mathbf{k}}^{(i-1)}$ as

$$\begin{aligned}
\omega_I^{(i)}(\mathbf{k}) &= \sum_{\mathbf{b}} w_b \; T_{\mathbf{k},\mathbf{b}}^{(i)} \\
&= N \sum_{\mathbf{b}} w_b - \sum_{\mathbf{b}} w_b \sum_{m=1}^{N} \sum_{n=1}^{N} \langle u_{m\mathbf{k}}^{(i)} | u_{n\mathbf{k}+\mathbf{b}}^{(i-1)} \rangle \langle u_{n\mathbf{k}+\mathbf{b}}^{(i-1)} | u_{m\mathbf{k}}^{(i)} \rangle \\
&= N \sum_{\mathbf{b}} w_b - \sum_{m=1}^{N} \langle u_{m\mathbf{k}}^{(i)} | \sum_{\mathbf{b}} w_b P_{\mathbf{k}+\mathbf{b}}^{(i-1)} | u_{m\mathbf{k}}^{(i)} \rangle \\
&= N \sum_{\mathbf{b}} w_b - \sum_{m=1}^{N} \langle u_{m\mathbf{k}}^{(i)} | \hat{Z}_{\mathbf{k}}^{(i-1)} | u_{m\mathbf{k}}^{(i)} \rangle \\
&\overset{(5.18)}{=} N \sum_{\mathbf{b}} w_b - \sum_{m=1}^{N} \langle u_{m\mathbf{k}}^{(i)} | \lambda_{m\mathbf{k}}^{(i)} | u_{m\mathbf{k}}^{(i)} \rangle \\
&= N \sum_{\mathbf{b}} w_b - \sum_{m=1}^{N} \lambda_{m\mathbf{k}}^{(i)}. \qquad (5.19)
\end{aligned}$$

As $0 \leq \lambda_{m\mathbf{k}}^{(i)} \leq \sum_{\mathbf{b}} w_b$, the minimal value for $\omega_I^{(i)}(\mathbf{k})$ is found by picking the leading[3] N out of $N_{\mathbf{k}}$ eigenvalues $\lambda_{m\mathbf{k}}^{(i)}$ of $\hat{Z}_{\mathbf{k}}^{(i-1)}$ and the optimal subspace $\mathcal{S}^{(i)}(\mathbf{k})$ is spanned by the corresponding eigenvectors.

At this point it should be mentioned that the equations derived here differ slightly from the ones in [47]. In particular, as, according to Eq. (5.5), Ω_I does not explicitly depend on $\omega_I(\mathbf{k}+\mathbf{b})$, the occurrence of terms containing functional derivatives of $\omega_I(\mathbf{k}+\mathbf{b})$ as in Eq. (13) of [47] is questionable. However, the additional terms occurring there are yielding an additional factor of 2 which is re-absorbed into matrices $\tilde{\Lambda}_{\mathbf{k}} \propto \frac{1}{2}\Lambda_{\mathbf{k}}$. Since these matrices are only Lagrangian multipliers, the correctness of the algorithm itself is not affected, though.

---

[3] Leading is meant in the sense of having the largest eigenvalue.

5.1.4 Implementation of the algorithm

In the following, the steps that have to be performed in the iterative minimization scheme described above are listed consecutively.

Calculation of inner product matrices $M_{mn}^{(\mathbf{k},\mathbf{b})(0)}$
At first the inner product matrices have to be calculated numerically:

$$M_{mn}^{(\mathbf{k},\mathbf{b})(0)} = \langle u_{m\mathbf{k}}^{(0)} | u_{n\mathbf{k}+\mathbf{b}}^{(0)} \rangle \ . \tag{5.20}$$

This step is computationally expensive, but has to be performed only once in the beginning.

Construction of the $Z(\mathbf{k})$ matrices
In the next step, the matrix elements of the $\hat{Z}(\mathbf{k})$-operator have to be determined. For this purpose, the set of functions $|u_{n\mathbf{k}}^{(0)}\rangle$, created by a preceding preprocessing step which will be described in Sec. 5.1.6, is chosen as a basis. These functions are supposed to correspond to the iteration step $i = 0$. In this basis the representation of the operator $\hat{Z}(\mathbf{k})$ is:

$$\begin{aligned}
Z_{mn}^{(i)}(\mathbf{k}) &= \langle u_{m\mathbf{k}}^{(0)} | \sum_{\mathbf{b}} w_b \left[P_{\mathbf{k}+\mathbf{b}}^{(i-1)} \right] | u_{n\mathbf{k}}^{(0)} \rangle & (5.21) \\
&= \sum_{\mathbf{b}} w_b \sum_{l=1}^{N} \langle u_{m\mathbf{k}}^{(0)} | u_{l\mathbf{k}+\mathbf{b}}^{(i-1)} \rangle \langle u_{l\mathbf{k}+\mathbf{b}}^{(i-1)} | u_{n\mathbf{k}}^{(0)} \rangle & (5.22) \\
&= \sum_{\mathbf{b}} w_b \sum_{l=1}^{N} M_{ml}^{(\mathbf{k},\mathbf{b})(0,i-1)} \left(M^{(\mathbf{k},\mathbf{b})(0,i-1)^{\dagger}} \right)_{ln} & (5.23) \\
&= \sum_{\mathbf{b}} w_b \sum_{l=1}^{N} \sum_{m',n'}^{N_{\mathbf{k}+\mathbf{b}}} \Bigg\{ M_{mm'}^{(\mathbf{k},\mathbf{b})(0)} \ A_{m'l}^{(i-1)}(\mathbf{k}+\mathbf{b}) & (5.24) \\
&\qquad\qquad A_{n'l}^{(i-1)^*}(\mathbf{k}+\mathbf{b}) \ M_{nn'}^{(\mathbf{k},\mathbf{b})(0)^*} \Bigg\} \ , &
\end{aligned}$$

where $A_{mn}^{(i)}(\mathbf{k}) = \langle u_{m\mathbf{k}}^{(0)} | u_{n\mathbf{k}}^{(i)} \rangle$ are the unitary matrices expressing states $|u_{n\mathbf{k}}^{(i)}\rangle$ at the i-th iteration step in the basis of initial functions $|u_{m\mathbf{k}}^{(0)}\rangle$ and $M_{mn}^{(\mathbf{k},\mathbf{b})(0,i)} = \langle u_{m\mathbf{k}}^{(0)} | u_{n\mathbf{k}+\mathbf{b}}^{(i)} \rangle$ are the inner product matrices between states at iteration step 0 and i. Please note, that the sum in Eqs. (5.23)–(5.24) runs only until $l = N$ and not $l = N_{\mathbf{k}}$ and, therefore, a complete matrix-matrix product is not given in these equations.

Determination of $\mathcal{S}^{(i)}(\mathbf{k})$ and convergence check
After the construction of the $Z(\mathbf{k})$-matrices, these matrices have to be diagonalized and the resulting composite matrices of eigenvectors[4] are represented by the

---

[4] Ordered by decreasing value of corresponding eigenvalue.

matrices $A^{(i)}_{mn}(\mathbf{k}) = \langle u^{(0)}_{m\mathbf{k}} | u^{(i)}_{n\mathbf{k}} \rangle$ introduced above. The leading N eigenvectors span the subspace $\mathcal{S}^{(i)}(\mathbf{k})$. It has to be checked whether $\mathcal{S}^{(i)}(\mathbf{k}) = \mathcal{S}^{(i-1)}(\mathbf{k})$ and if this is true, convergence is reached and the first N columns of the unitary matrices $A^{(i)}_{mn}(\mathbf{k})$ are utilized to construct the set of functions spanning the optimized subspace $\mathcal{S}^{(i)}(\mathbf{k}) = \mathcal{S}(\mathbf{k})$ according to

$$|u^{(i)}_{n\mathbf{k}}\rangle = \sum_{m=1}^{N} A^{(i)}_{mn}(\mathbf{k}) |u^{(0)}_{m\mathbf{k}}\rangle, \qquad n = 1 \dots N \, . \tag{5.25}$$

Updating of inner product matrices

If $\mathcal{S}^{(i)}(\mathbf{k}) \neq \mathcal{S}^{(i-1)}(\mathbf{k})$ still holds at iteration step i, the inner product matrices have to be updated using

$$M^{(\mathbf{k},\mathbf{b})(0,i-1)}_{mn} \quad = \quad \sum_{m'}^{N_{\mathbf{k}+\mathbf{b}}} M^{(\mathbf{k},\mathbf{b})(0)}_{mm'} A^{(i-1)}_{m'n}(\mathbf{k}+\mathbf{b}) \tag{5.26}$$

$$= \quad \left(M^{(\mathbf{k},\mathbf{b})(0)} A^{(i-1)}(\mathbf{k}+\mathbf{b}) \right)_{mn} \tag{5.27}$$

and the $Z(\mathbf{k})$-matrices of the subsequent iteration step $i+1$ can be calculated. The entire procedure is continued until convergence is obtained.

Matrix dimensions and initial values

For clarity, the dimensions of the occurring matrices are listed

$$M^{(\mathbf{k},\mathbf{b})(0,i-1)} : \; N_{\mathbf{k}} \times N_{\mathbf{k}+\mathbf{b}} \, , \tag{5.28}$$

$$A^{(i)}_{mn}(\mathbf{k}) : \; N_{\mathbf{k}} \times N_{\mathbf{k}} \, , \tag{5.29}$$

$$Z^{(i)}_{mn}(\mathbf{k}) : \; N_{\mathbf{k}} \times N_{\mathbf{k}} \, , \tag{5.30}$$

and also their initial values

$$A^{(0)}_{mn}(\mathbf{k}) \quad = \quad \delta_{mn} \, , \tag{5.31}$$

$$M^{(\mathbf{k},\mathbf{b})(0)}_{mn} \quad = \quad \langle u^{(0)}_{m\mathbf{k}} | u^{(0)}_{n\mathbf{k}+\mathbf{b}} \rangle \, , \tag{5.32}$$

$$Z^{(1)}_{mn}(\mathbf{k}) \quad = \quad \sum_{\mathbf{b}} w_b \sum_{l=1}^{N} M^{(\mathbf{k},\mathbf{b})(0)}_{ml} \; M^{(\mathbf{k},\mathbf{b})(0)*}_{nl} \, . \tag{5.33}$$

At this point it should be mentioned that for the calculation of the invariant part, Ω_{I}, of the spread functional Ω according to (5.2) the fully updated inner product matrices $M^{(\mathbf{k},\mathbf{b})(i)}_{mn} = A^{(i)\dagger}(\mathbf{k}) \; M^{(\mathbf{k},\mathbf{b})(0)} \; A^{(i)}(\mathbf{k}+\mathbf{b})$ rather than $M^{(\mathbf{k},\mathbf{b})(0,i)}_{mn}$ are required. Also the notations $A^{(\mathbf{k})(i)}_{mn}$ and $A^{(i)}_{mn}(\mathbf{k})$ are synonym. Latter one has been preferred here due to the additional iteration index i.

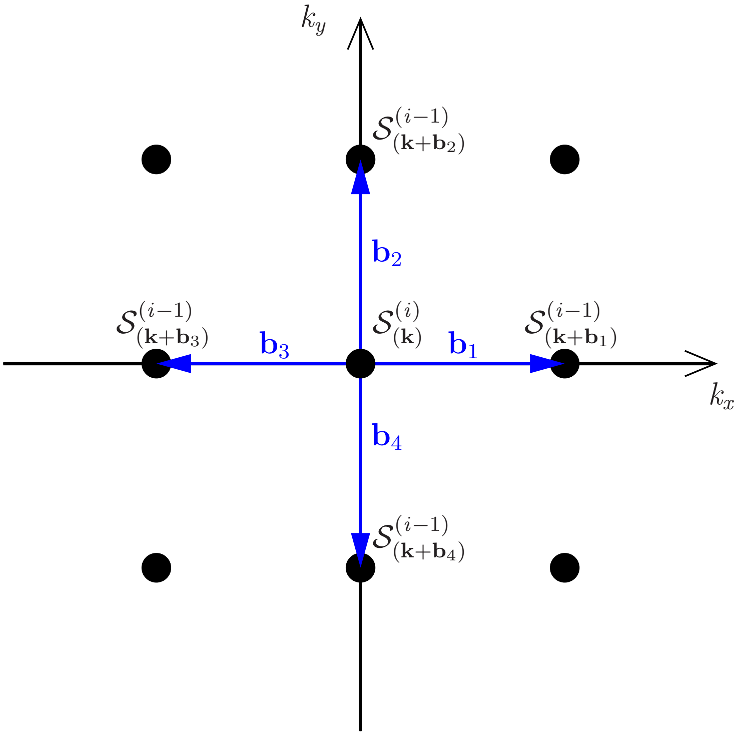


Figure 5.3: *Section of a Monkhorst-Pack discretization of the first Brillouin Zone of a square lattice PC around a general wave vector* $\mathbf{k}$. *The mesh is uniformly spaced with distance* $b = |\mathbf{b}|$ *and the vectors* $\mathbf{b}$ *point from* $\mathbf{k}$ *to its nearest neighbors. The figure also illustrates, that for the determination of the optimal subspace* $\mathcal{S}^{(i)}(\mathbf{k})$ *the "combined spillage" between the neighboring spaces* $\mathcal{S}^{(i-1)}(\mathbf{k}+\mathbf{b})$ *of the previous iteration step is calculated.*

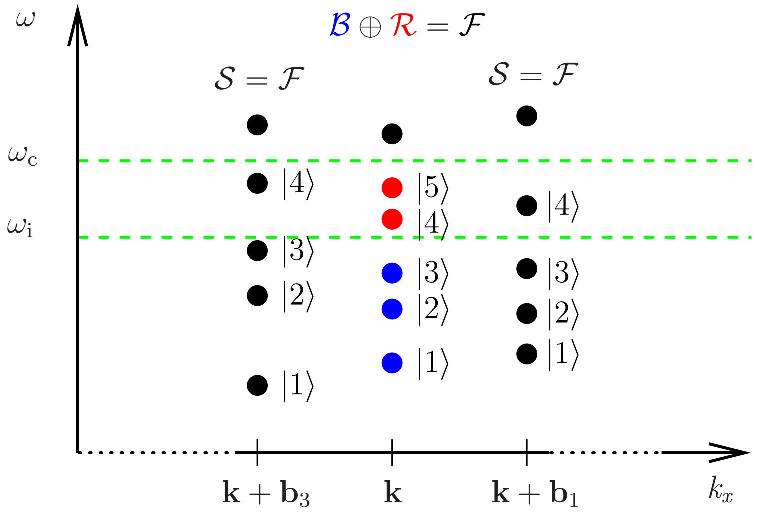


Figure 5.4: *Dispersion relation along the k_x-axis of the Brillouin zone as shown already in Fig. (5.2). At each k-point the additional cutoff frequency ω_i defines an inner frequency window containing an $M_\mathbf{k}$ dimensional vector space $\mathcal{B}(\mathbf{k})$ spanned by all states $|u_{n\mathbf{k}}\rangle$ that belong to eigenfrequencies $\omega_{n\mathbf{k}} \leq \omega_i$. The functions spanning $\mathcal{B}(\mathbf{k})$ are blocked and do not take part in the subspace optimization procedure. Therefore, the only functions which are "rotated" are the $N_\mathbf{k} - M_\mathbf{k}$ states spanning the space $\mathcal{R}(\mathbf{k})$ obeying $\mathcal{R}(\mathbf{k}) \oplus \mathcal{B}(\mathbf{k}) = \mathcal{F}(\mathbf{k})$. Metaphorically speaking, one could talk of sets of "hot" (red circles) and "frozen" (blue circles) bands.*

5.1.5 Inner window and frozen bands

To prevent the "pollution" of lower lying bands of $\mathcal{S}(\mathbf{k})$ with mode informations of higher bands, having, for instance, the same symmetry type, it is in general advisable to introduce an additional inner frequency window, limited by ω_i, in which $M_\mathbf{k}$, with $M_\mathbf{k} \leq N \leq N_\mathbf{k}$, bands are "blocked" or "frozen", i.e. they are not taking part in the subspace optimization procedure and remain completely unchanged. An illustration of this situation is given in Fig. 5.4. For this case the representation of the $\hat{Z}(\mathbf{k})$-operator reads

$$Z^{(i)}_{mn}(\mathbf{k}) = \langle u^{(0)}_{m\mathbf{k}}| \sum_{\mathbf{b}} w_b P^{(i-1)}_{\mathbf{k}+\mathbf{b}} |u^{(0)}_{n\mathbf{k}}\rangle, \tag{5.34}$$

with

$$P^{(i)}_{\mathbf{k}} = \sum_{l=M_\mathbf{k}+1}^{N} |u^{(i)}_{l\mathbf{k}}\rangle\langle u^{(i)}_{l\mathbf{k}}| \,, \qquad m,n = 1\ldots(N_\mathbf{k} - M_\mathbf{k}) \,. \tag{5.35}$$

Besides this change, the entire algorithm remains as described above. The newly introduced frequency ω_i, determining the inner window, should thereby be chosen in such a way, that is it relatively certain, that the blocked bands are not carrying any mode information of higher bands not contained in $\mathcal{S}(\mathbf{k})$. This condition has to hold for any wave vector $\mathbf{k}$ of the first Brillouin Zone.

5.1.6 Preprocessing

As the subspace optimization algorithm described above states a fixed point iteration problem, a good initial guess on the optimal subspaces is important. The procedure is similar to the preprocessing described in Sec. 4.3.3. At first, a set of N trial functions, $T_n(\mathbf{r})$, is projected onto all considered BFs at each wave vector $\mathbf{k}$,

$$|\phi_{n\mathbf{k}}\rangle = \sum_{m=1}^{N_\mathbf{k}} |\psi_{m\mathbf{k}}\rangle\langle\psi_{m\mathbf{k}}|T_n\rangle = \sum_{m=1}^{N_\mathbf{k}} A_{mn}^{(\mathbf{k})}\,|\psi_{m\mathbf{k}}\rangle\ , \qquad n = 1\ldots N\ . \tag{5.36}$$

The obtained functions $|\phi_{n\mathbf{k}}\rangle$ can be orthonormalized by Löwdin's symmetric orthogonalization via

$$|\widetilde{\psi}_{n\mathbf{k}}\rangle = \sum_{m=1}^{N}\left(S^{(-1/2)}\right)_{mn}|\phi_{m\mathbf{k}}\rangle = \sum_{m=1}^{N_\mathbf{k}}\left(AS^{-1/2}\right)_{mn}|\psi_{m\mathbf{k}}\rangle \tag{5.37}$$

where $S_{mn} = \langle\phi_{m\mathbf{k}}|\phi_{n\mathbf{k}}\rangle = (A^\dagger A)_{mn}$. The matrix $AS^{-1/2}$ is readily obtainable as $AS^{-1/2} = Z\mathbb{1}V$, where Z and V are unitary matrices of dimensions $N_\mathbf{k}\times N_\mathbf{k}$ and $N\times N$, respectively, resulting from the singular-value decomposition $A = ZDV$, and $\mathbb{1}$ represents the $N_\mathbf{k}\times N$ identity matrix. A vector space $\mathcal{G}(\mathbf{k})$ is now defined as the N-dimensional space spanned by the newly constructed states, $|\widetilde{\psi}_{n\mathbf{k}}\rangle$:

$$\mathcal{G}(\mathbf{k}) = \left\{\ |\widetilde{\psi}_{n\mathbf{k}}\rangle,\ n = 1\ldots N\right\}\ . \tag{5.38}$$

The projector onto $\mathcal{G}(\mathbf{k})$ is given by

$$P_\mathcal{G}(\mathbf{k}) = \sum_{n=1}^{N}|\widetilde{\psi}_{n\mathbf{k}}\rangle\langle\widetilde{\psi}_{n\mathbf{k}}|\ . \tag{5.39}$$

Its representation in the basis of the $N_\mathbf{k}$ initial states, $|\psi_{m\mathbf{k}}\rangle$, is given by

$$\begin{aligned}
\Big(\underbrace{P_\mathcal{G}(\mathbf{k})}_{N_\mathbf{k}\times N_\mathbf{k}}\Big)_{mn} &= \sum_{n\prime=1}^{N}\langle\psi_{m\mathbf{k}}|\widetilde{\psi}_{n'\mathbf{k}}\rangle\langle\widetilde{\psi}_{n'\mathbf{k}}|\psi_{m\mathbf{k}}\rangle &(5.40)\\
&= \sum_{n\prime=1}^{N} A_{mn\prime}^{(\mathbf{k})}\,{A_{nn\prime}^{(\mathbf{k})}}^{*} &(5.41)\\
&= \Big(\underbrace{A^{(\mathbf{k})}}_{N_\mathbf{k}\times N}\ \underbrace{A^{(\mathbf{k})\dagger}}_{N\times N_\mathbf{k}}\Big)_{mn}. &(5.42)
\end{aligned}$$

The projector onto the space $\mathcal{B}(\mathbf{k})$ of the $M_\mathbf{k}$ inner window states is now defined as

$$P_\text{inner}(\mathbf{k}) = \sum_{n=1}^{M_\mathbf{k}}|\psi_{n\mathbf{k}}\rangle\langle\psi_{n\mathbf{k}}|\ , \tag{5.43}$$

and

$$\begin{aligned}
Q_{\text{inner}}(\mathbf{k}) &= 1 - P_{\text{inner}}(\mathbf{k}) && (5.44)\\
&= 1 - \sum_{n=1}^{M_{\mathbf{k}}} |\psi_{n\mathbf{k}}\rangle\langle\psi_{n\mathbf{k}}| && (5.45)\\
&= \sum_{n=M_{\mathbf{k}}+1}^{N_{\mathbf{k}}} |\psi_{n\mathbf{k}}\rangle\langle\psi_{n\mathbf{k}}| && (5.46)
\end{aligned}$$

is the projector onto its complementary space $\mathcal{R}(\mathbf{k})$ defined in the caption of Fig. 5.4. In the basis of initial states, $|\psi_{m\mathbf{k}}\rangle$, these projectors adopt the form

$$\begin{aligned}
\Big(\underbrace{P_{\text{inner}}(\mathbf{k})}_{N_{\mathbf{k}}\times N_{\mathbf{k}}}\Big)_{mn} &= \langle\psi_{m\mathbf{k}}|\Big\{\sum_{n\prime=1}^{M_{\mathbf{k}}} |\psi_{n\prime\mathbf{k}}\rangle\langle\psi_{n\prime\mathbf{k}}|\Big\}|\psi_{n\mathbf{k}}\rangle && (5.47)\\
&= \delta_{mn}\,\Theta(M_{\mathbf{k}}, m) && (5.48)
\end{aligned}$$

and

$$\begin{aligned}
\Big(\underbrace{Q_{\text{inner}}(\mathbf{k})}_{N_{\mathbf{k}}\times N_{\mathbf{k}}}\Big)_{mn} &= \delta_{mn} - \delta_{mn}\,\Theta(M_{\mathbf{k}}, m) && (5.49)\\
&= \delta_{mn}\,\Theta(m, M_{\mathbf{k}}+1) && (5.50)
\end{aligned}$$

where $\Theta(m,n) = 1$ for $m \geq n$ and $\Theta(m,n) = 0$ otherwise. Choosing now the remaining $N - M_{\mathbf{k}}$ basis vectors of the envisaged subspace $\mathcal{S}^{(0)}(\mathbf{k})$ to be the eigenvectors, $|\lambda_{n\mathbf{k}}\rangle$, corresponding to the $N - M_{\mathbf{k}}$ largest eigenvalues of

$$Q_{\text{inner}}(\mathbf{k})P_{\mathcal{G}}(\mathbf{k})Q_{\text{inner}}(\mathbf{k})|\lambda_{n\mathbf{k}}\rangle = \lambda_{n\mathbf{k}}|\lambda_{n\mathbf{k}}\rangle\ , \qquad (5.51)$$

yields —together with the unmodified $M_{\mathbf{k}}$ bands inside the inner window— the initial guess $\mathcal{S}^{(0)}(\mathbf{k})$ on the optimal subspace $\mathcal{S}(\mathbf{k})$.

5.2 The complete algorithm at a glance

Finally, after all elements of the band disentangling procedure have been introduced, a brief summary over all presented steps might help to keep the overview. Fig. 5.2 shows the consecutive sequence of procedures that have to be applied to achieve maximally localized Wannier functions. The main building blocks are the subspace optimization [47] and the spread minimization algorithm [46]. The first poses a multidimensional fixed point problem and the latter a multidimensional, highly nonlinear minimization problem. Therefore, both of these procedures require preceding preprocessing algorithms to achieve good initial guesses on the fixed points and global minima.

The application of this entire complex of algorithms for the generation of MLWFs for the triangular lattice model system (see Sec. 2.3.2) is featured in the subsequent section.

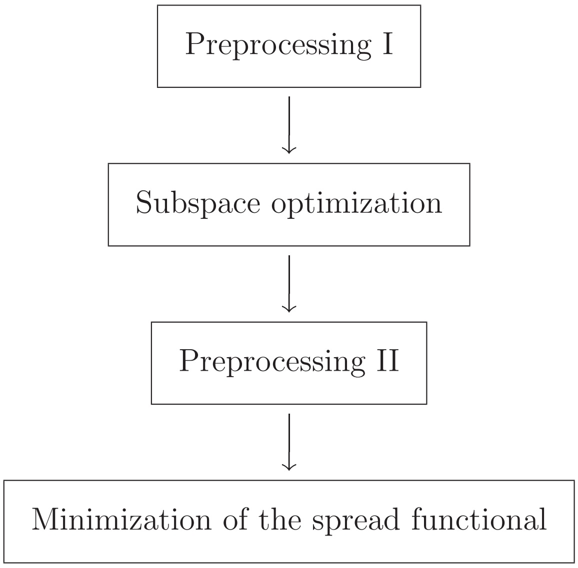


Figure 5.5: *Sequential structure of the complete algorithm, featuring the subspace optimization and spread minimization procedures with corresponding preprocessing steps. The subspace optimization artificially isolates a group of bands, for which maximally localized Wannier functions can be created by means of the spread minimization algorithm.*

5.3 Application to the triangular lattice model system

In the following, the procedures described above are used to construct MLWFs for the triangular lattice model system, i.e. for air pores etched into silicon and H-polarized light. This system probably states one of the most difficult cases for the construction of symmetric MLWF. This is due to the fact that for such inverse structures on a triangular lattice there is no high-symmetry point of the direct lattice in high index regions (i.e. in the silicon), which has the full symmetry of the lattice. Therefore, the WFs corresponding to material modes tend to hybridize and form tiny, well confined functions with centers on general positions inside the WSC. For this case, the minimum-spread trials described in Sec. 4.3.3.2 do not state a good guess on the MLWFs and are, therefore, less suited for preprocessing. In general, as it is completely unknown in which way the functions are going to hybridize, there are no good sets of trial functions available. However, it turned out that the symmetry of the trial functions is of paramount importance in these systems, and using the full-symmetric set of trial functions, introduced in Sec. 4.3.3.2, a failure of convergence to the global minimum was never observed. But compared to the square lattice system way more iteration steps, in general several thousands or tens of thousands, are required to align the hybridized WFs and reach the global minimum of the spread functional. The subspace optimization algorithm seems to be unaffected, though.

Another profound problem occurring systems is to determine a-priori which bands have to be grouped together such that symmetrical MLWF are obtained[5]. While “unsymmetric” functions, in principle, state an equally well-suited expansion basis for defect calculations, numerical inaccuracies and asymmetries in the overlap matrix elements make them less applicable. On the other hand, improper values for the cutoff frequencies ω_i and ω_c, can lead to asymmetric MLWFs, too.

One might, therefore, imagine that finding a proper grouping of bands, in combination with suitably defined inner and outer frequency windows is rather difficult and tricky.

Despite all these complications it was possible to find suitable choices for the grouping of bands and the cutoff frequencies for the considered model system. It turned out that treating the bands 2–9 as a composite group yields a set of symmetrical MLWFs. The reasons why symmetric functions are obtained right for this grouping of bands are given in Chap. 6.

In Fig. 5.6 the photonic band structure of the considered model system is depicted in a frequency range that encloses these bands. An outer frequency window is defined by $\omega_\mathrm{c} = 1.12$ and an inner frequency window by $\omega_\mathrm{i} = 0.88$. The maximally localized Wannier functions obtained for this band complex are exhibited in Fig. 5.8. This set of MLWFs consists of two dipole-like E_1-modes located at the origin and six highly localized functions with centers inside the background material which

---

[5] An example for this is given by the sp^2 hybrids displayed in Fig. 4.6(b) on p. 58: the hybrids correspond to the global minimum of the spread functional, but do not match the symmetry of the lattice. Therefore, this grouping of bands is considered to be inappropriate.

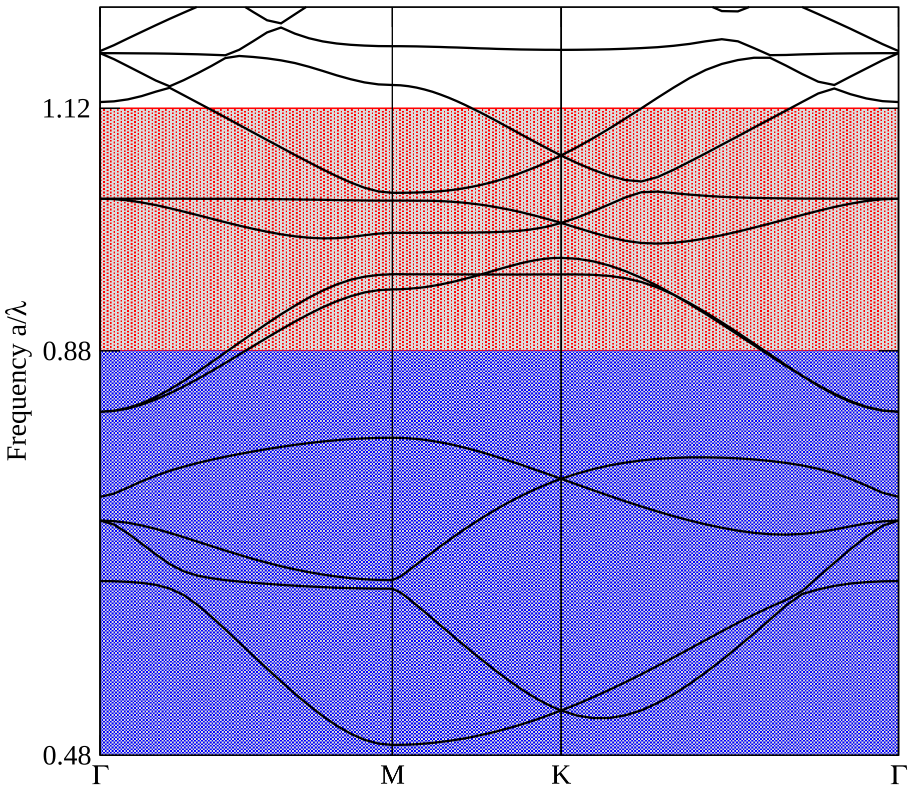


Figure 5.6: *Band structure for the triangular lattice model system in the frequency range* $a/\lambda = 0.48$–1.12*. The purpose is to construct MLWF for bands 2–9. An outer frequency window is defined by* $\omega_c = 1.12$ *and an inner frequency window by* $\omega_i = 0.88$*. The bands in the inner window (blue region) are blocked, i.e. they are not modified by the subspace optimization algorithm. Only the "hot" bands in the region between* ω_c *and* ω_i *(red region) are allowed to be mixed by means of unitary transformations. Metaphorically speaking, the band structure is melted in the red region, the bands are disentangled, and a set of bands that leads to the best global smoothness of connection is glued on top of the set of frozen bands.*

can be arranged in a star-like formation around the origin. For comparison, the corresponding WFs for this band complex, obtained by directly applying the LFT, Eq. (2.29), to the "bare" BFs as they are computed by standard band structure codes, are shown in Fig. 5.7 right above. It is not difficult to see that the localization has improved tremendously.

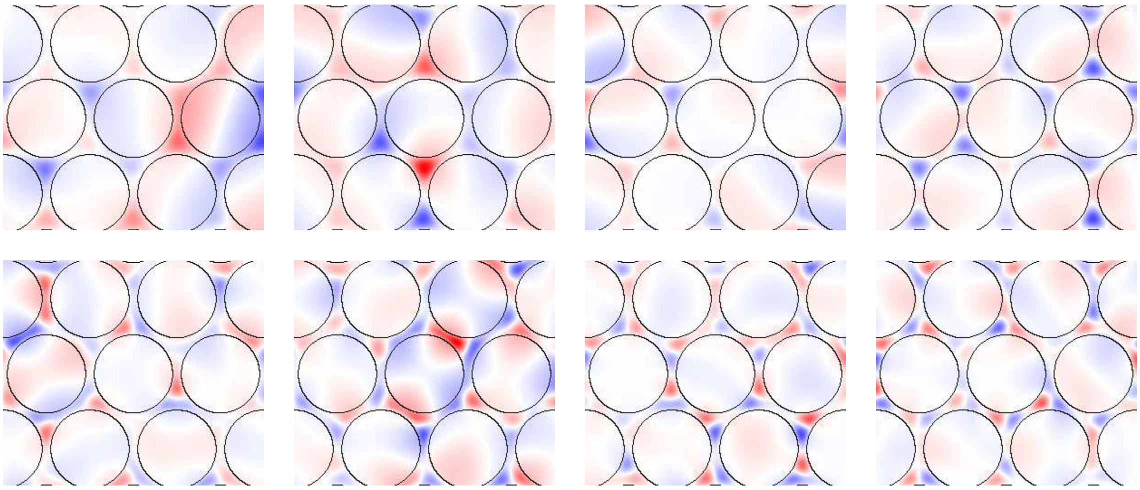


Figure 5.7: *Delocalized Wannier functions for the band range 2–9 of the triangular lattice model system, specified in Fig. 2.4(b), created by applying the Lattice Fourier Transform, Eq.* (2.29)*, to the "bare" Bloch functions as obtained by standard band structure computation. Though being by definition Wannier functions they show poor localization and are spread all over the computational domain.*

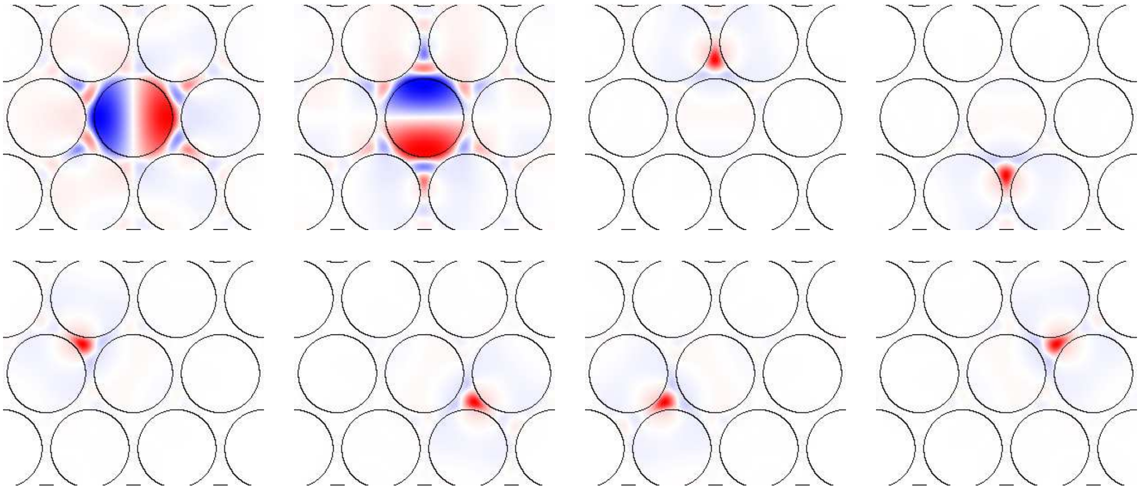


Figure 5.8: *Maximally localized Wannier functions for the band range 2–9 of the considered model system for a triangular lattice, specified in Fig. 2.4(b), featuring two dipole-like* E_1*-modes and six highly localized functions which can be arranged star-like around the origin. For the classification of functions according to irreducible representations of the point group of the underlying lattice, please see Sec. 2.3.3.*

6 Bottom-up approach

In the previous chapter finally symmetric and maximally localized Wannier functions have been obtained for the band complex 2–9 with the methods [46, 47] borrowed from solid state theory. However, in general and especially for the considered model systems, 10 Wannier functions are only sufficient to describe weak defects accurately. In contrast, in so-called direct structures, i.e. dielectric rods in air, an accurate description of stronger defects such as omitting rods completely is usually easily feasible, as demonstrated in Ref. [48]. This is due to the fact that many Bloch functions and, therefore, also the Wannier functions tend to be localized in the high-index material. When creating defects by omitting rods, these defects are created exactly at positions where many WFs have their center. In contrast to this, for the inverse structures considered here, i.e. pores etched into a high-index material, only a few WFs centered at the origin do in general exist. As infiltrating the air voids creates defects exactly at the origin and equivalent positions, way more bands have to be taken into account to get a sufficient number of WFs having centers inside the pore. It is important to emphasize at this point, that more WFs have to be constructed at first, but it is in general a good strategy to select a small subset out of the entire set of available functions that describes the given defect sufficiently well, out of the entire set of available functions.

Calculating a significantly large number of WFs is not appropriately feasible with the so far discussed methods. One could, of course, artificially isolate a very large group of bands from higher ones and would probably obtain MLWFs for this configuration. However, the obtained WFs will be wildly hybridized, building tiny functions centered at general positions in the WSC.

When using such hybrids for the description of defect structures, the contribution of all these functions to the obtained defect modes will be roughly of the same order. In addition, these hybrids do no longer transform under symmetry operations according to IREPs of the point group of the lattice. In contrast to this, the construction of WFs for a narrow frequency range prevents those functions located at the origin from hybridizing with higher order modes, and hence, they still belong to a certain IREP. As the defect modes of single cylindrical defects themselves transform according to IREPs, a WF centered at the defect site and belonging to the same IREP would state a more efficient basis for the calculation, as potentially higher order WFs belonging to different IREPs might be omitted completely.

In a nutshell one could say that strongly hybridized WFs abandon the possibility to describe a certain type of defect sufficiently well with a small subset out of the set of constructed Wannier functions. For accurate calculations all constructed functions would have to be taken into account to describe any kind of defect, no matter how weak. Therefore, these hybridized modes would completely kill the performance of the Wannier function approach.

In addition, a large group of bands slows down the construction time of WFs tremendously as the number of matrix elements grows quadratically with the number of composite bands leading to performance penalties, e.g. caused by missing cache optimizations for matrix-matrix products, as well as to numerical inaccuracies due to arising singularities of the involved unitary matrices.

6.1 Description and application

Due to the above mentioned complications it is desirable to keep the groups of bands as small as possible. Motivated by this fact the idea arose to apply the subspace optimization and spread minimization ascending step by step successively in the band structure. At each wave vector $\mathbf{k}$, the subspace optimization selects out of a larger set of $N_{\mathbf{k}}$ functions, N states spanning the optimal subspace $\mathcal{S}(\mathbf{k})$. The remaining $N_{\mathbf{k}} - N$ functions are in the original algorithm [47] effectively lost and thrown away. However, one could instead consider these “garbage” modes as the new $N_{\mathbf{k}} - N$ basis vectors of the remaining vector space, open to higher bands. For this space a new group of bands has to be found and artificially isolated by means of the subspace optimization algorithm, producing new “garbage” modes and so on and so forth.

Such an successive application of the subspace optimization and spread minimization methods is illustrated in Fig. 6.1. Blue areas indicate the “frozen” bands and reddish shaded regions represent bands that take part in the subspace optimization procedure. The corresponding cutoff frequencies are indicated on the frequency axis. A pictorial view of this algorithm would be to think of the connecting elements of neighbored vector spaces being cut and welded together such that the resulting spaces are not connected anymore.

The set of highly symmetric and maximally localized Wannier functions constructed by means of this novel bottom-up approach is depicted in Fig. 6.2. There are twelve functions which are centered at the origin and transform under the point group of the lattice according to irreducible representations of the group C_{6v}. In addition, there are four groups, each consisting of six functions which can be brought into a star-like arrangement around the origin and transform according to a reducible representation of C_{6v}. The last two WFs are highly localized functions centered at the K-point of the direct lattice. As their transformation under elements of C_{6v} involves their images in neighboring cells, a description solely with respect to the point group of the lattice is less suited.

This set of functions proves that a recursive application of the subspace optimization and spread minimization procedures is indeed successful, yielding a highly symmetric set of maximally localized Wannier functions with air-moded functions centered at the origin and transforming according to an IREP of the point group C_{6v} of the underlying triangular lattice.

When respecting the naturally given grouping of bands a set of symmetric functions will be obtained inherently. However, when the bands are entangled the choice of the last band contained in a group is completely at will of the user. Due to the improvements through special sets of trial functions and proper choice of the finite step width α for the spread minimization, it has become possible to almost always obtain a set of MLWFs for any grouping of bands, but only for certain combinations highly symmetric ones are obtained.

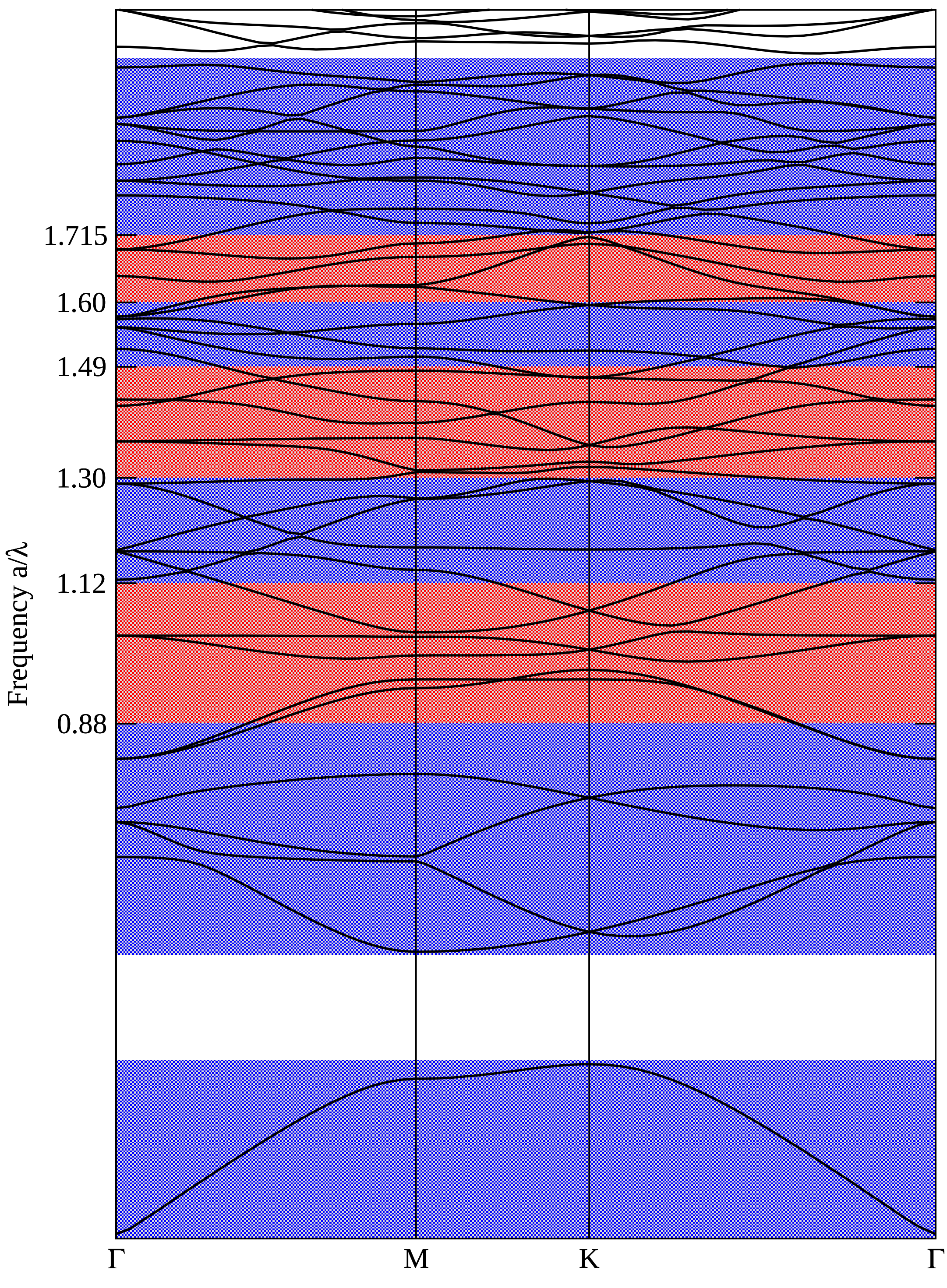


Figure 6.1: *Band structure for the triangular lattice model system in a frequency range enclosing the first 38 bands. It displays the positions of the inner and outer frequency windows. Metaphorically speaking, the bands in the reddish shaded frequency ranges are melted, disentangled, and afterwards glued together in a way such that the resulting vector spaces are quasi-isolated.*

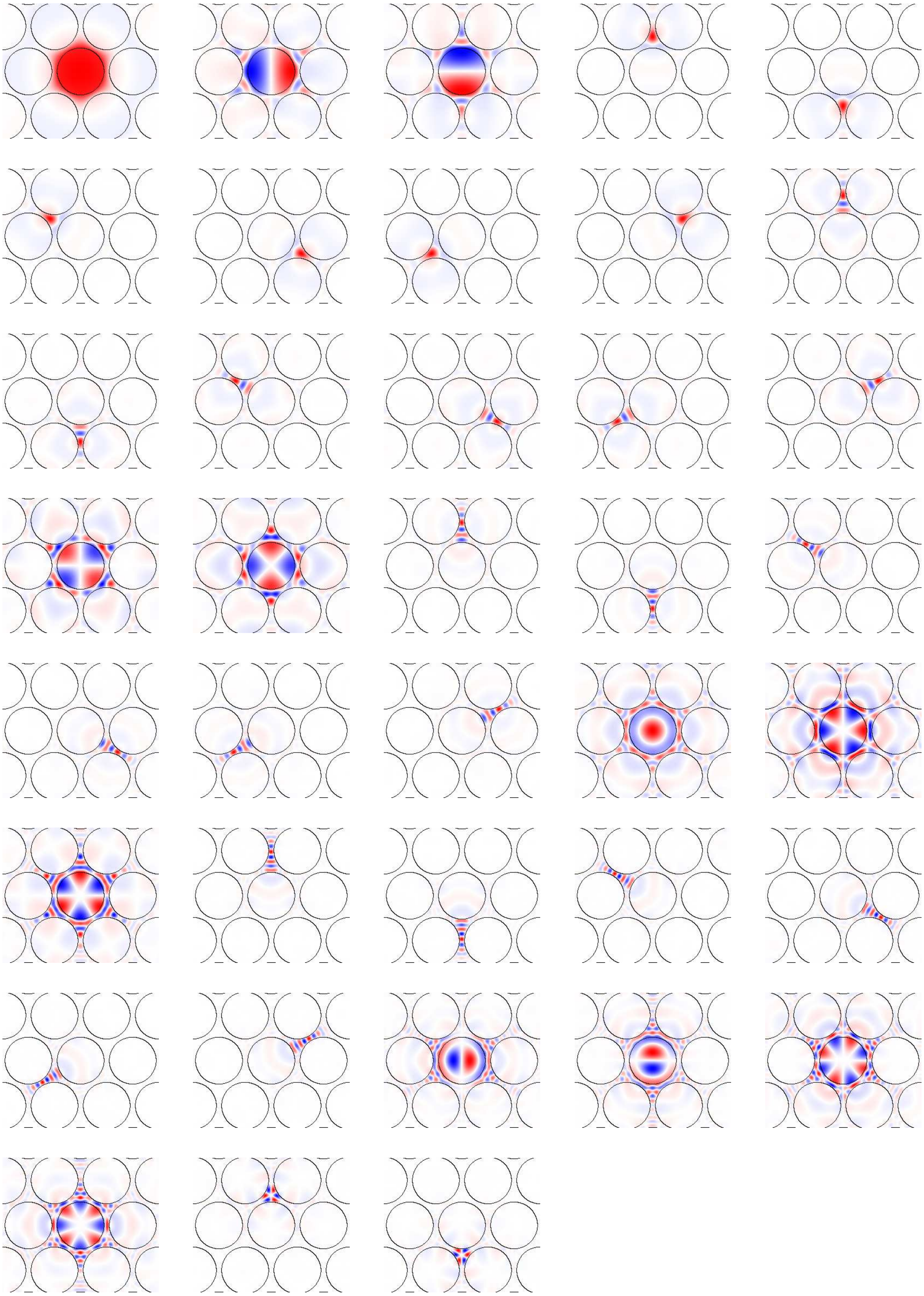


Figure 6.2: *The first 38 symmetric and maximally localized Wannier functions for the model system described in Fig. 2.4(b). The grouping of the bands and corresponding cutoff frequencies are given in Fig. 6.1.*

6.2 Band selection rules

To derive certain band selection rules for obtaining symmetric maximally localized Wannier functions, one first has to specify what "symmetric", "symmetry compliant", or "lattice compliant" means. For this purpose the following definitions are made where "group" is not meant in a mathematical sense unless explicitly stated:

Complete group of bands

A complete group of bands states a set of N composite bands so that

$$\sum_{n=1}^{N} \mathbf{r}_n = \mathbf{R} \tag{6.1}$$

holds for the obtained maximally localized Wannier functions, with $\mathbf{r}_n$ being the center of the n-th Wannier function and $\mathbf{R}$ a lattice vector of the direct lattice.

Complete and irreducible group of bands

If a complete group of bands contains at least two subgroups of bands with each of them being a complete group in the above defined sense, then such a group is called reducible. If such subgroups cannot be found the group is called irreducible.

As the only spatial point having the full symmetry of the lattice is the origin and its periodic images, it is observed and also known [104, 105] that the WFs belonging to the central WSC transform under elements R of the point group $\mathcal{M}$ of the lattice according to[1]

$$\hat{D}(R)\ W_n(\mathbf{r}) = \sum_{m=1}^{N} D_{mn}(R)\ W_m(\mathbf{r} - \mathbf{R}_m(R)) \ , \tag{6.2}$$

where $\mathbf{R}_m(R)$ is an optional lattice vector depending on the applied symmetry operation. For instance, for all WFs with centers inside the central WSC $\mathbf{R}_m(R) = \mathbf{0}$ holds, for all m and R. But for the WFs with indices 37 and 38, for instance, one finds

$$\hat{D}(C_6)\, W_{37}(\mathbf{r}) = \mathbf{W}_{38}(\mathbf{r} + \mathbf{a}_1) \ , \tag{6.3}$$

and thus $\mathbf{R}_{38}(C_6) = -\mathbf{a}_1$, with the primitive translation $\mathbf{a}_1$ as indicated in Fig. 2.2(b) on p. 8.

It is claimed now that for a complete and irreducible group of bands as defined above the vector space $\mathcal{S}$ spanned by the obtained MLWFs is decomposed by them into direct sums of spaces $\mathcal{S}_{\text{ir}}^{(\alpha)}$ on which the representation of C_{6v} is irreducible, and direct sums of hybridization spaces $\mathcal{S}_{\text{hyb}}^{(\beta)}$ on which the representations are six dimensional[2] and reducible, i.e.

$$\mathcal{S} = \bigoplus_{\alpha} \mathcal{S}_{\text{ir}}^{(\alpha)} \oplus \bigoplus_{\beta} \mathcal{S}_{\text{hyb}}^{(\beta)} \ . \tag{6.4}$$

---

[1] In formal agreement to [104, 105], but with $\boldsymbol{\varrho}_i = \mathbf{0}$ for all bands i, as the only possible symmetry center for a triangular lattice is given by the origin.

[2] For functions centered at points on the WSC boundary their translational images are taken into account to complete the six dimensional vector space.

In addition, the WFs spanning a hybridization space $\mathcal{S}_{\mathrm{hyb}}^{(\beta)}$ are generated by the cyclic group C_6 according to

$$W_n(\mathbf{r}) = C_6^{n-1} \, W_1(\mathbf{r}) \; , \tag{6.5}$$

where n labels all functions spanning the hybridization space $\mathcal{S}_{\mathrm{hyb}}^{(\beta)}$. If, for instance, one WF out of this space is known all other functions can be obtained by multiple rotations of 60°. Pictorially speaking, this requires that the functions — taking translational images into account if required— can be arranged in a star-like manner around the origin. Hence, their centers sum up to zero. As translational images are not taken into account when summing over the centers of the obtained WFs, (6.1) is zero up to a lattice vector $\mathbf{R}$. The space $\mathcal{S}_{\mathrm{ir}}^{(\alpha)}$ is built up from functions belonging to IREPs of C_{6v} and are, therefore, necessarily centered at the origin.

Functions of the hybridization space with centers inside the WSC show even symmetry with respect to σ_x mirror reflections and can be combined to six functions belonging to the irreducible representations A_1, B_2, E_1, E_2. Such Wannier functions belonging to IREPs of C_{6v} are shown in Fig. 6.3. The functions with centers at the high-symmetry point K of the direct lattice can, together with their translational images, be combined to functions belonging to the representations B_1 and A_2.

On the other hand it is known [89, 104] that the symmetry types of Wannier functions already determined by the symmetry types of the Bloch functions at k-points of highest symmetry, i.e. the Γ-point.

Hence, one might make the following definition:

Complete hybridization group of bands

A complete hybridization group is built by bands whose Bloch functions at the Γ*-point have pronounced intensities in the high-index regions of the WSC and consist exactly of six (Type I) or two bands (Type II) such that each symmetry type out of the set* A_1, B_2, E_1, E_2*, or* B_1, A_2*, respectively, is contained exactly once.*

Through this, it is possible to determine a complete and irreducible group of bands, yielding symmetric Wannier functions, by means of the symmetries of the BFs at the Γ-point as outlined in the following:

At first, all BFs, depicted in Fig. 6.4, which represent air-modes with obviously pronounced intensities[3] inside the pores, might be marked with a cross and disregarded. It is now claimed that a complete and irreducible group of bands is given as soon as all hybridization groups contained are complete.

The first band is isolated. Hence, the otherwise nontrivial determination of complete and irreducible groups starts with band 2. This band opens a hybridization group of type I and this group is closed by band 9. Therefore, the first complete and irreducible group of bands is found, and indeed, for this group symmetric WFs have

---

[3] The phase choice of the BFs at the Γ-point is such that the function value representing highest intensity is chosen to be purely real and positive. As positive values are indicated by red colors in all presented function plots, monopoles showing a blue color inside the pores do not have highest intensities there.

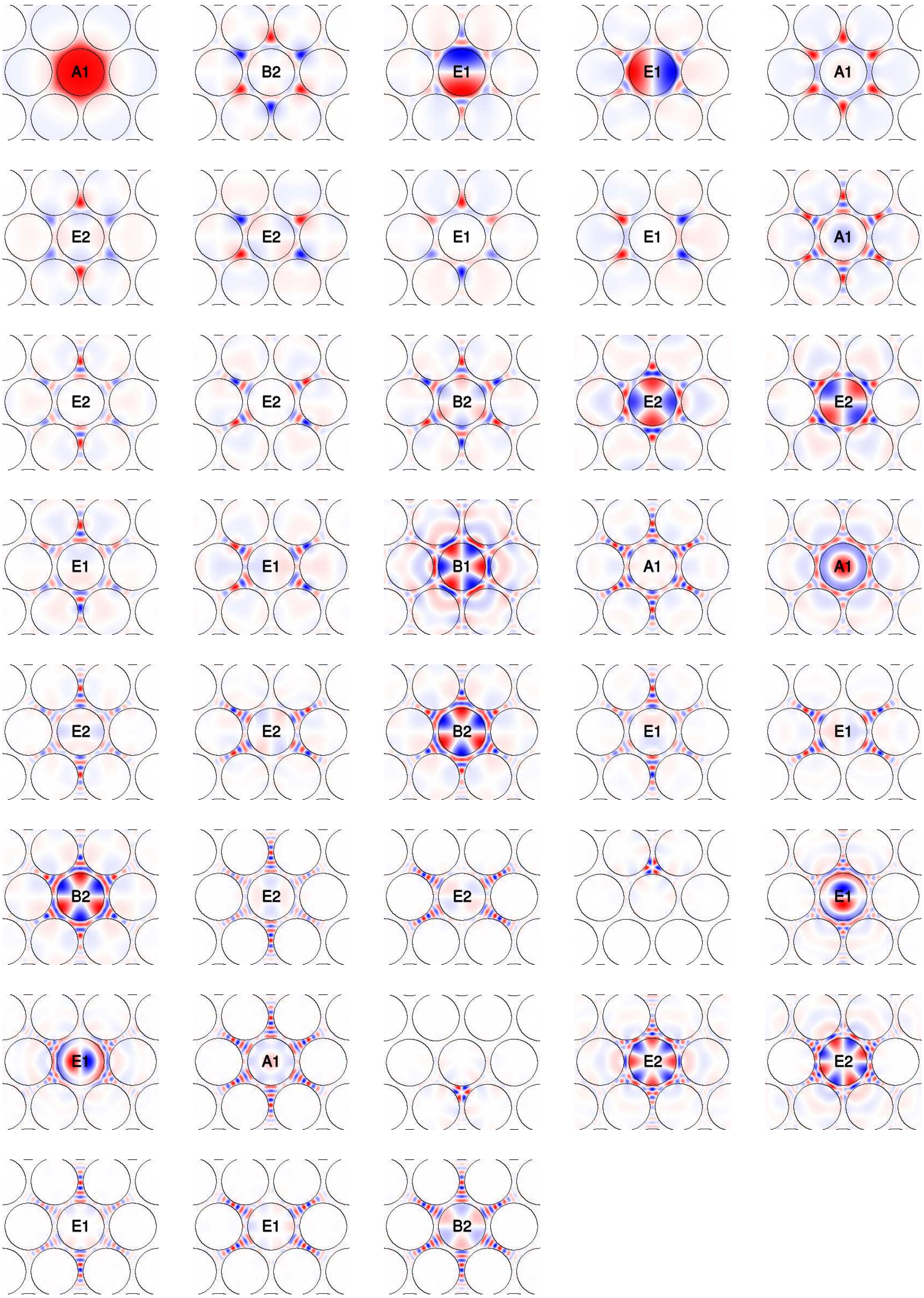


Figure 6.3: *Wannier functions transforming according to IREPs of the point group* C_{6v} *obtained from the maximally localized Wannier functions displayed in Fig. 6.2. The symmetry type is already determined by the Bloch functions at the* Γ*-point which are displayed in Fig. 6.4.*

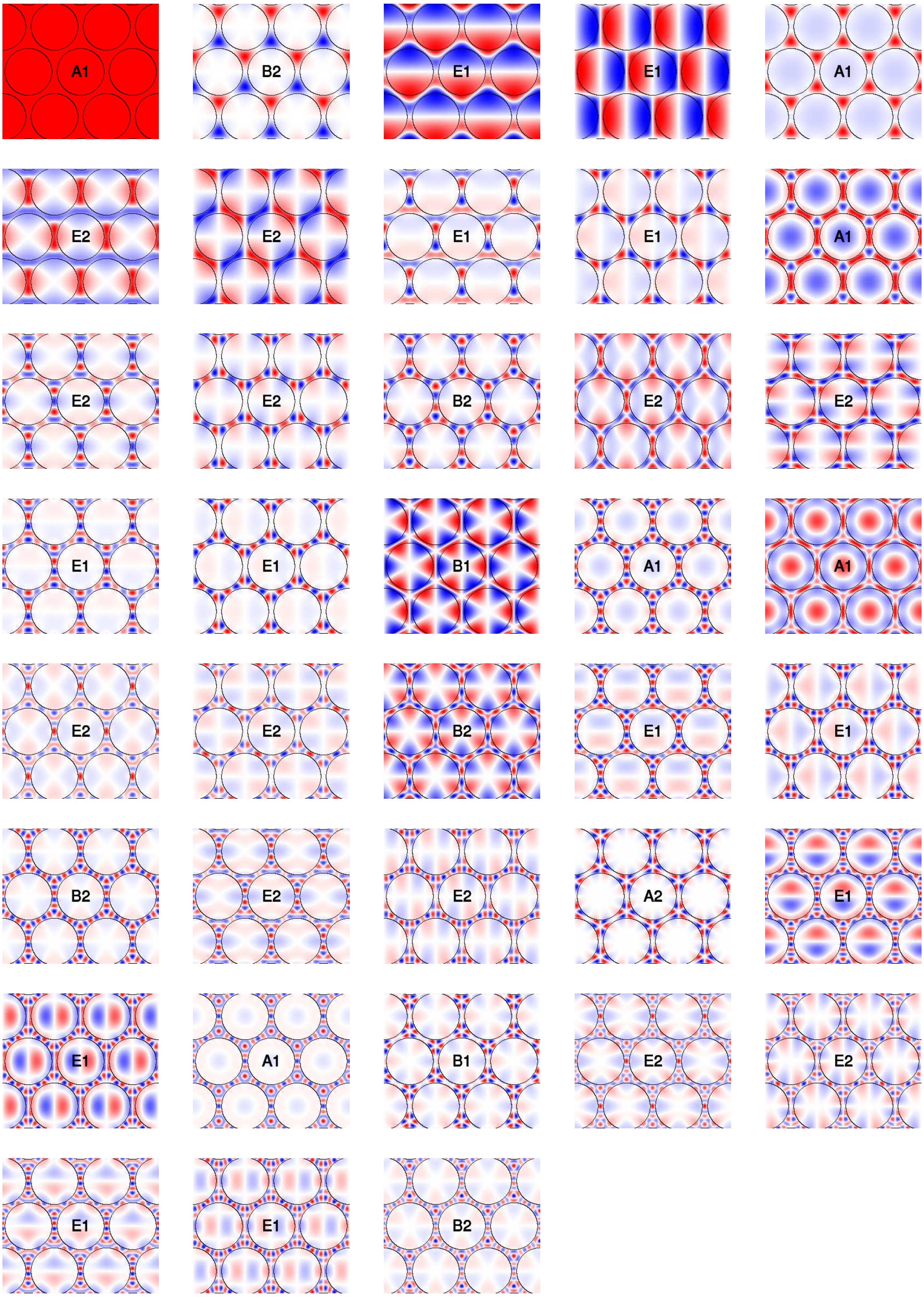


Figure 6.4: *The first 38 Bloch functions at the Γ-point classified with respect to the irreducible representations of the C_{6v} point group of the underlying triangular lattice.*

been obtained. Band 10 opens the next type I hybridization group, which is closed with band 17. The subsequent hexapole is an air-mode and the corresponding WF will belong to an IREP and, therefore, span a subspace of type $\mathcal{S}_{\mathrm{ir}}$. Hence, it might be absorbed into the group 10–17 or into the following group. This decision has to be made by inspecting the band structure diagram given in Fig. 6.1. As the potential dispersion curve of band 18 crosses only higher bands, band 18 is chosen to open the next group, which is later completed by band 26. The following group, which is also the last group considered, contains hybridization groups of type I and II and is completed by band 38, where also a band gap in the band structure appears. It is worth mentioning at this point that the group 2–38, defined naturally through band gaps, states a complete but reducible group of bands, which was successfully decomposed into four complete and irreducible groups of bands corresponding to the definitions given above.

Therefore, by means of the newly developed band selection rules, it is now possible to determine a priori an optimal grouping of bands for which the resulting maximally localized Wannier functions are perfectly reflecting the symmetry of the underlying lattice. These selection rules might not be unique, as the classification of BFs into air- and material-modes might be difficult. However, they reduce the number of possible band configurations significantly.

Through this, further group theoretical investigations and extensions to other lattice types seem to be very promising, especially with regard to three dimensional Photonic Crystals.

Summary and future prospects

Within this thesis the methods of Marzari and Vanderbilt [46] and Souza *et al.* [47] for the construction of maximally localized Wannier functions, developed in the context of electronic structure theory, have been recalculated, successfully implemented and transferred to periodically structured dielectric systems, i.e. Photonic Crystals. These methods represent multidimensional, highly nonlinear minimization and fixed point problems with in general several tens of thousands of unknowns. For these highly sensible procedures, novel sets of trial functions, based on the artificial localization of Bloch functions at the Γ-point, have been developed that have proven to yield adequate initial values for the unknowns and thus enabled the stable convergence to the fixed point and to the global minimum in any case and for both methods, and can even reach the global minimum directly for certain systems.

Furthermore, it was possible to extend the methods above to be recursively combinable, allowing the construction of arbitrarily many symmetry compliant maximally localized Wannier functions by successively ascending in the band structure as described in Chap. 6. This is demanded by the bosonic nature of photons, disallowing the accurate description of defect structures embedded in PCs by means of only a few MLWFs corresponding to a narrow frequency window around a characteristic energy. In contrast to this, the physical properties of crystalline solids, for which the methods mentioned above have been developed originally, are mostly determined by electronic states with energies close to the Fermi level.

The applicability of this novel bottom-up approach was proven by constructing 38 symmetry compliant maximally localized Wannier functions for a Photonic Crystal consisting of a triangular lattice of air pores etched in silicon where the magnetic field vector of the electromagnetic radiation was aligned parallel to the pore axes (H-polarization). It has been pointed out that maximally localized Wannier functions do not inherently reflect the symmetry of the underlying lattice. Only for certain band complexes containing selective symmetry types MLWFs can be obtained which are maximally localized and symmetry compliant at the same time. Explicit band selection rules for the a priori determination of these groupings of bands by means of the symmetries of the Bloch functions at the Γ-point were presented for the model system described above.

The efficiency and accuracy of Wannier function based lattice models for the theoretical description of defect structures embedded in PCs was demonstrated in Chap. 3, where novel designs for basic functional elements, such as waveguide bends, intersections and couplers, have been presented which were obtained in collaborative works [75, 84, 82].

From a practical point of view, a computer software package for the largely automated construction of maximally localized Wannier functions has been developed within this dissertation. It features the automatic refurbishing of Bloch functions at the Γ-point yielding purely real and nicely aligned sets of functions (Figs. 2.6 and 2.8), well suited for the subsequent creation of artificially localized sets of trial functions (Sec. 4.3.3.2). These steps, together with a symmetry analysis[4] of the Bloch functions at the Γ-point, are all performed automated in a first configuration step of the software package providing all information required to find appropriate choices for the grouping of the individual bands. After inspecting the corresponding band structure to determine adequate cutoff frequencies as outlined in Chap. 5, the construction of maximally localized Wannier functions is evoked with the band range and corresponding cutoff frequencies as command line arguments, performing the subspace optimization and spread minimization procedures with all required preprocessing steps and automatically constructing and plotting the obtained functions.

In this regard, promising future prospects are the extension of the existing software package to allow for the automated construction of fully vectorial 3D vector Wannier functions for three dimensional Photonic Crystals. The further group theoretical investigation and extension of the presented band selection rules seems to be of great advantage, particularly for 3D systems. An automated classification of the Bloch functions at the Γ-point according to irreducible representations of the point group of the underlying lattice together with an computer-aided determination of the "character" of these modes, i.e. whether they tend to have high intensities in the high-index or rather in the low-index material, could significantly ease the determination of the correct grouping of entangled bands, as the visualization of three-dimensional vector fields is rather complicated and the symmetries of the Bloch functions are not as obvious as in 2D Photonic Crystals.

As a final remark one might argue that the newly introduced bottom-up approach accompanied by explicit band selection rules, the automation of the construction process, together with several numerical subtleties which have been identified, might help to establish Wannier functions 70 years after their invention [43] as a standard numerical tool for the highly efficient computational description of physical phenomena in periodically structured systems, such as crystalline solids [46, 47], Photonic Crystals [48], and, potentially, ultra cold atomic and molecular gases in optical lattices [106].

---

[4] Implemented by Dipl. Phys. Patrick Mack as a part of his final year project [57].

Appendices

A Proofs

To improve the readability of this thesis all longer proofs have been sourced out into this appendix. Cross-references are given in the text.

A.1 Matrix elements of the position operator

Proposition A.1. $\langle r^2 \rangle_n$ *can be expressed in terms of periodic parts of Bloch functions as*

$$\langle r^2 \rangle_n = \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \ \left| |\nabla_{\mathbf{k}} u_{n\mathbf{k}} \rangle \right|^2 .$$

Proof. Using Eq. (4.13) and taking matrix elements only between Wannier functions located at the center cell yields

$$\langle r^2 \rangle_n = -\frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \ \langle u_{n\mathbf{k}} | \nabla^2_{\mathbf{k}} | u_{n\mathbf{k}} \rangle \ . \tag{A.2}$$

Integrating the right hand side by parts leads to

$$\begin{aligned} \langle r^2 \rangle_n &= -\frac{1}{V_{\mathrm{BZ}}} \left\{ \int\limits_{\partial\mathrm{BZ}} d^{(d-1)}\mathbf{k} \ \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n\mathbf{k}} \rangle - \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \ \langle \nabla_{\mathbf{k}} u_{n\mathbf{k}} | \nabla_{\mathbf{k}} u_{n\mathbf{k}} \rangle \right\} && \text{(A.3)} \\ &\overset{B.6}{=} \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \ \left| |\nabla_{\mathbf{k}} u_{n\mathbf{k}} \rangle \right|^2 . && \text{(A.4)} \end{aligned}$$

□

A.2 Expansion of overlap matrix elements

The minimization procedure is based on the expansion

$$M^{(\mathbf{k},\mathbf{b})}_{nn} = \langle u_{n\mathbf{k}} | u_{n\mathbf{k}+\mathbf{b}} \rangle = 1 + ixb + \frac{1}{2} y b^2 + O(b^3) \tag{A.5}$$

where x, y are supposed to be real numbers. To prove this, $\langle u_{n\mathbf{k}+\mathbf{b}}|$ and $|u_{n\mathbf{k}+\mathbf{b}}\rangle$ are expanded for small b

$$
\begin{aligned}
|u_{n\mathbf{k}+\mathbf{b}}\rangle &= |u_{n\mathbf{k}}\rangle + b\sum_i |\partial_{k_i} u_{n\mathbf{k}}\rangle\, e_i + \frac{1}{2}\, b^2 \sum_{i,j} |\partial_{k_i}\partial_{k_j} u_{n\mathbf{k}}\rangle\, e_i e_j + O(b^3) \quad \text{(A.6)}\\
\langle u_{n\mathbf{k}+\mathbf{b}}| &= \langle u_{n\mathbf{k}}| + b\sum_i \langle \partial_{k_i} u_{n\mathbf{k}}|\, e_i + \frac{1}{2}\, b^2 \sum_{i,j} \langle \partial_{k_i}\partial_{k_j} u_{n\mathbf{k}}|\, e_i e_j + O(b^3) \quad \text{(A.7)}
\end{aligned}
$$

where $\mathbf{e} = \mathbf{b}/b$ is the unit vector pointing along the direction of $\mathbf{b}$ and $\partial_{k_i} = \partial/\partial_{k_i}$ with $i = x, y, z$. Projecting (A.6) onto $|u_{n\mathbf{k}}\rangle$ leads to the expansion of $M_{nn}^{(\mathbf{k},\mathbf{b})}$

$$
\langle u_{n\mathbf{k}}|u_{n\mathbf{k}+\mathbf{b}}\rangle = \langle u_{n\mathbf{k}}|u_{n\mathbf{k}}\rangle + b\, \underbrace{\sum_i \langle u_{n\mathbf{k}}|\partial_{k_i} u_{n\mathbf{k}}\rangle e_i}_{ix} + \frac{1}{2}\, b^2\, \underbrace{\sum_{i,j} \langle u_{n\mathbf{k}}|\partial_{k_i}\partial_{k_j} u_{n\mathbf{k}}\rangle e_i e_j}_{y} + O(b^3)\ . \quad \text{(A.8)}
$$

To show that x, y are real numbers, $\langle u_{n\mathbf{k}+\mathbf{b}}|u_{n\mathbf{k}+\mathbf{b}}\rangle$ is expanded up to second order in b, as it is suggested in [46]:

$$
\begin{aligned}
\langle u_{n\mathbf{k}+\mathbf{b}}|u_{n\mathbf{k}+\mathbf{b}}\rangle = {} & \langle u_{n\mathbf{k}}|u_{n\mathbf{k}}\rangle \\
& + b \sum_i \left\{ \langle u_{n\mathbf{k}}|\partial_{k_i} u_{n\mathbf{k}}\rangle + \langle \partial_{k_i} u_{n\mathbf{k}}|u_{n\mathbf{k}}\rangle \right\} e_i \\
& + b^2 \sum_{i,j} \left\{ \langle \partial_{k_i} u_{n\mathbf{k}}|\partial_{k_j} u_{n\mathbf{k}}\rangle + \frac{1}{2}\langle u_{n\mathbf{k}}|\partial_{k_i}\partial_{k_j} u_{n\mathbf{k}}\rangle + \frac{1}{2}\langle \partial_{k_i}\partial_{k_j} u_{n\mathbf{k}}|u_{n\mathbf{k}}\rangle \right\} e_i e_j \\
& + O(b^3) \\
& = 1\ . \qquad \text{(A.9)}
\end{aligned}
$$

Equating the first order in b yields

$$
\sum_i \left\{ \langle u_{n\mathbf{k}}|\partial_{k_i} u_{n\mathbf{k}}\rangle + \langle \partial_{k_i} u_{n\mathbf{k}}|u_{n\mathbf{k}}\rangle \right\} e_i \quad = \quad 0 \qquad \text{(A.10)}
$$

which must hold for arbitrary directions $\mathbf{e}$, i.e.

$$
\Leftrightarrow \quad \langle u_{n\mathbf{k}}|\partial_{k_i} u_{n\mathbf{k}}\rangle + \langle u_{n\mathbf{k}}|\partial_{k_i} u_{n\mathbf{k}}\rangle^* = 0 \qquad \text{(A.11)}
$$

$$
\Leftrightarrow \quad \mathrm{Re}\left\{ \langle u_{n\mathbf{k}}|\partial_{k_i} u_{n\mathbf{k}}\rangle \right\} = 0\ , \qquad \text{(A.12)}
$$

which proves that the first order coefficient of $M_{nn}^{(\mathbf{k},\mathbf{b})}$ can be written as

$$
\sum_i \langle u_{n\mathbf{k}}|\partial_{k_i} u_{n\mathbf{k}}\rangle\ e_i = i\, x
$$

with x being a real number.

To prove that y is a real number, too, the definition of y in equation (A.8) is expressed as the trace of a matrix product

$$\begin{aligned} y &= \sum_{i,j} \underbrace{\langle u_{n\mathbf{k}} | \partial_{k_i} \partial_{k_j} u_{n\mathbf{k}} \rangle}_{=:S_{ij}} \underbrace{e_i\, e_j}_{:=E_{ij}} \\ &= \sum_{i,j} S_{ij}\, E_{ji} \\ &= \operatorname{Tr}(S\, E) \ , \end{aligned}$$

where it was used that E is symmetric. Again, due to this property of E, it is recognizable, that for $\operatorname{Tr}(S\, E)$ being a purely real number, S must be hermitian, i.e. $S_{ji} = S^*_{ij}$. But since the order of differentiation might be interchanged, it is already known that S is symmetric. Thus, finding S to be hermitian would also imply $\operatorname{Im} S_{ij} = 0$ for all i, j. To obtain a deterministic expression for S, which should state that $S_{ji} = S^*_{ij}$ or $\operatorname{Im} S_{ij} = 0$ for all i, j, the second orders in b of Eq. (A.9) are equated as suggested in [46], yielding

$$\begin{aligned} \sum_{i,j} \left\{ \langle \partial_{k_i} u_{n\mathbf{k}} | \partial_{k_j} u_{n\mathbf{k}} \rangle + \frac{1}{2} \langle u_{n\mathbf{k}} | \partial_{k_i} \partial_{k_j} u_{n\mathbf{k}} \rangle + \frac{1}{2} \langle \partial_{k_i} \partial_{k_j} u_{n\mathbf{k}} | u_{n\mathbf{k}} \rangle \right\} e_i\, e_j &= 0 \\ \sum_{i,j} \left\{ \underbrace{\langle \partial_{k_i} u_{n\mathbf{k}} | \partial_{k_j} u_{n\mathbf{k}} \rangle}_{=:H_{ij}} + \operatorname{Re} \left(\langle u_{n\mathbf{k}} | \partial_{k_i} \partial_{k_j} u_{n\mathbf{k}} \rangle \right) \right\} e_i\, e_j &= 0 \\ \sum_{i,j} \underbrace{(H_{ij} + \operatorname{Re} S_{ij})}_{=:A_{ij}} E_{ji} &= 0 \\ \operatorname{Tr}(A\, E) = \sum_i A_{ii} E_{ii} + \sum_i \sum_{j>i} (A_{ij} + A_{ji}) E_{ij} &= 0 \end{aligned}$$

which is exactly fulfilled if

$$\begin{aligned} A_{ij} &= - A_{ji} \\ \Leftrightarrow H_{ij} + \operatorname{Re} S_{ij} &= - H_{ji} - \operatorname{Re} S_{ji} \\ \Leftrightarrow 2 \operatorname{Re} S_{ij} &= - (H_{ij} + H_{ji}) \\ \Leftrightarrow \operatorname{Re} S_{ij} &= - \operatorname{Re} H_{ij} \ , \end{aligned} \tag{A.13}$$

where it was used that $S_{ji} = S_{ij}$ and $H_{ji} = H^*_{ij}$. However, in contrast to the expectation, the last equation does state only a relationship between the real part of S and the real part of the hermitian matrix H, and is, therefore, completely independent of the imaginary part of S. This means that the necessity of y being a real number for the replacements (4.41) on p. 46 to be valid as claimed in [46] could not be confirmed here. Instead, as the expansion (4.34) contains only a term proportional to the real part of y, the above restriction on y might easily be dropped[1]. In addition,

---

[1] This is also stressed in [97]

by re-expressing $H_{ij} = \langle \partial_{k_i} u_{n\mathbf{k}} | \partial_{k_j} u_{n\mathbf{k}} \rangle$ as

$$\begin{aligned} \langle \partial_{k_i} u_{n\mathbf{k}} | \partial_{k_j} u_{n\mathbf{k}} \rangle &= \partial_{k_j} \langle \partial_{k_i} u_{n\mathbf{k}} | u_{n\mathbf{k}} \rangle - \langle \partial_{k_j} \partial_{k_i} u_{n\mathbf{k}} | u_{n\mathbf{k}} \rangle && \text{(A.14)} \\ &= \partial_{k_j} c_i(\mathbf{k}) - \langle \partial_{k_j} \partial_{k_i} u_{n\mathbf{k}} | u_{n\mathbf{k}} \rangle && \text{(A.15)} \end{aligned}$$

the equations derived above are of further advantage. This becomes apparent when re-substituting the definitions of S_{ij} and H_{ij} into Eq. (A.13) again:

$$\mathrm{Re}\langle u_{n\mathbf{k}} | \partial_{k_i} \partial_{k_j} u_{n\mathbf{k}} \rangle = - \mathrm{Re}\langle \partial_{k_i} u_{n\mathbf{k}} | \partial_{k_j} u_{n\mathbf{k}} \rangle \; . \tag{A.16}$$

Recalling now that

$$\langle u_{n\mathbf{k}} | \nabla^2_{\mathbf{k}} | u_{n\mathbf{k}} \rangle = \sum_i \langle u_{n\mathbf{k}} | \partial_{k_i} \partial_{k_i} u_{n\mathbf{k}} \rangle \; ,$$

and

$$\frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \; \langle u_{n\mathbf{k}} | \nabla^2_{\mathbf{k}} u_{n\mathbf{k}} \rangle = \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \; \mathrm{Re}\langle u_{n\mathbf{k}} | \nabla^2_{\mathbf{k}} u_{n\mathbf{k}} \rangle \tag{A.17}$$

due to $|u_{n\,-\mathbf{k}}\rangle = |u_{n\mathbf{k}}\rangle^*$, one arrives at

$$\frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \; \langle u_{n\mathbf{k}} | \nabla^2_{\mathbf{k}} u_{n\mathbf{k}} \rangle = -\frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \; \sum_i \mathrm{Re}\langle \partial_{k_i} u_{n\mathbf{k}} | \partial_{k_i} u_{n\mathbf{k}} \rangle \; . \tag{A.18}$$

As a diagonal element of the hermitian matrix H, each scalar product $\langle \partial_{k_i} u_{n\mathbf{k}} | \partial_{k_i} u_{n\mathbf{k}} \rangle$ is a purely real quantity and, therefore, the Re can be dropped. This leads finally to

$$\begin{aligned} \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \; \langle u_{n\mathbf{k}} | \nabla^2_{\mathbf{k}} u_{n\mathbf{k}} \rangle &= - \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \; \sum_i \langle \partial_{k_i} u_{n\mathbf{k}} | \partial_{k_i} u_{n\mathbf{k}} \rangle \\ &= - \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \; \sum_i \left| |\partial_{k_i} u_{n\mathbf{k}}\rangle \right|^2 \\ &= - \frac{1}{V_{\mathrm{BZ}}} \int\limits_{\mathrm{BZ}} d^d\mathbf{k} \; \left| |\nabla_{\mathbf{k}} u_{n\mathbf{k}}\rangle \right|^2 \end{aligned}$$

which states another proof for Proposition A.1 on p. 93.

A.3 Behavior of the spread functional under special phase transformations

In the following the behavior under the transformation

$$|u_{n\mathbf{k}}\rangle \rightarrow |u_{n\mathbf{k}}\rangle\, e^{-i\mathbf{kR}}, \tag{A.19}$$

where $\mathbf{R}$ is a primitive translation vector of the underlying lattice, is listed for several quantities:

$$\langle u_{n\mathbf{k}}|u_{n\mathbf{k+b}}\rangle \xrightarrow{\text{(A.19)}} \langle u_{n\mathbf{k}}|u_{n\mathbf{k+b}}\rangle\, e^{-i\mathbf{bR}} \tag{A.20a}$$

$$M_{nn}^{(\mathbf{k},\mathbf{b})} \xrightarrow{\text{(A.19)}} M_{nn}^{(\mathbf{k},\mathbf{b})}\, e^{-i\mathbf{bR}} \tag{A.20b}$$

$$\ln M_{nn}^{(\mathbf{k},\mathbf{b})} \xrightarrow{\text{(A.19)}} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} - i\mathbf{bR} \tag{A.20c}$$

$$\operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})} \xrightarrow{\text{(A.19)}} \operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})} - \mathbf{bR}\ . \tag{A.20d}$$

Using the above, the Wannier centers $\mathbf{r}_n$, defined through Eq. (4.33), transform under this transformation according to

$$\begin{aligned}\mathbf{r}_n \quad \xrightarrow{\text{(A.19)}} \quad & -\frac{1}{N_{\text{kp}}}\sum_{\mathbf{k},\mathbf{b}} w_b\, \mathbf{b}\big[\operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})} - \mathbf{bR}\big] \\ & = \mathbf{r}_n + \sum_{\mathbf{b}} w_b\, \mathbf{b}\big(\mathbf{bR}\big) \\ & \overset{B.4}{=} \mathbf{r}_n + \mathbf{R}\ .\end{aligned} \tag{A.21}$$

The discretization of the quadratic term in the spread functional given in Eq. (4.37) transforms like

$$\begin{aligned}\langle r^2\rangle_n \quad \xrightarrow{\text{(A.19)}} \quad & -\frac{1}{N}\sum_{\mathbf{k},\mathbf{b}} w_b\, \Big\{1 - |M_{nn}^{(\mathbf{k},\mathbf{b})}|^2 + \big[\operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})}\big]^2 \\ & \qquad -2\,\mathbf{bR}\big[\operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})}\big] + \big(\mathbf{bR}\big)^2\Big\} \\ & = \langle r^2\rangle_n + 2\,\mathbf{r}_n\mathbf{R} + \sum_{\mathbf{b}} w_b\,\big(\mathbf{bR}\big)^2 \\ & = \langle r^2\rangle_n + 2\,\mathbf{r}_n\mathbf{R} + R^2\ .\end{aligned} \tag{A.22}$$

Putting all this together leads to the fact that the spread functional Ω stays invariant under (A.19):

$$\begin{aligned}\langle r^2\rangle_n - \mathbf{r}_n^2 \xrightarrow{\text{(A.19)}} \; & \langle r^2\rangle_n + 2\,\mathbf{r}_n\mathbf{R} + R^2 - \big(\mathbf{r}_n + \mathbf{R}\big)^2 = \\ & \langle r^2\rangle_n - \mathbf{r}_n^2\end{aligned}$$

$$\Omega \xrightarrow{\text{(A.19)}} \Omega\ .$$

Hence, the modified finite difference formulae (4.33) and (4.37) indeed yield a discrete formulation of the spread functional Ω that is invariant under the transformation (A.19).

A.4 Splitting of the spread functional into invariant, off-diagonal, and diagonal parts

Proposition A.23. *A splitting of the spread functional* Ω *according to*

$$\Omega = \Omega_I + \Omega_{OD} + \Omega_D \tag{A.24}$$

is possible. Here, Ω_I *,* Ω_{OD}*, and* Ω_D *are called the invariant, off-diagonal, and diagonal parts of the spread functional* Ω *and are defined in discretized form as*

$$\Omega_I = \frac{1}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \Big(N - \sum_{m,n} \big| M_{mn}^{(\mathbf{k},\mathbf{b})} \big|^2 \Big) \,, \tag{A.25}$$

$$\Omega_{OD} = \frac{1}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \sum_{m \neq n} \big| M_{mn}^{(\mathbf{k},\mathbf{b})} \big|^2 \,, \tag{A.26}$$

$$\Omega_D = \frac{1}{N_{kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \sum_{n} \Big(- \operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} - \mathbf{b}\mathbf{r}_n \Big)^2 \,. \tag{A.27}$$

Proof. Starting with

$$\begin{aligned} \Omega_{\mathrm{I,OD}} &:= \Omega_{\mathrm{I}} + \Omega_{\mathrm{OD}} \\ &= \frac{1}{N_{\mathrm{kp}}} \sum_{n} \sum_{\mathbf{k},\mathbf{b}} w_b \Big[1 - \big| M_{nn}^{(\mathbf{k},\mathbf{b})} \big|^2 \Big] \,, \end{aligned} \tag{A.28}$$

$$\Omega_{\mathrm{D}} = \frac{1}{N_{\mathrm{kp}}} \sum_{n} \sum_{\mathbf{k},\mathbf{b}} w_b \Big[\operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} + \mathbf{b}\, \mathbf{r}_n \Big]^2 \,, \tag{A.29}$$

one obtains

$$\begin{aligned}
\Omega &= \Omega_{\mathrm{I,OD}} + \Omega_{\mathrm{D}} \\
&= \frac{1}{N_{\mathrm{kp}}} \sum_n \sum_{\mathbf{k},\mathbf{b}} \Big\{ w_b \Big[1 - \big|M_{nn}^{(\mathbf{k},\mathbf{b})}\big|^2 \Big] \\
&\qquad\qquad + w_b \Big[\big(\operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})}\big)^2 + 2 \operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})} \mathbf{b}\mathbf{r}_n + \big(\mathbf{b}\mathbf{r}_n\big)^2 \Big] \Big\} \\
&= \frac{1}{N_{\mathrm{kp}}} \sum_n \sum_{\mathbf{k},\mathbf{b}} w_b \Big\{ \Big[1 - \big|M_{nn}^{(\mathbf{k},\mathbf{b})}\big|^2 \Big] + \Big[\operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})} \Big]^2 \Big\} \\
&\quad - 2 \Big\{ -\frac{1}{N_{\mathrm{kp}}} \sum_n \sum_{\mathbf{k},\mathbf{b}} w_b \mathbf{b} \operatorname{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})} \Big\} \mathbf{r}_n \\
&\quad + \sum_{\mathbf{b}} w_b \big(\mathbf{b}\mathbf{r}_n\big)^2 \\
&\overset{B.5}{=} \sum_n \Big(\langle r^2 \rangle_n - 2\mathbf{r}_n\mathbf{r}_n + \mathbf{r}_n^2 \Big) \\
&= \sum_n \Big(\langle r^2 \rangle_n - \mathbf{r}_n^2 \Big) .
\end{aligned}$$

$\square$

A.5 Derivation of the gradient

For convenience and in analogy to [46] the calculation of the gradient takes advantage of the decomposition of the spread functional Ω into

$$\Omega = \Omega_{\mathrm{I,OD}} + \Omega_{\mathrm{D}} , \tag{A.30}$$

derived in the previous section. The first order changes $d\Omega_{\mathrm{I,OD}}$ and $d\Omega_{\mathrm{D}}$ with respect to $dW_{mn}^{(\mathbf{k})}$ will be derived individually and recombined afterwards again to yield the total change $d\Omega$.

As both summands in Eq. (A.30) are functions of the overlap matrix elements $M_{nn}^{(\mathbf{k},\mathbf{b})}$, it has to be known how these matrix elements are changing with respect to the infinitesimal matrices $dW_{mn}^{(\mathbf{k})}$ introduced in Sec. 4.3.2.2. The first order changes $dM_{nn}^{(\mathbf{k},\mathbf{b})}$ of the overlap matrix elements $M_{nn}^{(\mathbf{k},\mathbf{b})}$ are in general defined by

$$M_{nn}^{(\mathbf{k},\mathbf{b})} = M_{nn}^{(\mathbf{k},\mathbf{b})^{(0)}} + dM_{nn}^{(\mathbf{k},\mathbf{b})} . \tag{A.31}$$

where $M_{nn}^{(\mathbf{k},\mathbf{b})^{(0)}}$ designates the initial matrices and $M_{nn}^{(\mathbf{k},\mathbf{b})}$ the ones after an infinitesimal "rotation" by means of the matrices $dW_{mn}^{(\mathbf{k})}$. Using their definition (4.24) and relating the rotated functions $|u_{n\mathbf{k}}\rangle$ to the initial ones $|u^{(0)}_{n\mathbf{k}}\rangle$ by means of Eq. (4.51) yields

$$
\begin{aligned}
M^{(\mathbf{k},\mathbf{b})}_{nn} &= \langle u_{n\mathbf{k}}|u_{n\mathbf{k}+\mathbf{b}}\rangle \\
&= \Big\{\langle u^{(0)}_{n\mathbf{k}}| + \sum_m \langle u^{(0)}_{m\mathbf{k}}| dW^{(\mathbf{k})\dagger}_{nm}\Big\}\Big\{|u^{(0)}_{n\mathbf{k}+\mathbf{b}}\rangle + \sum_{m'} dW^{(\mathbf{k}+\mathbf{b})}_{m'n}|u^{(0)}_{m'\mathbf{k}+\mathbf{b}}\rangle\Big\} \\
&= M^{(\mathbf{k},\mathbf{b})^{(0)}}_{nn} - \sum_m dW^{(\mathbf{k})}_{nm}\langle u^{(0)}_{m\mathbf{k}}|u^{(0)}_{n\mathbf{k}+\mathbf{b}}\rangle + \sum_{m'} dW^{(\mathbf{k}+\mathbf{b})}_{m'n}\langle u^{(0)}_{n\mathbf{k}}|u^{(0)}_{m'\mathbf{k}+\mathbf{b}}\rangle \\
&= M^{(\mathbf{k},\mathbf{b})^{(0)}}_{nn} - \Big[dW^{(\mathbf{k})}M^{(\mathbf{k},\mathbf{b})^{(0)}}\Big]_{nn} + \Big[M^{(\mathbf{k},\mathbf{b})^{(0)}}dW^{(\mathbf{k}+\mathbf{b})}\Big]_{nn} \\
&= M^{(\mathbf{k},\mathbf{b})^{(0)}}_{nn} - \Big[dW^{(\mathbf{k})}M^{(\mathbf{k},\mathbf{b})^{(0)}}\Big]_{nn} - \Big[dW^{(\mathbf{k}+\mathbf{b})}M^{(\mathbf{k}+\mathbf{b},-\mathbf{b})^{(0)}}\Big]^*_{nn} \qquad \text{(A.32)}
\end{aligned}
$$

where it was used that $dW^{(\mathbf{k})}_{mn}$ is anti-hermitian and that $M^{(\mathbf{k},\mathbf{b})}_{mn} = M^{(\mathbf{k}+\mathbf{b},-\mathbf{b})^\dagger}_{mn}$. From the latter equation the first order change $dM^{(\mathbf{k},\mathbf{b})}_{nn}$ for a general $M^{(\mathbf{k},\mathbf{b})}_{nn}$ is easily read off as

$$
dM^{(\mathbf{k},\mathbf{b})}_{nn} = -\Big[dW^{(\mathbf{k})}M^{(\mathbf{k},\mathbf{b})}\Big]_{nn} - \Big[dW^{(\mathbf{k}+\mathbf{b})}M^{(\mathbf{k}+\mathbf{b},-\mathbf{b})}\Big]^*_{nn} . \qquad \text{(A.33)}
$$

This expression is now used to derive the first order changes of $\Omega_{\mathrm{I,OD}}$ and Ω_{D}.

At first $d\Omega_{\mathrm{I,OD}}$ is going to be calculated. Based on its definition through Eq. (A.28) the first order change of $\Omega_{\mathrm{I,OD}}$ is given by

$$
d\Omega_{\mathrm{I,OD}} = -\frac{1}{N_{\mathrm{kp}}}\sum_n \sum_{\mathbf{k},\mathbf{b}} w_b \; d\Big(\big|M^{(\mathbf{k},\mathbf{b})}_{nn}\big|^2\Big) . \qquad \text{(A.34)}
$$

Hence, for the derivation of $d\Omega_{\mathrm{I,OD}}$ it has to be known how $\big|M^{(\mathbf{k},\mathbf{b})}_{nn}\big|^2$ changes based on (A.33). In this regard, latter quantity is expressed as

$$
\big|M^{(\mathbf{k},\mathbf{b})}_{nn}\big|^2 = M^{(\mathbf{k},\mathbf{b})}_{nn}M^{(\mathbf{k},\mathbf{b})^*}_{nn} \qquad \text{(A.35)}
$$

and considered to be a function of the two independent variables $M^{(\mathbf{k},\mathbf{b})}_{nn}$ and $M^{(\mathbf{k},\mathbf{b})^*}_{nn}$. Using Wirtinger calculus the change of this quantity is

$$
d\Big(M^{(\mathbf{k},\mathbf{b})}_{nn}M^{(\mathbf{k},\mathbf{b})^*}_{nn}\Big) = M^{(\mathbf{k},\mathbf{b})^*}_{nn}dM^{(\mathbf{k},\mathbf{b})}_{nn} + M^{(\mathbf{k},\mathbf{b})}_{nn}dM^{(\mathbf{k},\mathbf{b})^*}_{nn} , \qquad \text{(A.36)}
$$

and, therefore,

$$
\begin{aligned}
d\Omega_{\mathrm{I,OD}} =& -\frac{1}{N_{\mathrm{kp}}}\sum_n\sum_{\mathbf{k},\mathbf{b}} w_b\Big\{M_{nn}^{(\mathbf{k},\mathbf{b})}dM_{nn}^{(\mathbf{k},\mathbf{b})*}+M_{nn}^{(\mathbf{k},\mathbf{b})*}dM_{nn}^{(\mathbf{k},\mathbf{b})}\Big\}\\
=& -\frac{1}{N_{\mathrm{kp}}}\sum_n\sum_{\mathbf{k},\mathbf{b}} w_b\\
&\Bigg\{M_{nn}^{(\mathbf{k},\mathbf{b})}\Big(-\Big[dW^{(\mathbf{k})}M^{(\mathbf{k},\mathbf{b})}\Big]^*_{nn}-\Big[dW^{(\mathbf{k}+\mathbf{b})}M^{(\mathbf{k}+\mathbf{b},-\mathbf{b})}\Big]_{nn}\Big)\\
&+M_{nn}^{(\mathbf{k},\mathbf{b})*}\Big(-\Big[dW^{(\mathbf{k})}M^{(\mathbf{k},\mathbf{b})}\Big]_{nn}-\Big[dW^{(\mathbf{k}+\mathbf{b})}M^{(\mathbf{k}+\mathbf{b},-\mathbf{b})}\Big]^*_{nn}\Big)\Bigg\}\\
=&\frac{1}{N_{\mathrm{kp}}}\sum_n\sum_{\mathbf{k},\mathbf{b}} w_b\\
&\Bigg\{\Big(\sum_l dW_{nl}^{(\mathbf{k})}M_{ln}^{(\mathbf{k},\mathbf{b})}M_{nn}^{(\mathbf{k},\mathbf{b})*}\Big)^*+\Big(\sum_l dW_{nl}^{(\mathbf{k}+\mathbf{b})}M_{ln}^{(\mathbf{k}+\mathbf{b},-\mathbf{b})}M_{nn}^{(\mathbf{k},\mathbf{b})}\Big)\\
&+\Big(\sum_l dW_{nl}^{(\mathbf{k})}M_{ln}^{(\mathbf{k},\mathbf{b})}M_{nn}^{(\mathbf{k},\mathbf{b})*}\Big)+\Big(\sum_l dW_{nl}^{(\mathbf{k}+\mathbf{b})}M_{ln}^{(\mathbf{k}+\mathbf{b},-\mathbf{b})}M_{nn}^{(\mathbf{k},\mathbf{b})}\Big)^*\Bigg\}.
\end{aligned}
\tag{A.37}
$$

By introducing the abbreviation

$$
R_{ln}^{(\mathbf{k},\mathbf{b})} := M_{ln}^{(\mathbf{k},\mathbf{b})}M_{nn}^{(\mathbf{k},\mathbf{b})*}
\tag{A.38}
$$

and using that $M_{mn}^{(\mathbf{k},\mathbf{b})} = M_{nm}^{(\mathbf{k}+\mathbf{b},-\mathbf{b})*}$ one finally arrives at

$$
\begin{aligned}
d\Omega_{\mathrm{I,OD}} =&\frac{1}{N_{\mathrm{kp}}}\sum_n\sum_{\mathbf{k},\mathbf{b}} w_b\,\mathrm{Tr}\Big[dW^{(\mathbf{k})}\big(R^{(\mathbf{k},\mathbf{b})}-R^{(\mathbf{k},\mathbf{b})\dagger}\big)\Big]\\
&+\frac{1}{N_{\mathrm{kp}}}\sum_n\sum_{\mathbf{k},\mathbf{b}} w_b\,\mathrm{Tr}\Big[dW^{(\mathbf{k}+\mathbf{b})}\big(R^{(\mathbf{k}+\mathbf{b},-\mathbf{b})}-R^{(\mathbf{k}+\mathbf{b},-\mathbf{b})\dagger}\big)\Big]\\
=&\frac{2}{N_{\mathrm{kp}}}\sum_n\sum_{\mathbf{k},\mathbf{b}} w_b\,\mathrm{Tr}\Big[dW^{(\mathbf{k})}\big(R^{(\mathbf{k},\mathbf{b})}-R^{(\mathbf{k},\mathbf{b})\dagger}\big)\Big]
\end{aligned}
\tag{A.39}
$$

where it was used that the sums of the expressions in brackets are equal when the summation runs over all **k** and **b** which is due to the periodicity of $dW^{(\mathbf{k})}$ and $R^{(\mathbf{k},\mathbf{b})}$ in reciprocal space with respect to a reciprocal lattice vector and the fact that the set of vectors $\{-\mathbf{b}\}$ and $\{\mathbf{b}\}$ are identical[2]. Introducing the operation

$$\mathcal{A}\left[B\right] := \frac{1}{2}\big\{B - B^{\dagger}\big\} \,, \tag{A.40}$$

the final expression for $d\Omega_{\mathrm{I,OD}}$ reads

$$d\Omega_{\mathrm{I,OD}} = \frac{4}{N_{\mathrm{kp}}}\sum_{\mathbf{k},\mathbf{b}} w_b \,\mathrm{Tr}\left[dW^{(\mathbf{k})}\mathcal{A}\left[R^{(\mathbf{k},\mathbf{b})}\right]\right] . \tag{A.41}$$

In the following the change $d\Omega_{\mathrm{D}}$ will be derived with Ω_{D} defined in accordance to Eq. (A.27) as

$$\Omega_{\mathrm{D}} = \frac{1}{N_{\mathrm{kp}}}\sum_{n}\sum_{\mathbf{k},\mathbf{b}} w_b \left(q_n^{(\mathbf{k},\mathbf{b})}\right)^2 ,$$

where here the abbreviation

$$q_n^{(\mathbf{k},\mathbf{b})} := \mathrm{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})} + \mathbf{b}\mathbf{r}_n \tag{A.42}$$

was used. In this notation the change $d\Omega_{\mathrm{D}}$ is given by

$$d\Omega_{\mathrm{D}} = \frac{2}{N_{\mathrm{kp}}}\sum_{n}\sum_{\mathbf{k},\mathbf{b}} w_b \, q_n^{(\mathbf{k},\mathbf{b})} dq_n^{(\mathbf{k},\mathbf{b})} \,, \tag{A.43}$$

which can be separated using Eq. (A.42) into

$$\begin{aligned} d\Omega_{\mathrm{D}} =& d\Omega_{\mathrm{D}}^{(\mathrm{I})} + d\Omega_{\mathrm{D}}^{(\mathrm{II})} \\ =& \frac{2}{N_{\mathrm{kp}}}\sum_{n}\sum_{\mathbf{k},\mathbf{b}} w_b \, q_n^{(\mathbf{k},\mathbf{b})} d\Big(\mathrm{Im}\ln M_{nn}^{(\mathbf{k},\mathbf{b})}\Big) \\ &+ \frac{2}{N_{\mathrm{kp}}}\sum_{n}\sum_{\mathbf{k},\mathbf{b}} w_b \, q_n^{(\mathbf{k},\mathbf{b})} d\Big(\mathbf{b}\mathbf{r}_n\Big) \,. \end{aligned} \tag{A.44}$$

---

[2] The periodicity in $\mathbf{k}$-space and $\{-\mathbf{b}\} = \{\mathbf{b}\}$ allow for the index transformations $\mathbf{k}+\mathbf{b} \mapsto \mathbf{k}'$ and $-\mathbf{b} \mapsto \mathbf{b}'$ without adjustment of the summation boundaries.

Though $d\big(\mathbf{b}\mathbf{r}_n\big)$ is in general non-zero, the second part $d\Omega_{\mathrm{D}}^{(\mathrm{II})}$ vanishes due to

$$\begin{aligned}
d\Omega_{\mathrm{D}}^{(\mathrm{II})} =& \frac{2}{N_{\mathrm{kp}}} \sum_n \sum_{\mathbf{k},\mathbf{b}} w_b \, q_n^{(\mathbf{k},\mathbf{b})} \, d\big(\mathbf{b}\mathbf{r}_n\big) \\
=& \sum_n \left(\frac{2}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \, q_n^{(\mathbf{k},\mathbf{b})} \mathbf{b} \right) d\mathbf{r}_n \\
=& \sum_n \left(\frac{2}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \, \big(\operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} + \mathbf{b}\mathbf{r}_n \big) \mathbf{b} \right) d\mathbf{r}_n \\
=& \sum_n \left(\frac{2}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} \mathbf{b} + \frac{2}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b (\mathbf{b}\mathbf{r}_n) \mathbf{b} \right) \mathbf{r}_n \\
=& \sum_n \big(-2\mathbf{r}_n + 2\,\mathbf{r}_n \big) \, d\mathbf{r}_n = 0 \ ,
\end{aligned} \tag{A.45}$$

where the last step was performed by means of Eqs. (4.33) and (B.4). Consequently, $d\Omega_{\mathrm{D}}$ is solely determined by $d\Omega_{\mathrm{D}}^{(\mathrm{I})}$ for which the change of $\operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})}$ needs to be calculated. Considering

$$\operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} = \frac{1}{2i} \big[\ln M_{nn}^{(\mathbf{k},\mathbf{b})} - \ln M_{nn}^{(\mathbf{k},\mathbf{b})^*} \big] \tag{A.46}$$

as a function of two independent complex variables $M_{nn}^{(\mathbf{k},\mathbf{b})}$ and $M_{nn}^{(\mathbf{k},\mathbf{b})^*}$, and using Wirtinger calculus one obtains

$$d\big(\operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} \big) = \frac{1}{2i} \left(\frac{dM_{nn}^{(\mathbf{k},\mathbf{b})}}{M_{nn}^{(\mathbf{k},\mathbf{b})}} - \frac{dM_{nn}^{(\mathbf{k},\mathbf{b})^*}}{M_{nn}^{(\mathbf{k},\mathbf{b})^*}} \right) . \tag{A.47}$$

Based on this expression the change $d\Omega_{\mathrm{D}} = d\Omega_{\mathrm{D}}^{(\mathrm{I})}$ of the diagonal part of the spread functional is given by

$$\begin{aligned}
d\Omega_{\mathrm{D}}^{(\mathrm{I})} =& \frac{2}{N_{\mathrm{kp}}} \sum_n \sum_{\mathbf{k},\mathbf{b}} w_b \, q_n^{(\mathbf{k},\mathbf{b})} \, d\big(\operatorname{Im} \ln M_{nn}^{(\mathbf{k},\mathbf{b})} \big) \\
=& \frac{-i}{N_{\mathrm{kp}}} \sum_n \sum_{\mathbf{k},\mathbf{b}} w_b \, q_n^{(\mathbf{k},\mathbf{b})} \left(\frac{dM_{nn}^{(\mathbf{k},\mathbf{b})}}{M_{nn}^{(\mathbf{k},\mathbf{b})}} - \frac{dM_{nn}^{(\mathbf{k},\mathbf{b})^*}}{M_{nn}^{(\mathbf{k},\mathbf{b})^*}} \right) \\
=& \frac{-i}{N_{\mathrm{kp}}} \sum_n \sum_{\mathbf{k},\mathbf{b}} w_b \, q_n^{(\mathbf{k},\mathbf{b})} \Bigg\{ - \frac{\Big[dW^{(\mathbf{k})} M^{(\mathbf{k},\mathbf{b})} \Big]_{nn}}{M_{nn}^{(\mathbf{k},\mathbf{b})}} - \frac{\Big[dW^{(\mathbf{k}+\mathbf{b})} M^{(\mathbf{k}+\mathbf{b},-\mathbf{b})} \Big]_{nn}^*}{M_{nn}^{(\mathbf{k}+\mathbf{b},-\mathbf{b})^*}} \\
& \qquad + \frac{\Big[dW^{(\mathbf{k})} M^{(\mathbf{k},\mathbf{b})} \Big]_{nn}^*}{M_{nn}^{(\mathbf{k},\mathbf{b})^*}} + \frac{\Big[dW^{(\mathbf{k}+\mathbf{b})} M^{(\mathbf{k}+\mathbf{b},-\mathbf{b})} \Big]_{nn}}{M_{nn}^{(\mathbf{k}+\mathbf{b},-\mathbf{b})}} \Bigg\} .
\end{aligned} \tag{A.48}$$

Using now $q_n^{(\mathbf{k},\mathbf{b})} = q_n^{(\mathbf{k},\mathbf{b})*}$, $q_n^{(\mathbf{k}+\mathbf{b},-\mathbf{b})} = -q_n^{(\mathbf{k},\mathbf{b})}$, and the definition

$$T_{ln}^{(\mathbf{k},\mathbf{b})} := \frac{M_{ln}^{(\mathbf{k},\mathbf{b})}}{M_{nn}^{(\mathbf{k},\mathbf{b})}} q_n^{(\mathbf{k},\mathbf{b})} \; , \tag{A.49}$$

one finally arrives at

$$\begin{aligned} d\Omega_{\mathrm{D}} = d\Omega_{\mathrm{D}}^{(\mathrm{I})} =& \frac{i}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \operatorname{Tr}\left[dW^{(\mathbf{k})} \big(T^{(\mathbf{k},\mathbf{b})} + T^{(\mathbf{k},\mathbf{b})\dagger} \big) \right] \\ &+ \frac{i}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \operatorname{Tr}\left[dW^{(\mathbf{k}+\mathbf{b})} \big(T^{(\mathbf{k}+\mathbf{b},-\mathbf{b})} + T^{(\mathbf{k}+\mathbf{b},-\mathbf{b})\dagger} \big) \right] \\ =& \frac{2i}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \operatorname{Tr}\left[dW^{(\mathbf{k})} \big(T^{(\mathbf{k},\mathbf{b})} + T^{(\mathbf{k},\mathbf{b})\dagger} \big) \right] \end{aligned} \tag{A.50}$$

where it was used again that the sums of the expressions in brackets are equal when the summation runs over all **k** and **b**. By means of the definition

$$\mathcal{S}\left[B\right] := \frac{1}{2i}\left\{ B + B^{\dagger} \right\} \; . \tag{A.51}$$

the final expression for $d\Omega_{\mathrm{D}}$ reads

$$d\Omega_{\mathrm{D}} = -\frac{4}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \operatorname{Tr}\left[dW^{(\mathbf{k})} \mathcal{S}\left[T^{(\mathbf{k},\mathbf{b})} \right] \right] \; . \tag{A.52}$$

Recombining all individual expressions yields finally the total first order change

$$\begin{aligned} d\Omega =& d\Omega_{\mathrm{I,OD}} + d\Omega_{\mathrm{D}} \\ =& \frac{4}{N_{\mathrm{kp}}} \sum_{\mathbf{k},\mathbf{b}} w_b \operatorname{Tr}\left[dW^{(\mathbf{k})} \big(\mathcal{A}\left[R^{(\mathbf{k},\mathbf{b})} \right] - \mathcal{S}\left[T^{(\mathbf{k},\mathbf{b})} \right] \big) \right] \end{aligned} \tag{A.53}$$

from which the gradient can easily be determined by comparison of coefficients with Eq. (4.46) yielding the expression given in Eq. (4.58).

B Lemmata

This section provides useful lemmata for calculations related to the discretization of the Brillouin zone into an uniformly spaced Monkhorst-Pack mesh [66].

Lemma B.1. *For functions $f : \mathbb{R}^d \to \mathbb{R}$ which are even with respect to point reflections through the origin it holds that*

$$\sum_{\mathbf{b}} f(\mathbf{b})\,\mathbf{b} = \mathbf{0}\ .$$

Proof. Assume

$$f(\mathbf{b}) = f(-\mathbf{b})\ , \tag{B.2}$$

then

$$\sum_{\mathbf{b}} f(\mathbf{b})\,\mathbf{b} = -\sum_{\mathbf{b}} f(\mathbf{b})\big(-\mathbf{b}\big) \overset{\text{(B.2)}}{=} -\sum_{\mathbf{b}} f(-\mathbf{b})\big(-\mathbf{b}\big) = -\sum_{\mathbf{b}'} f(\mathbf{b}')\big(\mathbf{b}'\big)$$
$$\Longleftrightarrow \sum_{\mathbf{b}} f(\mathbf{b})\,\mathbf{b} = \mathbf{0}\ .$$

□

Corollary B.3. *For the special case $f(\mathbf{b}) \equiv 1$ one obtaines*

$$\Big(\sum_{\mathbf{b}} \mathbf{b}\Big) \overset{B.1}{=} \mathbf{0}\ .$$

Lemma B.4.

$$\sum_{\mathbf{b}} w_b\,\mathbf{b}\big(\mathbf{b}\mathbf{R}\big) = \mathbf{R}\ .$$

Proof. Using component notation with component index i running over all dimensions considered one finds

$$\Big[\sum_{\mathbf{b}} w_b\,\mathbf{b}\big(\mathbf{b}\mathbf{R}\big)\Big]_i = \sum_{\mathbf{b}} w_b\,b_i \sum_{\alpha} b_\alpha R_\alpha = \sum_{\alpha} R_\alpha \sum_{\mathbf{b}} w_b\,b_i b_\alpha = \sum_{\alpha} R_\alpha \delta_{i\alpha} = R_i\ .$$

□

Lemma B.5.

$$\sum_{\mathbf{b}} w_b\,\big(\mathbf{b}\mathbf{R}\big)^2 = R^2\ .$$

Proof. Again, using component notation one abtains

$$\sum_{\mathbf{b}} w_b \left(\mathbf{bR}\right)^2 = \sum_{\mathbf{b}} w_b \sum_{\alpha,\beta} b_\alpha b_\beta R_\alpha R_\beta = \sum_{\alpha,\beta} R_\alpha R_\beta \sum_{\mathbf{b}} w_b b_\alpha b_\beta$$

$$= \sum_{\alpha,\beta} R_\alpha R_\beta \, \delta_{i\alpha} = \sum_{\alpha} R_\alpha R_\alpha = R^2 \; ,$$

where the component indices α, β run over all dimensions. □

Lemma B.6. *Surface integrals over matrix elements of* $\nabla_{\mathbf{k}}$ *between the periodic parts of Bloch functions over the boundary of the Brillouin zone* ∂BZ *vanish, i.e.*

$$I_n := \int_{\partial\mathrm{BZ}} d^{(d-1)}\mathbf{k} \, \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n\mathbf{k}} \rangle = 0 \; . \tag{B.7}$$

Proof. The integral I_n is separable into N_p pairs of surface integrations over planes with anti-parallel normal vectors $O_p^{\uparrow}$, $O_p^{\downarrow}$ where the number N_p of these pairs depends on the dimension d and the explicit shape of the Brillouin zone

$$I_n = \sum_{p=1}^{N_p} \left\{ \int_{O_p^{\uparrow}} d\mathbf{O}_p^{\uparrow} \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n\mathbf{k}} \rangle + \int_{O_p^{\downarrow}} d\mathbf{O}_p^{\downarrow} \langle u_{n\mathbf{k}'} | \nabla_{\mathbf{k}'} | u_{n\mathbf{k}'} \rangle \right\} \text{ with } \mathbf{k} \in O_p^{\uparrow} \text{ and } \mathbf{k}' \in O_p^{\downarrow} \; . \tag{B.8}$$

The surfaces $O_p^{\uparrow}$ and $O_p^{\downarrow}$ are connected by a point reflection at the Γ-point and separated by a reciprocal lattice vector $\mathbf{G}$. Hence, one can express the integration over the surface $O_p^{\downarrow}$ as an integration over $O_p^{\uparrow}$ shifted by a lattice vector $\mathbf{G}$ and with opposite surface normal vector $d\mathbf{O}_p^{\downarrow} = -d\mathbf{O}_p^{\uparrow}$ according to

$$I_n = \sum_{p=1}^{N_p} \int_{O_p^{\uparrow}} d\mathbf{O}_p^{\uparrow} \left\{ \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n\mathbf{k}} \rangle - \langle u_{n\mathbf{k}+\mathbf{G}} | \nabla_{\mathbf{k}} | u_{n\mathbf{k}+\mathbf{G}} \rangle \right\} \quad \text{with } \mathbf{k} \in O_p^{\uparrow} \; . \tag{B.9}$$

Now using

$$|u_{n\mathbf{k}+\mathbf{G}}\rangle = e^{i\mathbf{Gr}} \, |u_{n\mathbf{k}}\rangle \qquad \langle u_{n\mathbf{k}+\mathbf{G}}| = \langle u_{n\mathbf{k}}| \, e^{-i\mathbf{Gr}} \tag{B.10}$$

$$|\nabla_{\mathbf{k}} u_{n\mathbf{k}+\mathbf{G}}\rangle = e^{i\mathbf{Gr}} \, |\nabla_{\mathbf{k}} u_{n\mathbf{k}}\rangle \qquad \langle \nabla_{\mathbf{k}} u_{n\mathbf{k}+\mathbf{G}}| = \langle \nabla_{\mathbf{k}} u_{n\mathbf{k}}| \, e^{-i\mathbf{Gr}} \tag{B.11}$$

one ends up with

$$I_n = \sum_{p=1}^{N_p} \int_{O_p^{\uparrow}} d\mathbf{O}_p^{\uparrow} \left\{ \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} u_{n\mathbf{k}} \rangle - \langle u_{n\mathbf{k}} | e^{-i\mathbf{Gr}} \, e^{i\mathbf{Gr}} \, |\nabla_{\mathbf{k}} u_{n\mathbf{k}} \rangle \right\} \quad \text{with } \mathbf{k} \in O_p^{\uparrow}, \tag{B.12}$$

$$= \sum_{p=1}^{N_p} \int_{O_p^{\uparrow}} d\mathbf{O}_p^{\uparrow} \left\{ \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} u_{n\mathbf{k}} \rangle - \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} u_{n\mathbf{k}} \rangle \right\} = 0 \; . \tag{B.13}$$

□

Acronyms

BF Bloch function

BZ first Brillouin Zone

FDTD Finite Difference Time Domain

FDC Folded Directional Coupler

IREP irreducible representation

RREP reducible representation

RLV reciprocal lattice vector

LFT Lattice Fourier Transform

MLWF maximally localized Wannier function

MLGWF maximally localized generalized Wannier function

PBG Photonic Band Gap

PBS Photonic Band Structure

PC Photonic Crystal

SD Steepest Descent

WF Wannier function

WSC Wigner-Seitz Cell

Bibliography

[1] E. Yablonovitch, “Inhibited spontaneous emission in solid-state physics and electronics”, *Phys. Rev. Lett.* **58**, 2059 (1987).

[2] S. John, “Strong localization of photons in certain disordered dielectric superlattices”, *Phys. Rev. Lett.* **58**, 2486 (1987).

[3] C. M. Soukoulis, editor, *Photonic Crystals and Light Localization in the 21st Century: Proceedings of the NATO Advanced Study Institute on Photonic Crystals and Light Localization, Limin Hersonissou, Crete, Greece, June 18-30, 2000*, vol. 563 of *NATO Science Series. Series C: Mathematical and Physical Sciences*, Kluwer, Dordrecht (2001).

[4] M. Ibanescu, Y. Fink, S. Fan, E. L. Thomas, and J. D. Joannopoulos, “An All-Dielectric Coaxial Waveguide”, *Science* **289**, 415 (2000).

[5] S. Johnson, M. Ibanescu, M. Skorobogatiy, O. Weisberg, T. Engeness, M. Soljacic, S. Jacobs, J. Joannopoulos, , and Y. Fink, “Low-loss asymptotically single-mode propagation in lemarge-core OmniGuide fibers”, *Optics Express* **9**, 748 (2001).

[6] A. Birner, R. B. Wehrspohn, U. M. Gosele, and K. Busch, “Silicon-based photonic crystals”, *Adv. Mater.* **13**, 377 (2001).

[7] T. F. Krauss and R. M. De la Rue, “Photonic crystals in the optical regime – past, present and future”, *Prog. Quantum Electron.* **23**, 51 (1999).

[8] A. Forchel, M. Kamp, T. Happ, J. P. Reithmaier, M. Bayer, J. Koeth, and R. Dietrich, “Photon confinement effects - from physics to applications”, *Microelectron. Eng.* **53**, 21 (2000).

[9] M. Loncar, T. Doll, J. Vuckovic, and A. Scherer, “Design and fabrication of silicon photonic crystal optical waveguides”, *J. Lightwave Technol.* **18**, 1402 (2000).

[10] H. Benisty, S. Olivier, C. Weisbuch, M. Agio, M. Kafesaki, C. M. Soukoulis, M. Qiu, M. Swillo, A. Karlsson, B. Jaskorzynska, A. Talneau, J. Moosburger, M. Kamp, A. Forchel, R. Ferrini, R. Houdre, and U. Oesterle, “Models and measurements for the transmission of submicron-width waveguide bends defined in two-dimensional photonic crystals”, *IEEE J. Quantum Electron.* **38**, 770 (2002).

[11] S. Noda, M. Imada, A. Chutinan, and N. Yamamoto, "III-V based-semiconductor photonic crystals", *Opt. Quantum Electron.* **34**, 723 (2002).

[12] C. Liguda, G. Bottger, A. Kuligk, R. Blum, M. Eich, H. Roth, J. Kunert, W. Morgenroth, H. Elsner, and H. G. Meyer, "Polymer photonic crystal slab waveguides", *Appl. Phys. Lett.* **78**, 2434 (2001).

[13] A. C. Edrington, A. M. Urbas, P. DeRege, C. X. Chen, T. M. Swager, N. Hadjichristidis, M. Xenidou, L. J. Fetters, J. D. Joannopoulos, Y. Fink, and E. L. Thomas, "Polymer-based photonic crystals", *Adv. Mater.* **13**, 421 (2001).

[14] O. J. A. Schueller, G. M. Whitesides, J. A. Rogers, M. Meier, and A. Dodabalapur, "Fabrication of photonic crystal lasers by nanomolding of solgel glasses", *Appl. Optics* **38**, 5799 (1999).

[15] A. Rosenberg, R. J. Tonucci, H. B. Lin, and E. L. Shirley, "Photonic-band-structure effects for low-index-contrast twodimensional lattices in the near-infrared", *Phys. Rev. B* **54**, R5195 (1996).

[16] P. Lalanne and H. Benisty, "Out-of-plane losses of two-dimensional photonic crystals waveguides: Electromagnetic analysis", *J. Appl. Phys.* **89**, 1512 (2001).

[17] M. Palamaru and P. Lalanne, "Photonic crystal waveguides: Out-of-plane losses and adiabatic modal conversion", *Appl. Phys. Lett.* **78**, 1466 (2001).

[18] H. Benisty, P. Lalanne, S. Olivier, M. Rattier, C. Weisbuch, C. J. M. Smith, T. F. Krauss, C. Jouanin, and D. Cassagne, "Finite-depth and intrinsic losses in vertically etched twodimensional photonic crystals", *Opt. Quantum Electron.* **34**, 205 (2002).

[19] S. Y. Lin, J. G. Fleming, D. L. Hetherington, B. K. Smith, R. Biswas, K. M. Ho, M. M. Sigalas, W. Zubrzycki, S. R. Kurtz, and J. Bur, "A three-dimensional photonic crystal operating at infrared wavelengths", *Nature* **394**, 251 (1998).

[20] S. Noda, K. Tomoda, N. Yamamoto, and A. Chutinan, "Full three-dimensional photonic bandgap crystals at nearinfrared wavelengths", *Science* **289**, 604 (2000).

[21] J. E. G. J. Wijnhoven and W. L. Vos, "Preparation of photonic crystals made of air spheres in titania", *Science* **281**, 802 (1998).

[22] A. Blanco, E. Chomski, S. Grabtchak, M. Ibisate, S. John, S. W. Leonard, C. Lopez, F. Meseguer, H. Miguez, J. P. Mondia, G. A. Ozin, O. Toader, and H. M. van Driel, "Large-scale synthesis of a silicon photonic crystal with a complete three-dimensional bandgap near 1.5 micrometres", *Nature* **405**, 437 (2000).

[23] Y. A. Vlasov, X.-Z. Bo, J. C. Sturm, and D. J. Norris, “On-chip natural assembly of silicon photonic bandgap crystals”, *Nature* **414**, 289.

[24] M. Campbell, D. N. Sharp, M. T. Harrison, R. G. Denning, and A. J. Turberfield, “Fabrication of photonic crystals for the visible spectrum by holographic lithography”, *Nature* **404**, 53 (2000).

[25] Y. V. Miklyaev, D. C. Meisel, A. Blanco, G. von Freymann, K. Busch, W. Koch, C. Enkrich, M. Deubel, and M. Wegener, “Three-dimensional face-centered-cubic photonic crystal templates by laser holography: fabrication, optical characterization, and band-structure calculations”, *Appl. Phys. Lett.* **82**, 1284 (2003).

[26] H. B. Sun, S. Matsuo, and H. Misawa, “Three-dimensional photonic crystal structures achieved with two-photon-absorption photopolymerization of resin”, *Appl. Phys. Lett.* **74**, 786 (1999).

[27] H. B. Sun, V. Mizeikis, Y. Xu, S. Juodkazis, J. Y. Ye, S. Matsuo, and H. Misawa, “Microcavities in polymeric photonic crystals”, *Appl. Phys. Lett.* **79**, 1 (2001).

[28] M. Straub and M. Gu, “Near-infrared photonic crystals with higher-order bandgaps generated by two-photon photopolymerization”, *Opt. Lett.* **27**, 1824 (2002).

[29] M. Deubel, G. v.Freymann, M. Wegener, S. Pereira, K. Busch, and C. M. Soukoulis, “Direct laser writing of three-dimensional photonic-crystal templates for telecommunications”, *Nature Materials* **3**, 444 (2004).

[30] S. Wong, M. Deubel, F. Prez-Willard, S. John, G. A. Ozin, M. Wegener, and G. von Freymann, “Direct Laser Writing of Three- Dimensional Photonic Crystals with a Complete Photonic Bandgap in Chalcogenide Glasses”, *Advanced Materials* **18**, 265 (2006).

[31] N. Tetreault, G. von Freymann, M. Deubel, M. Hermatschweiler, F. Perez-Willard, S. John, M. Wegener, and G. A. Ozin, “New Route to Three-Dimensional Photonic Bandgap Materials: Silicon Double Inversion of Polymer Templates”, *Advanced Materials* **18**, 457 (2006).

[32] A. Taflove, “Review of the formulation and applications of the finite-difference time-domain method for numerical modeling of electromagnetic-wave interactions with arbitrary structures”, *Wave Motion* **10**, 547 (1988).

[33] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House, Boston, MA (1995).

[34] C. A. J. Fletcher, *Computational Galerkin methods*, vol. XI of *Springer series in computational physics*, Springer, New York (1984).

[35] D. Mogilevtsev, T. A. Birks, and P. S. Russell, “Localized function method for modeling defect modes in 2-D photonic crystals”, *J. Lightwave Technol.* **17**, 2078 (1999).

[36] P. Russell, “Photonic Crystal Fibers”, *Science* **299**, 358 (2003).

[37] T. M. Monro, D. J. Richardson, N. G. R. Broderick, and P. J. Bennett, “Holey optical fibers: An efficient modal model”, *J. Lightwave Technol.* **17**, 1093 (1999).

[38] A. A. Asatryan, S. Fabre, K. Busch, R. C. McPhedran, L. C. Botten, C. M. de Sterke, and N. A. P. Nicorovici, “Two-dimensional local density of states in two-dimensional photonic crystals”, *Opt. Express* **8**, 191 (2001).

[39] E. Moreno, D. Erni, and C. Hafner, “Modeling of discontinuities in photonic crystal waveguides with the multiple multipole method”, *Phys. Rev. E* **66**, 036618 (2002).

[40] K. Ohtaka and Y. Tanabe, “Photonic band using vector spherical waves .1. Various properties of Bloch electric fields and heavy photons”, *J. Phys. Soc. Jpn.* **65**, 2265 (1996).

[41] K. Busch, “Photonic band structure theory: assessment and perspectives”, *Compte Rendus Physique* **3**, 53 (2002).

[42] K. Busch and S. John, “Photonic band gap formation in certain self-organizing systems”, *Phys. Rev. E* **58**, 3896 (1998).

[43] G. H. Wannier, “The structure of electronic excitation levels in insulating crystals”, *Phys. Rev.* **52**, 191 (1937).

[44] D. S. Bulyanitsa and Y. E. Svetlov, “Properties of the Bloch and Wannier Functions”, *Soviet Phys. – Solid State* **4**, 981 (1962).

[45] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, “Exponential localization of Wannier functions in insulators”, (2006), URL `arXiv:cond-mat/0606726`.

[46] N. Marzari and D. Vanderbilt, “Maximally localized generalized Wannier functions for composite energy bands”, *Phys. Rev. B* **56**, 12847 (1997).

[47] I. Souza, N. Marzari, and D. Vanderbilt, “Maximally localized Wannier functions for entangled energy bands”, *Phys. Rev. B* **65**, 035109 (2002).

[48] K. Busch, S. F. Mingaleev, M. Schillinger, and D. Hermann, “The Wannier function approach to photonic crystal circuits”, *Journal of Physics: Condensed Matter* **15**, R1233 (2003).

[49] D. R. Smith, J. B. Pendry, and M. C. K. Wiltshire, "Metamaterials and Negative Refractive Index", *Science* **305**, 788 (2006).

[50] V. G. Veselago, "The electrodynamics of substances with simultaneously negative values of ϵ and μ", *Sov. Phys. USPEKHI* **10**, 509 (1968).

[51] D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, "Composite Medium with Simultaneously Negative Permeability and Permittivity", *Phys. Rev. Lett.* **84**, 4184 (2000).

[52] J. B. Pendry, D. Schurig, and D. R. Smith, "Controlling Electromagnetic Fields", *Science* **312**, 1780 (2006).

[53] J. B. Pendry, "Negative refraction makes a perfect lens", *Phys. Rev. Lett.* **85**, 3966 (2000).

[54] G. Dolling, C. Enkrich, M. Wegener, C. M. Soukoulis, and S. Linden, "Simultaneous Negative Phase and Group Velocity of Light in a Metamaterial", *Science* **312**, 892 (2006).

[55] M. W. Klein, C. Enkrich, M. Wegener, and S. Linden, "Second-Harmonic Generation from Magnetic Metamaterials", *Science* **313**, 502 (2006).

[56] K. Busch and S. John, "Liquid-crystal photonic-band-gap materials: The tunable electromagnetic vacuum", *Phys. Rev. Lett.* **83**, 967 (1999).

[57] P. Mack, *Tunable anisotropic defect structures in 2D Photonic Crystals using the Wannier Function approach*, Master's thesis, Institut für Theoretische Festkörperphysik, Universität Karlsruhe (2006).

[58] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Thomson Learning (1976).

[59] L. Tkeshelashvili, J. Niegemann, S. Pereira, and K. Busch, "Nonlinear wave interaction in photonic band gap materials", *Photonics and Nanostructures - Fundamentals and Applications* **4**, 75 (2006).

[60] L. Tkeshelashvili and K. Busch, "Nonlinear three-wave interaction in photonic crystals", *Appl. Phys. B* **81**, 225 (2005).

[61] J. Niegemann, *Interaction of Solitons and Defects in one-dimensional Photonic Crystals*, Master's thesis, Universität Karlsruhe (2004).

[62] R. D. Meade, A. M. Rappe, K. D. Brommer, J. D. Joannopoulos, and O. L. Alerhand, "Accurate theoretical-analysis of photonic band-gap materials", *Phys. Rev. B* **48**, 8434 (1993).

[63] S. G. Johnson and J. D. Joannopoulos, “Block-iterative frequency-domain methods for Maxwell’s equations in a planewave basis”, *Optics Express* **8**, 173 (2001).

[64] D. Hermann, M. Frank, K. Busch, and P. Wölfle, “Photonic band structure computations”, *Opt. Express* **8**, 167 (2001).

[65] J. E. Sipe, “Vector $k \cdot p$ approach for photonic band structures”, *Phys. Rev. E* **62**, 5672 (2000).

[66] H. J. Monkhorst and J. D. Pack, “Special points for Brillouin-zone integrations”, *Phys. Rev. B* **13**, 5188 (1976).

[67] L. Mihaly and M. C. Martin, *Solid State Physics: Problems and Solutions*, Wiley & Sons, New York, NY, USA (1996).

[68] M. H. Schillinger and K. Busch, “Bottom-up approach to the construction of symmetrical maximally localized Wannier functions for bosonic systems”, *in preparation* (2006).

[69] T. Inui, Y. Tanabe, and Y. Onodera, *Group Theory and Its Applications in Physics*, no. 78 in Solid-State Sciences, Springer-Verlag, Berlin (1990).

[70] J. F. Cornwell, *Group Theory in Physics*, vol. 2, Harcourt Brace Jovanovich, London (1984).

[71] M. Hamermesh, *Group Theory and its Application to Physical Problems*, Dover Publications, New York, reprint edn. (1990).

[72] K. Sakoda, *Optical Properties of Photonic Crystals*, Springer, Berlin (2001).

[73] A. Mekis, J. C. Chen, I. Kurland, S. H. Fan, P. R. Villeneuve, and J. D. Joannopoulos, “High transmission through sharp bends in photonic crystal waveguides”, *Phys. Rev. Lett.* **77**, 3787 (1996).

[74] Y. Jiao, S. Fan, and D. A. B. Miller, “Demonstration of systematic photonic crystal device design and optimization by low-rank adjustments: an extremely compact mode separator”, *Opt. Lett.* **30**, 141 (2005).

[75] S. F. Mingaleev, M. Schillinger, D. Hermann, and K. Busch, “Tunable photonic crystal circuits: concepts and designs based on single-pore infiltration”, *Opt. Lett.* **29**, 2858 (2004).

[76] J. D. Joannopoulos, P. R. Villeneuve, and S. H. Fan, “Photonic crystals: Putting a new twist on light”, *Nature* **386**, 143 (1997).

[77] S. F. Mingaleev and K. Busch, “Scattering matrix approach to large-scale photonic-crystal circuits”, *Opt. Lett.* **28**, 619 (2003).

[78] H. Brand, *Schaltungslehre linearer Mikrowellennetze*, Hirzel, Stuttgart (1970).

[79] R. C. Jones, "New calculus for the treatment of optical systems", *J. Opt. Soc. Am.* **31**, 488 (1941).

[80] E. Hecht, *Optics*, Addison Wesley, Reading, MA, USA (1987).

[81] S. G. Johnson, C. Manolatou, S. H. Fan, P. R. Villeneuve, J. D. Joannopoulos, and H. A. Haus, "Elimination of cross talk in waveguide intersections", *Opt. Lett.* **23**, 1855 (1998).

[82] M. Schillinger, S. F. Mingaleev, D. Hermann, and K. Busch, "Highly localized Wannier functions for the efficient modeling of photonic crystal circuits", in A. Adibi, S.-Y. Lin, and A. Scherer, editors, "Photonic Crystal Materials and Devices III", vol. 5733 of *Proceedings of SPIE*, pp. 324–335, SPIE, Bellingham, WA (2005).

[83] T. P. White, L. C. Botten, R. C. McPhedran, and C. M. de Sterke, "Ultracompact resonant filters in photonic crystals", *Opt. Lett.* **28**, 2452 (2003).

[84] Y. Jiao, S. F. Mingaleev, M. Schillinger, D. A. B. Miller, S. Fan, and K. Busch, "Wannier Basis Design and Optimization of a Photonic Crystal Waveguide Crossing", *IEEE Photon. Technol. Lett.* **17**, 1875 (2005).

[85] Y. Jiao, S. Fan, and D. A. B. Miller, "Photonic crystal device sensitivity analysis with Wannier basis gradients", *Opt. Lett.* **30**, 302 (2005).

[86] Y. Jiao, S. Fan, and D. A. B. Miller, "Demonstration of systematic photonic crystal device design and optimization by low-rank adjustments: an extremely compact mode separator", *Opt. Lett.* **30**, 141 (2005).

[87] W. Kohn, "Analytic Properties of Bloch Waves and Wannier Functions", *Phys. Rev.* **115**, 809 (1959).

[88] W. Kohn and J. Onffroy, *Phys. Rev. B* **8**, 2485 (1973).

[89] J. Des Cloizeaux, "Orthogonal Orbitals and Generalized Wannier Functions", *Physical Review* **129**, 554 (1963).

[90] K. M. Leung, "Defect modes in photonic band structures - a green-function approach using vector wannier functions", *J. Opt. Soc. Am. B* **10**, 303 (1993).

[91] J. P. Albert, C. Jouanin, D. Cassagne, and D. Bertho, "Generalized Wannier function method for photonic crystals", *Phys. Rev. B* **61**, 4381 (2000).

[92] J. P. Albert, C. Jouanin, D. Cassagne, and D. Monge, "Photonic crystal modelling using a tight-binding Wannier function method", *Opt. Quantum Electron.* **34**, 251 (2002).

[93] M. L. D'Yerville, D. Monge, D. Cassagne, and J. P. Albert, "Tight-binding method modelling of photonic crystal waveguides", *Opt. Quantum Electron.* **34**, 445 (2002).

[94] A. Garcia-Martin, D. Hermann, K. Busch, and P. Wölfle, "Solid state theoretical methods for defect computations in photonic crystals", vol. 722 of *Mater. Res. Soc. Symp. Proc.*, pp. L1–7 (2002).

[95] A. Garcia-Martin, D. Hermann, K. Busch, and P. Wölfle, "Defect computation in photonic crystals: a solid state theoretical approach", *Nanotechnology* **14**, 177 (2003).

[96] D. M. Whittaker and M. P. Croucher, "Maximally localized Wannier functions for photonic lattices", *Phys. Rev. B* **67**, 085204 (2003).

[97] A. Klöckner, *On the Computation of Maximally Localized Wannier Functions*, Master's thesis, Institut für angewandte Mathematik I, Universität Karlsruhe (2004), `http://news.tiker.net/dl/academic/da/diplomarbeit.pdf`.

[98] D. Hermann, *Bandstrukturrechnung für Photonische Kristalle mittels Mehrgitter-Verfahren*, Master's thesis, Institut für Theorie der Kondensierten Materie, Universität Karlsruhe (2001).

[99] E. I. Blount, "Formalisms of Band Theory", *Solid State Phys.* **13**, 305 (1976).

[100] A. Ruszczynski, *Nonlinear Optimization*, Princeton University Press, Princeton, New Jersey, USA (2006).

[101] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing", *Science* **220**, 671 (1983).

[102] I. Nassi and B. Shneiderman, "Flowchart techniques for structured programming", *SIGPLAN Notices* **XII** (1973).

[103] E. Riedel, *Anorganische Chemie*, de Gruyter, New York (2002).

[104] E. Krüger, "Symmetrische verallgemeinerte Wannierfunktionen", *phys. stat. sol. (b)* **52**, 215 (1972).

[105] E. Krüger, "Symmetrische verallgemeinerte Wannierfunktionen", *phys. stat. sol. (b)* **52**, 519 (1972).

[106] I. Bloch, "Ultracold quantum gases in optical lattices", *Nature Physics* **1**, 23 (2005).

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Karlsruhe, 15. November 2006

Matthias Schillinger

Danksagung

Abschließend möchte ich mich bei all jenen bedanken, die zum erfolgreichen Entstehen dieser Arbeit beigetragen haben.

Herrn Prof. Busch bin ich dankbar für die Möglichkeit, mich mit dem faszinierenden Thema der Wannierfunktionen auseinandersetzen zu dürfen. Ebenso möchte ich mich für seine Unterstützung im Allgemeinen und speziell bei der Realisierung meines erfahrungsreichen Forschungsaufenthaltes in Orlando bedanken.

Herrn Prof. Wölfle danke ich für die Übernahme des Korreferats, seine Unterstützung bei der Realisierung des Auslandsaufenthalts und das freundliche Arbeitsklima.

Bei meinen Freunden und Kollegen der ersten Generation Marcus Diem, Meikel Frank, Dr. Frank Hagmann und Daniel Hermann möchte ich mich besonders für die langjährige gute und angenehme Zusammenarbeit, viele interessante Diskussionen, nützliche Tipps und die schöne gemeinsame Zeit bedanken.

Für das freundliche Arbeitsklima und hilfreiche Diskussionen möchte ich mich bei Sabine Essig, Julia Budde, Maria Dienerowitz, Jens Niegemann, Patrick Mack, Martin Pototschnig, Mauno Schelb, Johannes Hagmann und Dr. Lasha Tkeshelashvili bedanken.

Für das Korrekturlesen von Teilen dieser Arbeit danke ich ganz besonders Daniel Hermann, Martin Pototschnig, Marcus Diem, sowie Patrick Mack, dem ich zudem für seine Begeisterung und sein Engagement für das Wannierfunktionsprojekt sowie für die freundschaftliche Zusammenarbeit danken möchte.

I would like to thank Dr. Antonio Garcia-Martin for his support and the friendly working atmosphere in the early stage of my thesis.

I'm grateful to Dr. Sergei Mingaleev for interesting discussions and bringing me into touch with parallel object-oriented methods and applications.

I'm thankful for the fruitful cooperation with Prof. David A. B. Miller, Prof. Shanhui Fan, and Dr. Yang Jiao, Stanford University.

Mein ganz besonderer Dank gilt meiner Mutter, die mich immer unterstützt und gefördert hat.