Unified Orbital Description of the Envelope Dynamics in Two-Dimensional Simple Periodic Lattices

By M. J. Ablowitz and Y. Zhu

The propagation of wave envelopes in two-dimensional (2-D) simple periodic lattices is studied. A discrete approximation, known as the tight-binding (TB) approximation, is employed to find the equations governing a class of nonlinear discrete envelopes in simple 2-D periodic lattices. Instead of using Wannier function analysis, the orbital approximation of Bloch modes that has been widely used in the physical literature, is employed. With this approximation the Bloch envelope dynamics associated with both simple and degenerate bands are readily studied. The governing equations are found to be discrete nonlinear Schrödinger (NLS)-type equations or coupled NLS-type systems. The coefficients of the linear part of the equations are related to the linear dispersion relation. When the envelopes vary slowly, the continuous limit of the general discrete NLS equations are effective NLS equations in moving frames. These continuous NLS equations (from discrete to continuous) also agree with those derived via a direct multiscale expansion. Rectangular and triangular lattices are examples.

1. Introduction

In recent years, there has been enhanced interest in the study of wave propagation in nonlinear periodic lattices. The interplay between nonlinearity and periodicity has led researchers to discover new and interesting phenomena such as novel localized modes, i.e., solitons. A field where both nonlinearity and periodicity

Address for correspondence: Yi Zhu, Zhou Pei-Yuan Center for Applied Mathematics, Tsinghua University, Beijing 100084, China; e-mail: yizhu@tsinghua.edu.cn

DOI: 10.1111/j.1467-9590.2012.00575.x

STUDIES IN APPLIED MATHEMATICS 131:41–71

arise naturally is nonlinear optics. Ever since the theoretical prediction of discrete optical solitons [1] and its subsequent experimental realizations [2–4] in both one-dimensional (1-D) and two-dimensional (2-D) periodic lattices were reported, many localized structures have been predicted theoretically and demonstrated experimentally. Examples include dipole solitons [5], vortex solitons [6], soliton trains [7], etc. Another area of application is condensed matter physics where ultracold atoms, i.e., Bose–Einstein condensates (BEC), can be trapped in a periodic optical lattice. The experimental observation of gap solitons was reported [8] and theories has been developed [9] in BEC. With observations and theory in different fields, the study of the underlying phenomena and their properties has attracted significant scientific interest.

Periodic lattices are common in physical systems. The geometric distribution of local minima of the potentials, also called sites, can be used to classify the potentials. These sites are the positions of the potential wells. In optics, they have increased refractive index and the electromagnetic field is attracted to these sites. The distribution of the sites determines the properties of the associated Bloch waves. Generally speaking 1-D lattices are similar from a geometric point of view. Discrete evolution equations on 1-D lattices were studied by the so-called Wannier function approach in Refs. [10, 11]. However, there are significant differences between different 2-D periodic lattices which, in turn, can lead to significant dynamical effects. Roughly speaking, 2-D periodic lattices can be divided into two groups: simple and nonsimple lattices. Simple lattices only have one site in a unit cell while nonsimple lattices may have more than one site per cell. Examples of typical simple lattices are rectangular and triangular lattices. A typical nonsimple lattice is the honeycomb lattice, which has two sites in a unit cell and breaks up into two triangular sublattices. It leads to different results than those found in simple lattices. Due to the underlying symmetries in the honeycomb lattice, it is found that the dispersion relation of the associated Bloch theory may have isolated degenerate points where two dispersion surfaces touch each other. These points are called Dirac points and near these points the dispersion surface has conical structure. The evolution of the Bloch mode envelopes in the neighborhood of these points is governed by nonlinear Dirac systems [12, 13]. There are interesting phenomena associated with the Dirac system. An example in optics is conical diffraction—where a narrow beam transforms into bright expanding rings [14–16]. Honeycomb lattices also admit various types of band gap solitons, which like other 2-D periodic lattices is due to the effect of nonlinearity (cf. [17]). Another important application which exhibits a honeycomb lattice stucture is the material graphene [18]. In BECs, background honeycomb lattices can also lead to interesting phenomena [19].

Generally, there are two main approaches to study the underlying envelope equations. One is a discrete approach based upon the tight-binding (TB) approximation, which is asymptotically valid when the potential intensity is

very strong (cf. [11, 20, 21]). The other approach is a purely continuous derivation, usually obtained by multiscale expansion methods, which assumes that the envelope varies slowly compared to the lattice scale (cf. [22–24]). However, the two approaches have not been closely connected. We find a number of new connections between the discrete and continuous approaches, which have not been known previously. Our results include the following:

1. We employ a discrete approach in the full Brillouin zone. We approximate the Bloch functions in terms of orbital functions to find a general discrete NLS equation—see Section 3. Use of this orbital ansatz is not the same as the "Wannier function decomposition" previously used in the literature. Usually, Wannier function analysis is used in simple band dynamics and in 1-D problems. Here we study 2-D problems and find the envelope dynamics associated with simple and degenerate bands. With this orbital ansatz, we find, as an example, the "hyperbolic" discrete NLS equation (one direction is focusing while the other is defocusing). With the TB approach the dispersion relation surfaces can be approximated analytically as well as the orbitals. Hence all coefficients of associated discrete equation can be calculated. Examples of square and triangular lattices are included; the above analysis shows how the dispersion surface and all coefficients of the nonlinear wave equation can be obtained.
2. The general connection between the discrete and continuous approaches has not been previously discussed. In Section 4 we develop a unified description of how the general discrete and continuous evolution equations relate to each other. By introducing the limit where the envelope scale is much larger than the lattice scale we find the continuous **k**-dependent scalar and coupled NLS limit from the discrete equations.
   To further clarify the discrete to continuous limit we show the results are the same as those found directly by the continuous approach which is derived in the Appendix. In this sense we show the two approaches are naturally connected; they are not two distinct limits.

2. Preliminaries

Here we consider the 2-D lattice nonlinear Schrödinger (NLS) equation, written in dimensionless form:

$$i\psi_z + \nabla^2\psi - V(\mathbf{r})\psi + \sigma|\psi|^2\psi = 0, \tag{1}$$

where $\mathbf{r} = (x, y)$, $V(\mathbf{r})$ is the periodic potential and σ is a constant which is positive for focusing nonlinearity and negative for defocusing nonlinearity. This model arises in light propagation in a periodic Kerr nonlinear medium and in BEC trapped in a 2-D optical lattice.

In this paper, the potential $V(\mathbf{r})$ is a 2-D periodic bounded real-valued function whose two primitive lattice vectors are $\mathbf{v}_1$ and $\mathbf{v}_2$; we denote the set of lattice vectors by: $\mathbb{P} = \{m\mathbf{v}_1 + n\mathbf{v}_2 : m, n \in \mathbb{Z}\}$. We will also need the primitive reciprocal-lattice vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ and its corresponding set: $\mathbb{G} = \{m\mathbf{k}_1 + n\mathbf{k}_2 : m, n \in \mathbb{Z}\}$. The unit cell of the physical lattice is the parallelogram with $\mathbf{v}_1$ and $\mathbf{v}_2$ as its two sides and the unit cell of the reciprocal lattice, usually called the Brillouin zone, is the parallelogram determined by $\mathbf{k}_1$ and $\mathbf{k}_2$. The relation between the lattice and its reciprocal lattice is $\mathbf{v}_m \cdot \mathbf{k}_n = 2\pi\delta_{mn}$.

For the problems we are interested in, the local minima of the potential, which are called sites, are very important. Physically, local minima are the positions of the potential wells. In optics they have increased refractive index and the electric field is attracted to these sites. It is natural to use the distribution of the sites to classify the lattices. In this paper, we consider simple periodic lattices, which only have one local minimum in a unit cell. With a starting point and the lattice vectors, all the positions of the sites can be constructed. These sites form a discrete lattice in the $\mathbf{r}$ plane. We use $S_{\mathbf{v}}$ to denote the position of the site with index $\mathbf{v}$ and note that $S_{\mathbf{v}} = S_{\mathbf{0}} + \mathbf{v}$, where $S_{\mathbf{0}}$ is the starting point of the site lattice. Associated with the periodic functions, due to translational symmetry, one unit cell contains all needed information. For simplicity, we place $S_{\mathbf{0}} = \mathbf{0}$ and call the parallelogram determined by $\mathbf{v}_1$ and $\mathbf{v}_2$ whose center is $S_{\mathbf{0}}$ as the primitive unit cell Ω. We also choose the parallelogram determined by $\mathbf{k}_1$ and $\mathbf{k}_2$ whose center is $\mathbf{k} = \mathbf{0}$ as the primitive reciprocal unit cell, the Brillouin zone Ω'. It is noted that a nonsimple lattice has more than one site in any unit cell. In this case we need more than one starting point to construct the lattice. A detailed study of nonsimple lattices is outside the scope of this paper.

2.1. Bloch theory for 2-D lattices

When the nonlinear coefficient σ is very small, or equivalently, the envelope function $\psi(\mathbf{r})$ is infinitesimal, Equation (1) reduces to a linear Schrödinger equation with a periodic potential

$$i\psi_z + \nabla^2\psi - V(\mathbf{r})\psi = 0. \tag{2}$$

According to the Bloch theory, this linear Schrödinger equation can be solved by considering its dynamics of eigenstates. Namely, $\psi(\mathbf{r}, z) = \varphi(\mathbf{r})e^{-i\mu z}$, where $\varphi(\mathbf{r})$ satisfies the eigenvalue problem

$$\mu\varphi + \nabla^2\varphi - V(\mathbf{r})\varphi = 0. \tag{3}$$

It is well known that the eigenfunction $\varphi(\mathbf{r})$, called the Bloch mode, has the $\mathbf{k}$-dependent form

$$\varphi(\mathbf{r};\mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{r}}U(\mathbf{r};\mathbf{k}), \tag{4}$$

where $\mathbf{k} = (k_x, k_y)$ is the wave vector and $u(\mathbf{r};\mathbf{k})$ has the same periodicity as the potential $V(\mathbf{r})$ for any $\mathbf{k}$. For simplicity, we introduce the following two operators:

$$\mathcal{H} = -\nabla^2 + V(\mathbf{r}), \qquad \mathcal{H}_\mathbf{k} = -\nabla^2 - 2i\mathbf{k}\cdot\nabla + |\mathbf{k}|^2 + V(\mathbf{r}),$$

where $\mathcal{H}$ is the Schrödinger operator with a periodic potential and $\mathcal{H}_\mathbf{k}$ is a $\mathbf{k}$-dependent operator whose eigenfunctions are square integrable periodic functions in Ω; hence $U(\mathbf{r};\mathbf{k})$ satisfies the following eigenvalue problem,

$$\mathcal{H}_\mathbf{k}U(\mathbf{r};\mathbf{k}) = \mu U(\mathbf{r};\mathbf{k}); \quad U(\mathbf{r}+\mathbf{v}_j) = U(\mathbf{r};\mathbf{k}); \quad j = 1, 2,$$

where $\mu = \mu(\mathbf{k})$ is called the dispersion relation. The Bloch mode $\varphi(\mathbf{r};\mathbf{k})$ satisfies the eigenproblem

$$\mathcal{H}\varphi(\mathbf{r};\mathbf{k}) = \mu\varphi(\mathbf{r};\mathbf{k}); \quad \varphi(\mathbf{r}+\mathbf{v}_j) = e^{i\mathbf{k}\cdot\mathbf{v}_j}\varphi(\mathbf{r};\mathbf{k}); \quad j = 1, 2. \tag{5}$$

One can check that $\varphi(\mathbf{r};\mathbf{k})$ and $\varphi(\mathbf{r};\mathbf{k}+\mathbf{g})$ satisfy the same eigenvalue problem with boundary condition (5) due to $\mathbf{v}\cdot\mathbf{g} = 2m\pi$, where $\mathbf{v} \in \mathbb{P}$ and $\mathbf{g} \in \mathbb{G}$. Thus for simplicity we can treat them as the same function. Thus for any $\mathbf{r}$, $\varphi(\mathbf{r};\mathbf{k})$ is periodic with respect to $\mathbf{k}$.

For each $\mathbf{k}$, the operator $\mathcal{H}_\mathbf{k}$ has infinitely many discrete eigenvalues $\mu^{(j)}(\mathbf{k})$, $j = 0, 1, 2, \ldots$, tending to positive infinity. The dispersion relation μ and the associated Bloch modes may have a superscript j to indicate correspondence to different eigenvalues. In this paper, for simplicity, we often omit the superscript j. As $\mathbf{k}$ varies, the discrete eigenvalue $\mu(\mathbf{k})$ and the corresponding eigenfunctions $U(\mathbf{r};\mathbf{k})$ as functions of $\mathbf{k}$ are assumed to be regular over $\mathbf{k}$. The spectrum of the Schrödinger operator $\mathcal{H}$ generally has multiple band structures and there may exist band gaps between two dispersion surfaces where bounded Bloch modes are not allowed.

Due to the periodicity of $\varphi(\mathbf{r};\mathbf{k})$ over $\mathbf{k}$, one can represent $\varphi(\mathbf{r};\mathbf{k})$ as a Fourier series

$$\varphi(\mathbf{r};\mathbf{k}) = \sum_\mathbf{v} \hat{\varphi}_\mathbf{v}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{v}}, \tag{6}$$

where $\hat{\varphi}_\mathbf{v}(\mathbf{r})$ defined as

$$\hat{\varphi}_\mathbf{v}(\mathbf{r}) = \frac{1}{|\Omega'|}\int_{\Omega'} \varphi(\mathbf{r};\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{v}}d\mathbf{k} \tag{7}$$

is the so-called Wannier function [25]; here and afterwards, the sum over $\mathbf{v}$ means $\mathbf{v}$ takes all values in $\mathbb{P}$, i.e., $\mathbf{v} = m\mathbf{v}_1 + n\mathbf{v}_2$, for all $m, n \in \mathbb{Z}$. Wannier

functions contain all the information in the Bloch modes. If one has complete information regarding the Wannier functions the exact Bloch mode can be constructed through Equation (6) or vice versa. Unfortunately, in general it is not possible to analytically compute either Bloch modes or Wannier functions. However, under some limits such as TB limit, i.e., when the potential is large, they can be constructed allowing many useful results to be obtained.

2.2. Dispersion relations in the TB limit

A goal of this paper is to understand and develop a unified approach to study envelope dynamics in weakly nonlinear periodic media. To deal with these problems, we need a good understanding of the associated linear problem. The linear problem is governed by a linear Schrödinger equation with a periodic potential; the dispersion relation of this problem plays a key role.

The TB approximation has been widely used in solid state physics to calculate electronic band structure [26]. The TB limit corresponds to $V_0 \gg 1$, which leads to the Bloch wave function (4) being localized in the neighborhood of the potential wells (sites). In turn, this leads us to be able to carry out explicit calculations. The orbital method provides a natural framework for the discrete approach.

The method begins by approximating the potential as follows:

$$V(\mathbf{r}) \approx \sum_{\mathbf{v}} V_s(\mathbf{r} - \mathbf{v}), \tag{8}$$

where $V_s(\mathbf{r})$ denotes the potential at the site $S_\mathbf{0}$ and it only has a global minimum at $S_\mathbf{0}$. We also introduce

$$\Delta V(\mathbf{r}) = V(\mathbf{r}) - V_s(\mathbf{r}),$$

where $\Delta V(\mathbf{r})$ is zero in Ω. The rigorous construction of the potential $V_s(\mathbf{r})$ associated with the orbital (see later) can be found in [27]. Because the overall value of the potential is not important, we take the potential to satisfy $\max_\mathbf{r}\{V(\mathbf{r})\} = 0$.

In the TB limit, Bloch modes associated with lowest few bands can be approximated by orbitals which are defined as

$$\mathcal{H}_s\phi(\mathbf{r}) \equiv \left(-\nabla^2 + V_s(\mathbf{r})\right)\phi(\mathbf{r}) = E\phi(\mathbf{r}), \tag{9}$$

where E and $\phi(\mathbf{r})$ are the eigenvalue (called orbital energy) and corresponding eigenfunction (orbital) of the operator $\mathcal{H}_s$. $\mathcal{H}_s$ usually has finite number of discrete eigenvalues because $V_s(\mathbf{r})$ is bounded. In this paper, we concentrate on the lowest several bands in the TB limit. For convenience, we require that the orbitals are real and have norm 1, i.e., $\int \phi^2(\mathbf{r})d\mathbf{r} = 1$.

Next, a discrete approach is used to compute the dispersion relation. The Bloch modes are represented by

$$\varphi(\mathbf{r};\mathbf{k}) \approx \sum_{\mathbf{v}} \phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}}. \tag{10}$$

We note that orbitals and Wannier functions (7) are not equivalent. Wannier functions provide an exact representation of the Bloch modes (see Equation (6)). However, they are in general difficult to explicitly obtain, especially, in higher dimensions or in nonsimple lattices. On the other hand, orbitals can be used to approximate Bloch modes in a very straightforward way. This approximation also provides clear and intuitive insight into lattice problems.

We assume the nullspace of the operator $(\mathcal{H}_s + E)$ is 1-D; later we discuss the case when there is a multidimensional nullspace. Substituting the decomposition of the Bloch mode (10) into the eigenvalue problem (3) leads to

$$\sum_{\mathbf{v}} (\mathcal{H}_s + E)\phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}} = (E - \mu + \Delta V(\mathbf{r}))\sum_{\mathbf{v}} \phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}}. \tag{11}$$

Multiplying both sides by $\phi(\mathbf{r})$ and using

$$\int [(\mathcal{H}_s - E)\phi(\mathbf{r}-\mathbf{v})]\,\phi(\mathbf{r})d\mathbf{r} = 0,$$

we obtain the dispersion relation

$$\mu = E + \frac{\sum_{\mathbf{v}} \lambda_{\mathbf{v}} e^{i\mathbf{k}\cdot\mathbf{v}}}{\sum_{\mathbf{v}} \kappa_{\mathbf{v}} e^{i\mathbf{k}\cdot\mathbf{v}}}, \tag{12}$$

where

$$\lambda_{\mathbf{v}} = \int \phi(\mathbf{r})\Delta V(\mathbf{r})\phi(\mathbf{r}-\mathbf{v})d\mathbf{r} \quad \text{and} \quad \kappa_{\mathbf{v}} = \kappa_{-\mathbf{v}} = \int \phi(\mathbf{r})\phi(\mathbf{r}-\mathbf{v})d\mathbf{r}.$$

The dispersion relation in Equation (12) can be simplified in the TB limit. Note that $\kappa_{\mathbf{0}} = 1$, $\lambda_{\mathbf{v}} \ll 1$ and $\kappa_{\mathbf{v}} \ll 1$ when $\mathbf{v} \neq \mathbf{0}$ because $\phi(\mathbf{r})$ is highly localized. To leading order, the dispersion relation is: $\mu_0 = E + \lambda_{\mathbf{0}}$ which is actually the mean value of μ. And because $\sum_{\mathbf{v}\neq\mathbf{0}} \kappa_{\mathbf{v}} e^{i\mathbf{k}\cdot\mathbf{v}} \ll 1$, we have

$$\mu \approx \mu_0 + \sum_{\mathbf{v}} (\lambda_{\mathbf{v}} - \lambda_{\mathbf{0}}\kappa_{\mathbf{v}})e^{i\mathbf{k}\cdot\mathbf{v}}. \tag{13}$$

We have actually found the Fourier coefficients of the dispersion relation $\mu = \mu(\mathbf{k})$ in terms of integrals of orbitals.

Note that both $\lambda_{\mathbf{v}}$ and $\kappa_{\mathbf{v}}$ decay rapidly as $|\mathbf{v}| \to \infty$. Thus we only need to consider the nearest neighbor and onsite interactions in the dominant contribution. Furthermore, we will only need the leading order terms of $\sum_{\mathbf{v}} C_{\mathbf{v}} e^{i\mathbf{k}\cdot\mathbf{v}}$ in our calculations later.

If only nearest neighbor interactions are considered, we get the dispersion relation

$$\mu = E + \lambda_{\mathbf{0}} + \sum_{<\mathbf{v}>} [\lambda_{\mathbf{v}} - \lambda_{\mathbf{0}}\kappa_{\mathbf{v}}] e^{i\mathbf{k}\cdot\mathbf{v}}. \tag{14}$$

Here and afterwards $<\mathbf{v}>$ means the sum over $\mathbf{v}$ only takes nearest (nonzero) neighbor shift vectors.

All nearest neighbor interaction coefficients are small and of the same order of smallness, which we label as τ. Let $C_{\mathbf{v}} = (\lambda_{\mathbf{v}} - \lambda_{\mathbf{0}}\kappa_{\mathbf{v}})/\tau$ which is $O(1)$. Then we have the dispersion relation

$$\mu(\mathbf{k}) = \mu_0 + \tau\omega(\mathbf{k}), \tag{15}$$

where $\mu_0 = E + \lambda_0$ is the mean value of the dispersion relation, $\omega(\mathbf{k}) = \sum_{<\mathbf{v}>} C_{\mathbf{v}} e^{i\mathbf{k}\cdot\mathbf{v}}$ is the effective dispersion relation and τ measures the thickness of the dispersion relation band.

The coefficients $C_{\mathbf{v}}$ are given in terms of integrals that depend on orbitals. In general these integrals cannot be computed in closed form. A simple estimate can be obtained by using the quantum harmonic oscillator solutions [12, 28]; A rigorous approximation can be obtained via the WKB method [27].

As $V_0 \to \infty, \tau \to 0$ and consequently the dispersion surface becomes flatter and flatter. On the contrary, the two nearest orbital energy difference $E^{(j+1)} - E^{(j)} \sim O(\sqrt{V_0})$, thus $E^{(j+1)} - E^{(j)} \to \infty$. Consequently, there may exist a gap between $\mu^{(j+1)}(\mathbf{k})$ and $\mu^{(j)}(\mathbf{k})$.

Note that the ground state (lowest eigenfunction) of the operator $\mathcal{H}_s$ is taken to be simple in the above calculation; however, the eigenvalues associated with the higher excited states can be degenerate; i.e., there can be multiple eigenfunctions corresponding to one eigenvalue. The above is consistent with the TB limit. Suppose there exist two independent eigenfunction $\phi_1(\mathbf{r})$ and $\phi_2(\mathbf{r})$ corresponding to the eigenvalue E. We can always make $\phi_1(\mathbf{r})$ and $\phi_2(\mathbf{r})$ be orthogonal to each other, i.e., $\int \phi_1(\mathbf{r})\phi_2(\mathbf{r})\, d\mathbf{r} = 0$. Each eigenfunction $\phi_j(\mathbf{r})$, $j = 1, 2$ has one associated Bloch mode $\varphi_j(\mathbf{r}; \mathbf{k})$. In other words,

$$\varphi_j(\mathbf{r}; \mathbf{k}) = \sum_{\mathbf{v}} \phi_j(\mathbf{r} - \mathbf{v}) e^{i\mathbf{k}\cdot\mathbf{v}}. \tag{16}$$

Then each Bloch mode $\varphi_j(\mathbf{r}; \mathbf{k})$ corresponds to one dispersion relation branch $\mu_j(\mathbf{k})$, i.e., $\mathcal{H}\varphi_j(\mathbf{r}; \mathbf{k}) = \mu_j(\mathbf{k})\varphi_j(\mathbf{r}; \mathbf{k})$. From the above analysis, we can obtain the dispersion relation (12) for each $\mu_j(\mathbf{k})$,

$$\mu_j(\mathbf{k}) = E + \lambda_{j,\mathbf{0}} + \tau\omega_j(\mathbf{k}), \quad j = 1, 2, \tag{17}$$

where $\omega_j, \lambda_{j,\mathbf{0}}$ are similarly defined as above and the subindex corresponds to the orbital $\phi_j(\mathbf{r})$. The orders of the nearest interactions for two different orbitals are the same. Each dispersion relation branch $\mu_j(\mathbf{k})$ as a function of $\mathbf{k}$ is single-valued and we assume has certain regularity such as smoothness,

which holds in the TB limit. The thickness of each branch which is measured by τ is the same order. Because $C_{s,\mathbf{v}}$ is not the same for different branches, i.e., usually $\mu_1(\mathbf{k}) \neq \mu_2(\mathbf{k})$. However, there may be some values at which $\mu_1(\mathbf{k})$ and $\mu_2(\mathbf{k})$ are equal. Then the two branches intersect with each other and hence can be thought of as belonging to the same dispersion relation band. In such a case the band is degenerate and has two dispersion relation branches that intersect with each other. If the multiplicity of the nullspace of the operator $\mathcal{H}_s + E$ is greater than two, then there will be more branches.

3. Discrete envelope dynamics

In the linear limit, the dynamics of Bloch modes are determined by the dispersion relation. Similar to Fourier modes, the Bloch modes form a complete set in $L^2(\mathbb{R}^2)$[29], thus a function in $L^2(\mathbb{R}^2)$ can be decomposed of Bloch mode components. Due to the superposition principle of linear problems, different Bloch modes have different dynamics and they do not interfere with each other; thus each Bloch mode component obeys its own dynamics. However, when nonlinearity is present, the dynamics is more complex. In this section, we will derive the governing equation of the envelope dynamics with weak nonlinearity by employing a discrete approach.

We first study the case when $\mu(\mathbf{k})$ has a single dispersion relation branch and assume to leading order

$$\psi \sim \left(\sum_{\mathbf{v}} a_{\mathbf{v}}(Z)\phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}} \right) e^{-i\mu z}. \tag{18}$$

Here $a_{\mathbf{v}}$ represents the Bloch wave envelope at the site $S_{\mathbf{v}}$ and varies slowly under evolution; here $Z = \varepsilon z$; the small parameter ε will be determined later.

Substituting the envelope representation (18) into the lattice NLS Equation (1), one obtains

$$\begin{aligned} &\left(-\nabla^2 + V_s(\mathbf{r}-\mathbf{p}) - E\right) \left(\sum_{\mathbf{v}} a_{\mathbf{v}}(Z)\phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}} \right) \\ =&\sum_{\mathbf{v}} \left(\varepsilon i \frac{da_{\mathbf{v}}}{dZ} + a_{\mathbf{v}}\left(\mu - E - \Delta V(\mathbf{r}-\mathbf{p})\right) \right) \phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}} \\ &+ \sigma \left(\sum_{\mathbf{v}} a_{\mathbf{v}}\phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}} \right)^2 \left(\sum_{\mathbf{v}} a_{\mathbf{v}}\phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}} \right)^*, \end{aligned} \tag{19}$$

where $\mathbf{p} \in \mathbb{P}$. Multiplying both sides by $\phi(\mathbf{r}-\mathbf{p})$ yields

$$i\varepsilon \frac{da_{\mathbf{p}}}{dZ} + \sum_{\mathbf{v}} a_{\mathbf{p}+\mathbf{v}}((\mu - E)\kappa_{\mathbf{v}} - \lambda_{\mathbf{v}})e^{i\mathbf{k}\cdot\mathbf{v}} + \sigma \sum_{\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3} \gamma_{\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3} a_{\mathbf{p}_1} a_{\mathbf{p}_2} a^*_{\mathbf{p}_3} = 0, \tag{20}$$

where we use the definitions in the previous section, consider only leading order terms and define

$$\gamma_{\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3} = \int \phi(\mathbf{r}-\mathbf{p}_1)\phi(\mathbf{r}-\mathbf{p}_2)\phi(\mathbf{r}-\mathbf{p}_3)\phi(\mathbf{r}-\mathbf{p})\,d\mathbf{r}e^{i\mathbf{k}\cdot(\mathbf{p}_1+\mathbf{p}_2-\mathbf{p}_3-\mathbf{p})}. \quad (21)$$

When only onsite and nearest neighbor interactions are taken into account, then the governing equation is

$$i\varepsilon\frac{da_\mathbf{p}}{dZ} + \tau\omega(\mathbf{k})a_\mathbf{p} - \tau\sum_{<\mathbf{v}>} a_{\mathbf{p}+\mathbf{v}}C_\mathbf{v}e^{i\mathbf{k}\cdot\mathbf{v}} + \sigma g|a_\mathbf{p}|^2a_\mathbf{p} = 0, \quad (22)$$

where $g = \gamma_{\mathbf{ppp}}$ is the only onsite interaction term taken for the nonlinear term. Here we assume that ε, σ, and τ all have the same order to ensure maximal balance.

Finally, we obtain the following discrete NLS-type equation

$$i\frac{da_\mathbf{p}}{dZ} + \omega(\mathbf{k})a_\mathbf{p} - \sum_{<\mathbf{v}>} a_{\mathbf{p}+\mathbf{v}}C_\mathbf{v}e^{i\mathbf{k}\cdot\mathbf{v}} + s(\sigma)g|a_\mathbf{p}|^2a_\mathbf{p} = 0, \quad (23)$$

where we have taken $\varepsilon = \tau = |\sigma|$ for convenience, $s(\sigma)$ is the sign of σ. Equation (23) is the unified discrete system, which describes the dynamics of a single envelope in any simple nonlinear periodic lattice. Note that the coefficients of the linear terms of the dynamic equation are directly related to the coefficients of the linear dispersion relation in the TB limit. Later we show that this system reduces to the NLS equation in the continuous limit.

Next we consider the case when $\mu(\mathbf{k})$ at some $\mathbf{k}$ is degenerate; e.g., the case when it has two dispersion relation branches and they intersect with each other at the $\mathbf{k}$ values we are studying. To leading order, we assume the field is a combination of two envelopes corresponding to the two linearly independent Bloch modes

$$\psi \sim \left(\sum_\mathbf{v}(a_\mathbf{v}\phi_1(\mathbf{r}-\mathbf{v}) + b_\mathbf{v}\phi_2(\mathbf{r}-\mathbf{v}))e^{i\mathbf{k}\cdot\mathbf{v}}\right)e^{-i\mu z}, \quad (24)$$

where $\phi_1(\mathbf{r})$ and $\phi_2(\mathbf{r})$ are the two linearly independent eigenfunctions of the operator $\mathcal{H}_s$ and correspond to the the same eigenvalue E. The eigenfunctions can be made orthogonal, i.e.,

$$\int \phi_1(\mathbf{r})\phi_2(\mathbf{r})\,d\mathbf{r} = 0.$$

Substituting Equation (24) into Equation (1) yields

$$
\begin{aligned}
&(-\nabla^2+V_s(\mathbf{r}-\mathbf{p})-E)\sum_{\mathbf{v}}(a_{\mathbf{v}}\phi_1(\mathbf{r}-\mathbf{v})+b_{\mathbf{v}}\phi_2(\mathbf{r}-\mathbf{v}))e^{i\mathbf{k}\cdot\mathbf{v}}\\
&\quad=\sum_{\mathbf{v}}\left(\varepsilon i\frac{da_{\mathbf{v}}}{dZ}+a_{\mathbf{v}}\left(\mu-E-\Delta V(\mathbf{r}-\mathbf{p})\right)\right)\phi_1(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}}\\
&\qquad+\sum_{\mathbf{v}}\left(\varepsilon i\frac{db_{\mathbf{v}}}{dZ}+b_{\mathbf{v}}\left(\mu-E-\Delta V(\mathbf{r}-\mathbf{p})\right)\right)\phi_2(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}}\\
&\qquad+\sigma\left(\sum_{\mathbf{v}}(a_{\mathbf{v}}\phi_1(\mathbf{r}-\mathbf{v})+b_{\mathbf{v}}\phi_2(\mathbf{r}-\mathbf{v}))e^{i\mathbf{k}\cdot\mathbf{v}}\right)^2\\
&\qquad\times\left(\sum_{\mathbf{v}}(a_{\mathbf{v}}\phi_1(\mathbf{r}-\mathbf{v})+b_{\mathbf{v}}\phi_2(\mathbf{r}-\mathbf{v}))e^{i\mathbf{k}\cdot\mathbf{v}}\right)^*.
\end{aligned}
$$

Due to the degeneracy, multiplying both sides by $\phi_1(\mathbf{r}-\mathbf{p})$ and $\phi_2(\mathbf{r}-\mathbf{p})$, respectively, yield the governing system

$$
i\varepsilon\frac{da_{\mathbf{p}}}{dZ}+\tau\omega_1(\mathbf{k})a_{\mathbf{p}}-\tau\sum_{<\mathbf{v}>}a_{\mathbf{p}+\mathbf{v}}C_{1,\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}}+\sigma\Psi_1=0, \tag{25a}
$$

$$
i\varepsilon\frac{db_{\mathbf{p}}}{dZ}+\tau\omega_2(\mathbf{k})b_{\mathbf{p}}-\tau\sum_{<\mathbf{v}>}b_{\mathbf{p}+\mathbf{v}}C_{2,\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}}+\sigma\Psi_2=0, \tag{25b}
$$

where the nonlinear terms are

$$
\begin{aligned}
\Psi_1&=g_1|a_{\mathbf{p}}|^2a_{\mathbf{p}}+g_3\left(b_{\mathbf{p}}^2a_{\mathbf{p}}^*+2|b_{\mathbf{p}}|^2a_{\mathbf{p}}\right)+g_4\left(a_{\mathbf{p}}^2b_{\mathbf{p}}^*+2|a_{\mathbf{p}}|^2b_{\mathbf{p}}\right)+g_5|b_{\mathbf{p}}|^2b_{\mathbf{p}},\\
\Psi_2&=g_2|b_{\mathbf{p}}|^2b_{\mathbf{p}}+g_3\left(a_{\mathbf{p}}^2b_{\mathbf{p}}^*+2|a_{\mathbf{p}}|^2b_{\mathbf{p}}\right)+g_5\left(b_{\mathbf{p}}^2a_{\mathbf{p}}^*+2|b_{\mathbf{p}}|^2a_{\mathbf{p}}\right)+g_4|a_{\mathbf{p}}|^2a_{\mathbf{p}}.
\end{aligned}
$$

Here $g_1=\int\phi_1^4d\mathbf{r}$, $g_2=\int\phi_2^4d\mathbf{r}$, $g_3=\int\phi_1^2\phi_2^2d\mathbf{r}$, $g_4=\int\phi_1^3\phi_2d\mathbf{r}$, and $g_5=\int\phi_1\phi_2^3d\mathbf{r}$.

Letting $\varepsilon=\tau=|\sigma|$, we have

$$
\begin{aligned}
i\frac{da_{\mathbf{p}}}{dZ}+\omega_1(\mathbf{k})a_{\mathbf{p}}-\sum_{<\mathbf{v}>}a_{\mathbf{p}+\mathbf{v}}C_{1,\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}}+s(\sigma)\Psi_1=0,\\
i\frac{db_{\mathbf{p}}}{dZ}+\omega_2(\mathbf{k})b_{\mathbf{p}}-\sum_{<\mathbf{v}>}b_{\mathbf{p}+\mathbf{v}}C_{2,\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}}+s(\sigma)\Psi_2=0.
\end{aligned}
$$

It is noted that the linear parts of the two components are decoupled and the coefficients of the linear parts of Equations (25) are directly related to the Fourier coefficients of the two underlying dispersion relations. The two components are coupled only with nonlinear terms. Usually $g_1=g_2$, and $g_4=g_5=0$ if ϕ_1 and ϕ_2 are odd over x or y which is the case of the second band.

The system (25) is a unified description of a coupled discrete NLS-type system that results when two dispersion branches intersecting at point $\mathbf{k}$ in a simple lattice. This description can be extended to higher order degeneracies.

4. Continuous limit of discrete systems

We consider a further limit, we assume that the envelope $a_{\mathbf{v}}$ varies slowly over $\mathbf{v}$ with a scale $\mathbf{R} = \nu\mathbf{r}$, $\nu \ll 1$. Thus, the envelope has the form

$$\psi \sim \sum_{\mathbf{v}} a_{\mathbf{v}}(Z)\phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}}e^{-i\mu z} \approx \sum_{\mathbf{v}} a(\mathbf{R}, Z)\phi(\mathbf{r}-\mathbf{v})e^{i\mathbf{k}\cdot\mathbf{v}}e^{-i\mu z}, \quad (26)$$

In this new spacial coordinate system, $a_{\mathbf{v}} \approx \int a(\mathbf{R})\phi^2(\mathbf{r}-\mathbf{v})d\mathbf{r} \approx a(\nu S_{\mathbf{v}})$, where $a_{\mathbf{v}}$ is defined at site points. We are going to find the evolution equations of $a(\mathbf{R}, Z)$ from the discrete equations.

Before proceeding, we assume each branch of the dispersion relation is sufficiently smooth at the $\mathbf{k}$ value we are studying. We also introduce the notations: $\partial_m = \frac{\partial}{\partial \mathbf{r}_m}$ and $\nabla = (\partial_1, \partial_2)$; $\tilde{\partial}_m = \frac{\partial}{\partial \mathbf{R}_m}$ and $\tilde{\nabla} = (\tilde{\partial}_1, \tilde{\partial}_2)$; $\bar{\partial}_m = \frac{\partial}{\partial k_m}$ and $\bar{\nabla} = (\bar{\partial}_1, \bar{\partial}_2)$; $\bar{\partial}_{m,n} = \bar{\partial}_m\bar{\partial}_n$ and $\tilde{\partial}_{m,n} = \tilde{\partial}_n\tilde{\partial}_m$.

Using Taylor expansion, we get

$$a_{\mathbf{p}+\mathbf{v}} \approx a_{\mathbf{p}} + \nu\mathbf{v}\cdot\tilde{\nabla}a + \frac{\nu^2}{2}\mathbf{v}\tilde{\mathbf{H}}\mathbf{v}^T a,$$

where $\tilde{\mathbf{H}} = \begin{pmatrix} \tilde{\partial}_{11} & \tilde{\partial}_{12} \\ \tilde{\partial}_{21} & \tilde{\partial}_{22} \end{pmatrix}$ is the Hessian matrix operator.

Then

$$\begin{aligned} \sum_{<\mathbf{v}>} a_{\mathbf{p}+\mathbf{v}}C_{\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}} &\approx a_{\mathbf{p}}\sum_{<\mathbf{v}>} C_{\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}} + \nu\tilde{\nabla}a\cdot\sum_{<\mathbf{v}>}\mathbf{v}C_{\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}} + \frac{\nu^2}{2}\sum_{<\mathbf{v}>}C_{\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}}\mathbf{v}\tilde{\mathbf{H}}\mathbf{v}^T a \\ &= a_{\mathbf{p}}\omega(\mathbf{k}) - i\nu\bar{\nabla}\omega\cdot\tilde{\nabla}a - \frac{\nu^2}{2}\sum_{m,n=1}^{2}\bar{\partial}_{m,n}\omega\tilde{\partial}_{m,n}a. \end{aligned} \quad (27)$$

Substituting Equation (27) into Equation (22) yields

$$i\varepsilon\frac{\partial a}{\partial Z} + i\nu\tau\bar{\nabla}\omega\cdot\tilde{\nabla}a + \frac{\nu^2\tau}{2}\sum_{m,n=1}^{2}\bar{\partial}_{m,n}\omega\tilde{\partial}_{m,n}a + \sigma g|a|^2 a = 0. \quad (28)$$

Note that we use Equation (22) instead of Equation (23) to derive the continuous equation because the maximal balance is modified due to the new scale ν.

The above equation governs the nonlinear dynamics in simple periodic media of a single Bloch mode envelope in the continuous limit from the discrete system. It is valid for any value of $\mathbf{k}$. In analogy to homogeneous media, $\bar{\nabla}\omega$ plays the role of the group velocity; it is the velocity of the envelope. In special cases, $\bar{\nabla}\omega = \mathbf{0}$. This condition gives the extreme points of the dispersion surface. At these points, the group velocity is zero, and the

envelope will remain at the center of the cross-section. The envelope has a spatial shift in the cross-section when propagating along z direction if $\bar{\nabla}\omega \neq \mathbf{0}$. However, introducing a moving frame variable and $\bar{\mathbf{R}} = \mathbf{R} - \bar{\nabla}\omega Z$ (we drop the bar on $\mathbf{R}$) and rescaling yield the equation

$$i\frac{\partial a}{\partial Z} + \frac{1}{2}\sum_{m,n=1}^{2} \bar{\partial}_{m,n}\omega\tilde{\partial}_{m,n}a + s(\sigma)g|a|^2 a = 0, \tag{29}$$

where we have taken the maximal balance condition $\varepsilon = \nu^2\tau = |\sigma|$. The above Equation (29) is a 2-D NLS equation. At different values of $\mathbf{k}$, the dispersive terms may be elliptic, hyperbolic, or even parabolic.

However, when the eigenspace corresponding to $\mu(\mathbf{k})$ is multidimensional, the envelope may belong to different Bloch modes and they may interact with each other. The envelope is assumed to have the form

$$\begin{aligned}\psi &\sim \sum_{\mathbf{v}}(a_{\mathbf{v}}(Z)\phi_1(\mathbf{r}-\mathbf{v}) + b_{\mathbf{v}}(Z)\phi_2(\mathbf{r}-\mathbf{v}))e^{i\mathbf{k}\cdot\mathbf{v}}e^{-i\mu z}\\ &\approx \sum_{\mathbf{v}}(a(\mathbf{R}, Z)\phi_1(\mathbf{r}-\mathbf{v}) + b(\mathbf{R}, Z)\phi_2(\mathbf{r}-\mathbf{v}))e^{i\mathbf{k}\cdot\mathbf{v}}e^{-i\mu z}.\end{aligned}$$

Similar to the one single Bloch mode envelope case, the system (25) becomes

$$i\varepsilon\frac{da}{dZ} + i\nu\tau\bar{\nabla}\omega_1\cdot\tilde{\nabla}a + \frac{\nu^2\tau}{2}\sum_{m,n=1}^{2}\bar{\partial}_{m,n}\mu_1\tilde{\partial}_{m,n}a + \sigma\overline{\Psi}_1 = 0, \tag{30a}$$

$$i\varepsilon\frac{db}{dZ} + i\nu\tau\bar{\nabla}\omega_2\cdot\tilde{\nabla}b + \frac{\nu^2\tau}{2}\sum_{m,n=1}^{2}\bar{\partial}_{m,n}\mu_2\tilde{\partial}_{m,n}b + \sigma\overline{\Psi}_2 = 0. \tag{30b}$$

Here $\overline{\Psi}_j$ is the continuous version of Ψ_j, $j = 1, 2$,

$$\overline{\Psi}_1 = g_1|a|^2a + g_3(b^2a^* + 2|b|^2a) + g_4(a^2b^* + 2|a|^2b) + g_5|b|^2b, \tag{31a}$$

$$\overline{\Psi}_2 = g_2|b|^2b + g_3(a^2b^* + 2|a|^2b) + g_5(b^2a^* + 2|b|^2a) + g_4|a|^2a. \tag{31b}$$

The system (30) is the continuous system governing the propagation of two Bloch mode envelopes corresponding to the same μ. It is noted that if $|\bar{\nabla}\omega_j| = 0$, i.e., both μ_1 and μ_2 reach the same extreme value at the same value of $\mathbf{k}$, the above system becomes (after rescaling),

$$i\frac{\partial a}{\partial Z} + \frac{1}{2}\sum_{m,n=1}^{2}\bar{\partial}_{m,n}\omega_1\tilde{\partial}_{m,n}a + s(\sigma)\overline{\Psi}_1 = 0, \tag{32a}$$

$$i\frac{\partial b}{\partial Z} + \frac{1}{2}\sum_{m,n=1}^{2}\bar{\partial}_{m,n}\omega_2\tilde{\partial}_{m,n}b + s(\sigma)\overline{\Psi}_2 = 0, \tag{32b}$$

where we have taken $\varepsilon = \tau\nu^2 = |\sigma|$.

As was the case for the discrete couple NLS-type system, the above system (32) is a system of coupled NLS equations with unusual nonlinear couplings (even if $g_1 = g_2$ and $g_4 = g_5 = 0$). If $\mathbf{k}$ is in the vicinity of those points with $|\bar{\nabla}\omega_2 - \bar{\nabla}\omega_1| \sim O(\nu)$ and $\mu(\mathbf{k})$ remains degenerate, a detuning phenomenon occurs.

After rescaling and absorbing the transport term of the a component, we get

$$i\frac{\partial a}{\partial Z} + \frac{1}{2}\sum_{m,n=1}^{2} \bar{\partial}_{m,n}\omega_1 \tilde{\partial}_{m,n} a + s(\sigma)\overline{\Psi}_1 = 0, \tag{33a}$$

$$i\frac{\partial b}{\partial Z} + i\mathbf{g}\cdot\tilde{\nabla} b + \frac{1}{2}\sum_{m,n=1}^{2} \bar{\partial}_{m,n}\omega_2 \tilde{\partial}_{m,n} b + s(\sigma)\overline{\Psi}_2 = 0, \tag{33b}$$

where $\mathbf{g} = \frac{1}{\nu}\left(\bar{\nabla}\omega_2 - \bar{\nabla}\omega_1\right)$.

In the previous sections, we derived the governing equations of the Bloch envelope dynamics for general simple lattices. First from the TB approximation we obtained discrete evolution equations for all values of $\mathbf{k}$ in the Brillouin zone. From these discrete equations the corresponding continuous equations are derived naturally in the continuous limit. These continuous equations can also be derived directly by multiscale expansion methods (see the Appendix).

As mentioned earlier, the discrete approach is from the TB limit, i.e., $V_0 \gg 1$. This limit ensures the dispersion relation thickness τ is small. We have two other small scales: the slow time evolution scale ε and the nonlinear scale σ. The maximal balance in these discrete systems is $\varepsilon \sim O(\tau) \sim O(\sigma)$. In discrete systems, the spatial scale is the order of the lattice constant which is assumed to be $O(1)$. Thus the discrete system describes the spatial dynamics on the lattice scale. To study phenomena on a large spatial scale, the envelopes are assumed to vary slowly on a scale $O(\nu)$ in space. Then the discrete systems can be transformed to continuous systems. In the continuous system, the maximal balance then becomes $\varepsilon \sim O(\tau\nu^2) \sim O(\sigma)$. In this case, effective NLS-type equations are derived in a moving frame.

5. Typical simple lattices

In the above-mentioned sections we have derived the dispersion relation for arbitrary simple lattices and the corresponding dynamics of Bloch mode envelopes. In this section, we will study two typical simple lattices which are special cases of the above general analysis.

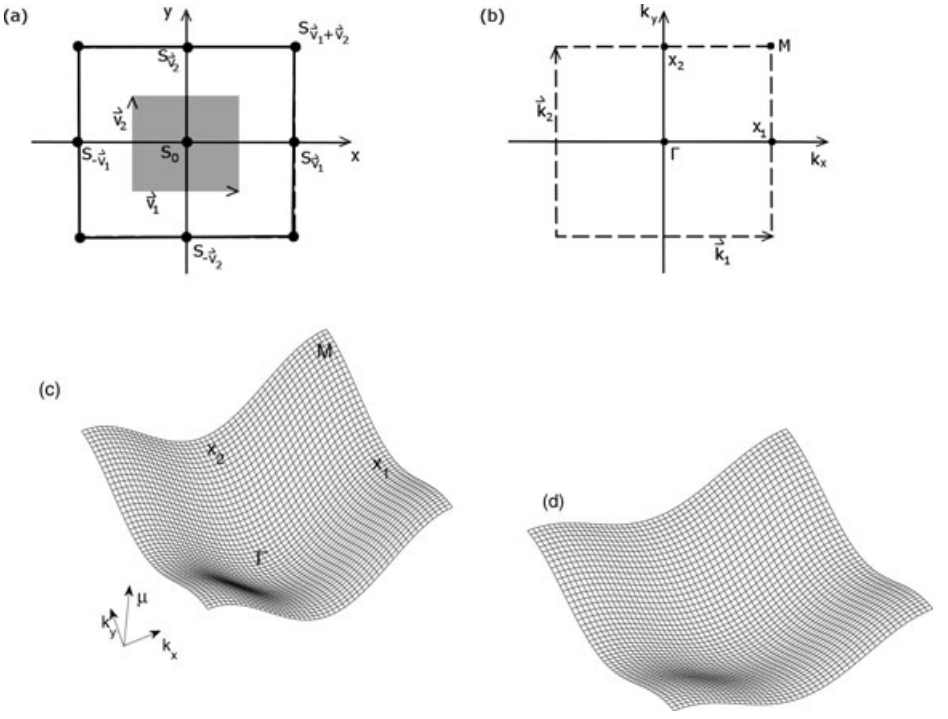

Figure 1. (a) The site distributions of the square lattice (34). The shadow region can be considered as a primitive unit cell. (b) The Brillouin zone of the square lattice is also a square denoted here by dashed lines. The first dispersion relation band (c) from direction simulation and (d) from the analytical formula (35).

5.1. Square lattices

Square 2-D periodic structures are common in nature and can be readily engineered in optics (cf. [4]). A typical square lattice is

$$V(x, y) = \frac{V_0}{2}\left(\sin^2(k_0 x) + \sin^2(k_0 y) - 2\right), \tag{34}$$

where $V_0 > 0$ is the lattice intensity; here k_0 corresponds to the scaled wave length of the interfering plane waves.

Its characteristic vectors are

$$\mathbf{v}_1 = l\,(1, 0)\,, \quad \mathbf{v}_2 = l\,(0, 1)\,,$$
$$\mathbf{k}_1 = \frac{2\pi}{l}\,(1, 0)\,, \quad \mathbf{k}_2 = \frac{2\pi}{l}\,(0, 1)\,,$$

where $l = \frac{\pi}{k_0}$ is the length of the periods (i.e., the lattice constant). In this paper we let $l = 1$ for simplicity.

The site distribution is displayed in Figure 1(a). For this potential, one site has four nearest neighbors. The nearest neighbor shift vectors are $\mathbf{v}_1$, $-\mathbf{v}_1$, $\mathbf{v}_2$,

$-\mathbf{v}_2$. They all have the same length. Note that $S_{\mathbf{v}_1+\mathbf{v}_2}$ is not one of the nearest neighbors of $S_{\mathbf{0}}$.

For the first band, we know that $C_{\mathbf{v}_1} = C_{\mathbf{v}_2}$. Thus the dispersion relation of the first band is

$$\mu(\mathbf{k}) = E + \lambda_{\mathbf{0}} + 2\tau(\cos(k_x) + \cos(k_y)). \tag{35}$$

The Brillouin zone is also a square; it is displayed in Figure 1(b) as well as special points. The dispersion relation obtained by direct numerical simulation ($V_0 = 100$) of the eigenproblem (5) is displayed in Figure 1(c) and it agrees very well with the dispersion relation obtained by the formula (35) both qualitatively and quantitatively.

From the analytical formula (35), we can readily get that

$$\bar{\mathbf{H}}\mu = 2\tau \begin{pmatrix} \cos(k_x) & 0 \\ 0 & \cos(k_y) \end{pmatrix},$$

where $\bar{\mathbf{H}} = \begin{pmatrix} \bar{\partial}_{11} & \bar{\partial}_{12} \\ \bar{\partial}_{21} & \bar{\partial}_{22} \end{pmatrix}$ is the Hessian matrix operator with respect to $\mathbf{k}$.

The Hessian matrices at the Γ point ($\mathbf{k} = (0, 0)$) and the M point ($\mathbf{k} = \frac{\pi}{l}(1, 1)$) are, respectively,

$$\Gamma : 2\tau \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\mathrm{M} : 2\tau \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The determinants of the matrices are both positive. Thus we obtain elliptic NLS equations; one is focusing and the other is defocusing. Γ point is the band edge, which corresponds to the semi-infinite gap and M point is the band edge which corresponds to the first band gap. These elliptic NLS equations can describe the band-gap soliton modes [22, 23].

We also note that the corresponding discrete dynamic equations are also "elliptic" discrete NLS equations. For example, at the Γ point the discrete NLS equation is

$$i\frac{da_{m,n}}{dZ} + (a_{m+1,n} + a_{m,n+1} + a_{m-1,n} + a_{m,n-1} - 4a_{m,n}) + s(\sigma)g|a_{m,n}|^2 a_{m,n} = 0.$$

The existence of the discrete solitons and their properties have been well studied (cf. [30]).

However, the Hessian matrices at the X_1 point ($\mathbf{k} = (\pi, 0)$) and the X_2 point ($\mathbf{k} = (0, \pi)$) are, respectively,

$$X_1 : 2\tau \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$X_2 : 2\tau \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The determinants of the matrices are both negative. X_1 and X_2 points are saddle points of the dispersion surface. They are not at the band edges. The envelope dynamics at these points are governed by hyperbolic NLS equations. Correspondingly, the discrete dynamic equations are also "hyperbolic." For example, at the X_1 point the governing equation is

$$i\frac{da_{m,n}}{dZ} + (-a_{m+1,n} - a_{m-1,n} + a_{m,n+1} + a_{m,n-1}) + s(\sigma)g|a_{m,n}|^2 a_{m,n} = 0.$$

Unlike the "elliptic" discrete NLS equations, this equation has not been well investigated. It is noted that defining $(-1)^m a_{m,n}$ can convert "elliptic" to "hyperbolic" discrete NLS equations. It means that the above equation admits "staggered" soliton solutions.

The second band of the dispersion relation for a square lattice is more complicated. It is noted that $\mu(\mathbf{k})$ has two dispersion surfaces $\mu_1(\mathbf{k})$ and $\mu_2(\mathbf{k})$ because the corresponding orbital energy is degenerate. From symmetries, we obtain that

$$\lambda_{1,\mathbf{0}} = \lambda_{2,\mathbf{0}}, \quad g = g_1 = g_2 = 3g_3,$$

and

$$C_{1,\mathbf{v}_1} = C_{2,\mathbf{v}_2} > 0; \quad C_{2,\mathbf{v}_1} = C_{1,\mathbf{v}_2} < 0.$$

Then the two dispersion relation branches are

$$\mu_1(\mathbf{k}) = E + \lambda_{1,\mathbf{0}} + 2\tau(\cos(k_x) - h\cos(k_y)), \tag{36a}$$

$$\mu_2(\mathbf{k}) = E + \lambda_{1,\mathbf{0}} + 2\tau(-h\cos(k_x) + \cos(k_y)), \tag{36b}$$

where $h = \frac{|C_{1,\mathbf{v}_1}|}{|C_{1,\mathbf{v}_2}|} > 0$.

The dispersion relation obtained from the direct numerical simulation is displayed in Figure 2(a) and it agrees very well with the dispersion relation obtained by the formulas (36) given in Figure 2(b). The two different branches obtained analytically are also displayed in Figures 2(c) and (d). It is seen that $\mu_1(\mathbf{k}) \neq \mu_2(\mathbf{k})$ except $k_x = \pm k_y$. Later we study special points where NLS-type equation results. However the theory is valid along the curves of

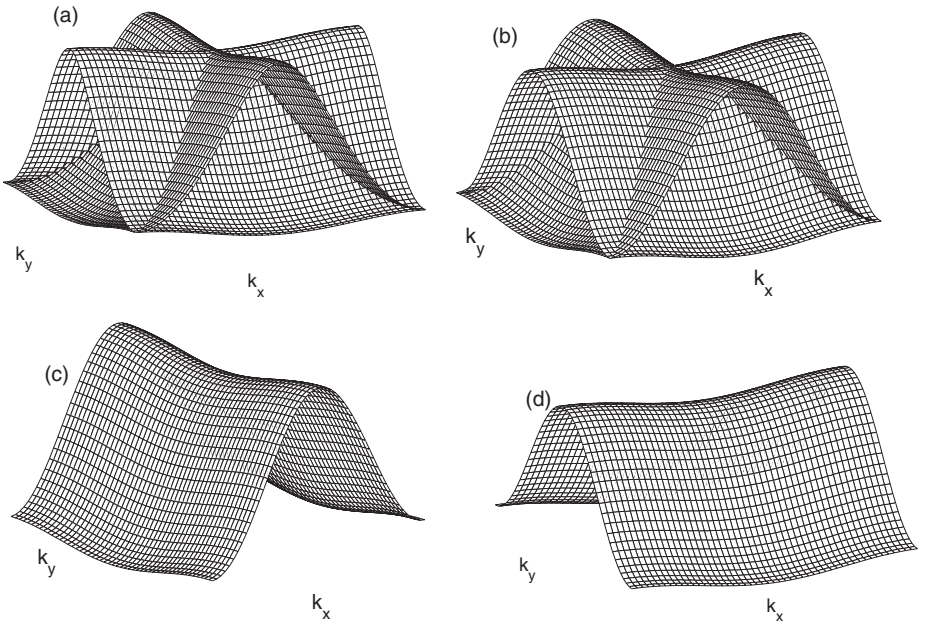

Figure 2. The second dispersion relation band obtained (a) from direct simulation and (b) from the analytical formulas (35). The second band is degenerate which has two branches. (c) The branch $\mu_1(\mathbf{k})$; (d) the branch $\mu_2(\mathbf{k})$.

intersection (here $k_x = \pm k_y$). However, in general, along these curves, one only has transport-type equations to leading order; there is not a maximal balance leading to NLS-type equations.

From the dispersion relations (36), we get

$$\bar{\mathbf{H}}\mu_1 = -2\tau \begin{pmatrix} \cos(k_x) & 0 \\ 0 & -h\cos(k_y) \end{pmatrix};$$

$$\bar{\mathbf{H}}\mu_2 = -2\tau \begin{pmatrix} -h\cos(k_x) & 0 \\ 0 & \cos(k_y) \end{pmatrix}.$$

The Hessian matrices associated with the first and second branches at the special points are

$$\Gamma: \bar{\mathbf{H}}\mu_1 = -2\tau \begin{pmatrix} -1 & 0 \\ 0 & h \end{pmatrix}, \quad \bar{\mathbf{H}}\mu_2 = -2\tau \begin{pmatrix} h & 0 \\ 0 & -1 \end{pmatrix},$$

$$\mathrm{M}: \bar{\mathbf{H}}\mu_1 = -2\tau \begin{pmatrix} -1 & 0 \\ 0 & h \end{pmatrix}, \quad \bar{\mathbf{H}}\mu_2 = -2\tau \begin{pmatrix} h & 0 \\ 0 & -1 \end{pmatrix},$$

$$\mathrm{X}_1: \bar{\mathbf{H}}\mu_1 = -2\tau \begin{pmatrix} -1 & 0 \\ 0 & -h \end{pmatrix}, \quad \bar{\mathbf{H}}\mu_2 = 2\tau \begin{pmatrix} h & 0 \\ 0 & 1 \end{pmatrix},$$

$$\mathrm{X}_2: \bar{\mathbf{H}}\mu_1 = -2\tau \begin{pmatrix} -1 & 0 \\ 0 & -h \end{pmatrix}, \quad \bar{\mathbf{H}}\mu_2 = 2\tau \begin{pmatrix} h & 0 \\ 0 & 1 \end{pmatrix}.$$

From the Figure 2 or the above Hessian matrices, we can see that X_1 (X_2) is the maximum point of the second (first) branch and the minimum point of the first (second) branch. And there is no degeneracy at X_1 and X_2 points. The dynamic equations of the corresponding Bloch mode envelope at these points are just the scalar NLS equations. However, μ_1 equals to μ_2 at the Γ and M points, thus they are degenerate points. The corresponding eigenspaces are 2-D. It is also noted that Γ and M points are both saddle points at the two branches. Thus coupled hyperbolic NLS equations are used to govern the Bloch mode envelopes. For example, at the Γ point, the governing system in the continuous limit is

$$i\frac{\partial a}{\partial Z} + (\partial_{XX} - h\partial_{YY})a + s(\sigma)g\left(|a|^2 a + \frac{1}{3}(b^2 a^* + 2|b|^2 a)\right) = 0, \quad (37a)$$

$$i\frac{\partial b}{\partial Z} + (-h\partial_{XX} + \partial_{YY})b + s(\sigma)g\left(|b|^2 b + \frac{1}{3}(a^2 b^* + 2|a|^2 b)\right) = 0. \quad (37b)$$

This is a coupled hyperbolic NLS system (recall $h > 0$).

5.2. *Triangular lattices*

An equilateral triangular lattice is also a common simple lattice in nature and has recently attracted interests in nonlinear optics [31]. In this paper we consider the triangular lattice in Figure 3. The unit cell (see the shadow region in Figure 3a) is composed of two equilateral triangles pointing left and right. Its reciprocal lattice is shown in Figure 3(b). The reciprocal unit cell is also composed of two equilateral triangles, now pointing up and down. In both the physical and reciprocal lattices, the unit cells which are parallelograms can be grouped to form hexagonal tiles. For example in Figure 3(b), the shadow region is the reciprocal unit cell. It is equivalent to the hexagon surrounded by dotted lines. The equivalence means the two tiles give the same information of the lattice due to the periodicity. Thus equilateral triangular lattices are also called hexagonal lattices.

A typical triangular lattice which has the structures in Figure 3 is

$$V(\mathbf{r}) = -\frac{V_0}{9}|e^{ik_0\mathbf{b}_1\cdot\mathbf{r}} + e^{ik_0\mathbf{b}_2\cdot\mathbf{r}} + e^{ik_0\mathbf{b}_3\cdot\mathbf{r}}|^2, \quad (38)$$

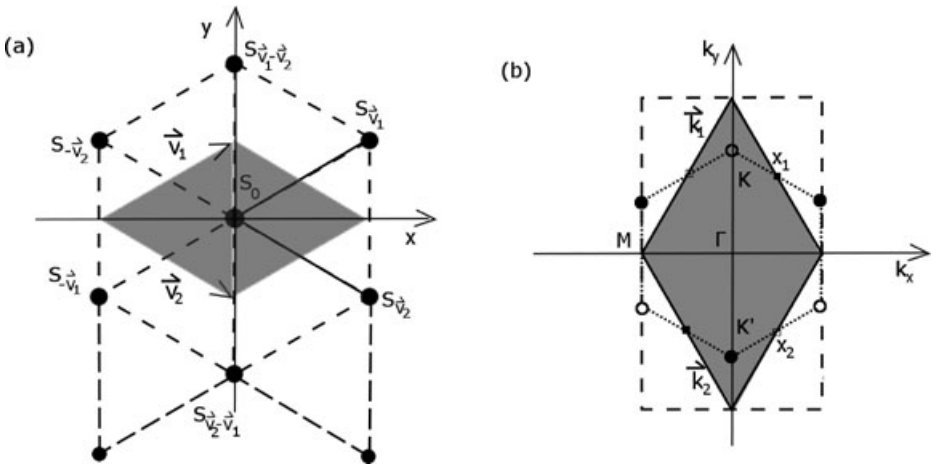

Figure 3. (a) The triangular lattice (38). The shadow region is the primitive unit cell. (b) The extended Brillouin zone of the triangular lattice. The shadow region is the Brillouin zone.

where $\mathbf{b}_n = (\sin(\frac{2n\pi}{3}), \cos(\frac{2n\pi}{3}))$, $n = 1, 2, 3$. The characteristic vectors are

$$\mathbf{v}_1 = l\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right), \quad \mathbf{v}_2 = l\left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right),$$
$$\mathbf{k}_1 = \frac{4\pi}{\sqrt{3}\,l}\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad \mathbf{k}_2 = \frac{4\pi}{\sqrt{3}\,l}\left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right),$$

where $l = \frac{4\pi}{3k_0}$ is the lattice constant. We let $l = 1$ in this paper for simplicity.

In the previous sections, we have shown that the envelope dynamics is closely related the dispersion relation, which motivates the study of the linear dispersion relation. Here we consider the first band of the dispersion relation. From Figure 3, we can see that one site has six nearest neighbors. For instance, the six nearest neighbor sites of site S_0 are $S_{\mathbf{v}_1}$, $S_{\mathbf{v}_2}$, $S_{-\mathbf{v}_1}$, $S_{-\mathbf{v}_2}$, $S_{\mathbf{v}_1-\mathbf{v}_2}$, and $S_{\mathbf{v}_2-\mathbf{v}_1}$. Thus the shift vectors from a site to its nearest neighbors are $\mathbf{v}_1$, $\mathbf{v}_2$, $-\mathbf{v}_1$, $-\mathbf{v}_2$, $\mathbf{v}_1 - \mathbf{v}_2$, $\mathbf{v}_2 - \mathbf{v}_1$.

In this section, we only consider the dispersion relation and the corresponding dynamics of the first band. The dispersion relation is

$$\mu(\mathbf{k}) = E + \lambda_0 - 2\tau(\cos(\mathbf{k}\cdot\mathbf{v}_1) + \cos(\mathbf{k}\cdot\mathbf{v}_2) + \cos(\mathbf{k}\cdot(\mathbf{v}_1 - \mathbf{v}_2))). \quad (39)$$

The dispersion relation of the first band is displayed in Figure 4. Figure 4(a) is from the direct numerical simulation of Equation (3) and Figure 4(b) is from the analytical formula (39). For convenience, the dispersion relation is showed in an extended Brillouin zone which is the rectangular region surrounded by

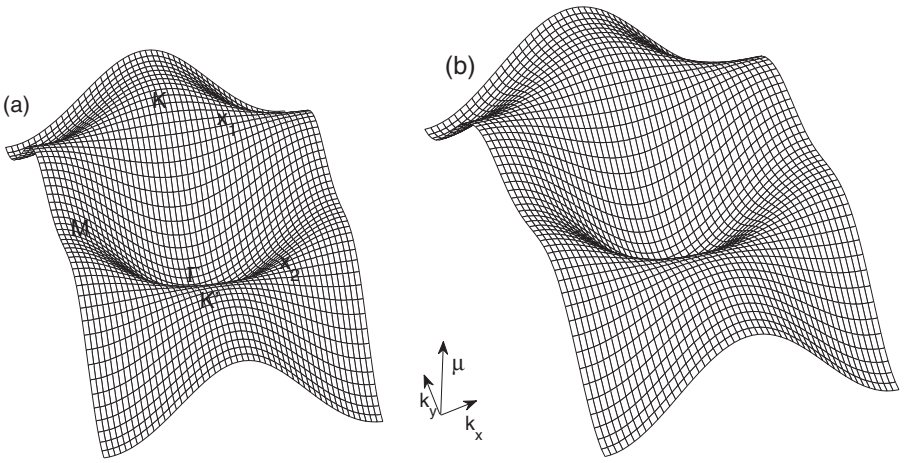

Figure 4. The first band of the dispersion relation: (a) from direct simulation; (b) from the analytical formula (39).

dashed lines in Figure 3. Our analysis is in good agreement with the direct numerical simulation.

From the dispersion (39) we can easily get that

$$\bar{\mathbf{H}}\mu = \frac{\tau}{2}\begin{pmatrix} 3(\cos(\mathbf{k}\cdot\mathbf{v}_1)+\cos(\mathbf{k}\cdot\mathbf{v}_2)) & \sqrt{3}(\cos(\mathbf{k}\cdot\mathbf{v}_1)-\cos(\mathbf{k}\cdot\mathbf{v}_2)) \\ \sqrt{3}(\cos(\mathbf{k}\cdot\mathbf{v}_1)-\cos(\mathbf{k}\cdot\mathbf{v}_2)) & \cos(\mathbf{k}\cdot\mathbf{v}_1)+\cos(\mathbf{k}\cdot\mathbf{v}_2)+4\cos(\mathbf{k}\cdot(\mathbf{v}_1-\mathbf{v}_2)) \end{pmatrix},$$

where $\bar{\mathbf{H}} = \begin{pmatrix} \bar{\partial}_{11} & \bar{\partial}_{12} \\ \bar{\partial}_{21} & \bar{\partial}_{22} \end{pmatrix}$ is the Hessian matrix operator with respect to $\mathbf{k}$.

At the Γ point ($\mathbf{k} = (0, 0)$), the Hessian matrix is

$$\bar{\mathbf{H}}\mu = 2\tau \begin{pmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}.$$

Thus Γ point is the minimum point and the envelope dynamics can be described by a focusing NLS equation if $\sigma > 0$.

At the K ($\mathbf{k} = \frac{4\pi}{3}(0, 1)$) point and K$'$ point ($\mathbf{k} = \frac{4\pi}{3}(0, -1)$), the Hessian matrices are the same,

$$\bar{\mathbf{H}}\mu = 2\tau \begin{pmatrix} -\frac{3}{4} & 0 \\ 0 & -\frac{3}{4} \end{pmatrix}.$$

Thus they are maximum points and the envelope dynamics can be described by the same defocusing NLS equation if $\sigma > 0$.

At the M point ($\mathbf{k} = \frac{4\pi}{\sqrt{3}}(1, 0)$), the Hessian matrix is

$$\bar{\mathbf{H}}\mu = 2\tau \begin{pmatrix} -\frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

Thus M point is a saddle point and the envelope dynamics is governed by a hyperbolic NLS equation.

At the X_1 point ($\mathbf{k} = \frac{2\pi}{\sqrt{3}} \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$), the Hessian matrix is

$$\bar{\mathbf{H}}\mu = 2\tau \begin{pmatrix} 0 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -1 \end{pmatrix}.$$

At the X_2 point($\mathbf{k} = \frac{2\pi}{\sqrt{3}}(\frac{1}{2}, -\frac{\sqrt{3}}{2})$), the Hessian matrix is

$$\bar{\mathbf{H}}\mu = 2\tau \begin{pmatrix} 0 & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -1 \end{pmatrix}.$$

It is seen that both X_1 and X_2 are saddle points but with different eigendirections. The envelope equations are also hyperbolic NLS equations.

We do not consider the higher band case for triangular lattices; the basic situation is explained in the square lattice example. An interested reader can find higher band results by applying the analysis presented in this paper.

6. Conclusion

In this paper, discrete Bloch mode envelopes and their continuous limits in general simple periodic media are studied. It is found that the dispersion relation of the underlying Bloch system plays a key role. Namely, all nonlinear evolution equation for the envelope are found to be deeply connected to the dispersion relation in both discrete and continuous limits. This is similar to nonlinear wave problems in constant bulk media. In the TB limit elementary and convenient analytical descriptions of the dispersion relation can be obtained.

After investigating the dispersion relation, the structure of the nonlinear envelope equation is found. First, the discrete wave system for a single Bloch mode envelope using the TB approximation is developed. Because the system is obtained in *general* simple periodic media the equation depends on the location of the Brillouin zone. Hence this is not the discrete NLS equation in standard form. This discrete nonlinear wave equation takes into account the dispersion relation and geometric structure of the potential. The coefficients of the linear part of the nonlinear discrete wave equation depends on the Fourier coefficients of the dispersion relation. Only under certain conditions or lattice arrangements, for example, a square lattice, does the wave equation

reduce to the standard 2-D discrete NLS equation. When the envelope is a combination of two discrete Bloch mode envelopes which correspond to two linearly independent Bloch modes, a coupled discrete wave system with unusual nonlinear couplings is obtained.

Then, in a further limit from the discrete system the envelopes are taken to vary slowly with respect to the lattice scale. In this continuous limit the discrete systems turn out to be effective NLS equations and the coefficients of the dispersive linear term is the second derivative of the dispersion relation. This is similar to what occurs in standard constant media. These effective NLS equations agree with those derived by a direct multiscale expansion method. Thus, we are able to make the connection between the two approaches.

In addition, general formulas are derived for the dispersion relation in the TB limit, which are valid for any simple periodic lattice. Due to the periodicity/symmetry of the potential, the dispersion relation can multiple band structures. For the lowest bands the parameters are estimated. Details for two simple lattices are given as examples. Direct numerics and analytical results show close agreement.

In this paper, only simple lattices were considered. If the lattice is not simple, the approach needs to be modified appropriately. Results for a honeycomb lattice (a typical nonsimple lattice) can be found in Ref. [12].

Acknowledgments

This research was partially supported by the U.S. Air Force Office of Scientific Research, under grant FA9550-09-1-0250 and by NSF under grant DMS-0905779, Tsinghua University Initiative Scientific Research Program, and NSFC under grant no. 11204155.

Appendix

In this appendix we derive the dynamics of Bloch mode envelopes with a multi-scale expansion method. We assume the input envelopes are much wider than the lattice scale. Under evolutions, envelopes vary slowly both in time and in space. We also assume the nonlinearity is weak. Thus there are many scales involving in the evolutions. Maximal balance conditions will be used to derive the governing equations.

We study the easier case where the eigenspace corresponding to μ is 1-D while the multidimensional eigenspace case can be extended. We also assume $\mu(\mathbf{k})$ is sufficiently smooth at the $\mathbf{k}$ value we are studying. We assume the solution of Equation (1) can be expanded in a multiscale perturbation series

$$\psi = \left(A(\mathbf{R}, Z)U(\mathbf{r};\mathbf{k}) + \nu\psi_1 + \nu^2\psi_2 + \cdots\right) e^{i\mathbf{k}\cdot\mathbf{r} - i\mu z}, \tag{A.1}$$

where $Z = \varepsilon z$ and $\mathbf{R} = \nu \mathbf{r}$. The orders of ε and ν will be determined later via maximal balance; $u(\mathbf{r}; \mathbf{k})$ is taken to have the unit norm, i.e., $\int_\Omega U U^* d\mathbf{r} = 1$.

Substituting Equation (A.1) into the Equation (1) yields that the leading order is satisfied automatically. At order $O(\nu)$, the equation for ψ_1 is

$$(\mathcal{H}_\mathbf{k} - \mu)\psi_1 = i\frac{\partial A}{\partial Z}U + 2\tilde{\nabla}A \cdot (i\mathbf{k} + \nabla)U, \tag{A.2}$$

where we recall that $\tilde{\nabla}A$ corresponds to derivatives with respect to $\mathbf{R}$. $(\mathcal{H}_\mathbf{k} + \mu)$ has a 1-D nullspace, which contains U. Thus the Fredholm condition yields

$$\int_\Omega \left(i\frac{\partial A}{\partial Z}U + 2\tilde{\nabla}A \cdot (i\mathbf{k} + \nabla)U \right) U^* d\mathbf{r} = 0. \tag{A.3}$$

Here we have assumed $\varepsilon = \nu$ and the nonlinearity is of higher order $\sigma \sim O(\nu^2)$.

Note that because

$$(\mathcal{H}_\mathbf{k} - \mu)U = 0,$$

taking derivative over k_m yields

$$(\mathcal{H}_\mathbf{k} - \mu)\bar{\partial}_m U = 2i(ik_m + \partial_m)U + \bar{\partial}_m \mu U \equiv 2i\mathcal{L}_m U, \tag{A.4}$$

where the operator $\mathcal{L}_m = ik_m + \partial_m - \frac{i}{2}\bar{\partial}_m \mu$. The Fredholm condition on equation (A.4)

$$2i \int_\Omega (\mathcal{L}_m U) U^* d\mathbf{r} = 0$$

gives

$$\int_\Omega (ik_m U + \partial_m U) U^* d\mathbf{r} = \frac{i}{2}\bar{\partial}_m \mu. \tag{A.5}$$

Substituting Equation (A.5) to Equation (A.3) yields

$$\frac{\partial A}{\partial Z} + \bar{\nabla}\mu \cdot \tilde{\nabla}A = 0. \tag{A.6}$$

This is the leading order equation for the envelope which is a transport equation. Note $\bar{\nabla}\mu$ plays the role of the group velocity. Because we assumed the nonlinear term is of the next higher order, we have to expand to the next order. An easy way to obtain the correction to Equation (A.6) is to insert a correction into the above equation, i.e.,

$$\frac{\partial A}{\partial Z} + \bar{\nabla}\mu \cdot \tilde{\nabla}A = \nu h_1 + \cdots. \tag{A.7}$$

This treatment can avoid introducing a new scale.

At the order $O(\nu^2)$, the equation is

$$(\mathcal{H}_\mathbf{k} - \mu)\psi_2 = \tilde{\nabla}^2 A U + 2\tilde{\nabla} \cdot (i\mathbf{k} + \nabla)\psi_1 + s(\sigma)|A|^2 A |U|^2 U + i h_1 U.$$

Here we have taken $|\sigma| = \nu^2$ to ensure the maximal balance. Using the Fredholm condition we get

$$\int_\Omega \left(\tilde{\nabla}^2 AU + 2\tilde{\nabla}\cdot(i\mathbf{k}+\nabla)\psi_1 + s(\sigma)|A|^2 A|U|^2 U + ih_1 U\right) U^* d\mathbf{r} = 0. \quad \text{(A.8)}$$

To compute the above integral, we take the derivative over k_n in Equation (A.4) and get

$$(\mathcal{H}_\mathbf{k} - \mu)\bar{\partial}_{m,n} U = 2i\mathcal{L}_n \bar{\partial}_m U + 2i\mathcal{L}_m \bar{\partial}_n U + 2\delta_{mn} U + \bar{\partial}_{m,n}\mu U.$$

The Fredholm condition ensures that

$$2i\int_\Omega (\mathcal{L}_n \bar{\partial}_m U + \mathcal{L}_m \bar{\partial}_n U) U^* d\mathbf{r} = 2\delta_{mn} - \bar{\partial}_{m,n}\mu. \quad \text{(A.9)}$$

Note that $\bar{\partial}_m U$ satisfies Equation (A.4). Thus $\bar{\partial}_m U = -2i(\mathcal{H}_\mathbf{k} + \mu)^{-1}\mathcal{L}_m U + MU$, where M is a function independent of $\mathbf{r}$. Here the operator $(\mathcal{H}_\mathbf{k} + \mu)^{-1}$ exists because $\mathcal{L}_m U$ is orthogonal to the nullspace of $\mathcal{H}_\mathbf{k} + \mu$. MU is the homogeneous solution and does not make any contribution to the integrals in Equation (A.9). We can make it zero for simplicity. It is also noted that both $\mathcal{L}_m$ and $(\mathcal{H}_\mathbf{k} + \mu)^{-1}$ are Hermitian operators, then we get

$$\begin{aligned}\int_\Omega (\mathcal{L}_n \bar{\partial}_m U) U^* d\mathbf{r} &= -2i\int_\Omega (\mathcal{L}_n (\mathcal{H}_\mathbf{k} + \mu)^{-1}\mathcal{L}_m U) U^* d\mathbf{r} \\ &= -2i\int_\Omega (\mathcal{L}_m (\mathcal{H}_\mathbf{k} + \mu)^{-1}\mathcal{L}_n U) U^* d\mathbf{r} = \int_\Omega (\mathcal{L}_m \bar{\partial}_n U) U^* d\mathbf{r}.\end{aligned}$$

Thus we have

$$2i\int_\Omega \left(\mathcal{L}_n \bar{\partial}_m U\right) U^* d\mathbf{r} = \delta_{mn} - \frac{1}{2}\bar{\partial}_{m,n}\mu. \quad \text{(A.10)}$$

Substituting Equation (A.6) into Equation (A.2) yields

$$(\mathcal{H}_\mathbf{k} - \mu)\psi_1 = 2\tilde{\nabla} A \cdot \left(i\mathbf{k} + \nabla - \frac{i}{2}\bar{\nabla}\mu\right) U. \quad \text{(A.11)}$$

Comparing Equations (A.11) and (A.4), we find that ψ_1 and $-i\tilde{\nabla} A \cdot \bar{\nabla} U$ satisfies the same equation, thus

$$\psi_1 = -i\tilde{\nabla} A \cdot \bar{\nabla} U + dU. \quad \text{(A.12)}$$

Here dU is the homogeneous solution and it can be seen that it does not enter the final result, thus we take it to be zero for simplicity.

From (A.10) and (A.12), we get

$$\begin{aligned}
&\int_\Omega \left(2\tilde{\nabla}\cdot(i\mathbf{k}+\nabla-\frac{i}{2}\bar{\nabla}\mu)\psi_1\right)U^* d\mathbf{r}\\
&\quad=\int_\Omega \left(2\tilde{\nabla}\cdot(i\mathbf{k}+\nabla-\frac{i}{2}\bar{\nabla}\mu)(-i\tilde{\nabla}A\cdot\bar{\nabla}U)\right)U^* d\mathbf{r}\\
&\quad=-2i\sum_{m,n}\tilde{\partial}_{m,n}A\int_\Omega(\mathcal{L}_m\bar{\partial}_n U)U^* d\mathbf{r}\\
&\quad=-\sum_{m,n}\tilde{\partial}_{m,n}A\left(\delta_{mn}-\frac{1}{2}\bar{\partial}_{m,n}\mu\right).
\end{aligned}\tag{A.13}$$

Substituting Equation (A.13) into Equation (A.8), we obtain the equation for the second-order $O(\nu^2)$:

$$ih_1+\frac{1}{2}\sum_{m,n}\bar{\partial}_{m,n}\mu\tilde{\partial}_{m,n}A+s(\sigma)\gamma|A|^2A=0,$$

where $\gamma=\int_\Omega|U|^4d\mathbf{r}$.

Then we have the final equation with the first two orders is

$$i\frac{\partial A}{\partial Z}+i\bar{\nabla}\mu\cdot\tilde{\nabla}A+\nu\left(\frac{1}{2}\sum_{m,n}\bar{\partial}_{m,n}\mu\tilde{\partial}_{m,n}A+s(\sigma)\gamma|A|^2A\right)=0.$$

To observe the higher order phenomenon, one need longer time to see it. By introducing a moving coordinate system and rescale the space coordinate $\mathbf{R}$ which is similar to the previous section, the final equation is

$$i\frac{\partial A}{\partial Z}+\frac{1}{2}\sum_{m,n}\bar{\partial}_{m,n}\mu\tilde{\partial}_{m,n}A+s(\sigma)\gamma|A|^2A=0.$$

This is an NLS equation for the envelope dynamics. We note that this equation has the *same* form as the general NLS equation that is derived in homogenous media with dispersion relation $\mu=\mu(\mathbf{k})$ [32]. It is noted that a similar derivation can be found in Ref. [33].

Next we will discuss the case that the eigenspace corresponding to μ is 2-D. Suppose $\mu(\mathbf{k})$ at some values of $\mathbf{k}$ is degenerate. We assume the degeneracy is due to the intersection of two different dispersion surfaces $\mu_j(\mathbf{k})$, $j=1,2$, and we also assume each dispersion surface $\mu_j(\mathbf{k})$ is sufficiently smooth. Then the envelope has the form

$$\psi(\mathbf{r},Z)=\left(A(\mathbf{R})U_1(\mathbf{r};\mathbf{k})+B(\mathbf{R})U_2(\mathbf{r};\mathbf{k})+\nu\psi_1+\nu^2\psi_2+\cdots\right)e^{i\mathbf{k}\cdot\mathbf{r}-i\mu z}.$$

Here $U_j(\mathbf{r};\mathbf{k})$ belongs to the dispersion branch $\mu_j(\mathbf{k})$, $j = 1, 2$ and we require that the two eigenmodes are orthogonal which means

$$\int_\Omega U_1 U_2^* d\mathbf{r} = 0. \tag{A.14}$$

At order $O(\nu)$, the equation for ψ_1 is

$$(\mathcal{H}_\mathbf{k} + \mu)\psi_1 = -i\frac{\partial A}{\partial Z}U_1 - 2\tilde{\nabla}A \cdot (i\mathbf{k} + \nabla)U_1 - i\frac{\partial B}{\partial Z}U_2 - 2\tilde{\nabla}B \cdot (i\mathbf{k} + \nabla)U_2. \tag{A.15}$$

Because $\mu = \mu_1 = \mu_2$ at the $\mathbf{k}$ value we are studying, we have two Fredholm conditions; i.e., calling the RHS of the above equation F then the conditions are $\int_\Omega F U_j^* d\mathbf{r} = 0$, $j = 1, 2$. Due to the orthogonality of U_1 and U_2, we obtain two equations for A and B which are decoupled

$$\frac{\partial A}{\partial Z} + \bar{\nabla}\mu_1 \cdot \tilde{\nabla}A = 0; \tag{A.16a}$$

$$\frac{\partial B}{\partial Z} + \bar{\nabla}\mu_2 \cdot \tilde{\nabla}B = 0. \tag{A.16b}$$

At order $O(\nu^2)$, the equation for ψ_2 is

$$\begin{aligned}(\mathcal{H}_\mathbf{k} + \mu)\psi_2 \\ = -\tilde{\nabla}^2 A u_1 - \tilde{\nabla}^2 B u_2 - 2\tilde{\nabla} \cdot (i\mathbf{k} + \nabla)\psi_1 - s(\sigma)\Phi - ih_1 U_1 - ih_2 U_2,\end{aligned} \tag{A.17}$$

where Φ is the nonlinear term which is of the form

$$\begin{aligned}\Phi = {} & |A|^2 A|U_1|^2 U_1 + B^2 A^* U_2^2 U_1^* + 2|B|^2 A|U_2|^2 U_1 + A^2 B^* U_1^2 U_2^* \\ & + 2|A|^2 B|U_1|^2 U_2 + |B|^2 B|U_2|^2 U_2;\end{aligned}$$

and h_1, h_2 are two corrections of the first-order equations.

Similar to the one component case, we only need to be concerned with the terms corresponding to ∇U_j, $j = 1, 2$. Substituting Equation (A.16) into Equation (A.15), we obtain

$$\begin{aligned}(\mathcal{H}_\mathbf{k} + \mu)\psi_1 = {} & -2\tilde{\nabla}A \cdot \left(i\mathbf{k} + \nabla - \frac{i}{2}\bar{\nabla}\mu_1\right) U_1 \\ & - 2\tilde{\nabla}A \cdot \left(i\mathbf{k} + \nabla - \frac{i}{2}\bar{\nabla}\mu_2\right) U_2.\end{aligned} \tag{A.18}$$

It can be checked that

$$\psi_1 = -i\tilde{\nabla}A \cdot \bar{\nabla}U_1 - i\tilde{\nabla}B \cdot \bar{\nabla}U_2 + d_1 U_1 + d_2 U_2.$$

The coefficients of the homogeneous solution d_1 and d_2 can be treated as zero for simplicity.

Applying Fredholm condition on Equation (A.17) yields

$$\begin{aligned}
&\int_\Omega \left(2\tilde{\nabla}\cdot(i\mathbf{k}+\nabla)\psi_1\right) U_j^* d\mathbf{r} \\
&= \int_\Omega \left(2\tilde{\nabla}\cdot(i\mathbf{k}+\nabla\mu)(-i\tilde{\nabla}A\cdot\bar{\nabla}U_1 - i\tilde{\nabla}B\cdot\bar{\nabla}U_2)\right) U_j^* d\mathbf{r} \\
&= -2i\delta_{1j}\sum_{m,n}\tilde{\partial}_{m,n}A\int_\Omega(\mathcal{L}_m\bar{\partial}_n U_1)U_j^* d\mathbf{r} - 2i\delta_{2j}\sum_{m,n}\tilde{\partial}_{m,n}A\int_\Omega(\mathcal{L}_m\bar{\partial}_n U_2)U_j^* d\mathbf{r} \\
&= -\delta_{1j}\sum_{m,n}\tilde{\partial}_{m,n}A\left(\delta_{mn}-\frac{1}{2}\bar{\partial}_{m,n}\mu\right) - \delta_{2j}\sum_{m,n}\tilde{\partial}_{m,n}B\left(\delta_{mn}-\frac{1}{2}\bar{\partial}_{m,n}\mu\right),
\end{aligned}$$

where we have used the identities

$$2i\int_\Omega(\mathcal{L}_m\bar{\partial}_n U_j)U_j^* d\mathbf{r} = 2i\int_\Omega(\mathcal{L}_n\bar{\partial}_m U_j)U_j^* d\mathbf{r} = \delta_{mn}-\frac{1}{2}\mu_j, \quad \text{(A.19)}$$

and

$$2i\int_\Omega(\mathcal{L}_m\bar{\partial}_n U_j)U_{3-j}^* d\mathbf{r} = 2i\int_\Omega(\mathcal{L}_n\bar{\partial}_m U_j)U_{3-j}^* d\mathbf{r} = 0, \quad \text{(A.20)}$$

which can be readily obtained by employing the two Fredholm condition on the equation

$$(\mathcal{H}_\mathbf{k}+\mu)\bar{\partial}_{m,n}U_j = -2i\mathcal{L}_n\bar{\partial}_m U_j - 2i\mathcal{L}_m\bar{\partial}_n U_j + 2\delta_{mn}U_j - \bar{\partial}_{m,n}\mu_j U_j.$$

Then the next order equations are (we finally get the evolution equations of the first two orders):

$$\begin{aligned}
&i\frac{\partial A}{\partial Z} + i\bar{\nabla}\mu_1\cdot\tilde{\nabla}A + \nu\left(\frac{1}{2}\sum_{m,n}\bar{\partial}_{m,n}\mu_1\tilde{\partial}_{m,n}A + s(\sigma)\Phi_1\right) = 0, \\
&i\frac{\partial B}{\partial Z} + i\bar{\nabla}\mu_2\cdot\tilde{\nabla}B + \nu\left(\frac{1}{2}\sum_{m,n}\bar{\partial}_{m,n}\mu_2\tilde{\partial}_{m,n}B + s(\sigma)\Phi_2\right) = 0,
\end{aligned}$$

where the nonlinear terms are

$$\begin{aligned}
\Phi_1 &= \gamma_1^{(1)}|A|^2A + \gamma_2^{(1)}B^2A^* + 2\gamma_3^{(1)}|B|^2A + \gamma_4^{(1)}A^2B^* + 2\gamma_5^{(1)}|A|^2B + \gamma_6^{(1)}|B|^2B); \\
\Phi_2 &= \gamma_1^{(2)}|B|^2B + \gamma_2^{(2)}A^2B^* + 2\gamma_3^{(2)}|A|^2B + \gamma_4^{(1)}B^2A^* + 2\gamma_5^{(2)}|B|^2A + \gamma_6^{(2)}|A|^2A);
\end{aligned}$$

and

$$\begin{aligned}
&\gamma_1^{(j)} = \int_\Omega |U_j|^4 d\mathbf{r}; \quad \gamma_2^{(j)} = \int_\Omega U_j^{*2}U_{3-j}^2 d\mathbf{r}; \quad \gamma_3^{(j)} = \int_\Omega |U_j|^2|U_{3-j}|^2 d\mathbf{r}; \\
&\gamma_4^{(j)} = \int_\Omega |U_j|^2U_jU_{3-j}^* d\mathbf{r};\ \gamma_5^{(j)} = \int_\Omega |u_j|^2U_j^*U_{3-j} d\mathbf{r};\ \gamma_6^{(j)} = \int_\Omega |U_{3-j}|^2U_{3-j}U_j^* d\mathbf{r}.
\end{aligned}$$

In the above nonlinear terms, $\gamma_1^{(j)}$ and $\gamma_3^{(j)}$ are real. However, the reality of other parameters can not be obtained easily. If $U_1(\mathbf{r};\mathbf{k})$ and $U_2(\mathbf{r};\mathbf{k})$ have the same phase, for example, at Γ, M, X_1, and X_2 points U_1 and U_2 have the same phase $e^{-i\mathbf{k}\cdot\mathbf{r}}$, then the parameters $\gamma_m^{(j)},\ j=1,2; m=1,2,\ldots,6$ are all real.

It is noted that if the orbitals functions are highly localized, the above parameters can be related to the nonlinear coefficients in Equation (25). For example,

$$\gamma_1^{(1)} = \int_\Omega |u^{(1)}|^2 d\mathbf{r} = \int_\Omega \left| \sum_{\mathbf{v}} \phi_1(\mathbf{r}-\mathbf{v}) e^{i\mathbf{k}\cdot\mathbf{v}} \right|^4 d\mathbf{r} \approx \int \phi_1^4(\mathbf{r}) d\mathbf{r} = g_1.$$

Similarly, we can get

$$\begin{aligned}
&\gamma_2^{(1)} \approx \gamma_2^{(2)} \approx \gamma_3^{(1)} \approx \gamma_3^{(2)} \approx \gamma_4^{(2)} \approx g_3;\\
&\gamma_4^{(1)} \approx \gamma_6^{(2)} \approx \gamma_5^{(1)} \approx g_4;\\
&\gamma_4^{(2)} \approx \gamma_6^{(1)} \approx \gamma_5^{(2)} \approx g_5.
\end{aligned}$$

Thus in this limit, $\Phi_j \approx \overline{\Psi}_j,\ j=1,2$, and the two approaches reduce to the same equation.

Similar to the previous section, if $\bar{\nabla}\mu_1 = \bar{\nabla}\mu_2 = \mathbf{0}$, i.e., μ_1 and μ_2 reach the same extreme value, the above system becomes (after rescalling the distance $\bar{Z} = \nu Z$ and dropping the bar below)

$$\begin{aligned}
&i\frac{\partial A}{\partial Z} + \frac{1}{2}\sum_{m,n} \bar{\partial}_{m,n}\mu_1\tilde{\partial}_{m,n}A + s(\sigma)\Phi_1 = 0,\\
&i\frac{\partial B}{\partial Z} + \frac{1}{2}\sum_{m,n} \bar{\partial}_{m,n}\mu_2\tilde{\partial}_{m,n}B + s(\sigma)\Phi_2 = 0.
\end{aligned}$$

We can also obtain the detuning phenomenon if $|\bar{\nabla}\mu_j^{(2)}(\mathbf{k}) - \bar{\nabla}\mu_j^{(1)}(\mathbf{k})| \sim O(\nu)$. The system becomes

$$\begin{aligned}
&i\frac{\partial A}{\partial Z} + \frac{1}{2}\sum_{m,n} \bar{\partial}_{m,n}\mu_1\tilde{\partial}_{m,n}A + s(\sigma)\Phi_1 = 0,\\
&i\frac{\partial B}{\partial Z} + i\mathbf{g}\cdot\tilde{\nabla}b + \frac{1}{2}\sum_{m,n} \bar{\partial}_{m,n}\mu_2\tilde{\partial}_{m,n}B + s(\sigma)\Phi_2 = 0,
\end{aligned}$$

where $\mathbf{g} = \frac{1}{\nu}(\bar{\nabla}\mu_2 - \bar{\nabla}\mu_1))$.

References

1. D. Christodoulides and R. Joseph, Discrete self-focusing in nonlinear arrays of coupled wave-guides, *Opt. Lett.* 13: 794–796 (1988).
2. H. Eisenberg, Y. Silberberg, R. Morandotti, A. Boyd, and J. Aitchison, Discrete spatial optical solitons in waveguide arrays, *Phys. Rev. Lett.* 81: 3383–3386 (1998).
3. J. W. Fleischer, T. Carmon, M. Segev, N. K. Efremidis, and D. N. Christodoulides, Observation of discrete solitons in optically induced real time waveguide arrays, *Phys. Rev. Lett.* 90:023902 (2003).
4. J. W. Fleischer, M. Segev, N. K. Efremidis, and D. N. Christodoulides, Observation of two-dimensional discrete solitons in optically induced nonlinear photonic lattices, *Nature* 422: 147–150 (2003).
5. J. K. Yang, I. Makasyuk, A. Bezryadina, and Z. Chen, Dipole solitons in optically induced two-dimensional photonic lattices, *Opt. Lett.* 29: 1662–1664 (2004).
6. D. N. Neshev, T. J. Alexander, E. A. Ostrovskaya, Y. S. Kivshar, I. Martin, H. Makasyuk, and Z. G. Chen, Observation of discrete vortex solitons in optically induced photonic lattices, *Phys. Rev. Lett.* 92: 123903 (2004).
7. X. Wang, Z. Chen, J. Wang, and J. Yang, Observation of in-band lattice solitons, *Phys. Rev. Lett.* 99: 243901 (2007).
8. O. Morsch and M. Oberthaler, Dynamics of Bose-Einstein condensates in optical lattices, *Rev. Mod. Phys.* 1: 179–215 (2006).
9. A. Aftalion and B. Helffer, On mathematical models for Bose-Einstein condensates in opticallattices, *Rev. Math. Phys.* 21: 229–278 (2009).
10. P. Kevrekidis, K. Rasmussen, and A. Bishop, The discrete nonlinear Schrödinger equation: A survey of recent results, *Int. J. Mod. Phys. B* 15: 2833–2900 (2001).
11. G. L. Alfimov, P. G. Kevrekidis, V. V. Konotop, and M. Salerno, Wannier functions analysis of the nonlinear Schrödinger equation with a periodic potential, *Phys. Rev. E* 66: 046608 (2002).
12. M. J. Ablowitz and Y. Zhu, Evolution of Bloch-mode envelopes in two-dimensional generalized honeycomb lattices, *Phys. Rev. A* 82: 013840 (2010).
13. M. J. Ablowitz and Y. Zhu, Nonlinear waves in shallow honeycomb lattices, *SIAM J. Appl. Math.* 72: 240–260 (2012).
14. O. Peleg, G. Bartal, B. Freedman, O. Manela, M. Segev, and D. N. Christodoulides, Conical diffraction and gap solitons in honeycomb photonic lattices, *Phys. Rev. Lett.* 98: 103901 (2007).
15. M. Ablowitz, S. Nixon, and Y. Zhu, Conical diffraction in honeycomb lattices, *Phys. Rev. A* 79: 053830 (2009).
16. M. J. Ablowitz and Y. Zhu, Nonlinear diffraction in photonic graphene, *Opt. Lett.* 36: 3762–3764 (2011).
17. P. G. Kevrekidis, B. A. Malomed, and Y. B. Gaididei, Solitons in triangular and honeycomb dynamical lattices with the cubic nonlinearity, *Phys. Rev. E* 66: 016609 (2002).
18. A. K. Geim and K. S. Novoselov, The rise of graphene, *Nat. Mater.* 6: 1476–1122 (2007).
19. L. H. Haddad and L. C. Carr, The nonlinear dirac equation in Bose-Einstein condensates: Foundation and symmetries, *Physica D* 238: 1413–1421 (2009).
20. D. Pelinovsky and G. Schneider, Bounds on the tight-binding approximation for the Gross-Pitaevskii equation with a periodic potential, *J. Diff. Eqs.* 248: 837–849 (2010).
21. D. Pelinovsky, G. Schneider, and R. S. MacKay, Justification of the lattice equation for a nonlinear elliptic problem with a periodic potential, *Comm. Math. Phys.* 284: 803–831 (2008).

22. B. Ilan and M. I. Weinstein, Band-edge solitons, nonlinear Schrödinger/Gross-Pitaevskii equations and effective media, *SIAM Mult. Model. Simulation* 8: 1055–1101 (2010).
23. Z. Shi and J. Yang, Solitary waves bifurcated from Bloch-band edges in two-dimensional periodic media, *Phys. Rev. E* 75: 056602 (2007).
24. D. Dohnal and H. Uecker, Coupled-mode equations and gap solitons for the 2D Gross-Pitaevskii equation with a non-separable periodic potential, *Physica D* 238: 860–879 (2009).
25. G. H. Wannier, The structure of electronic excitation levels in insulating crystals, *Phys. Rev.* 52: 191–197 (1937).
26. J. Callaway, *Quantum Theory of the Solid State* (2nd ed.), Academic Press, San Diego, 1991.
27. M. J. Ablowitz, C. W. Curtis, and Y. Zhu, On tight binding approximations in optical lattices, *Stud. Appl. Math.*, to appear.
28. L. I. Schiff, *Quantum Mechanics* (3rd ed.), McGraw-Hill, New York, 1968.
29. F. Odeh and J. B. Keller, Partial differential equations with periodic coefficients and Bloch waves in crystals, *J. Math. Phys.* 5: 1499–1503 (1964).
30. P. G. Kevrekidis, *The Discrete Nonlinear Schrödinger Equation: Mathematical Analysis, Numerical Computations and Physical Perspectives*, Springer-Verlag, Berlin/ Heidelberg 2009.
31. C. R. Rosberg, D. N. Neshev, A. A. Sukhorukov, W. Krolikowski, and Y. S. Kivshar, Observation of nonlinear self-trapping in triangular photonic lattices, *Opt. Lett.* 32: 397–399 (2007).
32. B. D.J. and N. A.C., Propagation of nonlinear wave envelopes, *J. Math. Phys. (Stud. Appl. Math.)* 46: 133–139 (1967).
33. C. Sparber, Effective mass theorems for nonlinear Schrödinger equations, *SIAM J. Appl. Math.* 66: 820–842 (2006).

University of Colorado
Tsinghua University

(Received August 28, 2012)