

BAND-EDGE SOLITONS, NONLINEAR SCHRÖDINGER/GROSS–PITAEVSKII EQUATIONS, AND EFFECTIVE MEDIA*

B. ILAN[†] AND M. I. WEINSTEIN[‡]

Abstract. We consider a class of nonlinear Schrödinger/Gross–Pitaevskii (NLS/GP) equations with periodic potentials having an even symmetry. We construct “solitons,” centered about any point of symmetry of the potential. For focusing (attractive) nonlinearities, these solutions bifurcate from the zero state at the lowest band-edge frequency into the semi-infinite spectral gap. Our results extend to bifurcations into finite spectral gaps, for focusing or defocusing (repulsive) nonlinearities under more restrictive hypotheses.

Soliton nonlinear bound states with frequencies near a band edge are well approximated by a slowly decaying solution of a *homogenized NLS/GP equation*, with constant homogenized effective mass tensor and effective nonlinear coupling coefficient, modulated by a Bloch state.

For the critical NLS equation with a periodic potential, e.g., the cubic two-dimensional NLS/GP with a periodic potential, our results imply the following: (1) The limiting soliton squared L^2 norm, as the spectral band-edge frequency is approached, is equal to $\mathcal{P}_{edge} = \zeta_* \times \mathcal{P}_{cr}$, where \mathcal{P}_{cr} denotes the minimal mass soliton of the translation invariant critical NLS. \mathcal{P}_{cr} is also known as the Townes critical power for self-focusing of optical beams. (2) The constant ζ_* is expressible in terms of the band-edge Bloch eigenfunction and the determinant of the *effective mass* tensor. If the potential is nonconstant, then $0 < \zeta_* < 1$ and \mathcal{P}_{edge} is strictly less than \mathcal{P}_{cr} . The results are confirmed by numerical computation of bound states with frequencies near the spectral band edge.

Finally, these results have implications for the control of nonlinear waves using periodic structures.

Key words. multiple scales, Lyapunov–Schmidt reduction, nonlinear optics, Bose–Einstein condensates

AMS subject classifications. 35B27, 35B32, 35B35, 35B40

DOI. 10.1137/090769417

1. Introduction and outline. “Solitons” are spatially localized concentrations of energy, which are of great interest in many nonlinear wave systems. They arise from a balance of dispersion (or diffraction), which tends to spread energy and (focusing/attractive) nonlinearity which tends to concentrate energy. Although their importance was first recognized in the context of hydrodynamics [66, 2], soliton-like coherent structures are now understood to play a central role in contexts ranging from optical pulses (temporal solitons) to stationary beams (spatial solitons) of nonlinear optics [11, 40] to soliton matter waves in macroscopic quantum systems [46]. Advances in the design of micro- or nano-structured media have greatly enabled the control of optical and matter waves. Thus, it is of interest to develop a fundamental understanding of the effect of inhomogeneities in a medium on the dynamics of nonlinear dispersive waves and, in particular, on the dynamics of solitons. See, for example, [24] for an experimental investigation of solitons in periodic structures.

*Received by the editors August 30, 2009; accepted for publication (in revised form) February 16, 2010; published electronically April 30, 2010.

<http://www.siam.org/journals/mms/8-4/76941.html>

[†]School of Natural Sciences, University of California, Merced, CA (bilan@ucmerced.edu).

[‡]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY (miw2103@columbia.edu).

In this article we consider solitons in nonhomogeneous media governed by a class of nonlinear Schrödinger/Gross–Pitaevskii (NLS/GP) equations:

$$(1.1) \quad i\partial_t \psi = -\Delta \psi + V(\mathbf{x})\psi - |\psi|^{2\sigma} \psi .$$

Here $\psi = \psi(\mathbf{x}, t)$ denotes a complex-valued function of $(\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}^1$, $d \geq 1$. The potential $V(\mathbf{x})$ is assumed to be real-valued, smooth, periodic, and symmetric about one or more points.^{1,2}

NLS/GP is a Hamiltonian system, expressible as follows:

$$(1.2) \quad i\partial_t \psi = \frac{\delta \mathcal{H}}{\delta \psi^*}$$

$$\mathcal{H}[\psi, \psi^*] = \int \nabla \psi \cdot \nabla \psi^* + V(\mathbf{x})\psi\psi^* - \frac{1}{\sigma+1} \psi^{\sigma+1} (\psi^*)^{\sigma+1},$$

where ψ^* denotes the complex conjugate of ψ . By Noether’s theorem, the invariance $t \mapsto t + t_0$ implies the time-invariance of \mathcal{H} for solutions of NLS/GP. Furthermore, the invariance $\psi \mapsto e^{i\theta} \psi$ implies the additional time-invariant quantity

$$(1.3) \quad \mathcal{P}[\psi, \psi^*] = \int \psi \psi^* d\mathbf{x} .$$

The parameter $\sigma > 0$ allows for variation of the strength of the nonlinearity. In physical systems, we typically have $\sigma = 1$. Allowing σ to vary enables one to quantify the balance between nonlinear effects and dispersive/diffractive effects, which depend on spatial dimensionality, d . Local well posedness in time for the initial value problem for (1.1) with data

$$(1.4) \quad \psi(\mathbf{x}, t = 0) = \psi_0(\mathbf{x})$$

in $\psi_0 \in H^1(\mathbb{R}^d)$ (see, for example, [16, 59]) holds for all $\sigma > 0$ for $d = 1, 2$ and all $0 < \sigma < 2(d-2)^{-1}$ for $d \geq 3$. Global well posedness for arbitrary data holds for $\sigma < 2/d$. For well posedness for data in spaces of weaker regularity, see [10, 60].

The NLS/GP equation, (1.1), with $\sigma = 1$ governs the dynamics of a macroscopic quantum state, a Bose–Einstein condensate, comprised of a large collection of interacting bosons in the mean-field limit [46, 25, 15]. The attractive nonlinear potential, $-|\psi|^2$, corresponds to a species of bosons, whose two-particle interactions have a negative scattering length. A second important area of application of NLS/GP is its description of the evolution of the slowly varying envelope of a stationary and nearly monochromatic laser beam propagating through a nonlinear medium [40, 11]. Here the attractive nonlinear potential is due to the Kerr nonlinear effect; regions of higher electric field intensity have a higher refractive index. In this setting t denotes the distance along the direction of propagation and $\mathbf{x} \in \mathbb{R}^2$, the transverse dimensions. In the quantum mechanical setting the potential, $V(\mathbf{x})$, is determined by magnetic and optical effects that are used to confine a cloud of bosons. In optics, the potential is determined by the spatial variations of the background *linear refractive index* of the medium. The functional \mathcal{P} denotes the optical power or, in the quantum setting, the particle number.

¹Periodic potentials are often called “lattice” potentials.

²The main results of this paper extend to general nonlinearities of the form $f(|\psi|^2)\psi = g [1 + \mathcal{O}(|\psi|^2)] |\psi|^{2\sigma} \psi$.

Nonlinear bound states or *solitons* of NLS/GP are solutions of standing wave type

$$(1.5) \quad \psi(\mathbf{x}, t) = e^{-i\mu t} u(\mathbf{x}, \mu),$$

where μ denotes the frequency (propagation constant in optics, chemical potential in quantum many-body theory) and u is a real-valued solution of

$$(1.6) \quad (-\Delta_{\mathbf{x}} + V(\mathbf{x})) u(\mathbf{x}, \mu) - u^{2\sigma+1}(\mathbf{x}, \mu) = \mu u(\mathbf{x}, \mu), \quad u(\cdot, \mu) \in H^1(\mathbb{R}^d).$$

We shall construct solutions of (1.6) with μ located in a spectral gap of $-\Delta + V$.

The properties of solitons in *homogeneous* media, $V \equiv 0$, are reviewed in detail in section 2. Briefly, for $\sigma < 2/d$ (subcritical nonlinearities) dynamically stable solitons exist at any prescribed L^2 norm (in one-to-one correspondence with any $\mu < 0$). In the critical case and supercritical cases, $\sigma \geq 2/d$, solitons are unstable.

We raise several motivating questions below and outline our results in the next subsection.

- (Q1) *Persistence and stability*: What is the effect of a periodic potential on the existence and stability properties of solitons?
- (Q2) *Stabilization*: Can unstable solitons be stabilized by a potential, $V(\mathbf{x})$? This question was first addressed in [48] in the context of localized potentials and more recently for more general potentials [54], e.g., periodic and quasi-periodic.
- (Q3) *Excitation thresholds/minimal mass solitons*: How does periodic structure effect soliton excitation thresholds? For critical nonlinearity, $\sigma = 2/d$, and $V \equiv 0$, the soliton squared L^2 norm is independent of μ ; $\mathcal{P}[u(\cdot, \mu)] = \mathcal{P}_{cr}$. Thus, *there is an L^2 threshold below which there are no solitons*. This L^2 *excitation threshold* for soliton formation is of great physical interest [65, 26, 19, 41, 9]. In optics it corresponds to the critical *power* for self-focusing [18, 42]. Such solitons are also often called *minimal mass solitons*. See Remark 3.3.

1.1. Outline of results. In order to outline the results of this paper, we begin with very quick review of the spectral theory of Schrödinger operators, $-\Delta + V$, for V periodic [47, 23, 36]. If $V(\mathbf{x})$ is a periodic potential, then the spectrum of $-\Delta + V$ is real, bounded below, tends to positive infinity, is absolutely continuous, and consists of the union of closed intervals (*spectral bands*). The open intervals separating the spectral bands are called *spectral gaps*. One-dimensional Schrödinger operators with periodic potentials generically have infinitely many gaps. In dimensions $d \geq 2$, there are at most finitely many gaps.

We denote by E_* the lowest point in the spectrum, the left endpoint, or the edge of the first spectral band. E_* is simple and is the ground state (lowest) eigenvalue of $-\Delta + V$, subject to periodic boundary conditions on the basic period cell of V . The eigenspace associated with E_* is spanned by $w(\mathbf{x})$, a nontrivial solution of

$$(1.7) \quad (-\Delta + V(\mathbf{x})) w(\mathbf{x}) = E_* w(\mathbf{x}), \quad w(\mathbf{x}) \text{ periodic.}$$

For the case $V \equiv 0$, $E_* = 0$, and we can take $w(\mathbf{x}) \equiv 1$.

The present work considers the bifurcation and dynamic stability properties of families of solitons emerging from a spectral band edge. Such edge-bifurcating solitons have a multiscale character described below. Our results include the following.

1. *Theorem 3.1:* Let \mathbf{x}_0 denote any point of symmetry of $V(\mathbf{x})$.³ There is a family

$$(1.8) \quad \mu \mapsto u(\mathbf{x}, \mu) \approx (E_* - \mu)^{\frac{1}{2\sigma}} F\left(\sqrt{E_* - \mu}(\mathbf{x} - \mathbf{x}_0)\right) w(\mathbf{x}),$$

which bifurcates from the zero solution at energy E_* into the semi-infinite gap $(-\infty, E_*)$ for $0 < E_* - \mu$ sufficiently small. Here $\sigma \in \mathbb{N}$ for $d = 1, 2$ and $\sigma = 1$ for $d = 3$.

$F(\mathbf{y})$ denotes the soliton profile for an effective medium with effective mass tensor, A^{ij} , given by (3.6) and effective nonlinear coupling constant, γ_{eff} , given in (3.7), and it satisfies the following *homogenized soliton equation*:

$$(1.9) \quad - \sum_{i,j=1}^d \partial_{y_i} A^{ij} \partial_{y_j} F(\mathbf{y}) - \gamma_{\text{eff}} F^{2\sigma+1}(\mathbf{y}) = -F(\mathbf{y}),$$

$$F > 0, \quad F \in H^1(\mathbb{R}^d).$$

The leading order expansion is constructed via multiple-scale expansion. The error term is studied by decomposition of the corrector into spectral components near and far from the band edge and is estimated via a Lyapunov–Schmidt strategy; see also [14, 21, 22]. The results can be extended to solitons near edges of *finite* spectral bands for focusing and defocusing nonlinear potentials under more restrictive hypotheses on V ; see subsection 3.1. A variant of Theorem 3.1 holds in dimension one in any spectral gap near a “positive curvature” band edge; see Theorem 3.5.

2. *Corollary 3.1, part 1:* Consider the critical cases: $\sigma = 1, d = 2$ and $\sigma = 2, d = 1$. Near the band edge, i.e., for $0 < E_* - \mu \ll 1$, we have

$$(1.10) \quad \mathcal{P}[u(\cdot, \mu)] = \zeta_* \mathcal{P}_{cr} + (\mu - E_*) \zeta_{1*} + \mathcal{O}((\mu - E_*)^2).$$

Here $\mathcal{P}_{cr}, \zeta_*$, and ζ_{1*} are defined as follows.

- (a) $\mathcal{P}_{cr} = \mathcal{P}[R(\cdot, -1)]$, where $R(\cdot, -1)$ denotes the unique (up to translations) solution of

$$(1.11) \quad \Delta R - R + R^{\frac{4}{d}+1} = 0, \quad R > 0, \quad R \in H^1;$$

see (Q3) above.

- (b)

$$(1.12) \quad \mathcal{P}_{edge} \equiv \lim_{\mu \rightarrow E_*} \mathcal{P}[u(\cdot, \mu)] = \zeta_* \mathcal{P}_{cr}$$

is given by (3.17) and satisfies the inequality $0 < \zeta_* < 1$, unless V is identically constant. In the latter case, $\zeta_* = 1$.

- (c)

$$(1.13) \quad \zeta_{1*} \equiv \left. \frac{d}{d\mu} \right|_{\mu=E_*} \mathcal{P}[u(\cdot, \mu)]$$

is given by (3.18).

³That is, $f(\mathbf{x}) = f(x_1, \dots, x_d)$ is symmetric (about the origin) if $f(x_1, \dots, x_d) = f(s_1 x_1, \dots, s_d x_d)$, $s_j = \pm 1$. \mathbf{x}_0 is a point of symmetry of $V(\mathbf{x})$ if $V(\mathbf{z}) \equiv V(\mathbf{x}_0 + \mathbf{z})$ is symmetric. Thus, by translating coordinates, we can arrange for a point of symmetry to be at the origin.

(d) For periodic potentials V of the form $\delta \tilde{V}(x)$, where δ is sufficiently small, we show that $\zeta_{1*} > 0$; see (3.20).

(e) *Positive slope conjecture:* In general, $\zeta_{1*} > 0$.

Both ζ_* and ζ_{1*} depend on the edge (periodic) Bloch eigenstate and the Hessian matrix (of second partial derivatives) of the band dispersion function, $D^2 E_1$, near E_* . The latter is often called the *inverse effective mass tensor*.

3. *Instability for μ near the band edge:* Consider the critical cases $\sigma = 1, d = 2$ and $\sigma = 2, d = 1$. For V nonzero and μ near E_* , nonlinear bound states are linearly exponentially unstable, provided $\zeta_{1*} > 0$. We conjecture $\zeta_{1*} > 0$, in general, and have verified it for potentials $V = \delta \tilde{V}$, with δ sufficiently small. For $V \equiv 0$, the linear instability is algebraic, although for the nonlinear dynamics, solutions can blow up in finite time or decay to zero dispersively (diffractively) as t tends to infinity. In contrast, consider the case $V \neq 0$. For μ close enough to the spectral band edge, the curve $\mu \mapsto \mathcal{P}[u(\cdot, \mu)]$ lies below the line $\mathcal{P} = \mathcal{P}_{cr}$. Thus, for initial conditions $\psi_0 = u(\cdot, \mu) + \epsilon$ with $|\epsilon| \ll 1$, the solution of (1.1) exists globally in time in H^1 , even though it does not remain close to $u(\cdot, \mu)$ in H^1 .

4. *Section 6:* Numerical computations are used to illustrate the asymptotic results and to study the global behavior.

Figure 1.1 below summarizes a key consequence of our results.

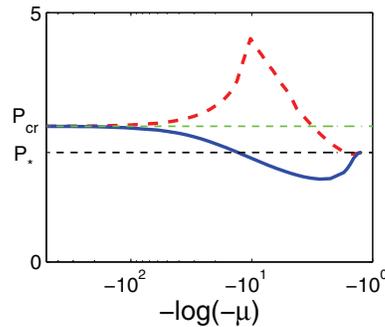


FIG. 1.1. Plot of power curves: $\mathcal{P}[u(\cdot, \mu)]$ vs. μ (using semi-log axis) for the quintic one-dimensional NLS/GP equation, (6.4), with $V_0 = 10$ and $K = 2\pi$ (here $E_* \approx -1.23$). Solid curve corresponds to power curve for soliton family centered at a local minimum. Dashed curve corresponds to centering at a local maximum. Agreement is shown between numerical computations and the analytically obtained value for the band-edge power (dashed lower horizontal line), $\mathcal{P}_{edge} = \lim_{\mu \rightarrow E_*} \mathcal{P}[u(\cdot, \mu)] = \zeta_* \times \mathcal{P}_{cr} \approx 2.2$ (3.15). For μ large and negative, $\mathcal{P}(\mu)$ converges to $\mathcal{P}_{cr} \approx 2.72$ (dashed horizontal line), which is the critical power of the Townes soliton in the translation invariant ($V \equiv \text{const}$) case.

For any nontrivial periodic $V(x)$ the limiting L^2 norm at the band edge is *strictly less* than that of the homogeneous medium. The slope of the curve $\mu \rightarrow \mathcal{P}[u(\cdot, \mu)]$ is strictly positive. As $-\mu = |\mu|$ increases, solitons become increasingly localized in space and thus depend more and more on the local properties of V . The limiting ($|\mu| \rightarrow \infty$) squared L^2 norm is \mathcal{P}_{cr} . The orbital stability theory, outlined in section 2, implies that solitons with energies μ near the band edge (where $\partial_\mu \mathcal{P}[u(\cdot, \mu)] > 0$) are unstable, while those which are centered and sufficiently concentrated ($-\mu$ sufficiently large, $\partial_\mu \mathcal{P}[u(\cdot, \mu)] < 0$) about a local minimum of V are stable. It is natural to conjecture that for localized initial conditions with L^2 norm strictly less than $\inf_{\mu \leq E_*} \mathcal{P}[u(\cdot, \mu)]$,

solutions to the initial value problem disperse to zero as $t \rightarrow \infty$; see the discussion in the proof of part 3 of Theorem 3.3 and [64].

Remark 1.1. Concerning the dependence of $\mu \mapsto \mathcal{P}[u(\cdot, \mu)]$ for μ near the band edge on the nonlinearity parameter, σ , and dimensionality d (see Theorem 3.2 and Corollary 3.1), it is useful to recall the analogous behavior in the translation invariant case: $V \equiv 0$. In this case, NLS is also invariant under dilation

$$(1.14) \quad \psi(\mathbf{x}, t) \mapsto \lambda^{\frac{1}{\sigma}} \psi(\lambda \mathbf{x}, \lambda^2 t) .$$

Let $R(\cdot, \mu)$ denote the positive (unique up to translation) solution of

$$(1.15) \quad -\Delta R - R^{2\sigma+1} = \mu R .$$

By uniqueness,

$$(1.16) \quad R(\mathbf{x}, \mu) = |\mu|^{\frac{1}{2\sigma}} R(|\mu|^{\frac{1}{2}} \mathbf{x}, -1) .$$

It follows that

$$(1.17) \quad \mathcal{P}[R(\cdot, \mu)] = \|R(\cdot, \mu)\|_2^2 = |\mu|^{\frac{1}{\sigma} - \frac{d}{2}} \|R(\cdot, -1)\|_2^2 ,$$

implying that as $\mu \rightarrow E_*$, $\mathcal{P}[R(\cdot, \mu)]$ tends

- to 0 for $\sigma < 2/d$,
- to $\|R(\cdot, -1)\|_2^2$ for $\sigma = 2/d$, and
- to $+\infty$ for $\sigma > 2/d$;

see Figure 1.2. In one space dimension, the family of solitons is given explicitly by

$$(1.18) \quad R(x, \mu) = [(\sigma + 1) |\mu|]^{\frac{1}{2\sigma}} \operatorname{sech}^{\frac{1}{\sigma}} \left(\sigma \sqrt{|\mu|} x \right) .$$

In the critical case, $\sigma = 2$,

$$(1.19) \quad \mathcal{P}[R(\cdot, \mu)] = \frac{\sqrt{3}}{2} \int_{\mathbb{R}} \operatorname{sech}(y) dy = \frac{\sqrt{3}}{2} \pi \sim 2.7207 ;$$

see Figure 1.2. Theorem 3.2 implies a similar trichotomy of behaviors for states bifurcating from the band edge, E_* , of a periodic potential. Also, for $\sigma = 2/d$, the curves $\mu \mapsto \mathcal{P}[u(\cdot, \mu)]$ in Figure 6.3 are seen to be deformations (for minimum and maximum centered solitons) of the horizontal line $\mu \mapsto \mathcal{P}_{cr}$ for the case $V \equiv 0$.

Previous work. Formal expansions and numerical approximation of nonlinear bound states near spectral band edges for periodic and aperiodic structures and their linearized stability properties were presented in [57, 7, 67, 45, 12, 8, 49, 50, 17, 1, 13, 5, 52, 56]. The band-edge limit of \mathcal{P} , for the case of a two-dimensional separable potential was obtained by formal perturbation theory and numerically in [51]. Two-scale convergence methods have been applied to rigorously derive homogenized effective equations, valid on large but finite time scales, in [6] for the linear Schrödinger equation and in [55] for the time-dependent NLS/GP, with two-scale-type initial conditions. Bifurcation of localized states from the continuous spectrum into spectral gaps has been considered in [37, 38, 30, 29, 4, 58, 43]. The connection with nonlinear coupled mode equations is explored in [14, 44, 21, 22]. The Lyapunov–Schmidt strategy applied herein is motivated by these latter approaches.

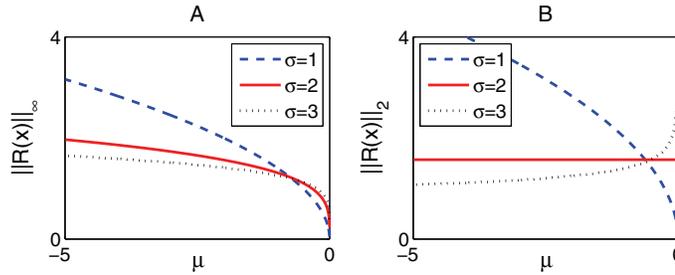


FIG. 1.2. (A) L^∞ norm and (B) squared L^2 norm (\mathcal{P}) as functions of frequency, μ , for the ground state solution of (1.6) in one dimension ($d = 1$) with three nonlinear exponents: subcritical ($\sigma d = 1 < 2$), critical ($\sigma d = 2$), and supercritical ($\sigma d = 3 > 2$); see legend. In L^∞ , bifurcation appears from a state with zero norm at $E_* = 0$. In L^2 , the limiting behavior as $\mu \rightarrow 0^-$ depends on σd .

Outline. The paper is structured as follows. In section 2 we discuss background for the formulation of our results. We state our main results in section 3. In section 4 a formal homogenization/two-scale expansion of solitons with frequencies near the band edge is derived. In subsection 4.3 we derive the consequences of our expansion of band-edge solitons for the character of $\mathcal{P}[u(\cdot, \mu)]$ as $\mu \rightarrow E_*$. The expansion and error estimates are proved in section 5. Section 6 contains a discussion of numerical simulations illustrating our main theorems. Section 7 contains a short summary and discussion. The latter sections of the paper are appendices containing technical results on the effective mass tensor.

1.2. Notation.

1. $\eta(\epsilon) = \mathcal{O}(\epsilon^\infty)$ if $\eta(\epsilon) = \mathcal{O}(\epsilon^q)$ for all $q \geq 1$
2. Fourier transform of G : $\hat{G}(\mathbf{k}) = \int e^{-2\pi i \mathbf{k} \cdot \mathbf{x}} G(\mathbf{x}) d\mathbf{x}$
3. $\chi(a \leq |\mathbf{k}| \leq b) =$ characteristic function of the set $\{\mathbf{k} : a \leq |\mathbf{k}| \leq b\}$
4. $\chi(|\nabla_{\mathbf{y}}| \leq a) G = \int e^{2\pi i \mathbf{k} \cdot \mathbf{y}} \chi(|\mathbf{k}| \leq a) \hat{G}(\mathbf{k}) d\mathbf{k}$
5. H^s , Sobolev space of order s ; H_{even}^s , space of even H^s functions

$$(1.20) \quad \|f\|_{H^s}^2 = \sum_{|\alpha| \leq s} \|\partial^\alpha f\|_{L^2}^2 \sim \|\hat{f}\|_{L^{2,s}}^2$$

6. H_{sym}^s , symmetric H^s functions, i.e., $f \in H_{sym}^s$ if $f \in H^s$ and $f(x_1, \dots, x_d) = f(s_1 x_1, \dots, s_d x_d)$, $s_j = \pm 1$
7. $\|f\|_{L^{2,s}(D)}^2 = \int_D |f(z)|^2 (1 + |z|^2)^s dz$
8. $C_\downarrow^m(\mathbb{R}^d)$, functions in $C^m(\mathbb{R}^d)$ with limit equal to zero as $|\mathbf{x}| \rightarrow \infty$
9. \mathcal{B} , the fundamental period cell; \mathcal{B}^* , the dual fundamental cell or the Brillouin zone.

2. Background.

2.1. Solitons and stability theory. We give a very brief review of the stability theory of solitons of NLS/GP (1.1).

DEFINITION 2.1. *The nonlinear bound state $u(\mathbf{x}, \mu)$ of NLS/GP is orbitally stable if for all $\epsilon > 0$, there is a $\delta > 0$ such that if the initial condition ψ_0 satisfies*

$$(2.1) \quad \inf_{\gamma \in [0, 2\pi)} \|\psi_0 - u(\cdot, \mu) e^{i\gamma}\|_{H^1} < \delta,$$

then the corresponding solution, $\psi(\cdot, t)$, satisfies

$$(2.2) \quad \inf_{\gamma \in [0, 2\pi)} \|\psi(\cdot, t) - u(\cdot, \mu)e^{i\gamma}\|_{H^1} < \varepsilon \text{ for all } t \neq 0.$$

This notion of soliton stability for NLS is natural since NLS/GP for V nonconstant is invariant under the group of phase translations, $\psi \mapsto e^{i\theta}\psi$, but not spatial translations.

A central role in the stability theory is played by the operator

$$(2.3) \quad L_+ \equiv -\Delta - \mu + V - (2\sigma + 1)u^{2\sigma},$$

which is the real part of the linearization of NLS/GP about $u(\cdot, \mu)$. Let $n_-(L_+)$ denote the number of negative eigenvalues of L_+ . If $u(\mathbf{x}, \mu)$ is a nonlinear bound state with $\mu < E_*$ (frequency lying in the semi-infinite gap), then $n_-(L_+) < \infty$ and the following nonlinear stability theorem holds [63, 48, 64, 28, 54].

THEOREM 2.2.

1. Let $u(\mathbf{x}, \mu)$ denote a positive soliton solution of NLS/GP with μ in the semi-infinite gap $(-\infty, E_*)$. The nonlinear bound state, $\psi(\mathbf{x}, t) = u(\mathbf{x}, \mu)e^{-i\mu t}$, is orbitally stable if the following two conditions hold:

(a) slope (VK) condition:

$$(2.4) \quad \frac{d}{d\mu} \mathcal{P}[u(\cdot, \mu)] < 0, \quad \text{and}$$

(b) spectral condition: L_+ has no zero eigenvalues and

$$(2.5) \quad n_-(L_+) = 1.$$

2. If either $\partial_\mu \mathcal{P}[u(\cdot, \mu)] > 0$ or $n_-(L_+) \geq 2$, then the soliton is unstable (nonlinearly unstable as well as linearly exponentially unstable).

Remark 2.1. As discussed in [54], the spectral condition can be associated with the suppression of a drift instability and the slope condition with the suppression of an energy-concentrating self-focusing instability.

2.2. Spectral theory for periodic potentials. We consider the Schrödinger operator $-\Delta + V(\mathbf{x})$ acting in $L^2(\mathbb{R}^d)$, where $V(\mathbf{x})$ is a smooth, real-valued potential which is periodic. That is, $V(\mathbf{x} + \mathbf{q}) = V(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^d$. Here $\mathbf{q} = \{\mathbf{q}_1, \dots, \mathbf{q}_d\}$ denotes a linearly independent set of vectors in \mathbb{R}^d that spans (over the integers) a lattice denoted by Γ . The set

$$(2.6) \quad \mathcal{B} = \left\{ \sum_{j=1}^d v_j \mathbf{q}_j : v_j \in \left[-\frac{1}{2}, \frac{1}{2} \right] \right\}$$

is called a fundamental period cell. The first Brillouin zone \mathcal{B}^* is generated by the dual basis $(\mathbf{r}_1, \dots, \mathbf{r}_d)$ given by $\mathbf{r}_j \cdot \mathbf{q}_k = 2\pi \delta_{jk}$; i.e.,

$$\mathcal{B}^* = \left\{ \sum_{j=1}^d v_j \mathbf{r}_j : v_j \in \left[-\frac{1}{2}, \frac{1}{2} \right] \right\},$$

and the dual lattice, Γ^* , is the integer span of \mathcal{B}^* .

It is useful to review some well-known results of Floquet–Bloch theory [35, 47, 23, 36]. The spectrum of $-\Delta + V$, denoted $\sigma(-\Delta + V)$, consists of a union of closed intervals called spectral *bands* separated by *gaps* (also known as band gaps and photonic band gaps). The spectral bands are characterized as follows.

For each $\mathbf{k} \in \mathcal{B}^*$ we seek solutions of the linear eigenvalue problem

$$(2.7) \quad (-\Delta + V(\mathbf{x})) u = E u$$

of the form $u(\mathbf{x}; \mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{x}}p(\mathbf{x}; \mathbf{k})$, where $p(\mathbf{x}; \mathbf{k})$ is periodic in \mathbf{x} :

$$\begin{aligned} \left[-(\nabla + i\mathbf{k})^2 + V(\mathbf{x}) \right] p(\mathbf{x}; \mathbf{k}) &= E(\mathbf{k})p(\mathbf{x}; \mathbf{k}), \\ p(\mathbf{x} + \mathbf{q}_j; \mathbf{k}) &= p(\mathbf{x}; \mathbf{k}), \quad j = 1, \dots, d. \end{aligned}$$

For each $\mathbf{k} \in \mathcal{B}^*$ this periodic elliptic boundary value problem has the following sequence of discrete eigenvalues, or *band dispersion functions*, tending to positive infinity: $E_1(\mathbf{k}) \leq E_2(\mathbf{k}) \leq \dots \leq E_m(\mathbf{k}) \leq \dots$. As \mathbf{k} varies over the Brillouin zone \mathcal{B}^* , each $E_m(\mathbf{k})$ sweeps out a closed subinterval of the real axis. The spectrum of $-\Delta + V$ acting on $L^2(\mathbb{R}^d)$ is the union of these subintervals,

$$(2.8) \quad \sigma(-\Delta + V) = \cup_{m \geq 1} \{E_m(\mathbf{k}) : \mathbf{k} \in \mathcal{B}^*\} \subset [\min_{\mathcal{B}} V, \infty),$$

and the states $\{u_m(\mathbf{x}; \mathbf{k}) = e^{i\mathbf{x}\cdot\mathbf{k}}p_m(\mathbf{x}; \mathbf{k})\}$ are complete in the sense that

$$(2.9) \quad f \in L^2(\mathbb{R}^d) \implies f(\mathbf{x}) = \sum_{m \geq 1} \int_{\mathcal{B}^*} \langle u_m(\cdot; \mathbf{k}), f(\cdot) \rangle_{L^2(\mathbb{R}^d)} u_m(\mathbf{x}; \mathbf{k}) d\mathbf{k}.$$

We denote the lowest point in the spectrum of (2.7) and corresponding periodic eigenstate by

$$(2.10) \quad E_* = E_1(0), \quad w(\mathbf{x}) = p(\mathbf{x}; \mathbf{k} = 0).$$

E_* is simple. We will often make use of the relation

$$(2.11) \quad L_* w = 0, \quad w > 0, \quad w(\mathbf{x} + \mathbf{q}_j) = w(\mathbf{q}_j),$$

where

$$(2.12) \quad L_* \equiv -\Delta + V - E_*.$$

Thus, w is the periodic ground state of L_* , $L_* \geq 0$, and 0 is a simple eigenvalue of L_* with kernel spanned by w . Note that if P^\perp is the orthogonal projection onto the subspace $\{w\}^\perp$,

$$(2.13) \quad P^\perp g = g - \left\langle \frac{w}{\|w\|}, g \right\rangle \frac{w}{\|w\|}, \quad \langle f, g \rangle = \int_{\mathcal{B}} \overline{f(\mathbf{x})} g(\mathbf{x}) dx,$$

then $L_*^{-1} P^\perp$ is bounded on the space of L^2 periodic functions with fundamental period cell \mathcal{B} .

Finally, note that we may, without loss of generality, restrict to the case where the fundamental period cell is $[-\pi, \pi]^d$. Indeed, if \mathcal{B} is the fundamental period cell

(see (2.6)), then define the constant matrix Q to be the matrix whose j th column is $(2\pi)^{-1}\mathbf{q}_j$. Then, under the change of coordinates $\mathbf{x} \mapsto \mathbf{z} = Q\mathbf{x}$,

$$\begin{aligned} & -\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} + V(\mathbf{x}) \text{ acting on } L_{per}^2(\mathcal{B}) \text{ transforms to} \\ & -\nabla_{\mathbf{z}} \cdot \alpha \nabla_{\mathbf{z}} + \tilde{V}(\mathbf{z}) \equiv - \sum_{i,j=1}^d \alpha_{ij} \frac{\partial^2}{\partial z_i \partial z_j} + \tilde{V}(\mathbf{z}) \\ & \text{acting on } L_{per}^2([-\pi, \pi]^d), \text{ where} \\ & \alpha = \frac{QQ^T}{|\det Q|}, \quad \tilde{V}(\mathbf{z}) = V(Q^{-1}\mathbf{z}), \quad \mathbf{x} = Q^{-1}\mathbf{z}. \end{aligned}$$

3. Main results. In this section we state our main results on bifurcation of solitons from the band edge, $E = E_*$, into the semi-infinite gap.

Hypotheses:

- (H1) Potential: $V(\mathbf{x})$ is smooth and periodic with $\mathcal{B} = [-\pi, \pi]^d$.
 (H2) Dimension/nonlinearity:⁴

$$(3.1) \quad d = 1, 2 : \sigma \in \mathbb{N}, \quad d = 3 : \sigma = 1.$$

THEOREM 3.1. *Let \mathbf{x}_0 denote any point of symmetry of $V(\mathbf{x})$.*

1. *For all μ less than and sufficiently near E_* , there is a family of nonlinear bound states of NLS/GP (“solitons”), $u(\cdot, \mu) \in H^s(\mathbb{R}^d)$, $s > d/2$, which is centered at \mathbf{x}_0 .*
2. *These solutions bifurcate from the zero solution at band-edge frequency $\mu = E_*$ into the semi-infinite gap. Specifically, this family is given by the two-scale expansion for small ϵ ,*

$$(3.2) \quad \mu_\epsilon = E_* - \epsilon^2,$$

$$(3.3) \quad u_\epsilon(\mathbf{x}, \mu_\epsilon) = \epsilon^{\frac{1}{\sigma}} [w(\mathbf{x})F(\epsilon(\mathbf{x} - \mathbf{x}_0)) + \epsilon U_1(\mathbf{x}, \epsilon(\mathbf{x} - \mathbf{x}_0)) \\ + \epsilon^2 U_2(\mathbf{x}, \epsilon(\mathbf{x} - \mathbf{x}_0)) + \eta(\mathbf{x}; \epsilon)],$$

where $\eta(\mathbf{x}; \epsilon)$ satisfies the estimate

$$(3.4) \quad |\eta(\cdot; \epsilon)|_{H^s} \leq C_s \epsilon^3, \quad s > d/2.$$

The terms in the expansion are given as follows: $w(\mathbf{x})$ is the band-edge Bloch state (see (2.11)), and $F(\mathbf{y})$ is the ground state solution of the NLS equation in an effective medium:

$$(3.5) \quad - \sum_{i,j=1}^d \partial_{y_i} A^{ij} \partial_{y_j} F(\mathbf{y}) - \gamma_{\text{eff}} F^{2\sigma+1}(\mathbf{y}) = -F(\mathbf{y}), \\ F > 0, \quad F \in H^1(\mathbb{R}^d).$$

⁴The assumption on the nonlinear term can be made less restrictive. However, since some of our results concerning the higher-order character of $\mu \mapsto \mathcal{P}[u(\cdot, \mu)]$ depend on the construction of a multiple-scale expansion to a sufficiently high order, we require a certain degree of smoothness of the nonlinear term in a neighborhood of zero. Note also that the methods and our results extend easily to more general nonlinearities, e.g., $\mathcal{K}[|u|^2]u$ (local or nonlocal).

The matrix A^{ij} is the inverse effective mass tensor [35], expressible in terms of the band dispersion function, $E_1(\mathbf{k})$, as

$$(3.6) \quad A^{ij} \equiv \delta_{ij} - \frac{4 \langle \partial_{x_j} w, L_*^{-1} \partial_{x_i} w \rangle}{\langle w, w \rangle} = \frac{1}{2} \frac{\partial^2 E_1}{\partial k_i \partial k_j}(\mathbf{k} = 0);$$

see Appendix A. The effective nonlinear coupling constant is given by

$$(3.7) \quad \gamma_{\text{eff}} = \frac{\int_{\mathcal{B}} w^{2\sigma+2}(\mathbf{x}) \, d\mathbf{x}}{\int_{\mathcal{B}} w^2(\mathbf{x}) \, d\mathbf{x}}.$$

A^{ij} is a symmetric, positive definite constant matrix. Its determinant, the product of inverse effective masses, is denoted by

$$(3.8) \quad \frac{1}{m_*} = \det(A^{ij}) \leq 1,$$

with $m_* = 1$ only if $V(\mathbf{x})$ is identically constant; see [34] and Appendix B.

3. $F(\mathbf{y})$ is a rescaled ground state of the NLS equation (1.11) as

$$(3.9) \quad F(\mathbf{y}) = \left(\frac{1}{\gamma_{\text{eff}}} \right)^{\frac{1}{2\sigma}} R(\Lambda^{-\frac{1}{2}} S \mathbf{y}, -1),$$

where S is an orthogonal matrix that diagonalizes the effective mass tensor; i.e.,

$$(3.10) \quad S_{ik} A^{kl} S_{lj} = \Lambda_{ij} \equiv \text{diag}(\lambda_1, \dots, \lambda_d),$$

where λ_i denote the eigenvalues of A^{ij} .

4. Combining (3.2), (3.3), and (3.9) gives for $0 < \epsilon^2 = E_* - \mu \ll 1$,

$$(3.11) \quad u(\mathbf{x}, \mu) = \left(\frac{\epsilon^2}{\gamma_{\text{eff}}} \right)^{\frac{1}{2\sigma}} \left[R\left(\Lambda^{-\frac{1}{2}} S \epsilon(\mathbf{x} - \mathbf{x}_0), -1\right) w(\mathbf{x}) + \mathcal{O}(\epsilon) \right].$$

5. The $\mathcal{O}(\epsilon)$ and $\mathcal{O}(\epsilon^2)$ corrections are given (using $\mathbf{y} = \epsilon(\mathbf{x} - \mathbf{x}_0)$ and summation over repeated indices) by

$$(3.12) \quad \begin{aligned} \mathcal{O}(\epsilon) : \quad U_1(\mathbf{x}, \mathbf{y}) &= 2L_*^{-1} [\partial_{x_i} w(\mathbf{x})] \partial_{y_i} F(\mathbf{y}), \\ \mathcal{O}(\epsilon^2) : \quad U_2(\mathbf{x}; \mathbf{y}) &= U_{2p}(\mathbf{x}, \mathbf{y}) + w(\mathbf{x}) F_{2h}(\mathbf{y}), \\ U_{2p}(\mathbf{x}, \mathbf{y}) &= L_*^{-1} \left[(\delta_{ij} + 4\partial_{x_j} L_*^{-1} \partial_{x_i} - A^{ij}) w(\mathbf{x}) \right] \partial_{y_i} \partial_{y_j} F(\mathbf{y}) \\ &\quad + L_*^{-1} [w^{2\sigma+1}(\mathbf{x}) - \gamma_{\text{eff}} w(\mathbf{x})] F^{2\sigma+1}(\mathbf{y}), \end{aligned}$$

$$(3.13) \quad L_+^A F_{2h}(\mathbf{y}) = S(\mathbf{y}),$$

where $S(\mathbf{y})$ is given by

$$(3.14) \quad \begin{aligned} S(\mathbf{y}) &= \langle w, w \rangle^{-1} \left[\langle w, (\Delta_{\mathbf{y}} - 1) U_{2p}(\cdot, \mathbf{y}) + (2\sigma + 1) U_0^{2\sigma} U_{2p}(\cdot, \mathbf{y}) \rangle \right. \\ &\quad \left. + \sigma(2\sigma + 1) \langle w, U_0^{2\sigma-1} U_1^2(\cdot, \mathbf{y}) \rangle + 2 \langle w, \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} \tilde{U}_3(\cdot, \mathbf{y}) \rangle \right], \end{aligned}$$

where \tilde{U}_3 is given in (4.19).

Theorem 3.1 is proved in sections 4 and 5.

Using expansion (3.2) we can derive the asymptotic behavior for $\mathcal{P}(\mu) = \mathcal{P}[u(\cdot, \mu)]$ as $\mu \rightarrow E_*$.

THEOREM 3.2. *Let \mathbf{x}_0 denote a point of symmetry of V and $u(\cdot, \mu)$ a soliton given in Theorem 3.1.*

1. For μ near the band edge $\mathcal{P}[u(\cdot, \mu)]$ is given by

$$(3.15) \quad \mathcal{P}[u(\cdot, \mu)] = |\mu - E_*|^{\frac{1}{\sigma} - \frac{d}{2}} \times [\zeta_* \mathcal{P}[R(\cdot, -1)] + \zeta_{1*}(\mu - E_*) + \mathcal{O}((\mu - E_*)^2)],$$

where $\mathcal{P}[R(\cdot, -1)]$, the optical power of the homogeneous NLS ground state, depends on σ and d ,

$$(3.16) \quad -\Delta R - R^{2\sigma+1} = -R, R > 0, R \in H^1,$$

and

$$(3.17) \quad 0 < \zeta_* \equiv \left(\frac{1}{m_*}\right)^{\frac{1}{2}} \left(\frac{(\int_{\mathcal{B}} w^2)^{\sigma+1}}{\int_{\mathcal{B}} w^{2\sigma+2}}\right)^{\frac{1}{\sigma}} \leq 1,$$

where $\int_{\mathcal{B}} g = \frac{1}{\text{vol}(\mathcal{B})} \int (\mathbf{x}) d\mathbf{x}$, and the slope is given by

$$(3.18) \quad \zeta_{1*} \equiv 4 \sum_{j=1}^d \int_{\mathcal{B}} |L_*^{-1} [\partial_{x_j} w(\mathbf{x})]|^2 d\mathbf{x} \cdot \int_{\mathbb{R}^d} |\partial_{y_j} F(\mathbf{y})|^2 d\mathbf{y} - \int_{\mathcal{B}} w^2(\mathbf{x}) d\mathbf{x} \int \left(\frac{1}{\sigma} F(\mathbf{y}) + \mathbf{y} \cdot \nabla_{\mathbf{y}} F(\mathbf{y})\right) S(\mathbf{y}) d\mathbf{y},$$

where $S(\mathbf{y})$ is given by (3.14). Note that to order $\mathcal{O}(|E_* - \mu|^1)$, the expansion is independent of \mathbf{x}_0 (the soliton centering).

2. Positive slope for small potentials: Let $V(x) = \delta V_1(x)$, where $|\delta|$ is sufficiently small and $V_1(x)$ is a smooth periodic function on \mathbb{R} with a zero cell average. Then, in the critical case $\sigma = 2$,

$$(3.19) \quad \zeta_* \sim 1 - 8\delta^2 \int_{\mathcal{B}} [(-\partial_{xx})^{-1} V_1]^2 dx ,$$

$$(3.20) \quad \zeta_{1*} \sim 34\sqrt{3} \pi \delta^2 \int_{\mathcal{B}} [(-\partial_{xxx})^{-1} V_1]^2 dx .$$

Here $(-\partial_{xx})^{-1}$ and $(-\partial_{xxx})^{-1}$ are, respectively, the second- and third-order integration operators in \mathcal{B} acting on the space of zero average functions to itself. Hence, $\zeta_{1*} > 0$ for small potentials.

3. Positive slope conjecture: $\zeta_{1*}[V] > 0$ if V is nonconstant.

Theorem 3.2 is proved in subsection 4.3, except for part 2, concerning small potentials, which is proved in Appendix D.

Remark 3.1. Concerning equality in (3.17), when $V(\mathbf{x})$ is constant, then so is $w(\mathbf{x})$. In that case $E_* = 0$, $E_1(\mathbf{k}) = \mathbf{k}^2$, and $m_*^{-1} = \det\{2^{-1} D_{k_i, k_j}^2 E_1(0)\} = 1$. Therefore, $\zeta_* = 1$.

Remark 3.2. That $\zeta_* \leq 1$ can be seen by considering each factor in the definition (3.17) separately. First, by Hölder’s inequality the quotient in the second factor of

(3.17) is bounded one with equality holding if and only if $w \equiv \text{constant}$. Furthermore, w is identically constant if and only if $V \equiv \text{constant}$. Concerning the first factor in (3.17), by Theorem 3.1, $0 < m_*^{-1} \leq 1$ with equality holding if $V \equiv \text{constant}$. Therefore, $0 < \zeta_* \leq 1$ with $\zeta_* = 1$ if and only if $V \equiv \text{constant}$.

In the critical case, an immediate consequence of Theorem 3.2 is the following result for critical nonlinearity ($\sigma = 2/d$).

COROLLARY 3.1. *Consider the critical case $\sigma = 2/d$; by hypotheses (H1) and (H2) this implies either $(d, \sigma) = (1, 2)$ or $(d, \sigma) = (2, 1)$.*

As $\mu \rightarrow E_*$ we have

$$(3.21) \quad \mathcal{P}[u(\cdot, \mu)] = \zeta_* \mathcal{P}_{cr} + \zeta_{1*} (\mu - E_*) + \mathcal{O}((\mu - E_*)^2) .$$

Here $\mathcal{P}_{cr} = \mathcal{P}[R(\cdot, -1)]$. Since $\zeta_* < 1$ for any nonconstant periodic potential, it follows that the limiting power at the band edge is strictly smaller than \mathcal{P}_{cr} ,

$$(3.22) \quad \mathcal{P}_{edge} \equiv \lim_{\mu \rightarrow E_*} \mathcal{P}[u(\cdot, \mu)] = \zeta_* \mathcal{P}_{cr} < \mathcal{P}_{cr} .$$

Corollary 3.1 is proved in subsection 4.3. The band-edge limiting behavior (3.22) is illustrated in Figure 1.1; see also Figure 6.3.

Concerning the NLS/GP dynamics near solitons, we have the following theorem.

THEOREM 3.3. *Consider the critical case $\sigma = 2/d$; by hypotheses (H1)–(H2) this implies either $(d, \sigma) = (1, 2)$ or $(d, \sigma) = (2, 1)$. If the positive slope conjecture of Theorem 3.2 holds, then*

1.

$$(3.23) \quad \left. \frac{d\mathcal{P}[u(\cdot; \mu)]}{d\mu} \right|_{\mu=E_*} > 0,$$

and it follows from Theorem 2.2 that for μ such that $E_* - \mu > 0$ and sufficiently small, $u(\cdot, \mu)$ is unstable.

2. In particular, for small periodic potentials, by Theorem 3.2, or μ such that $E_* - \mu > 0$ and sufficiently small, $u(\cdot, \mu)$ is unstable.

To complement this information about stability/instability of solitons, we remark on \mathcal{P}_{cr} and \mathcal{P}_{edge} as they relate to well posedness and blow-up/collapse.

THEOREM 3.4. *Let $\psi(\mathbf{x}, t)$ be a solution of the NLS/GP (1.1) with a smooth potential $V(\mathbf{x})$ and initial conditions $\psi_0(\mathbf{x})$ in H^1 . Denote by $R(\mathbf{x})$ the ground state (“Townes soliton”) corresponding to the translation invariant NLS equation (i.e., with $V \equiv 0$). Then if*

$$(3.24) \quad \mathcal{P}[\psi_0] = \int |\psi_0(\mathbf{x})|^2 d\mathbf{x} < \int R^2(\mathbf{x}) d\mathbf{x} \equiv \mathcal{P}_{cr},$$

then solutions of NLS/GP (1.1) exist globally in time—no singularity formation/no collapse.

Remark 3.3. Recall that in the spatially homogeneous case, $V \equiv 0$, if in addition to (3.24) we impose the additional assumption $|\mathbf{x}|\psi_0 \in L^2$, then $\psi(\mathbf{x}, t)$ tends to zero in L^p as $t \rightarrow \infty$ for a range of $p > 2$ [64]; see also [33] for scattering results in H^1 .

In view of Theorem 3.3 and Remark 3.3, \mathcal{P}_{cr} is thus called the *soliton excitation threshold* in the case of $V \equiv 0$. Excitation thresholds also play a role in systems without critical scaling symmetry. See, for example, [65, 26] and [19, 41, 9].

For V nonzero, the picture which emerges from the above theorems and numerics (see, for example, Figure 1.1) is quite different. The minimal mass (minimal power), band-edge power, and $V \equiv 0$ critical mass are related by:

$$(3.25) \quad \mathcal{P}_{\min} < \mathcal{P}_{edge} < \mathcal{P}_{cr}.$$

Here

$$(3.26) \quad \mathcal{P}_{\min} = \mathcal{P}[u(\cdot, \mu_{\min})] \equiv \min_{\mu \leq E_*} \mathcal{P}[u(\cdot, \mu)],$$

where in (3.26), $\mu \mapsto \mathcal{P}[u(\cdot, \mu)]$ is computed along the family of solitons centered at a local *minimum*; see the solid curve in Figure 1.1. Along this soliton curve, computations indicate that $u(\cdot, \mu) > 0$ and $n_-(L_+) = 1$. By Theorem 2.2 (applied for V periodic) there is an open set of initial data in the phase space H^1 ,

$$(3.27) \quad \{\psi_0 \in H^1 : \mathcal{P}_{\min} < \mathcal{P}[\psi_0] < \mathcal{P}_{edge} < \mathcal{P}_{cr}\},$$

within which there are coexisting unstable/“wide” and stable/“narrow” solutions. There is also an open set in H^1 ,

$$(3.28) \quad \{\psi_0 \in H^1 : \mathcal{P}_{edge} < \mathcal{P}[\psi_0] < \mathcal{P}_{cr}\},$$

where the only solitons are stable and “narrow.” The terms wide and narrow refer, respectively, to solitons with frequencies in an interval near (to the right of μ_{\min}) or far from (to the left of μ_{\min}) the band edge, E_* [54, 53].

Finally, we state a *soliton excitation threshold conjecture* (see also [64, 65, 33]). \mathcal{P}_{\min} is an *excitation threshold*: if $\mathcal{P}[\psi_0] < \mathcal{P}_{\min}$, then $\psi(\mathbf{x}, t)$ tends to zero in L^p as $t \rightarrow \infty$ for range of $p > 2$. Furthermore, there exists $f_{\pm} \in L^2(\mathbb{R}^d)$ such that $\psi(\mathbf{x}, t) - e^{i\Delta t} f_{\pm}$ tends to zero in H^1 as $t \rightarrow \pm\infty$.

Proof of Theorem 3.3. Part 1 follows from part 2 of Theorem 2.2, where we review results on the stability/instability of solitary waves. \square

Proof of Theorem 3.4. The proof follows from an application of the sharp Gagliardo–Nirenberg inequality; see [61, 64]. Specifically, for any function $f \in H^1(\mathbb{R}^d)$ we have

$$(3.29) \quad \left(1 - \frac{\|f\|_{L^2}}{\|R\|_{L^2}}\right)^{\frac{4}{d}} \int |\nabla f|^2 \leq \int \left(|\nabla f|^2 - \frac{1}{1 + \frac{2}{d}} |f|^{\frac{4}{d}+2}\right) \equiv \mathcal{H}_0[f],$$

where \mathcal{H}_0 denotes the conserved NLS/GP Hamiltonian for $V \equiv 0$. Estimate (3.29) was used in [61] to establish, for $V \equiv 0$, that if $\psi_0 \in H^1$ and $\|\psi_0\|_{L^2} < \|R\|_{L^2}$, then NLS has a global in time H^1 bounded solution. It was further used in [64] to show that if, in addition, we assume that $|\mathbf{x}|\psi_0 \in L^2$, then the solution decays to zero in L^p , for range of $p > 2$ (and therefore in L^2_{loc}).

To prove Theorem 3.4, note from (3.29) that

$$(3.30) \quad \left(1 - \left(\frac{\|f\|_{L^2}}{\|R\|_{L^2}}\right)^{\frac{4}{d}}\right) \int |\nabla f|^2 \leq \mathcal{H}[f] - \int V|f|^2.$$

Applying this inequality to a solution, $\psi(\mathbf{x}, t)$, of NLS/GP yields

$$(3.31) \quad \left(1 - \left(\frac{\|\psi_0\|_{L^2}}{\|R\|_{L^2}}\right)^{\frac{4}{d}}\right) \int |\nabla \psi(\mathbf{x}, t)|^2 \leq \mathcal{H}[\psi_0] + \|V\|_{L^\infty} \int |\psi_0|^2.$$

For initial data, ψ_0 , in small H^1 neighborhood of a soliton with frequency near the band edge, we have $\|\psi_0\|_{L^2} < \|R\|_{L^2}$. Estimate (3.31) implies a uniform bound on $\|\psi(\cdot, t)\|_{H^1}$ and therefore global existence (no singularity formation/no collapse). \square

3.1. Finite gaps—results for focusing and defocusing nonlinearities. In this subsection we remark on extensions of our results to solitons with frequencies in finite gaps (gap solitons). For this more general discussion, it is convenient to write NLS/GP and its nonlinear bound state equation in the form

$$(3.32) \quad i\partial_t \psi = -\Delta \psi + V(\mathbf{x})\psi + g|\psi|^{2\sigma} \psi,$$

$$(3.33) \quad (-\Delta + V)u + g u^{2\sigma+1} = \mu u,$$

where we have introduced a parameter g to encode the (i) focusing/attractive ($g = -1$) and the (ii) defocusing/repulsive ($g = +1$) cases.

Focusing nonlinearity, $g = -1$. Our results of the previous section applied to solitons with frequencies in the semi-infinite gap, $\mu < E_*$. The results on bifurcations of solutions from the spectral band edge can be extended to the case where E_* is replaced by E_{edge} , any band edge frequency. Here we consider the case where the following two conditions hold:

1. The space of \mathcal{B} -periodic solutions $(-\Delta + V)w(\mathbf{x}) = E_{edge}w(\mathbf{x})$ is one-dimensional, spanned by a function $w_{edge}(\mathbf{x})$; E_{edge} is attained by the band dispersion function at $\mathbf{k} = 0$.⁵
2. The inverse effective mass tensor, A^{ij} , is symmetric and positive definite.

In this case, we have solitons centered about any point of symmetry of $V(\mathbf{x})$, which in analogy to those described in Theorem 3.1, bifurcate from the left band edge toward lower frequencies into the spectral gap

$$(3.34) \quad \mu \mapsto u(\mathbf{x}, \mu) \approx (E_{edge} - \mu)^{\frac{1}{2\sigma}} w(\mathbf{x}) F\left(\sqrt{E_{edge} - \mu}(\mathbf{x} - \mathbf{x}_0)\right),$$

where $0 < E_* - \mu \ll 1$. Here F satisfies the effective medium nonlinear Schrödinger equation (3.5), whose inverse effective mass tensor, A^{ij} , is given by (3.6), with w replaced by w_{edge} . Alternatively, this is $(D^2 E_n(\mathbf{k}_0))_{ij}$, the Hessian matrix of a Bloch dispersion function, E_n , where $E_n(\mathbf{k}_0) = E_{edge}$, $\mathbf{k}_0 \in \mathcal{B}^*$.

Defocusing nonlinearity, $g = +1$. Here we consider the case where the following two conditions hold:

1. The space of \mathcal{B} -periodic solutions $(-\Delta + V)w(\mathbf{x}) = E_{edge}w(\mathbf{x})$ is one-dimensional, spanned by a function $w_{edge}(\mathbf{x})$; E_{edge} is attained by the band dispersion function at $\mathbf{k} = 0$.
2. The inverse effective mass tensor, $A^{ij} = -B^{ij}$, is symmetric and *negative* definite.

In this case, we have solitons centered about any point of symmetry of $V(\mathbf{x})$, bifurcating from the right band edge toward higher frequencies into the spectral gap.

Indeed, if we seek, along the lines of our previous analysis, soliton-like states with frequency

$$(3.35) \quad \mu = E_{edge} - \tau \epsilon^2,$$

our analysis near a band edge with negative definite effective mass tensor, $-B^{ij}$, yields an effective medium soliton equation:

$$-\sum_{i,j=1}^d \partial_{y_i} B^{ij} \partial_{y_j} F - \gamma_{\text{eff}} F^{2\sigma+1} = \tau F.$$

⁵In dimensions $d \geq 2$ band edges may be attained at $0 \neq \mathbf{k} \in \mathcal{B}^*$; see [22]. In this case, the corresponding solutions are complex-valued, and an extension of the present methods we use along the lines of [22] is necessary.

Thus, we can construct localized states for $\tau < 0$ and $\mu = E_{edge} + |\tau|\epsilon^2 > E_{edge}$.

Finally, we remark that all hypotheses concerning multiplicity of spectrum and curvature of band dispersion functions are verifiable in one space ($d = 1$) dimension. Thus, we have the following theorem.

THEOREM 3.5. *Let $V(x)$ denote a smooth, periodic, and even potential. Consider any finite width, nonempty, spectral gap, $-\infty < a < b < \infty$, of $-\partial_x^2 + V(x)$. The band dispersion curvature at $E = a$ is strictly negative and at $E = b$ is strictly positive; see Appendix C.*

1. *For focusing nonlinearity, $g = -1$, centered about any point of symmetry of V , there exists a family of solitons of NLS-GP (3.32), which bifurcates from the zero solution with frequencies in the gap less than $E = b$.*
2. *For defocusing nonlinearity, $g = +1$, centered about any point of symmetry of V , there exists a family of solitons of NLS-GP (3.32), which bifurcates from the zero state with frequencies bifurcating into the gap greater than $E = a$.*

These bifurcating branches have expansions and properties analogous to those described in Theorems 3.1 and 3.2.

Note. The results of this subsection indicate extensions to bifurcations into finite width gaps. In particular, for critical nonlinearities, we are able to analytically characterize the band-edge limit of the squared L^2 norm, \mathcal{P} . Note, however, that the factor, ζ_* , arising in finite gaps is associated with an excited Bloch state, i.e., a state $w_{edge}(\mathbf{x})$, which is not a positive ground state of the periodic boundary value problem. Since the estimate, $\zeta_* \leq 1$, hinged on the result [34], $(m_*)^{-1} \leq 1$, which makes use of the ground state property (in particular positivity), we do not have an estimate on the size of ζ_* in finite gap cases.

4. Homogenization/multiscale expansion. In this section we derive a formal multiple-scale expansion of solitons bifurcating from the band edge. In section 5 we prove an error estimate, thus completing the proof of Theorem 3.1 .

Without loss of generality we choose coordinates with $\mathbf{x}_0 = 0$. We seek a solution of the bound state equation (1.6), which bifurcates from the zero state at the band edge $\mu = E_*$, depending on a “fast” spatial scale \mathbf{x} and a slow spatial scale

$$(4.1) \quad \mathbf{y} = \epsilon(\mathbf{x} - \mathbf{x}_0) = \epsilon\mathbf{x}, \quad \epsilon \ll 1,$$

of the form

$$(4.2a) \quad \mu_\epsilon = E_* + \epsilon\mu_1 + \epsilon^2\mu_2 + \dots,$$

$$(4.2b) \quad u_\epsilon(\mathbf{x}) = \epsilon^{\frac{1}{\sigma}}U_\epsilon(\mathbf{x}, \mathbf{y}),$$

$$(4.2c) \quad U_\epsilon(\mathbf{x}, \mathbf{y}) = U_0(\mathbf{x}, \mathbf{y}) + \epsilon U_1(\mathbf{x}, \mathbf{y}) + \epsilon^2 U_2(\mathbf{x}, \mathbf{y}) + \dots .$$

We also impose periodicity in \mathbf{x} ; i.e.,

$$(4.3) \quad U_\epsilon(\mathbf{x} + \mathbf{q}_j, y) = U_\epsilon(\mathbf{x}, y), \quad j = 1, \dots, d.$$

Rewriting (1.6) by treating \mathbf{x} and \mathbf{y} as *independent* variables gives

$$-(\nabla_{\mathbf{x}} + \epsilon\nabla_{\mathbf{y}})^2 U_\epsilon + V(\mathbf{x})U_\epsilon - \epsilon^2 U_\epsilon^{2\sigma+1} = \mu_\epsilon U_\epsilon.$$

Using the expansion (4.2) and the operator L_* (see (2.11)), we obtain the following hierarchy of equations to $\mathcal{O}(\epsilon^4)$:

$$\begin{aligned} \mathcal{O}(\epsilon^0) : \quad & L_*U_0 = 0 , \\ \mathcal{O}(\epsilon^1) : \quad & L_*U_1 = (2\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} + \mu_1) U_0, \\ \mathcal{O}(\epsilon^2) : \quad & L_*U_2 = (2\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} + \mu_1) U_1 + (\Delta_{\mathbf{y}} + \mu_2)U_0 + U_0^{2\sigma+1}, \\ \mathcal{O}(\epsilon^3) : \quad & L_*U_3 = (2\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} + \mu_1) U_2 + (\Delta_{\mathbf{y}} + \mu_2)U_1 + (2\sigma + 1)U_0^{2\sigma}U_1 + \mu_3U_0, \\ \mathcal{O}(\epsilon^4) : \quad & L_*U_4 = (2\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} + \mu_1) U_3 + (\Delta_{\mathbf{y}} + \mu_2)U_2 \\ & + (2\sigma + 1)U_0^{2\sigma}U_2 + (2\sigma + 1)\sigma U_0^{2\sigma-1}U_1^2 + \mu_4U_0, \end{aligned}$$

where for each $k \geq 5$ we have

$$\begin{aligned} \mathcal{O}(\epsilon^k) : \quad & L_*U_k = \mu_k U_0 \\ & + (2\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} + \mu_1) U_{k-1}(\mathbf{x}, \mathbf{y}) \\ (4.4) \quad & + (\Delta_{\mathbf{y}} + \mu_2)U_{k-2} + \mathcal{F}_k[U_j(\mathbf{x}, \mathbf{y}), \mu_j : 1 \leq j \leq k - 2]. \end{aligned}$$

Note that L_* is self-adjoint with a one-dimensional null space spanned by w . In addition, μ_k is determined by a solvability condition of the form,

$$(4.5) \quad \mu_k \langle w(\cdot), U_0(\cdot, \mathbf{y}) \rangle + \langle w(\cdot), \tilde{\mathcal{F}}_k(\cdot, \mathbf{y}) \rangle = 0,$$

obtained by imposing orthogonality of w to the right-hand side of (4.4). Here, $\tilde{\mathcal{F}}_k$ denotes expression the sum of the last two lines on the right-hand side of (4.4). Condition (4.5) ensures the existence of a solution to (4.4) which is periodic in \mathbf{x} .

We now implement this procedure at successive orders in ϵ . In particular, we construct the terms $U_j(\mathbf{x}, \mathbf{y})$, $0 \leq j \leq 4$, as these are required in the proof of Theorem 3.1.

4.1. Solution at each $\mathcal{O}(\epsilon^k)$, $k = 0, 1, 2, 3, 4$.

$\mathcal{O}(\epsilon^0)$ terms. The $\mathcal{O}(\epsilon^0)$ equation is solved by the choice

$$(4.6) \quad U_0(\mathbf{x}, \mathbf{y}) = w(\mathbf{x})F(\mathbf{y}),$$

where w is the periodic Bloch state associated with the band edge, as defined in (2.11).

$\mathcal{O}(\epsilon^1)$ terms. The $\mathcal{O}(\epsilon)$ equation for U_1 , by (4.6), becomes

$$(4.7) \quad L_*U_1 = 2\nabla_{\mathbf{x}}w \cdot \nabla_{\mathbf{y}}F + \mu_1w F.$$

Orthogonality of the right-hand side of (4.7) to w implies $\mu_1 = 0$, from which we obtain (3.12):

$$(4.8) \quad U_1 = 2L_*^{-1}[\nabla_{\mathbf{x}}w] \cdot \nabla_{\mathbf{y}}F(\mathbf{y}) .$$

Remark 4.1. To be completely systematic, we should add to the right-hand side of (4.8) a term of the form $F_{1h}(\mathbf{y})w(\mathbf{x})$, which is in the null space of L_* , with $F_{1h}(\mathbf{y})$ to be determined. $F_{1h}(\mathbf{y})$ is determined via the solvability condition for U_3 . Symmetry considerations lead to $F_{1h}(\mathbf{y}) \equiv 0$ (see the discussion of U_3). We omit inclusion of this term to simplify the presentation. Note, however, that this degree of freedom is required at higher order. In particular, see the expression for $U_2(\mathbf{x}, \mathbf{y})$ and the role of $F_{2h}(\mathbf{y})$ in the solving for $U_4(\mathbf{x}, \mathbf{y})$.

$\mathcal{O}(\epsilon^2)$ terms. The $\mathcal{O}(\epsilon^2)$ equation for U_2 , by (4.6) and (3.12), becomes

$$\begin{aligned}
 L_* U_2 &= w(\mathbf{x}) (\Delta_{\mathbf{y}} + \mu_2) F(\mathbf{y}) + 4 \partial_{x_j} \partial_{y_j} L_*^{-1} [\partial_{x_i} w](\mathbf{x}) \partial_{y_i} F(\mathbf{y}) \\
 &\quad + w^{2\sigma+1}(\mathbf{x}) F^{2\sigma+1}(\mathbf{y}) \\
 &= w(\mathbf{x}) \Delta_{\mathbf{y}} F(\mathbf{y}) + 4 \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} [L_*^{-1} [\nabla_{\mathbf{x}} w](\mathbf{x}) \cdot \nabla_{\mathbf{y}} F] \\
 (4.9) \quad &\quad + \mu_2 w(\mathbf{x}) F(\mathbf{y}) + w^{2\sigma+1}(\mathbf{x}) F^{2\sigma+1}(\mathbf{y}).
 \end{aligned}$$

An equation for $F(\mathbf{y})$ is obtained by imposing orthogonality of the right-hand side of (4.9) to $w(\mathbf{x})$. It is convenient to formulate the following proposition.

PROPOSITION 4.1. Denote by \mathcal{L}_* the operator

$$\begin{aligned}
 (4.10) \quad \mathcal{L}_*(G) &\mapsto \mathcal{L}_*[G](\mathbf{x}, \mathbf{y}) \\
 &= w(\mathbf{x}) \Delta_{\mathbf{y}} G(\mathbf{y}) + 4 \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} [L_*^{-1} [\nabla_{\mathbf{x}} w](\mathbf{x}) \cdot \nabla_{\mathbf{y}} G(\mathbf{y})].
 \end{aligned}$$

Then

$$(4.11) \quad \langle w(\cdot), \mathcal{L}_*[G](\cdot, \mathbf{y}) \rangle = \partial_{y_i} A^{ij} \partial_{y_j} G(\mathbf{y}) \times \langle w, w \rangle.$$

Imposing orthogonality of the right-hand side of (4.9) to $w(\mathbf{x})$ and applying Proposition 4.1 yields (3.5) for $F = F(\mathbf{y}, \mu_2)$:

$$(4.12) \quad - \partial_{y_i} A^{ij} \partial_{y_j} F(\mathbf{y}, \mu_2) - \gamma_{\text{eff}} F^{2\sigma+1}(\mathbf{y}, \mu_2) = \mu_2 F(\mathbf{y}, \mu_2).$$

Here we consider only the positive decaying solution of (4.12), which by scaling and uniqueness can be expressed as

$$(4.13) \quad F(\mathbf{y}, \mu_2) = |\mu_2|^{\frac{1}{2\sigma}} F(|\mu_2|^{\frac{1}{2}} \mathbf{y}; -1).$$

We can, therefore, scale out $|\mu_2|$ and henceforth assume $\mu_2 = -1$.

Thus far, we have shown the following: To leading order, the slowly varying envelope function $F(\mathbf{y})$ of the nonlinear bound state of NLS/GP is composed of a nonlinear bound state of the NLS equation (3.5) for a homogeneous medium with effective mass tensor $(A^{ij})^{-1}$ (3.6) and effective nonlinearity γ_{eff} .

In subsection 4.2 we show that $F(\mathbf{y})$ is an appropriate scaling of $R(\mathbf{y})$, the Townes soliton, which is the ground state associated with an isotropic homogeneous medium.

We express the general solution of (4.9) in the form

$$(4.14) \quad U_2(\mathbf{x}, \mathbf{y}) = U_{2p}(\mathbf{x}, \mathbf{y}) + U_{2h}(\mathbf{x}, \mathbf{y}) = U_{2p}(\mathbf{x}, \mathbf{y}) + w(\mathbf{x}) F_{2h}(\mathbf{y}),$$

where U_{2p} denotes a particular solution of (4.9) and $w(\mathbf{x}) F_{2h}(\mathbf{y})$ lies in the kernel of L_* (recall $L_* w = 0$), with $F_{2h}(\mathbf{y})$ to be determined.

Using (4.12) to simplify the right-hand side of (4.9) gives

$$\begin{aligned}
 (4.15) \quad U_{2p}(\mathbf{x}, \mathbf{y}) &= \sum_{1 \leq i, j \leq d} L_*^{-1} [(\delta_{ij} + 4 \partial_{x_j} L_*^{-1} \partial_{x_i}) w(\mathbf{x}) \partial_{y_i} \partial_{y_j} F(\mathbf{y}) \\
 &\quad + w^{2\sigma+1}(\mathbf{x}) F^{2\sigma+1}(\mathbf{y}) - w(\mathbf{x}) F(\mathbf{y})] \\
 &= L_*^{-1} [(\delta_{ij} + 4 \partial_{x_j} L_*^{-1} \partial_{x_i} - A^{ij}) w(\mathbf{x})] \partial_{y_i} \partial_{y_j} F(\mathbf{y}) \\
 &\quad + L_*^{-1} [w^{2\sigma+1}(\mathbf{x}) - \gamma_{\text{eff}} w(\mathbf{x})] F^{2\sigma+1}(\mathbf{y})
 \end{aligned}$$

or

$$(4.16) \quad U_{2p}(\mathbf{x}, \mathbf{y}) \equiv \sum_{1 \leq i, j \leq d} L_*^{-1} X_{2p,1}^{ij}(\mathbf{x}) \partial_{y_i} \partial_{y_j} F(\mathbf{y}) + L_*^{-1} X_{2p,2}(\mathbf{x}) F^{2\sigma+1}(\mathbf{y}),$$

with $X_{2p,1}^{ij}$ and $X_{2p,2}$ given by the corresponding expressions in (4.15). To obtain (4.15), we use (4.12) for $F(\mathbf{y})$, in terms of the effective mass tensor, (3.6), and effective coupling, (3.7). This is a consequence of the solvability (orthogonality) condition for (4.9).

$\mathcal{O}(\epsilon^3)$ terms. Using (4.14), we obtain the following equation for $U_3(\mathbf{x}, \mathbf{y})$:

$$(4.17) \quad \begin{aligned} L_* U_3 &= 2 \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} (U_{2p} + w F_{2h}) \\ &+ (\Delta_{\mathbf{y}} - 1 + (2\sigma + 1) F^{2\sigma} w^{2\sigma}) 2L_*^{-1} \partial_{x_i} w \partial_{y_i} F + \mu_3 w F. \end{aligned}$$

Solvability of (4.17) requires orthogonality of the right-hand side to w . Since all terms, except the last, on the right-hand side of (4.17) are antisymmetric functions of \mathbf{x} , we have $\mu_3 = 0$. Thus, after substitution of the explicit expression for U_{2p} , we have

$$(4.18) \quad \begin{aligned} L_* U_3 &= 2 \partial_{x_i} w(\mathbf{x}) \partial_{y_i} F_{2h}(\mathbf{y}) \\ &+ 2 \partial_{x_i} L_*^{-1} X_{2p,1}^{ij}(\mathbf{x}) \partial_{y_i} \partial_{y_j} F(\mathbf{y}) + 2 \partial_{x_i} L_*^{-1} X_{2p,2}(\mathbf{x}) \partial_{y_i} F^{2\sigma+1}(\mathbf{y}) \\ &+ 2L_*^{-1} \partial_{x_i} w(\mathbf{x}) (\Delta_{\mathbf{y}} - 1) \partial_{y_i} F(\mathbf{y}) \\ &+ 2w^{2\sigma}(\mathbf{x}) L_*^{-1} \partial_{x_i} w(\mathbf{x}) \partial_{y_i} F^{2\sigma+1}(\mathbf{y}), \end{aligned}$$

with summation over repeated indices implied. Thus,

$$U_3 = 2L_*^{-1} [\nabla_{\mathbf{x}} w](\mathbf{x}) \cdot \nabla_{\mathbf{y}} F_{2h} + \tilde{U}_3,$$

where we have introduced

$$(4.19) \quad \begin{aligned} \tilde{U}_3 &\equiv 2L_*^{-1} [\nabla L_*^{-1} [X_{2p,1}^{ij}] \cdot \nabla_{\mathbf{y}} \partial_{y_i} \partial_{y_j} F + \nabla L_*^{-1} [X_{2p,2}] \cdot \nabla_{\mathbf{y}} F^{2\sigma+1} \\ &+ L_*^{-1} [\nabla_{\mathbf{x}} w] \cdot (\Delta_{\mathbf{y}} - 1) \nabla_{\mathbf{y}} F + 2w^{2\sigma} L_*^{-1} [\nabla_{\mathbf{x}} w] \cdot \nabla_{\mathbf{y}} F^{2\sigma+1}]. \end{aligned}$$

$\mathcal{O}(\epsilon^4)$ terms. For U_4 we have

$$(4.20) \quad \begin{aligned} L_* U_4 &= (2 \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} + \mu_1) U_3 + (\Delta_{\mathbf{y}} - 1) U_2 + (2\sigma + 1) U_0^{2\sigma} U_2 \\ &+ \sigma(2\sigma + 1) U_0^{2\sigma-1} U_1^2 + \mu_4 U_0 \\ &= \mathcal{L}_* [F_{2h}](\mathbf{x}, \mathbf{y}) + (2\sigma + 1) w^{2\sigma+1}(\mathbf{x}) F^{2\sigma} \\ &+ (\Delta_{\mathbf{y}} - 1) U_{2p} + (2\sigma + 1) U_0^{2\sigma} U_{2p} + 2 \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} \tilde{U}_3 \\ &+ \sigma(2\sigma + 1) U_0^{2\sigma-1} U_1^2 + \mu_4 w(\mathbf{x}) F(\mathbf{y}). \end{aligned}$$

The operator $\mathcal{L}_*[\cdot](\mathbf{x}, \mathbf{y})$, appearing in (4.20), is defined in Proposition 4.1. Imposing orthogonality of the right-hand side of (4.20) and applying Proposition 4.1 gives the following equation for F_{2h} :

$$(4.21) \quad \begin{aligned} L_+^A F_{2h}(\mathbf{y}) &= \langle w, w \rangle^{-1} [\langle w, (\Delta_{\mathbf{y}} - 1) U_{2p}(\cdot, \mathbf{y}) + (2\sigma + 1) U_0^{2\sigma} U_{2p}(\cdot, \mathbf{y}) \rangle \\ &+ \sigma(2\sigma + 1) \langle w, U_0^{2\sigma-1} U_1^2(\cdot, \mathbf{y}) \rangle + 2 \langle w, \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}} \tilde{U}_3(\cdot, \mathbf{y}) \rangle] + \mu_4 F(\mathbf{y}) \\ &\equiv S(\mathbf{y}), \end{aligned}$$

where L_+^A is the second-order linear Schrödinger operator

$$(4.22) \quad L_+^A \equiv -\partial_{y_i} A^{ij} \partial_{y_j} + 1 - (2\sigma + 1) \gamma_{\text{eff}} F^{2\sigma+1}(\mathbf{y}).$$

We now show that we can take $\mu_4 = 0$. Equation (4.21) can be solved in $L^2(\mathbb{R}^d)$ for $F_{2h}(\mathbf{y})$ if and only if $S(\mathbf{y})$ is L^2 -orthogonal to the kernel of L_+^A . The kernel of L_+^A has dimension d and is generated by translations; i.e., $\text{Ker}(L_+^A) = \text{span}\{\partial_{y_j} F(\mathbf{y}), j = 1, \dots, d\}$ [62, 39]. Since $F(\mathbf{y})$ is even, the kernel of L_+^A consists of functions which are antisymmetric in one coordinate direction. Moreover, it is easy to see that all terms in $S(\mathbf{y})$ are symmetric and therefore orthogonal to the kernel of L_+^A . Thus, we set $\mu_4 = 0$.

4.2. $F(\mathbf{y})$ is a scaled Townes soliton. Thus far, we have constructed the formal expansion (3.2) of $(u_\epsilon, \mu_\epsilon)$ through $\mathcal{O}(\epsilon^2)$. The proof of its validity, in particular the error estimate (3.4), is given in section 5.

We conclude this subsection by relating the *effective medium soliton* $F(\mathbf{y})$, which solves the NLS bound state equation with effective medium parameters A^{ij} and γ_{eff} to the unique ground state of the uniform-medium NLS equation,

$$(4.23) \quad -\Delta R - R^{2\sigma+1} = \mu R, \quad R > 0, \quad R \in H^1(\mathbb{R}^d).$$

Let $A = (A^{ij})$ and $\Lambda \equiv \text{diag}(\lambda_1, \dots, \lambda_d)$ denote the diagonal matrix whose diagonal entries are the eigenvalues of A . Let S denote an orthogonal matrix for which

$$(4.24) \quad SAS^T = \Lambda \equiv \text{diag}(\lambda_1, \dots, \lambda_d).$$

Then, under the change of coordinates $\mathbf{y} \mapsto \mathbf{z} = \Lambda^{-\frac{1}{2}} S \mathbf{y}$, $F_1(\mathbf{z}) = F(\mathbf{y}, -1)$ solves (4.23) with $\mu = -1$. By uniqueness up to translations, the solution to the isotropic NLS equation (4.23) is given by

$$(4.25) \quad F(\mathbf{y}) = F_1(\mathbf{z}) = R(\Lambda^{-\frac{1}{2}} S \mathbf{y}, -1) = \left(\frac{1}{\gamma_{\text{eff}}}\right)^{\frac{1}{2\sigma}} R(\Lambda^{-\frac{1}{2}} S \mathbf{y}, -1).$$

Note that our expansion gives

$$(4.26) \quad \mu_\epsilon = E_* - \epsilon^2 + \mathcal{O}(\epsilon^5) \approx E_* - \epsilon^2.$$

In fact, as shown in the proof (see (5.3)) we can take $\mu_\epsilon = E_* - \epsilon^2$. Substitution of (4.25) and (4.26) into the expansion (4.2) and using (4.6) and $\mathbf{y} = \epsilon \mathbf{x}$ yields the leading order expansion of $u(\mathbf{x}, \mu)$ displayed in (3.3).

4.3. $\mathcal{P}[u(\cdot, \mu)]$ near the band edge. To prove Theorem 3.2 and Corollary 3.1, we evaluate $\int |u_\epsilon(\mathbf{x})|^2 d\mathbf{x}$, where u_ϵ is given by the two-scale expansion plus error term (3.3) of Theorem 3.1. We obtain (again recalling the choice of coordinates so that $\mathbf{y} = \epsilon(\mathbf{x} - \mathbf{x}_0) = \epsilon \mathbf{x}$)

$$(4.27) \quad \int_{\mathbb{R}^d} |u_\epsilon(\mathbf{x})|^2 d\mathbf{x} = \epsilon^{\frac{2}{\sigma}} \left[\int \underbrace{|U_0(\mathbf{x}, \epsilon \mathbf{x})|^2}_{\alpha_0} + \underbrace{2\epsilon U_0(\mathbf{x}, \epsilon \mathbf{x}) U_1(\mathbf{x}, \epsilon \mathbf{x})}_{\alpha_1} + \epsilon^2 \left(\underbrace{|U_1(\mathbf{x}, \epsilon \mathbf{x})|^2 + 2U_0(\mathbf{x}, \epsilon \mathbf{x}) U_2(\mathbf{x}, \epsilon \mathbf{x})}_{\alpha_2} \right) d\mathbf{x} + \mathcal{O}(\epsilon^3) \right] = \epsilon^{\frac{2}{\sigma}} [\mathcal{I}_0^\epsilon + \mathcal{I}_1^\epsilon + \mathcal{I}_2^\epsilon + \mathcal{O}(\epsilon^3)].$$

Each of the three terms on the right-hand side will be treated below using the following general averaging method.

LEMMA 4.1. *Let $p(\mathbf{x})$ be periodic on the lattice Γ having the fundamental period cell \mathcal{B} . Let \mathcal{B}^* denote the dual fundamental cell (first Brillouin zone) which spans the dual lattice Γ^* . Assume that $\sum_{\mathbf{k} \in \Gamma^*} |p_{\mathbf{k}}| < \infty$, where $\{p_{\mathbf{k}}\}$ denotes the set of Fourier coefficients of p . Let $G \in L^1(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)$. Then, as $\epsilon \rightarrow 0$,*

$$\left| \epsilon^d \int_{\mathbb{R}^d} p(\mathbf{x}) G(\epsilon \mathbf{x}) d\mathbf{x} - \int_{\mathcal{B}} p(\mathbf{x}) d\mathbf{x} \cdot \int_{\mathbb{R}^d} G(\mathbf{y}) d\mathbf{y} \right| = \mathcal{O}(\epsilon^\infty) \times \sum_{\mathbf{k} \in \Gamma^*} |p_{\mathbf{k}}| = \mathcal{O}(\epsilon^\infty),$$

where the cell average $f_{\mathcal{B}}$ is defined by $f_{\mathcal{B}} p = \frac{1}{|\mathcal{B}|} \int_{\mathcal{B}} p$.

Proof of Lemma 4.1. $p(\mathbf{x})$ has the Fourier representation

$$(4.28) \quad p(\mathbf{x}) = \sum_{\mathbf{k} \in \Gamma^*} p_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{x}}, \quad \text{where } p_{\mathbf{k}} = \int_{\mathcal{B}} e^{-i\mathbf{k} \cdot \mathbf{x}} p(\mathbf{x}) d\mathbf{x}.$$

Therefore,

$$\begin{aligned} \epsilon^d \int_{\mathbb{R}^d} p(\mathbf{x}) G(\epsilon \mathbf{x}) d\mathbf{x} &= \epsilon^d \sum_{\mathbf{k} \in \Gamma^*} p_{\mathbf{k}} \int_{\mathbb{R}^d} e^{i\mathbf{k} \cdot \mathbf{x}} G(\epsilon \mathbf{x}) d\mathbf{x} \\ &= p_0 \int_{\mathbb{R}^d} G(\mathbf{y}) d\mathbf{y} + \sum_{0 \neq \mathbf{k} \in \Gamma^*} p_{\mathbf{k}} \int_{\mathbb{R}^d} e^{i\frac{\mathbf{k}}{\epsilon} \cdot \mathbf{y}} G(\mathbf{y}) d\mathbf{y} \\ &= \int_{\mathcal{B}} p(\mathbf{x}) d\mathbf{x} \int_{\mathbb{R}^d} G(\mathbf{y}) d\mathbf{y} + \sum_{0 \neq \mathbf{k} \in \Gamma^*} p_{\mathbf{k}} \hat{G}\left(\frac{\mathbf{k}}{2\pi\epsilon}\right). \end{aligned}$$

By smoothness of G , for all $q \geq 1$ and $\xi \in \mathbb{R}^d$ there is a positive constant, $r_{G,q}$, such that $|\hat{G}(\xi)| \leq r_{G,q}(1 + |\xi|)^{-q}$. The required estimate of the remainder term follows. This completes the proof of the lemma. \square

We now proceed with proof of the Corollary 3.1 by evaluating the terms \mathcal{I}_j^ϵ , $j = 0, 1, 2$, in (4.27).

Claim 1.

$$(4.29) \quad \mathcal{I}_0^\epsilon = \epsilon^{-d} \zeta_* \mathcal{P}_{cr} + \mathcal{O}(\epsilon^\infty), \quad \text{where}$$

$$(4.30) \quad \zeta_* = \left(\frac{\left(\int_{\mathcal{B}} w^2 \right)^{\sigma+1}}{\int_{\mathcal{B}} w^{2\sigma+2}} \right)^{\frac{1}{\sigma}} m_*^{-\frac{1}{2}}.$$

Proof. By (4.6) and Lemma 4.1 one has

$$\mathcal{I}_0^\epsilon = \int \alpha_0 d\mathbf{x} = \int_{\mathbb{R}^d} w^2(\mathbf{x}) F^2(\epsilon \mathbf{x}) d\mathbf{x} = \epsilon^{-d} \int_{\mathcal{B}} w^2(\mathbf{x}) d\mathbf{x} \int_{\mathbb{R}^d} F^2(\mathbf{y}) d\mathbf{y} + \mathcal{O}(\epsilon^\infty).$$

Using expression (3.9) for $F(\mathbf{y})$ as a scaling of $R(\mathbf{y}; -1)$, we get

$$\begin{aligned} \int_{\mathbb{R}^d} F^2(\mathbf{y}) d\mathbf{y} &= (\gamma_{\text{eff}})^{-\frac{1}{\sigma}} m_*^{-\frac{1}{2}} \int_{\mathbb{R}^d} R^2(\mathbf{y}; -1) d\mathbf{y} \\ &= \left(\frac{\int_{\mathcal{B}} w^2}{\int_{\mathcal{B}} w^{2\sigma+2}} \right)^{\frac{1}{\sigma}} m_*^{-\frac{1}{2}} \mathcal{P}_{cr}. \quad \square \end{aligned}$$

Claim 2. $\mathcal{I}_1^\epsilon = \mathcal{O}(\epsilon^\infty)$.

Proof. We proceed similarly by using (3.12) and Lemma 4.1. We obtain

$$\begin{aligned} \mathcal{I}_1^\epsilon &= \int \alpha_1 \, d\mathbf{x} = 2\epsilon \sum_{j=1}^d \int_{\mathbb{R}^d} w(\mathbf{x}) F(\epsilon\mathbf{x}) \cdot 2L_*^{-1} [\partial_{x_j} w](\mathbf{x}) \partial_{y_j} F(\epsilon\mathbf{x}) \, d\mathbf{x} \\ &= 4\epsilon^{1-d} \int_{\mathcal{B}} w(\mathbf{x}) L_*^{-1} (\partial_{x_j} w)(\mathbf{x}) \, d\mathbf{x} \int F(\mathbf{y}) \partial_{y_j} F(\mathbf{y}) \, d\mathbf{y} + \epsilon \mathcal{O}(\epsilon^\infty) = \mathcal{O}(\epsilon^\infty), \end{aligned}$$

since $\int_{\mathbb{R}^d} F(\mathbf{y}) \partial_{y_i} F(\mathbf{y}) \, d\mathbf{y} = 0$. \square

Finally, we turn to \mathcal{I}_2^ϵ .

Claim 3.

$$\begin{aligned} \mathcal{I}_2^\epsilon &= \epsilon^{-d} \cdot \epsilon^2 \left[4 \sum_{j=1}^d \int_{\mathcal{B}} |L_*^{-1} [\partial_{x_j} w(\mathbf{x})]|^2 \, d\mathbf{x} \int_{\mathbb{R}^d} |\partial_{y_j} F(\mathbf{y})|^2 \, d\mathbf{y} \right. \\ &\quad \left. + \int_{\mathcal{B}} w^2(\mathbf{x}) \, d\mathbf{x} \int \partial_\Omega F(\mathbf{y}) S(\mathbf{y}) \, d\mathbf{y} \right] + \mathcal{O}(\epsilon^\infty), \end{aligned}$$

(4.31)

where $S(\mathbf{y})$ is explicitly displayed in (4.21).

Proof.

$$\begin{aligned} \mathcal{I}_2^\epsilon &= \int \alpha_2 \, d\mathbf{x} = \epsilon^2 \int [|U_1(\mathbf{x}, \epsilon\mathbf{x})|^2 + 2U_0(\mathbf{x}, \epsilon\mathbf{x}) (U_{2p}(\mathbf{x}, \epsilon\mathbf{x}) + w(\mathbf{x})F_{2h}(\epsilon\mathbf{x}))] \, d\mathbf{x} \\ &= \mathcal{I}_{2,a} + \mathcal{I}_{2,b} + \mathcal{I}_{2,c}, \end{aligned}$$

and by Lemma 4.1, $\mathcal{I}_{2,a}^\epsilon$ is given by

$$\begin{aligned} \mathcal{I}_{2,a}^\epsilon &\equiv \epsilon^2 \int |U_1(\mathbf{x}, \epsilon\mathbf{x})|^2 \, d\mathbf{x} = 4\epsilon^2 \int |L_*^{-1} [\partial_{x_j} w(\mathbf{x})] \partial_{y_j} F(\epsilon\mathbf{x})|^2 \, d\mathbf{x} \\ &= \epsilon^{2-d} \left(4 \sum_{j=1}^d \int_{\mathcal{B}} |L_*^{-1} [\partial_{x_j} w(\mathbf{x})]|^2 \, d\mathbf{x} \cdot \int_{\mathbb{R}^d} |\partial_{y_j} F(\mathbf{y})|^2 \, d\mathbf{y} + \mathcal{O}(\epsilon^\infty) \right). \end{aligned}$$

Concerning $\mathcal{I}_{2,b}$, we assert the following:

$$(4.32) \quad \mathcal{I}_{2,b}^\epsilon = 2\epsilon^2 \int_{\mathbb{R}^d} w(\mathbf{x}) F(\epsilon\mathbf{x}) U_{2p}(\mathbf{x}, \epsilon\mathbf{x}) \, d\mathbf{x} = \mathcal{O}(\epsilon^\infty). \quad \square$$

Proof of Theorem 3.2. To prove (4.32) we note that U_2 (4.16) is of a sum of terms that have the factored form

$$(4.33) \quad U_{2p}(\mathbf{x}, \epsilon\mathbf{x}) = \sum_j G_j(\epsilon\mathbf{x}) \cdot L_*^{-1} P^\perp g_j(\mathbf{x}).$$

Here P^\perp denotes the projection onto the orthogonal complement of w in $L^2_{\text{periodic}}(\mathcal{B})$; see (2.13). Substitution of (4.33) gives

$$(4.34) \quad \mathcal{I}_{2,b}^\epsilon = 2\epsilon^2 \sum_j \int_{\mathbb{R}^d} w(\mathbf{x}) L_*^{-1} P^\perp g_j(\mathbf{x}) \cdot F(\epsilon\mathbf{x}) G_j(\epsilon\mathbf{x}) \, d\mathbf{x},$$

which by Lemma 4.1 implies

$$(4.35) \quad \mathcal{I}_{2,b}^\epsilon = \sum_j \int_{\mathcal{B}} w(\mathbf{x}) L_*^{-1} P^\perp g_j(\mathbf{x}) d\mathbf{x} \int F(\mathbf{y}) G_j(\mathbf{y}) d\mathbf{y} + \mathcal{O}(\epsilon^\infty).$$

Since P^\perp commutes with functions of L_* and $P^\perp w = 0$, we have

$$\langle w(\mathbf{x}), L_*^{-1} P^\perp g_j \rangle = \langle w, L_*^{-1} P^\perp P^\perp g_j \rangle = \langle P^\perp w, L_*^{-1} P^\perp g_j \rangle = 0.$$

It remains to calculate $\mathcal{I}_{2,c}$:

$$\begin{aligned} \mathcal{I}_{2,c} &= 2\epsilon^2 \int U_0(\mathbf{x}, \epsilon\mathbf{x}) U_{2h}(\mathbf{x}, \epsilon\mathbf{x}) d\mathbf{x} \\ &= 2\epsilon^2 \int w(\mathbf{x}) F(\epsilon\mathbf{x}) \cdot w(\mathbf{x}) F_{2h}(\epsilon\mathbf{x}) d\mathbf{x} \\ &= 2\epsilon^{2-d} \int_{\mathcal{B}} w^2 \cdot \int F(\mathbf{y}) F_{2h}(\mathbf{y}) d\mathbf{y} + \mathcal{O}(\epsilon^\infty) \\ &= 2\epsilon^{2-d} \int_{\mathcal{B}} w^2 \cdot \int F(\mathbf{y}) (L_+^A)^{-1} S(\mathbf{y}) d\mathbf{y} + \mathcal{O}(\epsilon^\infty) \\ &= 2\epsilon^{2-d} \int_{\mathcal{B}} w^2 \cdot \int (L_+^A)^{-1} F(\mathbf{y}) S(\mathbf{y}) d\mathbf{y} + \mathcal{O}(\epsilon^\infty) \\ (4.36) \quad &= -\epsilon^{2-d} \int_{\mathcal{B}} w^2 \cdot \left(\frac{1}{\sigma} F(\mathbf{y}) + \mathbf{y} \cdot \nabla_{\mathbf{y}} F(\mathbf{y}) \right) S(\mathbf{y}) d\mathbf{y} + \mathcal{O}(\epsilon^\infty). \end{aligned}$$

Here we have used the relation

$$(L_+^A)^{-1} F = \partial_{\mu_2} F(\cdot, \mu_2)|_{\mu_2=-1} = -\frac{1}{2} \left(\frac{1}{\sigma} F(\mathbf{y}) + \mathbf{y} \cdot \nabla_{\mathbf{y}} F(\mathbf{y}) \right),$$

which follows from differentiation of the equation for $F = F(\mathbf{x}; \mu_2)$ with respect to μ_2 ; see (4.12) and (4.13).

Therefore, summing up the terms we have

$$\begin{aligned} \int_{\mathbb{R}^d} |u_\epsilon(\mathbf{x})|^2 d\mathbf{x} &= \epsilon^{\frac{2}{\sigma}} [\mathcal{I}_0^\epsilon + \mathcal{I}_1^\epsilon + \mathcal{I}_2^\epsilon + \mathcal{O}(\epsilon^3)] \\ &= (\epsilon^2)^{\frac{1}{\sigma} - \frac{d}{2}} \zeta_* \mathcal{P}_{cr} \\ &\quad + (\epsilon^2)^{\frac{1}{\sigma} - \frac{d}{2} + 1} \left[4 \sum_{j=1}^n \int_{\mathcal{B}} |L_*^{-1} [\partial_{x_j} w(\mathbf{x})]|^2 d\mathbf{x} \int_{\mathbb{R}^d} |\partial_{y_j} F(\mathbf{y})|^2 d\mathbf{y} \right. \\ &\quad \left. - \int_{\mathcal{B}} w^2 d\mathbf{x} \int \left(\frac{1}{\sigma} F(\mathbf{y}) + \mathbf{y} \cdot \nabla_{\mathbf{y}} F(\mathbf{y}) \right) S(\mathbf{y}) d\mathbf{y} \right] + \mathcal{O}(\epsilon^\infty) \\ &= (\epsilon^2)^{\frac{1}{\sigma} - \frac{d}{2}} (\zeta_* \mathcal{P}_{cr} + \epsilon^2 \zeta_{1*} + \mathcal{O}(\epsilon^\infty)) \end{aligned}$$

Recall $S(\mathbf{y})$ is displayed in (4.21). □

This concludes the proof of Theorem 3.2.

5. The error estimate (3.4). In this section we prove the error estimate (3.4), which concludes the proof of Theorem 3.1. For ease of presentation, we focus on the cubic ($\sigma = 1$) one-dimensional case ($d = 1$):

$$(5.1) \quad (-\partial_x^2 + V(x)) u - u^3 = \mu u.$$

The proof carries over to the more general setting in the statement of Theorem 3.1. After the proof, we indicate the modifications required for the proof to go through in general dimension $d = 1, 2, 3$; see Remark 5.2 below.

We shall construct a solution $(u, E) = (u_\epsilon, \mu_\epsilon)$, using the formal multiple-scale expansion of section 4:

$$(5.2) \quad u_\epsilon = \epsilon U^\epsilon(x) = \epsilon \left[\sum_{k=0}^4 \epsilon^k U_k(x, y) + \epsilon^3 U_5^\epsilon(x) \right],$$

$$(5.3) \quad \mu_\epsilon = E_* - \epsilon^2.$$

The expansion includes an error term, $\epsilon^3 U_5^\epsilon(x)$, which must be estimated. The equation for $\epsilon^3 U_5^\epsilon$ is

$$(5.4) \quad \begin{aligned} & (-\partial_x^2 + V(x) - 3\epsilon^2 U_0^2(x, \epsilon x) - E_* + \epsilon^2) U_5^\epsilon(x) \\ &= \epsilon^2 R^\epsilon[U_j; 0 \leq j \leq 4, U_5^\epsilon(x)] \\ &\equiv \epsilon^2 R_0^\epsilon[U_j; 0 \leq j \leq 4] + \epsilon^3 R_2^\epsilon[U_j; 0 \leq j \leq 4] U_5^\epsilon \\ &+ \epsilon^5 R_2^\epsilon[U_j; 0 \leq j \leq 4] (U_5^\epsilon)^2 + \epsilon^8 (U_5^\epsilon)^3, \end{aligned}$$

where $R_k^\epsilon[U_j; 0 \leq j \leq 4]$ denotes the coefficient of the k th power of U_5^ϵ and is a polynomial in the previously constructed functions $U_j, j = 0, 1, 2, 3, 4$.

The scaling of the error term in (5.2) is motivated as follows. Formally, the correction to the leading order sum in (5.2) will be of order ϵ^5 . In our analysis, we find that the frequency components of the corrector (to the truncated multiple-scale expansion near the band edge) are of order $\mathcal{O}(\epsilon^3)$. Therefore, anticipate this result in (5.3). We will, in fact, show that for $s > d/2, \|U_5^\epsilon\|_{H^s}$ is bounded uniformly in ϵ . This implies the error bound of Theorem 3.1.

In particular,

$$(5.5) \quad R_0^\epsilon = 2\partial_x \partial_y U_4 + (\partial_y^2 - 1)U_3 + 3U_0^2 + 6U_0 U_1 U_2 + U_1^3 + \mathcal{O}(\epsilon).$$

Our goal is to estimate U_5^ϵ , and to do this we employ the spectral (Floquet–Bloch) decomposition of the operator $-\Delta + V(x)$.

5.1. Floquet–Bloch theory and the Bloch transform. See [23, 47, 36] for basic results on the spectral theory of operators with periodic coefficients.

Assume $V(x + 2\pi) = V(x)$. For each $k \in \mathbb{T} = [-\frac{1}{2}, \frac{1}{2}]$ we seek solutions of the eigenvalue equation for the operator $(-\partial_x^2 + V(x))$ of the form:

$$(5.6) \quad u(x; k) = e^{ikx} p(x; k), \quad p(x + 2\pi; k) = p(x; k), \quad x \in \mathbb{R}.$$

This yields the periodic elliptic eigenvalue problem for $p(x; k)$:

$$(5.7) \quad (-\partial_x + ik)^2 + V(x) p(x; k) = E p(x; k), \quad p(x + 2\pi; k) = p(x; k).$$

For each $k \in \mathbb{T}$ the spectrum is discrete, giving rise to eigenpairs $(E_m(k), p_m(x; k))_{m \geq 1}$ and a complete orthonormal set $\{p_m(x; k)\}$ in L^2_{per} with respect to the inner product:

$$(5.8) \quad \langle f, g \rangle_{L^2_{per}} = \int_0^{2\pi} \overline{f(x)} g(x) dx.$$

$(E_n(k), u_n(x; k))$, $n \geq 1$, $k \in [-1/2, 1/2]$ are solutions of the eigenvalue problem,

$$\begin{aligned} (-\partial_x^2 + V(x)) u_n(x; k) &= E_n(k) u_n(x; k), \\ u_n(x + 2\pi; k) &= e^{2\pi i k} u_n(x; k), \quad x \in \mathbb{R}, \end{aligned}$$

where $k \mapsto E_m(k)$ sweeps out the m^{th} spectral band, and yield a complete set of states in $L^2(\mathbb{R})$; see (2.9).

Note. In this section we assume that $w(x)$ is normalized, $\langle w, w \rangle = 1$. Thus, $w(x)$ is the unique normalized ground state of the periodic boundary value problem and

$$(5.9) \quad (p_1(x; 0), E_1(0)) = (p_1(x; 0), E_1(0)) = (w(x), E_*).$$

Furthermore, for each $k \in \mathbb{T}$, the set $\{p_n(x; k)\}$ is an orthonormal set in $L^2_{\text{per}}([0, 2\pi])$.

Introduce the Gelfand–Bloch transform, $(\mathcal{T}\phi)(x; k) = \tilde{\phi}(x; k)$, and its inverse, \mathcal{T}^{-1} :

$$(5.10) \quad \begin{aligned} (\mathcal{T}\phi)(x; k) &= \tilde{\phi}(x; k) = \sum_{m \in \mathbb{Z}^d} e^{im \cdot x} \hat{\phi}(k + m), \\ (\mathcal{T}^{-1}\tilde{\phi})(x) &= \int_{[-\frac{1}{2}, \frac{1}{2}]^d} e^{ik \cdot x} \tilde{\phi}(x; k) dk, \end{aligned}$$

where $\hat{\phi}(k)$ denotes the Fourier transform of $\phi(x)$. Clearly, we have

$$(5.11) \quad \tilde{\phi}(x + 2\pi; k) = \tilde{\phi}(x; k) \quad \text{and} \quad \tilde{\phi}(x; k + 1) = e^{-ix} \tilde{\phi}(x; k).$$

One can check that

$$(5.12) \quad \mathcal{T} \mathcal{T}^{-1} = \text{identity on } L^2(\mathbb{R}).$$

Another important property of \mathcal{T} is that it commutes with multiplication by a periodic function

$$(5.13) \quad f(x + 2\pi) = f(x) \implies (\mathcal{T}fg)(x; k) = f(x) (\mathcal{T}g)(x).$$

Since $\tilde{\phi}(x; k)$ is 2π -periodic in x , we have

$$(5.14) \quad \tilde{\phi}(x; k) = \sum_{m \geq 1} \langle p_m(\cdot; k), \tilde{\phi}(\cdot; k) \rangle p_m(x; k).$$

We conclude this subsection with some basic definitions and results required below; see, for example, [22] and references cited therein.

THEOREM 5.1.

1. *There exist positive constants c_1, c_2 , and band dispersion functions $E_n(k)$, $n \geq 1$, satisfy the bounds⁶*

$$(5.15) \quad c_1 n^2 \leq E_n(k) \leq c_2 n^2, \quad |k| \leq 1/2;$$

see [20, 31].

⁶In dimension d , n^2 is replaced by $n^{\frac{2}{d}}$.

2. *The mapping*

$$(5.16) \quad \phi(x) \mapsto \left(\left\langle \tilde{\phi}(\cdot, k), p_n(\cdot, k) \right\rangle \right)_{n \geq 1} \equiv \left(\tilde{\phi}_n(k) \right)_{n \geq 1}$$

is an isomorphism of $H^s(\mathbb{R}^1)$ with $\mathcal{X}^s = L^2(\mathbb{T}^1; l^{2,s})$, with the following norm:

$$(5.17) \quad \begin{aligned} \left\| \left(\tilde{\phi}_n(k) \right)_{n \geq 1} \right\|_{\mathcal{X}^s}^2 &\equiv \left\| \left(\left\langle \tilde{\phi}(\cdot, k), p_n(\cdot, k) \right\rangle \right)_{n \geq 1} \right\|_{\mathcal{X}^s}^2 \\ &= \int_{\mathbb{T}} dk \sum_{n \geq 1} (1 + |n|^2)^s \left| \left\langle \tilde{\phi}(\cdot, k), p_n(\cdot, k) \right\rangle \right|^2. \end{aligned}$$

3. *Moreover, there exist positive constants C_1, C_2 such that we have the norm equivalence*

$$(5.18) \quad C_1 \|\phi\|_{H^s} \leq \left\| \left\langle \tilde{\phi}(\cdot, k), p_n(\cdot, k) \right\rangle_{n \geq 1} \right\|_{\mathcal{X}^s} \leq C_2 \|\phi\|_{H^s}.$$

4. *Assume $\phi, \psi \in H^s(\mathbb{R}^d)$.*

(a) *If $s > q + d/2$, then $\phi \in C^q_{\downarrow}(\mathbb{R}^d)$, with the space of C^q functions, f , with $|\partial^\alpha f(x)| \rightarrow 0$ as $x \rightarrow \infty$, $|\alpha| \leq q$.*

(b) *If $s > d/2$, then H^s is an algebra; i.e., $\phi\psi \in H^s$ and $\|\phi\psi\|_{H^s} \leq C \|\phi\|_{H^s} \|\psi\|_{H^s}$.*

Remark 5.1. The bounds (5.15) are well known; see [20, 31]. To prove the isomorphism, recall the operator $L_* = -\Delta + V - E_* \geq 0$; see (2.12). Standard elliptic theory implies that $\phi \mapsto \|L_*^{\frac{s}{2}} \phi\|_{L^2}$ defines a norm equivalent to the H^s norm. Furthermore, by (5.14)

$$(5.19) \quad \begin{aligned} \|\phi\|_{H^s}^2 &\sim \|(I + L_*)^{\frac{s}{2}} \phi\|_{L^2}^2 = \left\| \int_{[-\frac{1}{2}, \frac{1}{2}]} e^{ik \cdot} \sum_{n \geq 1} \tilde{\phi}_n(k) (1 + E_* - E_n(k))^s p_j(\cdot, k) \right\|_{L^2}^2 \\ &= \sum_{n \geq 1} \int_{[-\frac{1}{2}, \frac{1}{2}]} |\tilde{\phi}_n(k)|^2 |1 + E_* - E_n(k)|^s dk \\ &\sim \sum_{n \geq 1} \int_{[-\frac{1}{2}, \frac{1}{2}]} |\tilde{\phi}_n(k)|^2 (1 + |n|^2)^s dk \\ &\equiv \left\| \left(\tilde{\phi}_n(k) \right)_{n \geq 1} \right\|_{\mathcal{X}^s}^2. \end{aligned}$$

5.2. Corrector equation and localization in Bloch variables. In this subsection we express the equation for the corrector

$$(5.20) \quad \Psi^\epsilon(x) \equiv U_5^\epsilon(x)$$

in Floquet–Bloch variables and, in particular, decompose this equation into spectral components near and away from the band edge E_* ; see, for example, [14, 21, 22].

Applying the Bloch transform, \mathcal{T} , to (5.4), we obtain an equation for $\tilde{\Psi}^\epsilon(x; k) = (\mathcal{T}\Psi^\epsilon)(x; k)$:

$$(5.21) \quad \begin{aligned} &\left[-(\partial_x + ik)^2 + V(x) - E_* + \epsilon^2 \right] (\mathcal{T}\Psi^\epsilon)(x; k) - 3\epsilon^2 w^2(x) \mathcal{T} [F^2(\epsilon \cdot) \Psi^\epsilon(\cdot)](x; k) \\ &= \epsilon^2 (\mathcal{T}R^\epsilon)(x; k), \end{aligned}$$

where $\epsilon^2 R^\epsilon$ is defined in (5.4). Here we have used that $U_0(x, y) = w(x)F(y)$.

Now $\tilde{\Psi}^\epsilon(x; k)$ is periodic in x . Therefore,

$$(5.22) \quad \tilde{\Psi}^\epsilon(x; k) = \sum_{m=1}^{\infty} \tilde{\Psi}_m^\epsilon(k) p_m(x; k), \quad \tilde{\Psi}_m^\epsilon(k) \equiv \langle p_m(\cdot; k), \tilde{\Psi}^\epsilon(\cdot; k) \rangle.$$

We introduce a decomposition of Ψ^ϵ into spectral components near the band edge $E_1(k = 0) = E_*$ (*low frequencies*) and spectral components away from E_* (*high frequencies*) as follows. Let 1_A denote the characteristic function for the set A and define

$$(5.23) \quad \chi(a \leq k \leq b) \equiv 1_{\{k: a \leq k \leq b\}}.$$

Express $\tilde{\Psi}^\epsilon(x, k)$ as

$$(5.24) \quad \begin{aligned} \tilde{\Psi}^\epsilon(x; k) &= \underbrace{\chi(|k| \leq \epsilon^r) \Psi_1^\epsilon(k) p_1(x; k)}_{\tilde{\Psi}_{low}^\epsilon(x; k)} \\ &+ \underbrace{\chi(\epsilon^r \leq |k| \leq 2^{-1}) \Psi_1^\epsilon(k) p_1(x; k) + \sum_{m \geq 2} \tilde{\Psi}_m^\epsilon(k) p_m(x; k)}_{\tilde{\Psi}_{high}^\epsilon(x; k)}, \end{aligned}$$

where r is chosen to satisfy

$$(5.25) \quad 0 < r < 1.$$

Using the inverse Bloch transform, we obtain

$$\begin{aligned} \Psi^\epsilon(x) &= \mathcal{T}^{-1} \Psi_{low}^\epsilon(x; \cdot) + \mathcal{T}^{-1} \Psi_{high}^\epsilon(x; \cdot) \\ &= \Psi_{low}^\epsilon(x) + \Psi_{high}^\epsilon(x). \end{aligned}$$

Taking the inner product of (5.21) with $p_j(\cdot; k)$, we obtain

$$(5.26) \quad \begin{aligned} [E_j(k) - E_* + \epsilon^2] \tilde{\Psi}_j^\epsilon(k) - 3\epsilon^2 \langle p_j(\cdot; k), w^2(\cdot) \mathcal{T} [F^2(\epsilon \cdot) \Psi^\epsilon(\cdot; k)](\cdot; k) \rangle \\ = \epsilon^2 \langle p_j(\cdot, k), (\mathcal{T} R^\epsilon)(\cdot, k) \rangle \equiv \epsilon^2 (\mathcal{T} R^\epsilon)_j(k), j \geq 1. \end{aligned}$$

The system (5.26) can be viewed as two coupled systems for the following low and high frequencies:

$$(5.27) \quad \begin{aligned} \tilde{\Psi}_{1,low}^\epsilon(k) &\equiv \chi(|k| \leq \epsilon^r) \Psi_1^\epsilon(k) \text{ and} \\ \tilde{\Psi}_{high}^\epsilon(k) &\equiv \left(\chi(\epsilon^r \leq |k| \leq 2^{-1}) \tilde{\Psi}^\epsilon(k), \{ \tilde{\Psi}_j^\epsilon(k) \}_{j \geq 2} \right). \end{aligned}$$

Low-frequency components.

$$(5.28) \quad \begin{aligned} [E_1(k) - E_* + \epsilon^2] \tilde{\Psi}_{1,low}^\epsilon(k) - 3\epsilon^2 \chi(|k| \leq \epsilon^r) \langle p_1(\cdot; k), w^2(\cdot) \mathcal{T} [F^2(\epsilon \cdot) \Psi_{1,low}^\epsilon(\cdot)](\cdot; k) \rangle \\ = 3\epsilon^2 \chi(|k| \leq \epsilon^r) \langle p_1(\cdot; k), w^2(\cdot) \mathcal{T} [F^2(\epsilon \cdot) \Psi_{high}^\epsilon(\cdot)](\cdot; k) \rangle \\ + \epsilon^2 \tilde{R}_{1,low}^\epsilon, \end{aligned}$$

High-frequency components.

$$\begin{aligned}
 & [E_1(k) - E_* + \epsilon^2] \chi(\epsilon^r \leq |k| \leq 2^{-1}) \tilde{\Psi}_1^\epsilon(k) \\
 (5.29) \quad & = 3\epsilon^2 \chi(\epsilon^r \leq |k| \leq 2^{-1}) \langle p_1(\cdot; k), w^2(\cdot) \mathcal{T} [F^2(\epsilon \cdot) \Psi^\epsilon(\cdot)](\cdot) \rangle + \epsilon^2 \tilde{R}_{1,high}^\epsilon, \\
 & [E_j(k) - E_* + \epsilon^2] \tilde{\Psi}_j^\epsilon(k) \\
 (5.30) \quad & = 3\epsilon^2 \langle p_j(\cdot; k), w^2(\cdot) \mathcal{T} [F^2(\epsilon \cdot) \Psi^\epsilon(\cdot)](\cdot; k) \rangle + \epsilon^2 \tilde{R}_{j,high}^\epsilon, \quad j \geq 2.
 \end{aligned}$$

Here $\tilde{R}_{1,low}^\epsilon$ and $\tilde{R}_{high}^\epsilon = (\tilde{R}_{j,high}^\epsilon)_{j \geq 1}$ are given by

$$(5.31) \quad \tilde{R}_{1,low}^\epsilon = \chi(|k| \leq \epsilon^r) \langle p_1(\cdot, k), (\mathcal{T}R^\epsilon)(\cdot, k) \rangle,$$

$$(5.32) \quad \tilde{R}_{1,high}^\epsilon = \chi(\epsilon^r \leq |k| \leq 2^{-1}) \langle p_1(\cdot, k), (\mathcal{T}R^\epsilon)(\cdot, k) \rangle,$$

$$(5.33) \quad \tilde{R}_{j,high}^\epsilon = \langle p_j(\cdot, k), (\mathcal{T}R^\epsilon)(\cdot, k) \rangle, \quad j \geq 2,$$

where R^ϵ is defined in (5.4). We study the system for $\tilde{\Psi}_{1,low}^\epsilon(k)$, $\tilde{\Psi}_{high}^\epsilon(k)$ using the following.

Lyapunov–Schmidt reduction strategy.

1. Using the implicit function theorem, solve the infinite system of high-frequency component equations for $\tilde{\Psi}_{high}^\epsilon$ as a functional of $\tilde{\Psi}_{1,low}^\epsilon$: $\tilde{\Psi}_{high}^\epsilon = \tilde{\Psi}_{high}^\epsilon[\tilde{\Psi}_{1,low}^\epsilon]$ with an appropriate bound on this mapping.
2. Substitute $\tilde{\Psi}_{high}^\epsilon = \tilde{\Psi}_{high}^\epsilon[\tilde{\Psi}_{1,low}^\epsilon]$ into (5.28) to obtain a closed equation for the low-frequency components, which is solved via fixed-point iteration.

We now embark on implementing this strategy. Our first step is to rewrite the low-frequency equation (5.28) in appropriately rescaled variables.

5.3. Closed equation for the low-frequency components $\tilde{\Psi}_{1,low}^\epsilon$. With a view toward obtaining a closed equation for $\tilde{\Psi}_{1,low}^\epsilon$, we begin with several observations.

1. From our formal multiscale construction, we expect $\Psi_{1,low}^\epsilon(x) \sim \Phi(\epsilon x) w(x)$. This motivates the following *ansatz*: seek the low-frequency components in the form

$$(5.34) \quad \tilde{\Psi}_{low}^\epsilon(k) = \chi(|k| \leq \epsilon^r) \frac{1}{\epsilon} \hat{\Phi}\left(\frac{k}{\epsilon}\right).$$

Thus,

$$(5.35) \quad \tilde{\Psi}_{low}^\epsilon(x; k) = \chi(|k| \leq \epsilon^r) \frac{1}{\epsilon} \hat{\Phi}\left(\frac{k}{\epsilon}\right) p_1(x; k).$$

Using the definition of \mathcal{T}^{-1} and that $p_1(x; k) = p_1(x; 0) + \mathcal{O}(k) = w(x) + \mathcal{O}(\epsilon^r)$, $|k| \leq \epsilon^r$, we have

$$(5.36) \quad \Psi_{low}^\epsilon(x) = \Phi(\epsilon x) w(x) + \mathcal{O}(\epsilon^r).$$

2. Note that for $|k| < \epsilon^r$,

$$(5.37) \quad E_1(k) - E_* - \frac{1}{2} \partial_{\tilde{k}}^2 E_1(0) k^2 = \frac{1}{6} \partial_{\tilde{k}}^3 E_1(\tilde{k}) \tilde{k}^3, \quad 0 \leq \tilde{k} \leq \epsilon^r.$$

Thus,

$$\begin{aligned}
 (E_1(k) - E_* - \epsilon^2) \tilde{\Psi}_{1,low}^\epsilon(k) &= \left(\frac{1}{2} \partial_k^2 E_1(0) k^2 - \epsilon^2 \right) \chi(|k| \leq \epsilon^r) \frac{1}{\epsilon} \hat{\Phi} \left(\frac{k}{\epsilon} \right) \\
 &+ \mathcal{O} \left(\|\partial_k^3 E_1\|_\infty \epsilon^{3r} \frac{1}{\epsilon} \chi(|k| \leq \epsilon^r) \hat{\Phi} \left(\frac{k}{\epsilon} \right) \right).
 \end{aligned}
 \tag{5.38}$$

This and the ansatz (5.34) suggest the scaling

$$\kappa \equiv \frac{k}{\epsilon}.
 \tag{5.39}$$

In this scaling (5.38) becomes

$$\begin{aligned}
 (E_1(k) - E_* - \epsilon^2) \tilde{\Psi}_{1,low}^\epsilon(k) &= \epsilon^2 \left(\frac{1}{2} \partial^2 E_1(0) \kappa^2 - 1 \right) \chi(|\kappa| \leq \epsilon^{r-1}) \frac{1}{\epsilon} \hat{\Phi}(\kappa) \\
 &+ \mathcal{O} \left(\epsilon^{3r} \chi(|\kappa| \leq \epsilon^{r-1}) \frac{1}{\epsilon} \hat{\Phi}(\kappa) \right).
 \end{aligned}
 \tag{5.40}$$

3. Consider the last term on the left-hand side of (5.28). We have

$$\begin{aligned}
 &- 3\epsilon^2 \chi(|k| \leq \epsilon^r) \langle p_1(\cdot; k), w^2(\cdot) \mathcal{T} [F^2(\epsilon) \Psi_{1,low}^\epsilon(\cdot)](\cdot; k) \rangle \\
 &= - 3\epsilon^2 \chi(|k| \leq \epsilon^r) \langle p_1(\cdot; k), w^2(\cdot) \mathcal{T} [F^2(\epsilon) \chi(|\nabla_x| \leq \epsilon^r) \Phi(\epsilon) w(\cdot)](\cdot; k) \rangle + \mathcal{O}(\epsilon^{r+2}) \\
 &= - 3\epsilon^2 \chi(|k| \leq \epsilon^r) \langle p_1(\cdot; k), w^3(\cdot) \mathcal{T} [F^2(\epsilon) \chi(|\nabla_x| \leq \epsilon^r) \Phi(\epsilon)](\cdot; k) \rangle + \mathcal{O}(\epsilon^{r+2}) \\
 &= - 3\epsilon^2 \int w^4 dx \cdot \chi(|k| \leq \epsilon^r) [F^2(\epsilon) \chi(|\nabla_x| \leq \epsilon^r) \Phi(\epsilon)]^\wedge(k) + \mathcal{O}(\epsilon^{r+2}) \\
 &= - 3\epsilon^2 \int w^4 dx \cdot \chi(|k| \leq \epsilon^r) \frac{1}{\epsilon} [F^2 \chi(|\nabla_y| \leq \epsilon^{r-1}) \Phi]^\wedge \left(\frac{k}{\epsilon} \right) + \mathcal{O}(\epsilon^{r+2}) \\
 &= - 3\epsilon^2 \int w^4 dx \cdot \chi(|\kappa| \leq \epsilon^{r-1}) \frac{1}{\epsilon} [F^2 \chi(|\nabla_y| \leq \epsilon^{r-1}) \Phi]^\wedge(\kappa) + \mathcal{O}(\epsilon^{r+2})
 \end{aligned}$$

Use of (5.34) and (5.39) in (5.28) yields the following *closed equation for $\hat{\Phi}(\kappa)$* :

$$\begin{aligned}
 &\left(\frac{1}{2} E_1''(0) \kappa^2 + 1 \right) \chi(|\kappa| \leq \epsilon^{r-1}) \hat{\Phi}(\kappa) - 3\gamma_{\text{eff}} [F^2 \chi(|\nabla_y| \leq \epsilon^{r-1}) \Phi]^\wedge(\kappa) \\
 &= \chi(|\kappa| \leq \epsilon^{r-1}) \left[\hat{R}_{\text{rescaled low}}^\epsilon(\kappa; \Phi, \Psi_{\text{high}}^\epsilon) + \mathcal{O} \left(\|\partial_k^3 E_1\|_\infty \epsilon^{3r} |\hat{\Phi}(\kappa)| \right) \right].
 \end{aligned}
 \tag{5.41}$$

Here $E_1 : [-1/2, 1/2] \rightarrow [E_*, E_1(\frac{1}{2})]$, $k \mapsto E_1(k)$ denotes the band dispersion function for the first spectral band, and γ_{eff} is given by the expression in (3.7). Since we have assumed $p_1(x, 0) = w(x)$ to be normalized, we have $\gamma_{\text{eff}} = \int_0^{2\pi} w^4$.

We summarize the arguments of this subsection, which lead to the system we'll study.

PROPOSITION 5.1. *The coupled system, consisting of (5.41) for the rescaled low-frequency components, $\hat{\Phi}(\kappa)$, $|\kappa| \leq \epsilon^{r-1}$ ($|k| \leq \epsilon^r$), and the high-frequency equations, (5.29) and (5.30), is equivalent to the original system.*

5.4. Proof by Lyapunov–Schmidt reduction. We estimate the right-hand sides of the high-frequency equations (5.29) and (5.30).

PROPOSITION 5.2. *Let $s > d/2$. For some positive constants C_1 and C_2 , we have*

$$\begin{aligned}
 &\| 3\epsilon^2 \langle p_j(\cdot; \star), w^2(\cdot) \mathcal{T} [F^2(\epsilon) \Psi^\epsilon(\cdot)](\cdot; \star) \rangle \|_{\mathcal{X}^s} \leq C_1 \epsilon^2 \|\tilde{\Psi}^\epsilon\|_{\mathcal{X}^s}, \\
 &\| \hat{R}_{j,high}^\epsilon \|_{\mathcal{X}^s} \leq C_2 \left(\mathcal{O}(\epsilon^\infty) + \epsilon^3 \|\tilde{\Psi}^\epsilon\|_{\mathcal{X}^s} + \epsilon^5 \|\tilde{\Psi}^\epsilon\|_{\mathcal{X}^s}^2 + \epsilon^8 \|\tilde{\Psi}^\epsilon\|_{\mathcal{X}^s}^3 \right).
 \end{aligned}$$

PROPOSITION 5.3. *The system for $\tilde{\Psi}_{high}^\epsilon$ can be solved in terms of $\tilde{\Psi}_{1,low}^\epsilon = \chi(|k| \leq \epsilon^r) \tilde{\Psi}_{1,low}^\epsilon$, and we have the following estimate:*

$$\begin{aligned}
 & \left\| \tilde{\Psi}_{high}^\epsilon \left[\tilde{\Psi}_{1,low}^\epsilon \right] \right\|_{\mathcal{X}^s} \\
 & \leq C \left(\mathcal{O}(\epsilon^\infty) + \epsilon^{3-2r} \|\chi(|k| \leq \epsilon^r) \tilde{\Psi}_{1,low}^\epsilon\|_{\mathcal{X}^s} + \epsilon^{5-2r} \|\chi(|k| \leq \epsilon^r) \tilde{\Psi}_{1,low}^\epsilon\|_{\mathcal{X}^s}^2 \right. \\
 (5.42) \quad & \left. + \epsilon^{8-2r} \|\chi(|k| \leq \epsilon^r) \tilde{\Psi}_{1,low}^\epsilon\|_{\mathcal{X}^s}^3 \right).
 \end{aligned}$$

Proof. Consider the system for $\tilde{\Psi}_{high}^\epsilon$ (5.29–5.30). The result follows from direct estimation using

$$\begin{aligned}
 & |E_j(k) - E_* + \epsilon^2| \geq c > 0, \quad j \geq 2, \\
 (5.43) \quad & |E_1(k) - E_* + \epsilon^2| \geq \epsilon^{2r}, \quad \epsilon^r \leq |k| \leq 1/2,
 \end{aligned}$$

and applying the implicit function theorem. \square

Using Proposition 5.3 and the rescaled low-frequency equation (5.41) yields the following.

PROPOSITION 5.4. *Equation for $\hat{\phi}(\kappa)$ follows.*

$$\left(\frac{1}{2} E_1''(0) \kappa^2 + 1 \right) \chi(|\kappa| \leq \epsilon^{r-1}) \hat{\Phi}(\kappa) - 3\gamma_{\text{eff}} [F^2 \chi(|\nabla_y| \leq \epsilon^{r-1}) \Phi] (\hat{\kappa}) = \hat{H}^\epsilon,$$

where $\hat{H}^\epsilon = \chi(|\kappa| \leq \epsilon^{r-1}) \hat{H}^\epsilon$ and satisfies the bound

$$\begin{aligned}
 & \|\hat{H}^\epsilon\|_{L^{2,s}} \leq \|\hat{\mathcal{G}}_{\text{sym}}\|_{L^{2,s}} + \mathcal{O} \left(\|\partial^3 E_1\|_\infty \epsilon^{3r} \|\chi(|\kappa| \leq \epsilon^{r-1}) \hat{\Phi}\|_{L^{2,s}} \right) \\
 (5.44) \quad & + \epsilon^\sigma \left(\|\chi(|\kappa| \leq \epsilon^{r-1}) \hat{\Phi}\|_{L^{2,s}} + \|\chi(|\kappa| \leq \epsilon^{r-1}) \hat{\Phi}\|_{L^{2,s}}^3 \right),
 \end{aligned}$$

where $\sigma > 0$ and $\mathcal{G}_{\text{sym}} \in H_{\text{sym}}^s$.

Equivalently, by Fourier transforming one gets

$$(5.45) \quad (-\partial_y A \partial_y - 3\gamma_{\text{eff}} F^2 + 1) \chi(|\nabla_y| \leq \epsilon^{r-1}) \Phi = \chi(|\nabla_y| \leq \epsilon^{r-1}) H^\epsilon,$$

where $A = \frac{1}{2} E_1''(0)$ and

$$\begin{aligned}
 \|H^\epsilon\|_{H^s} & \leq \|\mathcal{G}_{\text{sym}}\|_{H^s} + \mathcal{O} \left(\|\partial^3 E_1\|_\infty \epsilon^{3r} \|\chi(|\nabla_y| \leq \epsilon^{r-1}) \Phi\|_{H^s} \right) \\
 & + \epsilon^\sigma \left(\|\chi(|\nabla_y| \leq \epsilon^{r-1}) \Phi\|_{H^s} + \|\chi(|\nabla_y| \leq \epsilon^{r-1}) \Phi\|_{H^s}^3 \right).
 \end{aligned}$$

We now complete the proof of Theorem 3.1. Denote by L_+^A the operator

$$(5.46) \quad L_+^A \equiv -\partial_y A \partial_y - 3\gamma_{\text{eff}} F^2(y) + 1$$

and

$$(5.47) \quad \chi_\epsilon = \chi(|\nabla_y| \leq \epsilon^{r-1}), \quad \overline{\chi}_\epsilon = 1 - \chi_\epsilon = \chi(|\nabla_y| \geq \epsilon^{r-1}).$$

We recall (see (5.25))

$$(5.48) \quad 0 < r < 1.$$

Equation (5.45) for Φ can be written as

$$(5.49) \quad \chi_\epsilon L_+^A \chi_\epsilon \Phi = \chi_\epsilon H^\epsilon[\Phi].$$

Since F is chosen to be centered at local extremum of the *symmetric potential*, $V(\mathbf{x})$, we have that the mapping

$$(5.50) \quad \Phi \mapsto \chi_\epsilon H^\epsilon[\Phi]$$

maps H_{even}^s to itself.

We claim that for some ϵ_0 , if $\epsilon \leq \epsilon_0$, then the operator $\chi_\epsilon L_+^A \chi_\epsilon : H_{sym}^{s+2} \rightarrow H_{sym}^s$ has an inverse with norm bound which depends only on ϵ_0 . Thus, for $0 \leq \epsilon < \epsilon_1$, we can reformulate (5.49) as

$$(5.51) \quad \Phi = (\chi_\epsilon L_+^A \chi_\epsilon)^{-1} \chi_\epsilon H^\epsilon[\Phi]$$

and show by fixed-point iteration that for some $0 < \epsilon_1 \leq \epsilon_0$ sufficiently small, equation (5.49) has a unique H^{s+2} solution, which is bounded uniformly for $\epsilon \leq \epsilon_1$. This then implies Theorem 3.1.

Therefore, the proof boils down to establishing the invertibility of $\chi_\epsilon L_+^A \chi_\epsilon : H_{sym}^{s+2} \rightarrow H_{sym}^s$. We first prove that $L_+^A : H_{sym}^{s+2} \rightarrow H_{sym}^s$ has a bounded inverse.

Now the operator L_+^A , acting in $L^2(\mathbb{R}^1)$, has a one-dimensional kernel, spanned by the function $\partial_y F$. To see this, differentiate the equation for $F(y)$

$$(5.52) \quad -\partial_y A \partial_y F + F - \gamma_{eff} F^3 = 0$$

and obtain $L_+^A F' = 0$. Moreover, $\text{Ker}(L_+^A) = \text{span}\{F'\}$, since the eigenvalues of a Sturm–Liouville operator are simple. Since F' is odd, L_+^A is an invertible and bounded map from $H_{even}^{s+2}(\mathbb{R}^1)$ to $H_{even}^s(\mathbb{R}^1)$.

Finally, we turn to the invertibility of $\chi_\epsilon L_+^A \chi_\epsilon : H_{sym}^{s+2} \rightarrow H_{sym}^s$ for ϵ sufficiently small. We begin by expressing $\chi_\epsilon L_+^A \chi_\epsilon$ as a perturbation of L_+^A :

$$\begin{aligned} \chi_\epsilon L_+^A \chi_\epsilon &= L_+^A + Q_\epsilon \\ Q_\epsilon &= -(\overline{\chi_\epsilon} L_+^A + L_+^A \overline{\chi_\epsilon}) + \overline{\chi_\epsilon} L_+^A \overline{\chi_\epsilon}. \end{aligned}$$

Therefore, it suffices to prove that

$$(5.53) \quad L_+^A + Q_\epsilon = L_+^A (I + (L_+^A)^{-1} Q_\epsilon)$$

has a bounded inverse defined on H_{sym}^s . A bounded inverse

$$(5.54) \quad (L_+^A + Q_\epsilon)^{-1} = (I + (L_+^A)^{-1} Q_\epsilon)^{-1} (L_+^A)^{-1}$$

exists provided the norm of $(L_+^A)^{-1} Q_\epsilon$ can be made smaller than one. This can be achieved by choosing ϵ sufficiently small as we show. First,

$$(5.55) \quad (L_+^A)^{-1} Q_\epsilon = -(L_+^A)^{-1} \overline{\chi_\epsilon} L_+^A \chi_\epsilon - \overline{\chi_\epsilon}.$$

Concerning the second term, the mapping $f \mapsto \overline{\chi_\epsilon} f$ maps H^s to H^s . If $s > 0$, the operator norm tends to zero as $\epsilon \rightarrow 0$ by explicit calculation using the Fourier transform.

Finally, consider the mapping $f \mapsto (L_+^A)^{-1} \overline{\chi}_\epsilon L_+^A f$. We prove that this mapping is bounded from H^k to $H^{k-\delta}$ for any $\delta > 0$. We see this as follows. Denote by $\langle y \rangle = (1 + |y|^2)^{\frac{1}{2}}$ and define the operator $\langle D \rangle^a$ via

$$(5.56) \quad \langle D \rangle^a f = \int e^{ik \cdot x} \langle k \rangle^a \hat{f}(k) dk.$$

Now, for any $a > 0$, we write

$$(L_+^A)^{-1} \overline{\chi}_\epsilon L_+^A = (L_+^A)^{-1} \langle D \rangle^a \cdot \langle D \rangle^{-a} \overline{\chi}_\epsilon \cdot L_+^A$$

and estimate the norm as follows:

$$\begin{aligned} & \| (L_+^A)^{-1} \overline{\chi}_\epsilon L_+^A \|_{H^{s-a} \leftarrow H^s} \\ & \leq \| (L_+^A)^{-1} \langle D \rangle^a \|_{H^{s-a} \leftarrow H^{s-2}} \| \langle D \rangle^{-a} \overline{\chi}_\epsilon \|_{H^{s-2} \leftarrow H^{s-2}} \cdot \| L_+^A \|_{H^{s-2} \leftarrow H^s}. \end{aligned}$$

Note that the first and third factors are bounded independently of ϵ . We claim that $\| \langle D \rangle^{-a} \overline{\chi}_\epsilon \|_{H^{s-2} \leftarrow H^{s-2}} \rightarrow 0$ as $\epsilon \rightarrow 0$. To see this, calculate as follows:

$$\begin{aligned} \| \langle D \rangle^{-a} \overline{\chi}_\epsilon f \|_{H^\tau}^2 &= \int \langle \kappa \rangle^{-2a} \mathbf{1}_{\{|\kappa| \leq \epsilon^{r-1}\}} \langle \kappa \rangle^\tau |\hat{f}(\kappa)|^2 d\kappa \\ &\leq \epsilon^{2a(1-r)} \| f \|_{H^\tau}^2. \end{aligned}$$

Thus for any $a > 0$, we have $\| \langle D \rangle^{-a} \overline{\chi}_\epsilon \|_{H^\tau \leftarrow H^\tau} \rightarrow 0$. This completes the proof of Theorem 3.1.

Remark 5.2 (general spatial dimensions $d \geq 1$). The proof given readily extends to general dimension $d \geq 1$. One works in spaces $H^s(\mathbb{R}^d)$, $s > d/2$. \mathcal{X}^s is constructed taking into account the behavior of the dispersion functions, $E_n(k)$, in dimension d . As before, we choose $F(\mathbf{y}) = F(\epsilon(\mathbf{x} - \mathbf{x}_0))$, with \mathbf{x}_0 a point of symmetry of $V(\mathbf{x})$. The kernel of $L_+^A = -\partial_{y_i} A^{ij} \partial_{y_j} + F - 3F^2(y)$ has dimension d and is generated by translations; i.e., $\text{Ker}(L_+^A) = \text{span}\{\partial_{y_j} F(y), j = 1, \dots, d\}$ [62, 39]. Since $\text{Ker}(L_+^A)$ is orthogonal to $H_{\text{even}}^2(\mathbb{R}^d)$, L_+^A is invertible mapping from $H^{s+2}(\mathbb{R}^n)$ to H^s .

6. Numerical computations in the semi-infinite gap. Our analytical results apply to solitons with frequencies in a spectral gap, which are also sufficiently close to a spectral band edge. In this section we present the results of numerical computations corroborating the rigorous asymptotic results near the spectral band edge and illustrating their approximate validity further away from the band edge and well into the spectral gap. The details of the numerical methods are discussed in subsection 6.3.

The particular rigorous asymptotic results we explore numerically in detail are the following:

1. the asymptotic structure of soliton’s lying near the edge of the spectral gap (Theorem 3.1):

$$(6.1) \quad u(x, \mu) \approx (E_* - \mu)^{\frac{1}{\sigma}} w(\mathbf{x}) F(\epsilon(\mathbf{x} - \mathbf{x}_0)),$$

where \mathbf{x}_0 is a local extremum of $V(\mathbf{x})$,

2. the asymptotic behavior of the soliton (nonlinear bound state) power, $\mathcal{P}(\mu)$, along minima- and maxima-centered solitons as μ approaches E_* ; see Theorem 3.2 and Corollary 3.1—in particular, in the critical case $\sigma = 2/d$, we have

$$(6.2) \quad \mathcal{P}(\mu) \approx \zeta_* \mathcal{P}_{cr}.$$

We focus on the one-dimensional NLS/GP equation (1.1) with critical nonlinearity and periodic potential governing $\psi(x, t)$ and nonlinear bound states: $\psi(x, t) = e^{-i\mu t}u(x, \mu)$.

For $d = 1, \sigma = 2$,

$$(6.3) \quad i\partial_t\psi = -\partial_x^2\psi + V_0 \cos^2(2\pi x)\psi - |\psi|^4\psi,$$

$$(6.4) \quad \mu u = -\partial_x^2u + V_0 \cos^2(2\pi x)u - |u|^4u,$$

where V_0 is the variation or contrast of the potential.

We have observed similar results to those presented below for the two-dimensional critical NLS/GP with periodic potential: $d = 2, \sigma = 1$:

$$(6.5) \quad i\partial_t\psi_t = -(\partial_x^2 + \partial_y^2)\psi + \frac{V_0}{2} [\cos^2(2\pi x) + \cos^2(2\pi y)]\psi - |\psi|^2\psi,$$

$$(6.6) \quad \mu u = -(\partial_x^2 + \partial_y^2)u + \frac{V_0}{2} [\cos^2(2\pi x) + \cos^2(2\pi y)]u - |u|^2u.$$

Theorems 3.1 and 3.2 and Corollary 3.1 apply to (6.4) with states centered at a minimum: $x_0 = .25$ or maximum: $x_0 = 0$. These results also apply to (6.6) with states centered at a minimum: $\mathbf{x}_0 = (.25, .25)$, at a maximum: $\mathbf{x}_0 = (0, 0)$, and at a saddle point: $\mathbf{x}_0 = (.25, 0)$ or $\mathbf{x}_0 = (0, .25)$.

6.1. Soliton profiles: Asymptotic theory and computation. Figures 6.1 and 6.2 display nonlinear bound state profiles of the one-dimensional NLS/GP equation (1.1) for values of μ in the semi-infinite gap of the Schrödinger operator: $\partial_x^2 + V_0 \cos(Kx)$; i.e.,

$$(6.7) \quad \mu \in (-\infty, E_*), \quad \mu < E_* = E_*(V_0, K),$$

both near and far from the band edge.

Plots A1, A2, and A3 in Figure 6.1 display the case of solitons centered at local minima of the potential with, from left to right, frequency μ approaching E_* , at distances $E_* - \mu = 10, 1, \text{ and } 0.01$, respectively. Plots B1, B2, and B3 in Figure 6.1 correspond to the case of solitons centered at local maxima of the potential.

We first note that the figures show the expected trend toward increased localization as $\mu < 0$ is decreased. For μ large and negative the solitons centered at maximum or minima approach a scaled $V \equiv 0$ soliton; see (1.18).

Our main analytical results apply to solitons whose frequencies lie near the band edge, although numerical studies indicate their approximate validity some distance away from the band edge.

Theorem 3.1 implies that nonlinear bound states are, to leading order in the distance to the spectral band edge, a product of a linear Bloch state with band-edge energy and a soliton in an effective homogeneous medium; see (3.3) and (1.18):

$$(6.8) \quad \begin{aligned} u(x, \mu) &\approx w(x)F(y), \\ F(y) &= \left(\frac{E_* - \mu}{\gamma_{\text{eff}}}\right)^{1/4} \text{sech}^{\frac{1}{2}}\left(2\sqrt{m_*(E_* - \mu)}(x - x_0)\right). \end{aligned}$$

The centering of the soliton is x_0 , a point of symmetry of the potential, $V(\mathbf{x})$.

The maximum of the Bloch modes $w(x)$ (normalized to be positive and with unit mass) occurs at the minimum of the potential. However, depending on the centering

point x_0 , $F(y)$ has a maximum (minimum) at the potential maximum (minimum). Thus, bound states centered on a potential minimum are approximately a product of functions that peak at the same values of x yielding a more peaked bound state; compare the top and bottom panels of Figure 6.1 with $E_* - \mu = 10$.

Figure 6.1 (A2, B2) shows that for $E_* - \mu = 1$, the bound states have discernible oscillations about a positive envelope, reflecting the solutions leading order behavior (6.8). These oscillations can be understood as a result of the “underlying” Bloch modes in (6.8). Here, as above, the asymptotic theory appears to capture the structure of the bound states even when μ is not very close to the band edge.

We note as well: for the soliton centered at the potential’s local maximum, a transition in the profile from *single-humped* to *double-humped* (having a dimple at $x = 0$) as μ decreases through $\mu = \mu_\#$, the value at which $\mathcal{P}[u(\cdot, \mu)]$, *along the branch of solitons centered at a local maximum of V* , achieves its maximum; see Figure 6.3. A related observation is made in [1].

Comparing Figure 6.1 panels (A3) and (B3) shows that near the band edge ($E_* - \mu = 0.01$) there is almost no visible difference between the bound states centered at potential minima and those centered at potential maxima. This is clear from (6.8), since in this regime $F(y) \sim$ decays only on a length scale much larger than the period of $V(x)$ and thus, for both maxima- and minima-centered solitons, $u(x) \sim$ constant $\times w(x)$.

Figure 6.2 shows a direct comparison between the asymptotic theory, i.e., the leading order solution (3.3) near the band edge, and the “actual” bound state profiles, computed by solving the bound state differential equation (6.4) with high accuracy.

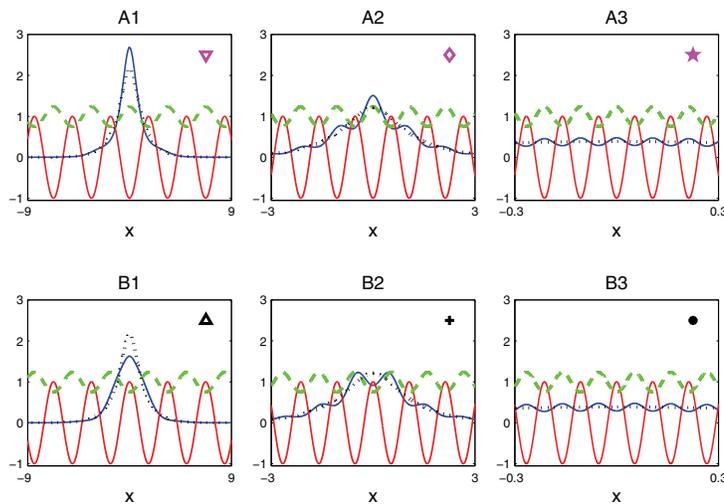


FIG. 6.1. Leading order asymptotic profiles obtained via Theorem 3.1 (solid lines decaying in amplitude) for bound states that are centered on the potential minimum (top panel) and on the potential maximum (bottom panel). Here (A1, B1) $E_* - \mu = 10$, (A2, B2) $E_* - \mu = 1$, and (A3, B3) $E_* - \mu = 0.01$. For clarity only a small portion of the domain is shown and the x -axes are zoomed in as $E_* - \mu$ decreases. Note that the asymptotic profiles in (A3, B3) decay in amplitude on a large scale (only visible outside the domain shown). Also shown in all six panels are the scaled potential $V(x)/V_0$ (solid lines that oscillate periodically in amplitude between -1 and 1), the Bloch wave $w(x)$ (dashed lines), and the rescaled homogeneous ground state $F(y) = F[\epsilon(x - x_0)]$ ((6.8), dotted lines). Geometric shapes correspond to those depicted in Figures 6.2 and 6.3.

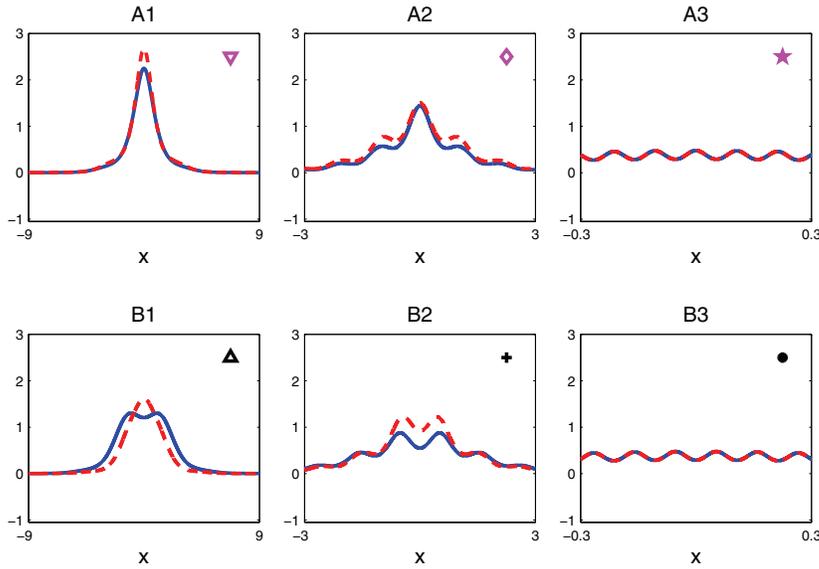


FIG. 6.2. Bound state profiles ((1.6), solid) computed using the renormalization method compared with the leading order asymptotic theory ((3.3), dashes) for the same parameters as in Figure 6.1.

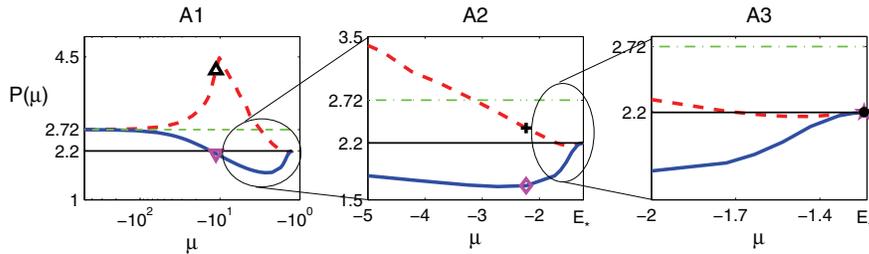


FIG. 6.3. Power μ plot for (6.4) with $V_0 = 10$ and $K = 2\pi$ for bound states centered on a maximum (dashes) and minimum (solid) of the potential. (A1), Wide view of the semi-infinite gap (semi-log μ -axis). (A2) and (A3), Zooming in near the band edge. The asymptotically computed value at the band edge, $P_{\text{edge}} = \zeta_* \times \mathcal{P}_{cr}$ ((3.15), solid line) and critical power \mathcal{P}_{cr} for the homogeneous (translation invariant) equation (dots) are shown as well. Geometric shapes correspond to the cases whose bound states are depicted in Figure 6.1.

6.2. Effective mass and the power curve $\mu \mapsto \mathcal{P}[u(\cdot, \mu)]$. Each plot in Figure 6.3 shows two curves of bound state power, $\mathcal{P}[u(\cdot, \mu)]$, in the semi-infinite gap as a function of μ for (6.4) with $V_0 = 10, K = 2\pi$. The solid curve corresponds to the variation of $\mathcal{P}[u(\cdot, \mu)]$ for the family of solitons centered at the potential's local minimum and the dashed curve for the family centered at the potential's local maximum.

We observe the following.

1. Panel (A1) of Figure 6.3 shows the variation of \mathcal{P} over a wide range of μ in the semi-infinite gap. As $\mu \rightarrow -\infty$, i.e., in the semi-classical limit, the power approaches the homogeneous ($V \equiv 0$) power of the ground state, $\mathcal{P}_{cr} = \frac{\sqrt{3}}{2}\pi \approx 2.72$.

2. Panels (A2) and (A3) of Figure 6.3 show, as predicted by Corollary 3.1, that as $\mu \rightarrow E_*$ the $\mathcal{P}[u(\cdot, \mu)]$ approaches the value $\zeta_* \mathcal{P}_{cr} \approx 2.2$ strictly less than $\mathcal{P}_{cr} \approx 2.72$. Here $\zeta_* \approx 2.2/2.72 \approx 0.8$. This is true for solitons centered at either potential minima or potential maxima; see Figure 6.3 (A3).
3. Although the asymptotic behavior of the \mathcal{P} for maxima- and minima-centered bound states is the same, across most of the gap bound states centered on lattice minima (resp., maxima) have power below (resp., above) $\mathcal{P}_{cr} \approx 2.72$.
4. Panels (A1) and (A2) of Figure 6.3 show a transition in the slope of the minima- and maxima-centered power curves at the *same* value $\mu \equiv \mu_\#$. As discussed in section 2.1, the transition in slope of $\mu \mapsto \mathcal{P}[\mu]$ along the power curve for minima-centered solitons signals a transition from the unstable (positive slope) to stable (negative slope) regime. Maxima-centered solitons, as discussed, are unstable, and the transition in slope signals a change in the number of unstable eigenvalues of the linearized problem [32, 27].

6.3. Numerical methods. The computation of the asymptotic bound states and the “actual” bound states are carried out using Matlab and Octave.

Computation of the Bloch mode, inverse effective mass, and ζ_* . The Bloch mode at the band edge, w , is computed using an eigenvalue solver within a single lattice cell (see [53, appendix], on using Matlab’s eigenvalue solver). For convenience we normalize the Bloch mode to have unit mass; i.e., $\int w^2 = 1$. The inverse effective mass tensor A^{ij} is computed by employing Matlab’s linear system solver for L_*^{-1} . It is then straightforward to compute the inverse effective mass (curvature) m_*^{-1} , coupling constant γ_{eff} , and the band-edge power factor ζ_* using (3.8) and (3.17).

Computation of the bound state at the band edge. The asymptotic bound state is composed of the Bloch mode and the rescaled homogeneous solution. The Bloch mode is obtained by periodically extending w from one lattice cell to the domain over which the bound state is computed—typically several hundred lattice cells. The rescaled homogeneous ground state, F , and its power, \mathcal{P}_{cr} , are computed in one dimension using the explicit solution, i.e., (1.18) with the rescaling in (6.8). Finally, the asymptotic bound state is obtained by shifting F to be centered at point of symmetry of the potential and taking its product with the Bloch mode.

Computation of the “actual” bound states. The bound states of (1.6) are computed using renormalization method [3]. This method is based on fixed-point iterations coupled to an algebraic condition whose role is to constraint the solution to a suitable integral identity consistent with the bound state (otherwise, the iterative solution would diverge). The convergence is monitored by the L_∞ norm of successive iterations and by relative change of the renormalization constant. For example, for a one-dimensional computation with $|\Omega| = 0.01$, the domain size is a few hundred lattice cells. We use 2^{16} grid points to resolve well the oscillations on the scale of the potential period. The computation of the bound state is considered to have converged when the difference between successive iterations satisfies $\|u^{n+1}(\mathbf{x}) - u^n(\mathbf{x})\|_\infty < 10^{-8}$. This typically happens within fewer than 100 iterations (a few minutes).

The renormalization method needs to be seeded with an initial guess. Deep inside the gap the renormalization method converges when seeded by a Gaussian (or sech) profile. On the other hand, near the band edge the method diverges when seeded by a Gaussian or sech, which are apparently too far from the basin of attraction of the bound state. We overcome this difficulty by seeding the renormalization method with the asymptotic solution.

7. Summary and discussion. In this paper we have studied the bifurcation of small amplitude ($H^s(\mathbb{R}^d)$, $s > d/2$) nonlinear bound states (solitary waves or “solitons”) of the NLS/GP equation with a periodic and symmetric potential. Our results provide insight into questions (Q1–Q3) of the introduction. We now briefly summarize our results, with reference to (Q1–Q3).

Concerning (Q1):

1. A family of bifurcating solitons (spatially localized standing wave states) can be constructed centered at any point of symmetry, \mathbf{x}_0 , of $V(\mathbf{x})$.
2. Solitons with frequencies near a spectral band edge have a two-scale structure: $u_\epsilon(\mathbf{x}) \approx \epsilon^{\frac{1}{\sigma}} F(\epsilon(\mathbf{x} - \mathbf{x}_0)) w(\mathbf{x})$, where $\epsilon^2 = |E_* - \mu|$ is the distance of the frequency to the spectral band edge.

Concerning (Q2):

1. We prove, in general, that the limit of the soliton power, along any family of solitons centered at a point of symmetry of $V(\mathbf{x})$, is strictly less than the power of the Townes soliton:

$$(7.1) \quad \lim_{\mu \rightarrow E_*} \mathcal{P}[u(\cdot, \mu)] = \zeta_* \mathcal{P}_{cr} < \mathcal{P}_{cr}.$$

Note: This limit is independent of the centering of the soliton, \mathbf{x}_0 .

2. We prove a high-order expansion, which is necessary to capture information about the slope of the curve, $\mu \mapsto \mathcal{P}[u(\cdot, \mu)]$, near the band edge. Encoded in the slope of this curve is information on nonlinear dynamic stability. We conjecture that for critical nonlinearities ($\sigma = 2/d$), the curve has positive slope near the band edge; therefore, solitons with frequencies near the band edge are unstable. We have verified this analytically for low-contrast potentials and numerically for a range of potentials without a smallness constraint on the contrast.
3. Our analytical results concerning the multiple-scale structure of solitons of NLS/GP and the curve $\mu \mapsto \mathcal{P}[u(\cdot, \mu)]$ are corroborated through careful numerical experiments.

Concerning (Q3), see Remark 3.3. In particular, see Figure 1.1 and the soliton excitation threshold conjecture.

We conclude this section with a discussion the emergent parameter, ζ_* , appearing in (7.1). From (3.7) and (3.17) we have

$$(7.2) \quad \zeta_* = \left(\frac{1}{m_*}\right)^{\frac{1}{2}} \left(\frac{(\int_{\mathcal{B}} w^2)^{\sigma+1}}{\int_{\mathcal{B}} w^{2\sigma+2}}\right)^{\frac{1}{\sigma}} = \left(\frac{1}{m_*}\right)^{\frac{1}{2}} \left(\frac{1}{\gamma_{\text{eff}}}\right)^{\frac{1}{\sigma}} \frac{1}{\text{vol}(\mathcal{B})}.$$

Here m_* denotes the determinant of the *effective mass* tensor; $w(\mathbf{x})$, the \mathcal{B} -periodic Bloch (band edge) state; γ_{eff} , the effective nonlinear coupling; and $\text{vol}(\mathcal{B})$, the volume of the fundamental periodic cell.

A matter of practical/experimental interest is that the parameters γ_{eff} and m_* are tunable via appropriate design of periodic structure, $V(\mathbf{x})$, therefore making it possible to manipulate the power curve, \mathcal{P} vs. μ . Figure 7.1 displays $m_*^{-1/2}$ and ζ_* as functions of the potential contrast V_0 in one and two dimensions. All three quantities are bounded between 0 and 1 and decrease monotonically with V_0 . In particular, this means that $\mathcal{P}_{\text{edge}}/\mathcal{P}_{cr}$ decreases with V_0 . This decrease is “faster” in one dimension than in two dimensions, at least for $V_0 < 40$.

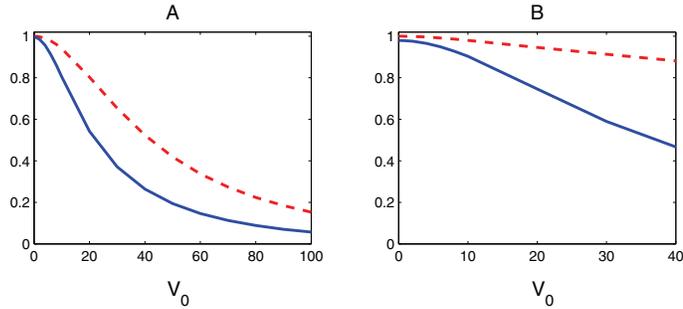


FIG. 7.1. Power factor ζ_* (solid) and $m_*^{-\frac{1}{2}}$ (dashes) as functions of the potential contrast V_0 in (A) one dimension with $\sigma = 2$ ((6.4) with $K = 2\pi$) and (B) two dimensions with $\sigma = 1$ ((6.6) with $K_x = K_y = 2\pi$).

Appendices.

A. Effective mass tensor. In this section we prove (3.6), which relates $D^2 E_1(0)$, the Hessian matrix of the band dispersion function E_1 , to the matrix A^{ij} arising in the multiple-scale analysis.

Denote by $e^{i\mathbf{k}\cdot\mathbf{x}}\phi(\mathbf{x}; \mathbf{k})$ the Bloch state associated with $E_1(\mathbf{k}) : \mathbf{k} \in \mathcal{B}^* \rightarrow \mathbb{R}$; see subsection 2.2. Here, $\phi(\mathbf{x}; \mathbf{k})$ is periodic. Thus,

$$(A.1) \quad \begin{aligned} (-\Delta - 2i\mathbf{k} \cdot \nabla + |\mathbf{k}|^2 + V) \phi &= E_1(\mathbf{k})\phi, \\ \phi(\mathbf{x} + \mathbf{q}_j; \mathbf{k}) &= \phi(\mathbf{x}; \mathbf{k}), \quad j = 1, \dots, d. \end{aligned}$$

At the band gap edge, one has

$$(A.2) \quad E_1(0) = E_*, \quad \phi(\mathbf{x}; 0) \equiv w(\mathbf{x}),$$

where $w(\mathbf{x})$ is the ground state of $-\Delta + V$ subject to periodic boundary conditions on \mathcal{B} .

We denote $f_{,k_j} \equiv \partial_{k_j} f$ and $f_{,r} \equiv \partial_{x_r} f$. Differentiation of (A.1) with respect to k_i gives

$$(A.3) \quad \begin{aligned} (-\Delta - 2ik_i \partial_{x_i} + |\mathbf{k}|^2 + V) \phi_{k_i} \\ = E_{1,k_i} \phi + E_1 \phi_{,k_i} + 2i\phi_{,x_i} - 2k_i \phi. \end{aligned}$$

At $\mathbf{k} = 0$, (A.2) and (A.3) yield

$$(A.4) \quad L_* \phi_{,k_i}(\mathbf{x}; 0) = E_{*,k_i} w + 2iw_{,i}.$$

Removing secular terms from (A.6) leads to

$$(A.5) \quad E_{*,k_i} = 0.$$

It follows from (A.4) and (A.5) that

$$(A.6) \quad \phi_{,k_i}(\mathbf{x}; 0) = 2iL_*^{-1}w_{,i}.$$

Differentiating (A.3) with respect to k_j , setting $\mathbf{k} = 0$ and using (A.2), we arrive at

$$\begin{aligned} L_* \phi_{,k_i k_j}(\mathbf{x}; 0) &= (E_{*,k_i k_j} - 2\delta_{ij}) w + [E_{*,k_i} \phi_{,k_j}(\mathbf{x}; 0) + E_{*,k_j} \phi_{,k_i}(\mathbf{x}; 0)] \\ &\quad + 2i [\phi_{,x_i k_j}(\mathbf{x}; 0) + \phi_{,x_j k_i}(\mathbf{x}; 0)]. \end{aligned}$$

Using (A.5) and (A.6) gives

$$L_* \phi_{*,k_i k_j}(\mathbf{x}; 0) = (E_{*,k_i k_j} - 2\delta_{ij}) w - 4(\partial_{x_i} L_*^{-1} \partial_{x_j} w + \partial_{x_j} L_*^{-1} \partial_{x_i} w).$$

Removing the secular growth requires that the inner product of the right-hand side with $w(\mathbf{x})$ vanish, i.e.,

$$\begin{aligned} (E_{*,k_i k_j} - 2\delta_{ij}) \langle w, w \rangle - 4\langle \partial_{x_i} L_*^{-1} \partial_{x_j} w, w \rangle \\ - 4\langle \partial_{x_j} L_*^{-1} \partial_{x_i} w, w \rangle = 0. \end{aligned}$$

Using the fact that L_*^{-1} is self-adjoint, the last two terms are equal to each other. Therefore, using integration by parts leads to

$$\frac{1}{2} \left. \frac{\partial^2 E_1}{\partial k_i \partial k_j} \right|_{\mathbf{k}=0} = \delta_{ij} - 4 \frac{\langle L_*^{-1} \partial_{x_j} w, \partial_{x_i} w \rangle}{\langle w, w \rangle} \equiv A^{ij}.$$

This proves the relation (3.6). \square

B. Bound on determinant of effective mass tensor.

PROPOSITION B.1. $m_*^{-1} \equiv \det(2^{-1} E_{1,k_i k_j}(0)) \leq 1$, with $m_* = 1$ only if $V(\mathbf{x})$ is constant.

Proof. For the proof we use $m_* > 0$; see [34]. Recall that

$$(B.1) \quad \frac{1}{2} E_{*,k_i k_j} = \delta_{ij} - B_{ij}, \quad B_{ij} \equiv \frac{4\langle L_*^{-1} w_{,j}, w_{,i} \rangle}{\langle w, w \rangle}.$$

We claim that B_{ij} is positive definite. To see this, first recall that $L_* \geq 0$ with one-dimensional $L^2(\mathbb{T}^d)$ kernel spanned by w . Clearly, $w \perp M \equiv \text{span}\{w_{,i} : i = 1, \dots, d\}$, and therefore, B is well defined.

Let $\mathbf{v} = (v_1, \dots, v_d) \in \mathbb{R}^d$ be arbitrary. Then

$$(B.2) \quad \mathbf{v} \cdot B \mathbf{v} = \langle L_*^{-1} \mathbf{v} \cdot \nabla w, \mathbf{v} \cdot \nabla w \rangle \geq \lambda_2^{-1} \|\mathbf{v} \cdot \nabla w\|^2 \geq C \|\mathbf{v}\|^2,$$

where $\lambda_2 > 0$ denotes the second eigenvalue of L_* acting on $L^2(\mathbb{T}^d)$.

The matrix $E_{*,k_i k_j}$ is positive definite [34]. Therefore, $E_{*,k_i k_j}$ can be diagonalized by a unitary transformation p_{ij} such that

$$p_{ri} E_{*,k_i k_j} p_{lj} = 2(1 - \beta_r) \delta_{rl},$$

where λ_i ($i = 1 \dots d$) are the eigenvalues of B_{ij} . It follows that

$$(B.3) \quad m_*^{-1} = \det\left(\frac{E_{*,k_i k_j}}{2}\right) = \prod_{i=1}^d (1 - \beta_i),$$

where $\beta_i > 0$, $i = 1, \dots, d$. In order to bound m_*^{-1} from above, we will show $\beta_i \in (0, 1)$, $i = 1, \dots, d$, and therefore $m_*^{-1} \leq 1$.

We argue by continuity. Consider the self-adjoint operator $L_*^\theta = -\Delta + V(\mathbf{x}; \theta)$ with the one-parameter family of potentials $V(\mathbf{x}; \theta) \equiv \theta V(\mathbf{x})$, where $\theta \in [0, 1]$, as well as the associated matrix $B_{ij}(\theta)$. Since L_*^θ is self-adjoint and $w(\mathbf{x}; \theta)$, its ground state, is simple, there are d continuous functions $\theta \rightarrow \beta_i(\theta)$, $i = 1, \dots, d$, defining the eigenvalues of $B_{ij}(\theta)$. For $\theta = 0$ (homogeneous medium), $E_* = 0$, $E_1(\mathbf{k}) = \mathbf{k}^2$, and $E_{*,k_i k_j} = 2\delta_{ij}$. Therefore, $B_{ij} = 0$ and $\beta_j(0) = 0$, $j = 1, \dots, d$. In this case (and

only in this case!), $m_*^{-1} = \det I = 1$. Next, consider $\theta = 1$, i.e., the original problem. We claim that $\beta_i(1) < 1$, $i = 1, \dots, d$. Otherwise, at some value of $\theta = \theta_* > 0$ an eigenvalue of $B(\theta_*)$ would attain the value 1. This would contradict the positive definiteness of E_{*,k_i,k_j}^θ . \square

C. Effective mass for $d = 1$ and the Floquet–Hill discriminant. In one space dimension the endpoints of the spectral bands are obtained by studying the periodic and antiperiodic eigenvalue problems [23]. Very briefly, for each E one constructs a 2×2 fundamental solution matrix, $M(x; E)$, and considers the values of E for which $M(q; E)$ has an eigenvalue $+1$ or -1 , corresponding to periodic or antiperiodic eigenvalues. This is equivalent to $\Delta(E) = \pm 2$, where

$$(C.1) \quad \Delta(E) \equiv \text{trace}(M(q; E)) \equiv 2 \cos[k(E)q]$$

is the Floquet discriminant.

The band edge, $E = E_*$, corresponds to $k = 0$, at which we have $\Delta(E_*) = 2$. Expanding (C.1) in Taylor series around $k = 0$ and $E = E_*$ gives

$$(C.2) \quad \Delta(E_*) + \Delta'(E_*)(E - E_*) + \mathcal{O}[(E - E_*)^2] = 2 \left(1 - \frac{k^2 q^2}{2} \right) + \mathcal{O}(k^4).$$

Using $\Delta(E_*) = 2$ and solving for the second term on the left-hand side gives to leading order

$$(C.3) \quad E - E_* = -\frac{k^2 q^2}{\Delta'(E_*)} + \mathcal{O}(k^4)$$

which yields the relation

$$(C.4) \quad m_*^{-1} = E_1''(0) = -\frac{2q^2}{\Delta'(E_*)}.$$

Since $\Delta'(E_*) < 0$, $m_* > 0$. More generally, we have that $E_j''(0) > 0$ at the *left* edge of each band and $E_j''(0) < 0$ at the *right* edge of each band.

D. Power and slope for small potentials. In this section we use a regular perturbation expansion to derive the power and slope constants near the band edge, i.e., ζ_* and ζ_{1*} , assuming a small potential. Such an expansion can be made rigorous by an argument based on the implicit function theorem. The derivation is composed of preliminary calculations in any dimension of the Bloch function, an inverse linear operator, and the inverse effective mass tensor and coupling constant. To simplify notation, subsequent calculations are carried out explicitly in the critical case ($d = 1, \sigma = 2$).

Remark D.1. In the derivation below δ is assumed to be a small constant independently of ϵ . The calculations are done to order $\mathcal{O}(\epsilon^2 \delta^m)$ for suitable m . For convenience, the ϵ^2 is suppressed from $\mathcal{O}(\cdot)$.

Let $V(\mathbf{x}) = \delta V_1(x)$, where $|\delta| \ll 1$. Without loss of generality we assume that $\langle V_1 \rangle \equiv \int_{\mathcal{B}} V_1(\mathbf{x}) d\mathbf{x} = 0$. Let $w_\delta \equiv w_*(\mathbf{x}; \delta)$ and $E_* = E_*(\delta)$ be the ground state eigenpair of

$$(D.1) \quad L_\delta w_*(\mathbf{x}; \delta) = 0, \quad L_\delta \equiv L_0 + \delta V_1 - E_*(\delta), \quad L_0 \equiv -\Delta,$$

with \mathcal{B} -periodic boundary conditions.

We expand $w_\delta(\mathbf{x})$ and $E_*(\delta)$ in a Taylor expansion in δ ,

$$w_\delta \equiv w_0 + \delta w_1 + \delta^2 w_2 + \dots, \quad E_*(\delta) = \delta E_1 + \delta^2 E_2 + \dots,$$

where $w_k \equiv w_k(\mathbf{x})$, and we set $E_0 = 0$ since we are interested in bifurcation from the lowest band edge. The first three terms in the hierarchy are

$$(D.2) \quad O(\delta^0) : L_0 w_0 = 0,$$

$$(D.3) \quad O(\delta^1) : L_0 w_1 = (E_1 - V_1) w_0,$$

$$(D.4) \quad O(\delta^2) : L_0 w_2 = E_2 w_0 + (E_1 - V_1) w_1.$$

Corresponding to the lowest band edge, (D.2) admits a constant solution $w_0(\mathbf{x}) = \text{const}$ that spans the kernel of L_0 . Without loss of generality we may choose this constant such that $\langle w_0(\mathbf{x}) \rangle = 1$. In order to remove secular growth the nonhomogeneous terms in (D.3) and (D.4) must be orthogonal to $w_0(\mathbf{x})$. Therefore, their cell-average must vanish. Removing secular terms at $O(\delta^1)$ gives

$$E_1 = \langle V_1 \rangle = 0, \quad w_1 = -L_0^{-1} V_1.$$

Substituting the above results into (D.4) and removing secular growth yields

$$E_2 = -\langle V_1 L_0^{-1} V_1 \rangle, \quad w_2 = \{ L_0^{-1} (V_1 L_0^{-1} V_1) \},$$

where the brace is a projection symbol defined as

$$\{ f \} \equiv f - \langle f \rangle.$$

Summarizing the above results gives

$$(D.5) \quad L_\delta = L_0 + \delta V_1 + O(\delta^2),$$

$$(D.6) \quad w_\delta = 1 - \delta L_0^{-1} V_1 + \delta^2 \{ L_0^{-1} (V_1 L_0^{-1} V_1) \} + O(\delta^3),$$

$$(D.7) \quad E_*(\delta) = -\delta^2 \langle V_1 L_0^{-1} V_1 \rangle + O(\delta^4).$$

We now approximate the inverse operator L_δ^{-1} . It is expedient to make the following definitions.

DEFINITION D.1. We denote the domains of L_0^{-1} and L_δ^{-1} as

$$K_0 \equiv \{ f(\mathbf{x}) | \mathbf{x} \in \mathcal{B}, f \text{ is periodic in } \mathcal{B}, f \in \text{Ker}^\perp L_0 \},$$

$$K_\delta \equiv \{ f(\mathbf{x}) | \mathbf{x} \in \mathcal{B}, f \text{ is periodic in } \mathcal{B}, f \in \text{Ker}^\perp L_\delta \},$$

respectively. We denote the projection operator into K_δ as

$$(D.8) \quad P_\delta F_\delta \equiv F_\delta - \frac{\langle F_\delta, w_\delta \rangle}{\langle w_\delta, w_\delta \rangle} w_\delta.$$

Using the above definitions we obtain the following.

LEMMA D.2. Let $F_\delta = F_0 + \delta F_1 + O(\delta^2)$. Then

$$L_\delta^{-1} P_\delta F_\delta = L_0^{-1} \{ F_0 \} + \delta [L_0^{-1} \{ F_1 \} + \langle F_0 \rangle L_0^{-2} V_1 - L_0^{-1} \{ V_1 L_0^{-1} \{ F_0 \} \}] + O(\delta^2).$$

Proof. The proof of Lemma D.2 follows directly from expanding the projection operator (D.8) in powers of δ , using (D.6), $L_0^{-1} : K_0 \rightarrow K_0$, and $\langle L_0^{-1} f \rangle = 0$. \square

In the derivations of ζ_* and ζ_{*1} , in each and every case that L_δ^{-1} is applied, one has $F_0 = 0$ and $\langle F_1 \rangle = 0$. Hence, we shall use Corollary D.1.

COROLLARY D.1. *It follows from Lemma D.2 that*

$$(D.9) \quad F_0 = 0, \langle F_1 \rangle = 0 \implies L_\delta^{-1}F_\delta = \delta L_0^{-1}F_1 + O(\delta^2).$$

In addition, the following approximations of the expressions related to the inverse effective mass tensor and effective nonlinear coupling constant are used in the approximation of ζ_{1*} :

$$(D.10) \quad \begin{aligned} X_1^{ij}(\mathbf{x}) &= (\delta^{ij} + 4\partial_{x_j}L_*^{-1}\partial_{x_i} - A^{ij})w(\mathbf{x}) \\ &\stackrel{(3.6)}{=} 4\left(\partial_{x_j}L_*^{-1}\partial_{x_i} + \frac{\langle \partial_{x_j}w, L_*^{-1}\partial_{x_i}w \rangle}{\langle w, w \rangle}\right)w(\mathbf{x}) \\ &\stackrel{(D.8)}{=} P_\delta(4\partial_{x_i}L_\delta^{-1}\partial_{x_j}w_\delta) \stackrel{(D.6),(D.9)}{=} -4\delta\partial_{x_i}\partial_{x_j}L_0^{-2}V_1 + O(\delta^2), \end{aligned}$$

$$(D.11) \quad \gamma_{\text{eff}} \stackrel{(3.7)}{=} \frac{\int_{\mathcal{B}} w^{2\sigma+2} d\mathbf{x}}{\int_{\mathcal{B}} w^2 d\mathbf{x}} \stackrel{(D.6)}{=} 1 + O(\delta^2),$$

and therefore,

$$(D.12) \quad X_2(\mathbf{x}) = w^{2\sigma+1} - \gamma_{\text{eff}}w \stackrel{(D.6),(D.11)}{=} 2\sigma\delta L_0^{-1}V_1 + O(\delta^2).$$

We proceed to calculate the power and its slope. For simplicity we consider the critical case ($d = 1, \sigma = 2$).

Calculation of the power constant ζ_* . Expanding, using (D.6) with ($d = 1, \sigma = 2$), gives

$$\begin{aligned} w_\delta^2 &= 1 - 2\delta L_0^{-1}V_1 + 2\delta^2\{L_0^{-1}(V_1L_0^{-1}V_1)\} + \delta^2(L_0^{-1}V_1)^2 + O(\delta^3), \\ w_\delta^6 &= 1 - 6\delta L_0^{-1}V_1 + 6\delta^2\{L_0^{-1}(V_1L_0^{-1}V_1)\} + 15\delta^2(L_0^{-1}V_1)^2 + O(\delta^3). \end{aligned}$$

When integrating these functions, the contributions of the second and third terms vanish, as they are in K_0 . Therefore, the first factor in ζ_* can be approximated by

$$(D.13) \quad \sqrt{\frac{(f_{\mathcal{B}}w^2)^3}{f_{\mathcal{B}}w^6}} \sim \sqrt{\frac{f_{\mathcal{B}}(1 + 3\delta^2L_0^{-1}V_1)}{f_{\mathcal{B}}(1 + 15\delta^2L_0^{-1}V_1)}} \sim 1 - 6\delta^2 \int_{\mathcal{B}} (L_0^{-1}V_1)^2.$$

Similarly, the inverse effective mass (Gaussian curvature) is

$$\begin{aligned} m_*^{-1} &= 1 - 4\frac{\langle \partial_x w, L_*^{-1}\partial_x w \rangle}{\langle w, w \rangle} \stackrel{(D.6)}{=} 1 - 4 \int_{\mathcal{B}} (-\delta\partial_x L_0^{-1}V_1)(-\delta\partial_x L_0^{-2}V_1) + O(\delta^3) \\ &\sim 1 - 4\delta^2 \int_{\mathcal{B}} (\partial_x L_0^{-3/2}V_1)^2, \end{aligned}$$

where in the last step we used the self-adjointness of L_0^{-1} . The operator in the above integral can be simplified as $\partial_x L_0^{-3/2} \equiv L_0^{-1}$. Thus, the second factor in ζ_* can be approximated with

$$\frac{1}{\sqrt{m_*}} \sim 1 - 2\delta^2 \int_{\mathcal{B}} (L_0^{-1}V_1)^2.$$

Combining with (D.13) yields (3.19).

Calculation of the slope constant ζ_{1*} . In one dimension (3.18) reduces to

$$(D.14) \quad \zeta_{1*} = \int_{\mathbb{R}} \int_{\mathcal{B}} |U_1(x, y)|^2 dx dy + \int_{\mathcal{B}} w^2 dx \int_{\mathbb{R}} S(y) \partial_{\Omega} F(y; \Omega)|_{\Omega=-1} dy .$$

As we shall see, $U_1 = O(\delta)$ and $S(y) = O(\delta^2)$. Therefore, both integral terms are $O(\delta^2)$. It follows from (3.5), (D.10), and (D.11) that to leading order in δ , $F(y)$ is the Townes mode; i.e.,

$$(D.15) \quad F(\mathbf{y}; \Omega, \delta) = R(\mathbf{y}; \Omega) + O(\delta^2) .$$

Furthermore, we shall use (1.18) with $\sigma = 2$ to explicitly evaluate the y -integrals.

The first integral term (D.14) depends on

$$(D.16) \quad \begin{aligned} U_1(x, y) &\stackrel{(3.12)}{=} 2L_{\delta}^{-1} \partial_x w \partial_y F \stackrel{(D.6)}{=} -2L_{\delta}^{-1} (\delta \partial_x L_0^{-1} V_1) R_y + O(\delta^2) \\ &\stackrel{(D.9)}{=} -2\delta \partial_x L_0^{-2} V_1 R_y + O(\delta^2) . \end{aligned}$$

Therefore,

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathcal{B}} |U_1(x, y)|^2 dx dy &= 4\delta^2 \int_{\mathcal{B}} (\partial_x L_0^{-2} V_1)^2 dx \int_{\mathbb{R}} R_y^2 dy + O(\delta^3) \\ &\stackrel{(1.18)}{=} \sqrt{3} \pi \delta^2 \int_{\mathcal{B}} (\partial_x L_0^{-2} V_1)^2 dx + O(\delta^3) . \end{aligned}$$

The operator in the above integral can be simplified as $\partial_x L_0^{-2} = (-\partial_{xxx})^{-1}$. Hence,

$$(D.17) \quad \int_{\mathbb{R}} \int_{\mathcal{B}} |U_1(x, y)|^2 dx dy = \sqrt{3} \pi \delta^2 \int_{\mathcal{B}} [(-\partial_{xxx})^{-1} V_1]^2 dx + O(\delta^3) .$$

For the second integral term in (D.14), we need to calculate $S(y)$. We use

$$(D.18) \quad U_0(x, y) \stackrel{(4.6)}{=} w(x) F(y) = R(y) + O(\delta^2) .$$

The first two terms in (3.14) are negligible. This follows from Lemma D.3.

LEMMA D.3. $\langle U_{2p}, w \rangle = 0$.

Proof. Equation (4.16) shows that the x -dependence of U_{2p} is of the form $L_*^{-1} X_k(x)$ for suitable $X_k(x)$, $k = 1, 2$. As L_{δ}^{-1} is into the orthogonal space to $w(x)$, the lemma follows. \square

That the first term in (3.14) is zero follows immediately from Lemma D.3. The second term in (3.14) has an additional $U_0^{2\sigma}$. However, in light of (D.18) the x -dependence of $U_0^{2\sigma}$ is constant to leading order. Therefore, the second term in $S(y)$ is $O(\delta^3)$. It remains to calculate the two last terms in $S(y)$. Note that the coefficient preceding the square brackets in (3.14) cancels with the w -integral in (D.14). This leaves (assuming $d = 1, \sigma = 2$)

$$(D.19) \quad S(y) = 10 \langle w, U_0^3 U_1^2 \rangle - 2 \langle \partial_x w, \partial_y \tilde{U}_3 \rangle + O(\delta^2) ,$$

where

$$\begin{aligned} \tilde{U}_3 &\stackrel{(4.19)}{=} 2L_0^{-1} \left[\partial_x L_0^{-1} X_1 \partial_{yyy} F + \partial_x L_0^{-1} X_2 \partial_y F^5 \right. \\ &\quad \left. + L_0^{-1} \partial_x w (\partial_{yy} - 1) \partial_y F + 2w^4 L_0^{-1} \partial_x w \partial_y F^5 \right] , \end{aligned}$$

and in one dimension $X_1 \equiv X_1^{ij}$. We denote by S_k the various terms in $S(y)$ when \tilde{U}_3 is explicitly inserted into it, and by I_k their corresponding contributions to (D.14).

Using (D.6), (D.16), and (D.18), the first term in (D.19) is

$$S_1 \equiv 10 \langle w, U_0^3 U_1^2 \rangle = 10 \langle U_1^2 \rangle + O(\delta^3) .$$

Therefore, its contribution to the slope is 10 times the first term in (D.14); i.e.,

$$(D.20) \quad I_1 \stackrel{(D.17)}{=} 10\sqrt{3} \pi \delta^2 \int_{\mathcal{B}} [(-\partial_{xxx})^{-1} V_1]^2 dx + O(\delta^3) .$$

The first term arising from substituting \tilde{U}_3 into $S(y)$ is

$$\begin{aligned} S_2 &\equiv -2 \langle \partial_x w, \partial_y 2L_0^{-1} \partial_x L_0^{-1} X_1 \partial_{yyyy} F \rangle \\ &\stackrel{(D.6),(D.10)}{=} 2 \langle -\delta L_0^{-1} \partial_x V_1, 2L_0^{-1} \partial_x L_0^{-1} (-4\delta \partial_{xx} L_0^{-2} V_1) \rangle \partial_{yyyy} F + O(\delta^3) \\ &\stackrel{(D.15)}{=} -16\delta^2 \int_{\mathcal{B}} (\partial_{xx} L_0^{-5/2} V_1)^2 dx \partial_{yyyy} R + O(\delta^3) , \end{aligned}$$

where in the last step we used the skew self-adjointness of ∂_x and the self-adjointness and positivity of L_0^{-1} . Simplifying the operator in the above integral, we obtain⁷

$$S_2 = -16\delta^2 \int_{\mathcal{B}} [(-\partial_{xxx})^{-1} V_1]^2 dx \partial_{yyyy} R + O(\delta^3) .$$

Substituting S_2 into (D.14) gives

$$I_2 = -16\delta^2 \int_{\mathcal{B}} [(-\partial_{xxx})^{-1} V_1]^2 dx \int_{\mathbb{R}} \partial_{yyyy} R \partial_{\Omega} R(y; \Omega)|_{\Omega=-1} dy + O(\delta^3) .$$

The following explicit integral can be obtained from (1.18):

$$\int_{\mathbb{R}} \partial_{yyyy} R \partial_{\Omega} R(y; \Omega)|_{\Omega=-1} dy = -\frac{11\sqrt{3} \pi}{16} .$$

Using this gives

$$(D.21) \quad I_2 = 11\sqrt{3} \pi \delta^2 \int_{\mathcal{B}} [(-\partial_{xxx})^{-1} V_1]^2 dx + O(\delta^3) .$$

Similar calculations can be carried out for I_3, I_4 , and I_5 using the explicit integral

$$\int_{\mathbb{R}} \partial_{yy} R^5 \partial_{\Omega} R(y; \Omega)|_{\Omega=-1} dy = \frac{13\sqrt{3} \pi}{16} .$$

Thus, to $O(\delta^3)$, we get

$$I_3 = 13\sqrt{3} \pi \delta^2 \int_{\mathcal{B}} [(-\partial_{xxx})^{-1} V_1]^2 dx$$

and $I_5 = -I_4 = \frac{1}{4} I_3$. Summing the contributions from $I_1 \dots I_5$ gives

$$\zeta_{1*} = 34\sqrt{3} \pi \delta^2 \int_{\mathcal{B}} [(-\partial_{xxx})^{-1} V_1]^2 dx + O(\delta^3) > 0 .$$

This concludes the proof of Corollary 3.1.

⁷Note that the Fourier representations of $\partial_x L_0^{-2}$ and $\partial_{xx} L_0^{-5/2}$ are the same only in $d = 1$.

Acknowledgments. The second author was supported in part by U.S. National Science Foundation grants DMS-04-12305 and DMS-07-07850. The authors thank Y. Sivan for discussions concerning this work and for comments on the manuscript. We would like to thank B. Altschuler, T. Dohnal, M. Hofer, A. Millis, P. Kuchment, and C. W. Wong for informative discussions. *We also thank the reviewers for a very careful reading of this work and detailed comments.*

REFERENCES

- [1] M. J. ABLOWITZ, B. ILAN, E. SCHONBRUN, AND R. PIESTUN, *Solitons in two-dimensional lattices possessing defects, dislocations, and quasicrystal structures*, Phys. Rev. E (3), 74 (2006), p. 035601(R).
- [2] M. ABLOWITZ AND H. SEGUR, *Solitons and the Inverse Scattering Transform*, Stud. Appl. Numer. Math. 4, Cambridge University Press, Cambridge, 2006.
- [3] M. J. ABLOWITZ AND Z. H. MUSSLIMANI, *Spectral renormalization method for computing self-localized solutions to nonlinear systems*, Optim. Lett., 30 (2005), pp. 2140–2142.
- [4] S. ALAMA AND Y. Y. LI, *Existence of solutions for semilinear elliptic equations with indefinite linear part*, J. Differential Equations, 96 (1992), pp. 89–115.
- [5] G. L. ALFIMOV, V. V. KONOTOP, AND P. PACCIANI, *Stationary localized modes of the quintic nonlinear Schrödinger equation with a periodic potential*, Phys. Rev. A (3), 75 (2007), p. 023624.
- [6] G. ALLAIRE AND A. PIATNITSKI, *Homogenization of the Schrödinger equation and effective mass theorems*, Comm. Math. Phys., 258 (2005), pp. 1–22.
- [7] B. B. BAIZAKOV, V. V. KONOTOP, AND M. SALERNO, *Regular spatial structures in arrays of Bose-Einstein condensates induced by modulational instability*, J. Phys. B, 35 (2002), pp. 5105–5119.
- [8] B. B. BAIZAKOV, B. A. MALOMED, AND M. SALERNO, *Multidimensional solitons in a low-dimensional periodic potential*, Phys. Rev. A (3), 70 (2004), p. 053613.
- [9] V. BANICA, R. CARLES, AND T. DUICKAERTS, *Minimal blow-up solutions to the mass-critical inhomogeneous NLS equation*, Preprint arXiv:0904.1317v1, 2009.
- [10] J. BOURGAIN, *Global Solutions of Nonlinear Schrödinger Equations*, AMS, Providence, RI, 1999.
- [11] R. W. BOYD, *Nonlinear Optics*, 3rd ed., Academic Press, New York, 2008.
- [12] V. A. BRAZHNYI AND V. V. KONOTOP, *Theory of nonlinear matter waves in optical lattices*, Modern Phys. Lett. B, 18 (2004), pp. 627–651.
- [13] V. A. BRAZHNYI, V. V. KONOTOP, AND M. PÉREZ-CARCÍA, *Driving defect modes of Bose-Einstein condensates in optical lattices*, Phys. Rev. Lett., 96 (2006), p. 060403.
- [14] K. BUSCH, G. SCHNEIDER, L. TKESHELASHVILI, AND H. UECKER, *Justification of the nonlinear Schrödinger equation in spatially periodic media*, Z. Angew. Math. Phys., 57 (2006), pp. 905–936.
- [15] R. CARRETERO-GONZALEZ, D. J. FRANTZESKAKIS, AND P. G. KEVREKIDES, *Nonlinear waves in Bose-Einstein condensates: Physical relevance and mathematical techniques*, Nonlinearity, 21 (2008), pp. R139–R202.
- [16] T. CAZENAVE, *Semilinear Schrödinger Equations*, Courant Lect. Notes Math. 10, Courant Institute Math. Sci., New York, 2003.
- [17] S.-C. CHENG, *An effective mass theory for bright gap solitons in optical lattices*, Chinese J. Phys., 43 (2005), pp. 835–846.
- [18] R. Y. CHIAO, E. GARMIRE, AND C. H. TOWNES, *Self-focusing of optical beams*, Phys. Rev. Lett., 13 (1964), pp. 479–482.
- [19] A. COMECH AND D. E. PELINOVSKY, *Purely nonlinear instability of standing waves with minimal energy*, Commun. Pure Appl. Math., 56 (2003), pp. 1565–1607.
- [20] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, Wiley-Interscience, New York, 1953.
- [21] T. DOHNAL, D. PELINOVSKY, AND G. SCHNEIDER, *Coupled-mode equations and gap solitons in a two-dimensional nonlinear elliptic problem with a separable potential*, J. Nonlinear Sci., 19 (2009), pp. 95–131.
- [22] T. DOHNAL AND H. UECKER, *Coupled mode equations and gap solitons for the 2D Gross–Pitaevskii equation with a non-separable periodic potential*, Phys. D, 238 (2009), pp. 860–879.

- [23] M. S. EASTHAM, *The Spectral Theory of Periodic Differential Equations*, Scottish Academic Press, Edinburgh, 1973.
- [24] N. K. EFREMIDIS, J. HUDOCK, D. N. CHRISTODOULIDES, J. W. FLEISCHER, O. COHEN, AND M. SEGEV, *Two-dimensional optical lattice solitons*, Phys. Rev. Lett., 91 (2003), p. 213906.
- [25] L. ERDŐS, B. SCHLEIN, AND H. T. YAU, *Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems*, Invent. Math., 59 (2007), pp. 1659–1741.
- [26] S. FLACH, K. KLADKO, AND R. S. MACKAY, *Energy thresholds for discrete breathers in one-, two-, and three-dimensional lattices*, Phys. Rev. Lett., 78 (1997), pp. 1207–1210.
- [27] M. G. GRILLAKIS, *Analysis of the linearization around a critical point of an infinite-dimensional hamiltonian system*, Comm. Pure Appl. Math., 43 (1990), pp. 299–333.
- [28] M. G. GRILLAKIS, J. SHATAH, AND W. A. STRAUSS, *Stability theory of solitary waves in the presence of symmetry*, I, J. Funct. Anal., 74 (1987), pp. 160–197.
- [29] H. P. HEINZ, T. KÜPPER, AND C. A. STUART, *Existence and bifurcation of solutions for nonlinear perturbations of the periodic Schrödinger equation*, J. Differential Equations, 100 (1992), pp. 341–354.
- [30] H. P. HEINZ AND C. A. STUART, *Solvability of nonlinear equations in spectral gaps of the linearization*, Nonlinear Anal., 19 (1992), pp. 145–165.
- [31] L. HÖRMANDER, *The Analysis of Linear Partial Differential Operators*, Vol. 3, Springer-Verlag, Berlin, 1985.
- [32] C. K. R. T. JONES, *An instability mechanism for radially symmetric standing waves of a nonlinear Schrödinger equation*, J. Differential Equations, 71 (1988), pp. 34–62.
- [33] R. KILLIP, M. VISAN, AND X. ZHANG, *The mass critical nonlinear Schrödinger equation with radial initial data in dimensions three and higher*, Preprint arXiv:0708.0849v1, 2007.
- [34] W. KIRSCH AND B. SIMON, *Comparison theorems for the gap of Schrödinger operators*, J. Funct. Anal., 75 (1987), pp. 396–410.
- [35] C. KITTEL, *Introduction to Solid State Physics*, 7th ed., Wiley, New York, 1995.
- [36] P. KUCHMENT, *The mathematics of photonic crystals*, in Mathematical Modeling in Optical Science, Frontiers Appl. Math. 22, SIAM, Philadelphia, 2001, pp. 207–266.
- [37] T. KUPPER AND C. A. STUART, *Gap-bifurcation for nonlinear perturbations of Hill’s equation*, J. Reine Angew. Math., 409 (1990), pp. 23–52.
- [38] T. KÜPPER AND C. A. STUART, *Necessary and sufficient conditions for gap-bifurcation*, Nonlinear Anal., 18 (1992), pp. 893–903.
- [39] M. K. KWONG, *Uniqueness of positive solutions of $\Delta u - u + u^p = 0$ in R^n* , Arch. Ration. Mech. Anal., 105 (1989), pp. 243–266.
- [40] J. V. MALONEY AND A. C. NEWELL, *Nonlinear Optics*, Westview Press, Boulder, CO, 2003.
- [41] J. MARZUOLA, S. RAYNOR, AND G. SIMPSON, *A system of ODEs for a perturbation of a minimal mass soliton*, J. Nonlinear Sci., to appear.
- [42] K. MOLL, G. FIBICH, AND A. GAETA, *Self-similar optical wave collapse: Observation of the Townes profile*, Phys. Rev. Lett., 90 (2003), p. 203902.
- [43] A. PANKOV, *Periodic nonlinear Schrödinger equation with application to photonic crystals*, Milan J. Math., 73 (2005), pp. 259–287.
- [44] D. E. PELINOVSKY AND G. SCHNEIDER, *Justification of the the coupled-mode approximation for a nonlinear elliptic problem with a periodic potential*, Appl. Anal., 86 (2007), pp. 1017–1036.
- [45] D. E. PELINOVSKY, A. A. SUKHORUKOV, AND Y. S. KIVSHAR, *Bifurcations and stability of gap solitons in periodic potentials*, Phys. Rev. E (3), 70 (2004), p. 036618.
- [46] L. P. PITAEVSKII AND S. STRINGARI, *Bose Einstein Condensation*, Oxford University Press, Oxford, 2003.
- [47] M. REED AND B. SIMON, *Methods of Modern Mathematical Physics IV: Analysis of Operators*, Academic Press, San Diego, CA, 1978.
- [48] H. A. ROSE AND M. I. WEINSTEIN, *On the bound states of the nonlinear Schrödinger equation with a linear potential*, Phys. D, 30 (1988), pp. 207–218.
- [49] H. SAKAGUCHI AND B. A. MALOMED, *Dynamics of positive- and negative-mass solitons in optical lattices and inverted traps*, J. Phys. B: At. Mol. Opt. Phys., 37 (2004), pp. 1443–1459.
- [50] H. SAKAGUCHI AND B. A. MALOMED, *Two-dimensional loosely and tightly bound solitons in optical lattices and inverted traps*, J. Phys. B: At. Mol. Opt. Phys., 37 (2004), pp. 2225–2239.
- [51] Z. SHI, J. WANG, Z. CHEN, AND J. YANG, *Linear instability of two-dimensional low-amplitude gap solitons near band edges in periodic media*, Phys. Rev. A (3), 78 (2008), p. 063812.
- [52] Z. SHI AND J. YANG, *Solitary waves bifurcated from Bloch-band edges in two-dimensional periodic media*, Phys. Rev. E (3), 75 (2007), p. 056602.

- [53] Y. SIVAN, G. FIBICH, N. K. EFREMIDIS, AND S. BAR-AD, *Analytic theory of narrow lattice solitons*, Nonlinearity, 21 (2008), pp. 509–536.
- [54] Y. SIVAN, G. FIBICH, B. ILAN, AND M. I. WEINSTEIN, *Qualitative and quantitative analysis of stability and instability dynamics of positive lattice solitons*, Phys. Rev. E (3), 78 (2008), p. 046602.
- [55] C. SPARBER, *Effective mass theorems for nonlinear Schrödinger equations*, SIAM J. Appl. Math., 66 (2006), pp. 820–842.
- [56] K. STALIUNAS, R. HERRERO, AND G. J. DE VALCÁRCEL, *Arresting soliton collapse in two-dimensional nonlinear Schrödinger systems via spatiotemporal modulation of the external potential*, Phys. Rev. E (3), 95 (2007), p. 011604.
- [57] M. J. STEEL AND W. ZHANG, *Bloch function description of a Bose-Einstein condensate in a finite optical lattice*, Preprint arXiv:cond-mat/9810284v1, 1998.
- [58] C. STUART, *Bifurcation from the essential spectrum*, in Topological Nonlinear Analysis, A. Matzeu and A. Vignoli, eds., Birkhauser, Basel, Switzerland, 1997.
- [59] C. SULEM AND P.-L. SULEM, *The Nonlinear Schrödinger Equation: Self-Focusing and Wave Collapse*, Applied Mathematical Sciences 139, Springer, Berlin, 1999.
- [60] T. TAO, *Nonlinear Dispersive Equations: Local and Global Analysis*, AMS, Providence, RI, 2006.
- [61] M. I. WEINSTEIN, *Nonlinear Schrödinger equations and sharp interpolation estimates*, Comm. Math. Phys., 87 (1983), pp. 567–576.
- [62] M. I. WEINSTEIN, *Modulational stability of ground states of nonlinear Schrödinger equations*, SIAM J. Math. Anal., 16 (1985), pp. 472–491.
- [63] M. I. WEINSTEIN, *Lyapunov stability of ground states of nonlinear dispersive evolution equations*, Comm. Pure Appl. Math., 39 (1986), pp. 51–68.
- [64] M. I. WEINSTEIN, *The Nonlinear Schrödinger Equation: Singularity Formation, Stability and Dispersion*, Contemp. Math., 99 (1989), pp. 213–232.
- [65] M. I. WEINSTEIN, *Excitation thresholds for nonlinear localized modes on lattices*, Nonlinearity, 12 (1999), pp. 673–691.
- [66] G. B. WHITHAM, *Linear and Nonlinear Waves*, Wiley-Interscience, New York, 1974.
- [67] J. YANG AND Z. H. MUSSLIMANI, *Fundamental and vortex solitons in a two-dimensional optical lattice*, Optim. Lett., 28 (2003), pp. 2094–2096.