Eigenvalues of a one-dimensional Dirac operator pencil

Daniel M. Elton   Michael Levitin   Iosif Polterovich

version September 28, 2013; LATEXed September 28, 2013

Abstract

We study the spectrum of a one-dimensional Dirac operator pencil, with a coupling constant in front of the potential considered as the spectral parameter. Motivated by recent investigations of graphene waveguides, we focus on the values of the coupling constant for which the kernel of the Dirac operator contains a square integrable function. In physics literature such a function is called a confined zero mode. Several results on the asymptotic distribution of coupling constants giving rise to zero modes are obtained. In particular, we show that this distribution depends in a subtle way on the sign variation and the presence of gaps in the potential. Surprisingly, it also depends on the arithmetic properties of certain quantities determined by the potential. We further observe that variable sign potentials may produce complex eigenvalues of the operator pencil. Some examples and numerical calculations illustrating these phenomena are presented.

1   Introduction and main results

1.1   Statement of the problem

Consider the system of differential equations

$$
\begin{aligned}
\big[V(x) - (\lambda - k)\big]\psi_1 - \frac{\mathrm{d}\psi_2}{\mathrm{d}x} &= 0, \\
\frac{\mathrm{d}\psi_1}{\mathrm{d}x} + \big[V(x) - (\lambda + k)\big]\psi_2 &= 0,
\end{aligned}
\tag{1}
$$

on $\mathbb{R}$, where k, λ are parameters and V is a potential. Equivalently one may define a self-adjoint operator by

$$
\mathrm{T}_V = \begin{pmatrix} V + k & -\nabla \\ \nabla & V - k \end{pmatrix} = -i\sigma_2 \nabla + k\sigma_3 + V,
$$

where $\nabla = \dfrac{\mathrm{d}}{\mathrm{d}x}$ and σ_2, σ_3 are Pauli matrices. Then (1) becomes the eigenvalue equation $\mathrm{T}_V \boldsymbol{\psi} = \lambda \boldsymbol{\psi}$, where $\boldsymbol{\psi} = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$.

For a given potential V let us set $\lambda = 0$ and introduce the γ-*spectrum* associated with V:

$$\Sigma_V = \{\gamma \in \mathbb{C} : 0 \in \mathrm{spec}(\mathrm{T}_{\gamma V})\}.$$

Equivalently Σ_V is the spectrum of the linear operator pencil $\gamma \mapsto \mathrm{T}_0 + \gamma V$. Our goal is to understand the properties of Σ_V, such as symmetries, existence of real and complex (non-real) eigenvalues, eigenvalue estimates and asymptotics. Similar problems, as well as some other related questions, have been studied in a variety of situations in mathematical literature — see, for instance, [BiLa], [GGHKSSV], [Kl], [Sa], [Sch].

Whilst the general asymptotic behaviour and estimates in our case are generally in line with earlier results (see Theorems 1.4 and 1.6; we should note that our methods allow the widest class of potentials), some unexpectedly subtle phenomena occur depending on the properties of V. In particular, Σ_V may have a totally different structure for single-sign and variable-sign potentials (compare Theorems 1.5 and 1.7), as well as for potentials having gaps (that is, whose support is not connected) and for no-gap potentials (see Examples 2.5 and 2.6). Also, variable-sign potentials can produce some (or even all) non-real eigenvalues, which have not been studied previously (see Theorem 1.7 and Example 2.4).

In physical literature this problem appears in the study of electron waveguides in graphene (see [HRP], [StDoPo] and references therein). Note that the electron dynamics in graphene is governed by the two-dimensional massless Dirac operator, and the one-dimensional system (1) is obtained as a result of the separation of variables: the parameter k corresponds to the frequency of the wavefunction ψ in the direction parallel to the waveguide. From the physical viewpoint solutions $\psi_1, \psi_2 \in L^2$ are of particular interest; these are called *confined modes*. Among them, especially important in the study of conductivity properties of graphene are *zero modes*: L^2-solutions corresponding to $\lambda = 0$. (See Section 1.8 for discussion of modes corresponding to $\lambda \neq 0$.) Zero-energy states in graphene have also been studied for potentials of other types — see, for instance, [BaTiBr], [BrFr]. It was shown in [HRP] that for the potential $V_{\mathrm{HRP}}(x) = -1/\cosh(x)$ the solutions of the system (1) can be found explicitly in terms of special functions. Moreover, there exists an infinite sequence of coupling constants γ such that 0 is an eigenvalue of the operator $\mathrm{T}_{\gamma V_{\mathrm{HRP}}}$. An attempt to formulate and prove precise mathematical statements confirming and generalising the results of [HRP] was the starting point of our research.

1.2 Basic results

To state precise results we need to make some basic restrictions on the local regularity and global decay of the potential V. We shall assume all potentials are real valued and locally L^2. Let $\mathbb{V}_0$ denote the class of such potentials which additionally satisfy

$$\|V\|_{L^2(x-1,x+1)} \to 0 \text{ as } |x| \to \infty;$$

roughly, $V \in \mathbb{V}_0$ if it decays at infinity. In the literature $\mathbb{V}_0$ is sometimes denoted as $c_0(L^2)$.

We can define the constant coefficient operator T_0 as a multiplication operator in Fourier space. If $V \in \mathbb{V}_0$ we show that V is a relatively compact perturbation of T_0, allowing us to define T_V as an unbounded self-adjoint operator on L^2 (see Section 3.1 for more details). The same construction can be used for complex-valued potentials (although, of course, the resulting operator will no longer be self-adjoint); this allows us to consider $\mathrm{T}_{\gamma V}$ for any $\gamma \in \mathbb{C}$. Further use of the relative compactness of V leads to the following:

Theorem 1.1. *If $V \in \mathbb{V}_0$ then Σ_V is a discrete subset of $\mathbb{C}$.*

Remark 1.2. *Standard spectrum.* The (usual) spectrum of the self-adjoint operator T_0 can be computed easily by considering it as a multiplication operator in Fourier space; we get

$$\mathrm{spec}(T_0) = \mathbb{R} \setminus (-|k|, |k|) =: \Lambda_k,$$

while this spectrum is purely absolutely continuous. Since $V \in \mathbb{V}_0$ is a relatively compact perturbation of T_0 the operators T_V and T_0 must have the same essential spectrum (see [ReSi, section XIII.4]); thus

$$\mathrm{spec}_{\mathrm{ess}}(T_V) = \mathrm{spec}_{\mathrm{ess}}(T_0) = \Lambda_k. \tag{2}$$

The operator T_V may have eigenvalues outside Λ_k but these must be isolated and of finite multiplicity (using the fact that we're dealing with a 1-dimensional problem it is not hard to show that these eigenvalues must in fact be simple; a somewhat restricted form of this result is given in Lemma 4.1).

In common with other Dirac operators, T_V possesses a number of elementary symmetries which lead to symmetries for the set Σ_V. In particular, if $V \in \mathbb{V}_0$ then $-\Sigma_V = \Sigma_V = \overline{\Sigma_V}$, while Σ_V is unchanged if we replace k with $-k$ in the definition of T_0. With this last symmetry in mind we shall henceforth assume $k > 0$; this will enable us to simplify the statement of some results.

To obtain estimates for the distribution of points in Σ_V we impose extra global decay conditions on the potential V. Let $\mathbb{V}_1$ denote the class of real valued locally L^2 potentials which satisfy

$$\int_{\mathbb{R}} |V(x)| \, \mathrm{d}x < +\infty;$$

that is, we require V to be integrable. Equivalently we can define $\mathbb{V}_1 = \mathbb{V}_0 \cap L^1$. The class $\mathbb{V}_1$ is sometimes denoted as $\ell^1(L^2)$.

Firstly we consider the number of points of Σ_V lying inside the disc $\{z \in \mathbb{C} : |z| \leq R\}$ of radius $R \geq 0$.

Theorem 1.3. *Suppose $V \in \mathbb{V}_1$. Then*

$$\#\big(\Sigma_V \cap \{z \in \mathbb{C} : |z| \leq R\}\big) \leq C \, \|V\|_{L^1} R$$

for any $R \geq 0$, where C is a universal constant (we can take $C = 4\mathrm{e}/\pi$).

This result can be generalised (using a rather different approach) to deal with potentials $V \in \mathbb{V}_0$ which have weaker decay than is required to be L^1; see Theorem 3.15.

Lower bounds which complement the upper bounds given by Theorem 1.3 can also be obtained. Restricting our attention to real points we have the following:

Theorem 1.4. *Suppose $V \in \mathbb{V}_1$. Then*

$$\#(\Sigma_V \cap [0, R]) \geq \frac{R}{\pi} \left| \int_{\mathbb{R}} V(x) \, \mathrm{d}x \right| + o(R)$$

as $R \to \infty$, while the same estimate holds for $\#(\Sigma_V \cap [-R, 0])$ (by symmetry). In particular, $\Sigma_V \cap \mathbb{R}$ contains infinitely many points if $\int_{\mathbb{R}} V(x) \, \mathrm{d}x \neq 0$.

1.3 Single-signed potentials

In general the set Σ_V may contain complex eigenvalues (see Section 2 for some examples of explicit potentials which illustrate various possible behaviours for complex points in Σ_V). Note that, even though the operator T_V is self-adjoint (recall that V is real valued), it does not follow in general that the corresponding operator pencil should have a purely real spectrum. However, if V does not change sign (as in the example considered in [HRP]) all eigenvalues of the operator pencil are real:

Theorem 1.5. *If $V \in \mathbb{V}_0$ is single-signed then $\Sigma_V \subset \mathbb{R}$.*

By symmetry we can write $\Sigma_V = \{\pm \gamma_n : n \in \mathbb{N}\}$ where $0 < \gamma_1 < \gamma_2 < \ldots$ denotes the sequence of positive points in Σ_V, arranged in order of increasing size. The bound in Theorem 1.4 can be turned into an asymptotics:

Theorem 1.6. *Suppose $V \in \mathbb{V}_1$ is single-signed. Then*

$$\#(\Sigma_V \cap [0, R]) = \frac{R}{\pi} \left| \int_{\mathbb{R}} V(x)\, \mathrm{d}x \right| + o(R) = \frac{\|V\|_{L^1}}{\pi} R + o(R)$$

as $R \to \infty$. If V is non-zero we can equivalently write

$$\gamma_n = \frac{\pi}{\|V\|_{L^1}} n + o(n)$$

as $n \to \infty$.

1.4 Anti-symmetric potentials

For potentials of variable sign the behaviour of the γ-spectrum may be different, in some cases quite drastically so. For anti-symmetric potentials we have the following:

Theorem 1.7. *If $V \in \mathbb{V}_0$ is anti-symmetric then $\Sigma_V \cap \mathbb{R} = \emptyset$.*

Note that, the γ-spectrum may still contain an infinite number of complex eigenvalues; see Example 2.4 below.

The absence of real points in the γ-spectrum together with Theorem 1.6 shows that the lower bound obtained in Theorem 1.4 is quite sharp.

Remark 1.8. It is easy to see that translating a potential V changes the operator $T_{\gamma V}$ to something which is unitarily equivalent. In particular, Theorem 1.7 also applies to potentials V satisfying the condition $V(a + x) = -V(a - x)$ for some $a \in \mathbb{R}$ and all $x \in \mathbb{R}$. The translation invariance of our problem will also be used to simplify the presentation of some arguments in Section 3.

1.5 Potentials without gaps

Let BV_0 denote the class of compactly supported real valued functions of (totally) bounded variation. Clearly $BV_0 \subset \mathbb{V}_1$ while BV_0 contains compactly supported piecewise constant potentials with a finite number of pieces, as well as compactly supported functions in C^1. We say that a potential $V \in BV_0$ *has no gaps* if

$$\left| \mathrm{co}(\mathrm{supp}(V)) \cap V^{-1}(0) \right| = 0,$$

where $|S|$ and $\mathrm{co}(S)$ denote the Lebesgue measure and convex hull of a set $S \subseteq \mathbb{R}$ respectively.

Theorem 1.9. *Suppose $V \in BV_0$ has no gaps. Then*

$$\#(\Sigma_V \cap [0, R]) = \frac{R}{\pi} \left| \int_{\mathbb{R}} V(x) \, dx \right| + O(1)$$

as $R \to \infty$. The same estimate holds for $\#(\Sigma_V \cap [-R, 0])$ (by symmetry).

Remark 1.10. When $\int_{\mathbb{R}} V(x) \, dx = 0$ this result simply states that $\Sigma_V \cap \mathbb{R}$ is finite.

1.6 Discussion

Our results give information about the asymptotics of the counting function $\#(\Sigma_V \cap [0, R])$ as $R \to \infty$. For any $V \in \mathbb{V}_1$ the results of Section 1.2 give asymptotic upper and lower bounds of

$$C \frac{R}{\pi} \int_{\mathbb{R}} |V(x)| \, dx \quad \text{and} \quad \frac{R}{\pi} \left| \int_{\mathbb{R}} V(x) \, dx \right| \tag{3}$$

respectively. Using Theorem 1.3 we can take $C = 2\mathrm{e}$, in which case the upper bound is actually uniform for $R \geq 0$. For an asymptotic upper bound the constant can be reduced to at least $C = \mathrm{e}$ in general (see Remark 3.14). Theorem 1.6 shows the constant can be reduced further to $C = 1$ for single-signed potentials; in this case the asymptotic upper and lower bounds agree and an asymptotic formula for the points in Σ_V is obtained. For variable-signed potentials the quantities $\int_{\mathbb{R}} |V(x)| \, dx = \|V\|_{L^1}$ and $\left| \int_{\mathbb{R}} V(x) \, dx \right|$ differ, leading to differences in the upper and lower bounds in (3) even if we could take $C = 1$. For no-gap potentials $V \in BV_0$ Theorem 1.9 shows that it is the lower bound that actually gives the leading order term in the asymptotics of $\#(\Sigma_V \cap [0, R])$ as $R \to \infty$.

The above results may lead to a hypothesis that, in fact, the lower bound always gives the leading order term in the asymptotics of the counting function of the γ-spectrum. However, as we show in the next section, this is not the case. Moreover, the precise asymptotic behaviour of $\#(\Sigma_V \cap [0, R])$ as $R \to \infty$ may depend on the properties of a variable–signed potential in a rather subtle way. In particular, it is sensitive to the presence of gaps, that is, intervals where $V \equiv 0$, appearing between components of $\mathrm{supp}(V)$. Even more surprisingly, the leading term of the asymptotics is affected by the arithmetic properties of certain quantities determined by the potential, such as the rationality of the ratio $\left| \int_{\mathbb{R}} V(x) \, dx \right| / \|V\|_{L^1}$.

1.7 One-gap potentials, zeros of trigonometric functions and arithmetic

Suppose $V \in BV_0$. We say V *has one gap* if we can write $V = V_1 + V_2$ for some non-zero $V_1, V_2 \in BV_0$ which have no gaps and disjoint supports. For $j = 1, 2$ the support of V_j is a closed bounded interval; write $\mathrm{supp}(V_j) = [a_j, b_j]$. Without loss of generality we may assume the support of V_1 lies to the left of that of V_2. Then $b_1 < a_2$ and the gap is the interval (b_1, a_2). Set

$$v_j = \int_{a_j}^{b_j} V(x) \, dx = \int_{\mathbb{R}} V_j(x) \, dx$$

for $j = 1, 2$. Thus $\int_{\mathbb{R}} V(x) \, dx = v_1 + v_2$ while $|v_1| + |v_2| \leq \|V\|_{L^1}$, with equality iff V_1 and V_2 are each single-signed.

Theorem 1.11. *If $\int_{\mathbb{R}} V(x) \, dx = 0$ then $\Sigma_V \cap \mathbb{R}$ contains only finitely many points.*

Remark 1.12. This result extends Remark 1.10 to one gap potentials. However the same result does *not* extend to zero integral potentials with two gaps; see Example 2.6.

5

We now suppose $v_1 + v_2 = \int_\mathbb{R} V(x)\,\mathrm{d}x \neq 0$. Set

$$\alpha = \tanh(k(a_2 - b_1)) \quad \text{and} \quad \beta = \left| \frac{v_1 - v_2}{v_1 + v_2} \right|.$$

Then $\alpha \in (0, 1)$ gives a measure of the gap length, while $0 \leq \beta < 1$ if $v_1 v_2 > 0$ and $\beta > 1$ if $v_1 v_2 < 0$. In particular, if V is single-signed then $\beta < 1$. When $\alpha\beta > 1$ we can further define

$$\nu_{\alpha,\beta} = \frac{2}{\pi} \left[\beta \arcsin \frac{\sqrt{\alpha^2\beta^2 - 1}}{\sqrt{\beta^2 - 1}} + \arcsin \frac{\sqrt{1 - \alpha^2}}{\alpha\sqrt{\beta^2 - 1}} \right]. \tag{4}$$

If we fix $\beta > 1$ and allow α to vary from $1/\beta$ to 1 it is easy to check that $\nu_{\alpha,\beta}$ varies continuously and monotonically from 1 to β.

If β is positive and rational write $\beta = p/q$ where $p, q \in \mathbb{N}$ are coprime. If p and q are both odd set $p_\beta = p$ and $q_\beta = q$; if p and q have opposite parity set $p_\beta = 2p$ and $q_\beta = 2q$.

Set

$$A(\alpha, \beta) = \begin{cases} 1 & \text{if } \alpha\beta < 1, \\ \nu_{\alpha,\beta} & \text{if } \alpha\beta > 1 \text{ and } \beta \notin \mathbb{Q}, \\ \dfrac{4}{q_\beta} \left\lfloor \dfrac{1}{4}(p_\beta + q_\beta \nu_{\alpha,\beta}) \right\rfloor - \dfrac{p_\beta}{q_\beta} + \dfrac{2}{q_\beta} & \text{if } \alpha\beta > 1 \text{ and } \beta \in \mathbb{Q}; \end{cases} \tag{5}$$

we are using $\lfloor x \rfloor$ to denote the largest integer which does not exceed x.

Theorem 1.13. *Suppose $\alpha\beta \neq 1$. If $\alpha\beta > 1$ and $\beta \in \mathbb{Q}$ suppose additionally that $p_\beta + q_\beta \nu_{\alpha,\beta} \notin 4\mathbb{Z}$. Then*

$$\#(\Sigma_V \cap [0, R]) = \frac{1}{\pi} A(\alpha, \beta)\, |v_1 + v_2|\, R + o(R)$$

as $R \to \infty$. The same estimate holds for $\#(\Sigma_V \cap [-R, 0])$ (by symmetry).

Remark 1.14. If $\alpha\beta > 1$ and $\beta \in \mathbb{Q}$ then the bounds $x - 1 \leq \lfloor x \rfloor \leq x$ give

$$\nu_{\alpha,\beta} - \frac{2}{q_\beta} \leq A(\alpha, \beta) \leq \nu_{\alpha,\beta} + \frac{2}{q_\beta}.$$

Therefore, for any sequence of rational numbers β_n converging to $\beta \notin \mathbb{Q}$, we have $A(\alpha, \beta_n) \to A(\alpha, \beta)$, and hence $A(\alpha, \beta)$ is continuous at irrational values of β. At the same time it is clear that $A(\alpha, \beta)$ has discontinuities at many rational values of β. Let us note that continuity at irrational values and discontinuity at rational values of a parameter was observed for other physically meaningful quantities — see, for instance, [GGL] (where the mathematical setting is somewhat similar to ours), as well as [AMS], [JM].

Theorem 1.13 comes almost directly from a result about the zeros of a perturbed trigonometric function. Consider the equation

$$\cos(x) + \alpha \cos(\beta x) + \phi(x) = 0 \tag{6}$$

where ϕ satisfies the decay condition

$$\phi \in C^2(\mathbb{R}), \quad \phi^{(n)}(x) = o(1) \text{ as } x \to \infty \text{ for } n = 0, 1, 2. \tag{7}$$

Theorem 1.15. *Let $0 \leq \alpha < 1$, $\beta \geq 0$ and $\alpha\beta \neq 1$. If $\alpha\beta > 1$ and $\beta \in \mathbb{Q}$ let us additionally assume that $p_\beta + q_\beta \nu_{\alpha,\beta} \notin 4\mathbb{Z}$. Also suppose that ϕ satisfies (7) and the solutions of (6) form a discrete subset of $\mathbb{R}$. Then*

$$\#\{x \in [0, R] : x \text{ satisfies } (6)\} = \frac{1}{\pi} A(\alpha, \beta)\, R + o(R)$$

as $R \to \infty$.

Note that, $n = 2$ in condition (7) is only needed in the case that $\alpha\beta > 1$ and $\beta \notin \mathbb{Q}$.

Consideration of Theorem 1.15 in the case $\phi \equiv 0$ goes back at least as far as [St] where the irrational case was established (a somewhat different problem was considered in the rational case).

Using Theorem 1.13 it is possible to show that we can obtain an asymptotic formula

$$\#(\Sigma_V \cap [0, R]) = \frac{C}{\pi} R + o(R) \tag{8}$$

as $R \to \infty$, where C can indeed take any value (strictly) between $\left|\int_{\mathbb{R}} V(x)\,\mathrm{d}x\right|$ and $\|V\|_{L^1}$; we state this as a separate result.

Theorem 1.16. *Let $0 < v < A < u$. Then there exists a piecewise constant one gap potential V such that $\left|\int_{\mathbb{R}} V(x)\,\mathrm{d}x\right| = v$, $\|V\|_{L^1} = u$ and (8) holds with $C = A$.*

1.8 Remarks on non-zero modes

If we consider the eigenvalues of $\mathrm{T}_{\gamma V}$ as functions of γ we can view Σ_V as the set of points at which these curves cross 0. One could equally consider crossings at any other point λ belonging to $(-k, k)$ (the spectral gap of the operator T_0). This leads to consideration of the set

$$\Sigma_{\lambda, V} = \big\{\gamma \in \mathbb{C} : \lambda \in \mathrm{spec}(\mathrm{T}_{\gamma V})\big\}.$$

With some straightforward modifications most of our analysis for Σ_V can be carried over to $\Sigma_{\lambda, V}$ for any $\lambda \in (-k, k)$. We now summarise the changes to the main results.

Theorem 1.1 holds for $\Sigma_{\lambda, V}$. For $V \in \mathbb{V}_0$, $\Sigma_{\lambda, V}$ is still symmetric under conjugation and unchanged if we replace k with $-k$; however, we cannot expect $\Sigma_{\lambda, V}$ to be symmetric about 0 in general (this symmetry generalises to $-\Sigma_{\lambda, V} = \Sigma_{-\lambda, V}$). Theorem 1.7 does not generalise.

Theorems 1.3, 1.4, 1.5 and 1.6 hold for $\Sigma_{\lambda, V}$ with two adjustments; firstly, the constant C in Theorem 1.4 may depend on λ, and secondly, the results for points in $\Sigma_{\lambda, V} \cap \mathbb{R}^-$ no longer follow "by symmetry" (but can be obtained by similar arguments).

The latter comment also applies to Theorem 1.9, which otherwise holds for $\Sigma_{\lambda, V}$ in the case that $\int_{\mathbb{R}} V(x)\,\mathrm{d}x \neq 0$. When $\int_{\mathbb{R}} V(x)\,\mathrm{d}x = 0$ we need to impose further conditions on λ (to ensure we avoid limiting values of the eigenvalues of $\mathrm{T}_{\gamma V}$ as $\gamma \to \pm\infty$; cf. [GGHKSSV], Theorem 8.2(i)).

Theorem 1.11 does not admit a straightforward generalisation to the case $\lambda \neq 0$. A generalisation of Theorem 1.13 will require Theorem 1.15 to be extended to cover equations of the form $\cos(x) + \alpha\cos(\beta x + \delta) + \phi(x) = 0$, where $\delta \in \mathbb{R}$ is an additional parameter (cf. [St] for the case $\phi \equiv 0$, $\beta \notin \mathbb{Q}$).

1.9 Organisation of the paper

Section 2 is devoted to examples. The main arguments, together with a number of auxiliary constructions and results, are collected in Section 3. Theorem 1.1 is essentially standard; its proof appears in Section 3.1. Some key ideas from the Prüfer method are introduced in Section 3.3. In particular, we re-characterise the set $\Sigma_V \cap \mathbb{R}$ in terms of a quantity Δ_V, which is closely related to the Prüfer argument (see Proposition 3.7). The asymptotic behaviour of Δ_V is described (Proposition 3.8) and leads directly to Theorem 1.4. Theorem 1.6 follows from a related argument, together with the additional monotonicity of Δ_V for single-signed V (as described in Proposition 3.10). An alternative approach based on the Birman-Schwinger principle that could be used in the case of single-signed potentials is discussed in Remark 3.3.

A uniform bound on Δ_V (given in Proposition 3.11) leads through several intermediate results to the proof of Theorem 1.3. Derivatives of Δ_V are considered in Sections 3.4 and 3.5. The justification of the monotonicity result (Proposition 3.10) appears in Section 3.4, while in Section 3.5 the proof of Theorem 1.9 is reduced to some technical estimates (given in Proposition 3.18). Theorem 1.11 is established in Section 3.6 while Theorem 1.13 is reduced to Theorem 1.15; Theorem 1.16 is then obtained as a straightforward consequence of the former.

For the sake of clarity the proofs of the results in Section 3 which require more technical arguments are deferred to Section 4. In Section 4.1 we consider Lemma 3.6, Section 4.2 deals with Propositions 3.8 and 3.11, and in Section 4.3 we establish Proposition 3.18.

The last part of the paper is devoted to the proof of Theorem 1.15, which is a variation on a classical theme of independent interest (cf. [St], [Kac], [KKW]). Some preliminary lemmas are established in Section 5.1, Section 5.2 contains the proof of Theorem 1.15 in the unperturbed case $\phi \equiv 0$, while the general case is completed in Section 5.3.

Acknowledgments

2 Examples

2.1 General description

The main purpose of this section is to illustrate the results stated above. We restrict our attention mostly to piecewise constant potentials with compact support; these allow the easiest analysis and already demonstrate the full range of effects. Consider points $a_0 < a_1 < \cdots < a_m$ which partition the real line into m finite intervals $I_j = (a_{j-1}, a_j)$, $j = 1, \ldots, m$, and two semi-infinite intervals $I_- = (-\infty, a_0)$ and $I_+ = (a_m, +\infty)$. Consider a potential

$$V(x) = W\big(x; [a_0, \ldots, a_m]; \{v_1, \ldots, v_m\}\big) := \begin{cases} v_j & \text{if } x \in I_j, \ j = 1, \ldots, m, \\ 0 & \text{if } x \in I_- \cup I_+, \end{cases} \qquad (9)$$

with some given real constants v_j. On each interval, we need to solve the equations

$$\begin{aligned} \nabla \psi_1 &= (k - \gamma V)\psi_2, \\ \nabla \psi_2 &= (k + \gamma V)\psi_1, \end{aligned} \qquad (10)$$

with $V(x) = v_j = \text{const}$, and then match the solutions to ensure continuity at the points a_j.

The following result is straightforward.

Lemma 2.1. *For a given constant potential $V(x) = v$ such that $k \neq \pm\gamma v$, the system (10) has the general solution*

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}(x) = C^{(1)} \begin{pmatrix} \sin\big(\sqrt{\gamma^2 v - k^2}x\big) \\ -\sqrt{\frac{\gamma v + k}{\gamma v - k}} \cos\big(\sqrt{\gamma^2 v - k^2}x\big) \end{pmatrix} + C^{(2)} \begin{pmatrix} \cos\big(\sqrt{\gamma^2 v - k^2}x\big) \\ \sqrt{\frac{\gamma v + k}{\gamma v - k}} \sin\big(\sqrt{\gamma^2 v - k^2}x\big) \end{pmatrix}.$$

If $v = 0$ this solution can be equivalently written as

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}(x) = C^{(1)} \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{kx} + C^{(2)} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-kx}. \qquad (11)$$

In both cases $C^{(1)}$, $C^{(2)}$ are arbitrary complex constants.

Remark 2.2. Lemma 2.1, or more precisely equation (11), immediately implies that for any compactly-supported (not necessarily piecewise-constant) potential $V(x)$, any eigenfunction $\psi \in L^2(\mathbb{R})$ satisfies, under the assumption $k > 0$,

$$\psi_1(\min \operatorname{supp}(V)) = \psi_2(\min \operatorname{supp}(V)), \qquad \psi_1(\max \operatorname{supp}(V)) = -\psi_2(\max \operatorname{supp}(V)), \quad (12)$$

in order to match the L^2 solutions at $\pm\infty$.

Let us return to the case of a piecewise constant potential (9). The solution on each interval I_j, $j = 1, \ldots, m$, can be written down using Lemma 2.1 with $v = v_j$ and $C^{(\ell)} = C_j^{(\ell)}$, $\ell = 1, 2$. By Remark 2.2, we have $\psi_1(a_0) = \psi_2(a_0)$, and $\psi_1(a_m) = -\psi_2(a_m)$. Together with continuity conditions at each a_j, $j = 1, \ldots, m-1$ this leads to the homogeneous linear system of $2m$ equations with respect to $2m$ unknowns $C_j^{(\ell)}$, $\ell = 1, 2$, $j = 1, \ldots, m$. Denote the determinant of the corresponding matrix of coefficients by $D_V(\gamma)$. As we are looking for a non-trivial L^2-solution ψ, we have $\gamma \in \Sigma_V$ if and only if

$$D_V(\gamma) = 0. \tag{13}$$

Thus, in each particular case our problem is reduced to constructing $D_V(\gamma)$ and finding its real or complex roots.

2.2 Calculations, graphs, and further observations

We visualise the real roots of $D_V(\gamma)$ by simply plotting its graph for real arguments. In the complex case we use the phase plot method (see [WeSe]) in which the value of $\arg D_V(\gamma) = -\mathrm{i} \log \dfrac{D_V(\gamma)}{|D_V(\gamma)|}$ is plotted using colours from a periodic scale. The roots of $D_V(\gamma)$ are singularities of $\arg D_V(\gamma)$ and appear on the phase plot as points at which all of the colours converge. The colour scale which we use in all such plots is shown in Figure 1.
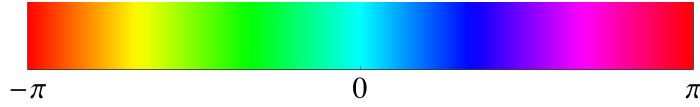

$$-\pi \qquad\qquad 0 \qquad\qquad \pi$$

Figure 1: Phase plot colour scale for the value of $\arg D_V(\gamma)$.

In the following examples it is convenient to set

$$\widetilde{\gamma} = \sqrt{\gamma^2 - 1}.$$

Also, we remark that our determinants $D_V(\gamma)$ are defined modulo a real or complex scaling constant, which we choose for convenience of presentation.

Example 2.3 (Illustration of Theorems 1.5 and 1.6)**.** Set $V_1(x) := W(x; [-1, 1]; \{1\})$. Then

$$D_{V_1}(\gamma) = \frac{2(\widetilde{\gamma} \cos(2\widetilde{\gamma}) + \sin(2\widetilde{\gamma}))}{\gamma - 1}$$

As the potential is single-signed, the spectrum Σ_V is real, as illustrated in Figures 2 and 3.

For large (positive) values of γ, the solutions of (13) with $V = V_1$ are approximately those of $\cos(2\gamma) = 0$, i.e.,

$$\gamma_n = \frac{\pi}{2} n + o(n) = \frac{\pi}{\|V_1\|_{L^1}} n + o(n),$$
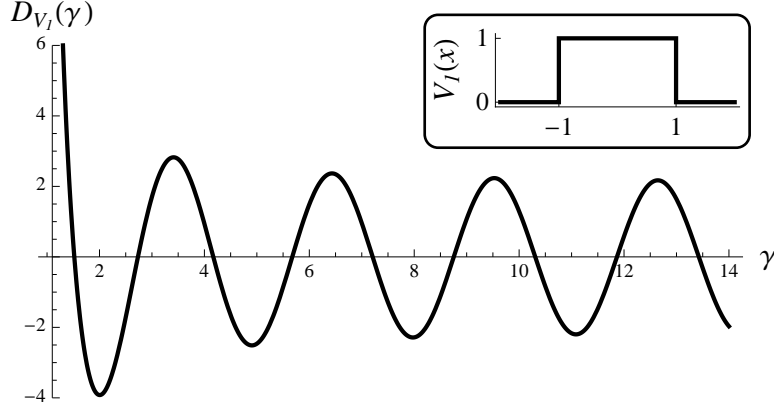
matching Theorem 1.6.

9

Figure 2: The graph of $D_{V_1}(\gamma) = \dfrac{2(\widetilde{\gamma}\cos(2\widetilde{\gamma}) + \sin(2\widetilde{\gamma}))}{\gamma - 1}$ against real γ for the potential $V_1(x) = V(x; [-1, 1]; \{1\})$.
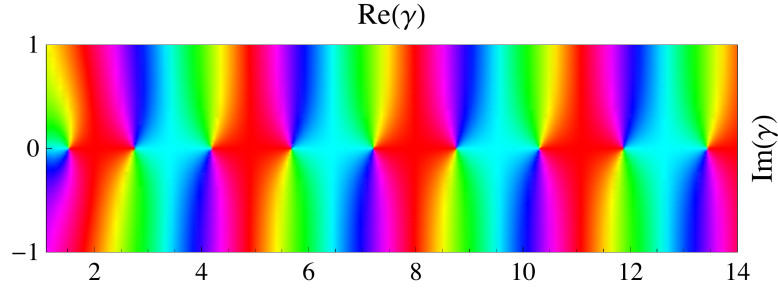

Figure 3: The phase plot of $D_{V_1}(\gamma)$ for complex γ.

Example 2.4 (Illustration of Theorem 1.7). Consider a class of anti-symmetric potentials $V_{2,g}(x) := W(x; [-1 - g/2, -g/2, g/2, g/2 + 1]; \{-1, 0, 1\})$ parametrised by the gap length $g \geq 0$. Then, up to a multiplication by a non-zero constant,

$$D_{V_{2,g}}(\gamma) = \frac{2\cosh(g)(\widetilde{\gamma}^2 + 1 - \cos(2\widetilde{\gamma}) + \widetilde{\gamma}\sin(2\widetilde{\gamma})) + 2\sinh(g)(\widetilde{\gamma}^2\cos(2\widetilde{\gamma}) + \widetilde{\gamma}\sin(\widetilde{\gamma}))}{\widetilde{\gamma}^2}.$$

For any g, the potential $V_{2,g}$ is anti-symmetric; hence the spectrum $\Sigma_{V_{2,g}}$ is purely non-real and $D_{V_{2,g}}(\gamma)$ does not have any real roots. This is illustrated in Figure 4 for $g = 0$ and $g = 1$.

It turns out that the behaviour of complex eigenvalues for the potentials $V_{2,g}$ differs substantially for zero and non-zero gaps g. By a rather intricate asymptotic analysis of the corresponding transcendental equations (which in a sense extends Theorem 1.15 to complex roots) we can show that the large eigenvalues with positive real parts are asymptotically located on the curves

$$\operatorname{Im}\gamma = \pm\frac{\ln\operatorname{Re}\gamma}{2} \qquad \text{if } g = 0 \tag{14}$$

and on the straight lines

$$\operatorname{Im}\gamma = \pm\operatorname{arcsinh}\left(\frac{1}{\sinh g}\right) \qquad \text{if } g > 0. \tag{15}$$

Figures 5 and 6 illustrate this behaviour of complex eigenvalues. For comparison, we also plot the corresponding curves (14) and (15); one can see that the asymptotics is accurate even for the low eigenvalues.
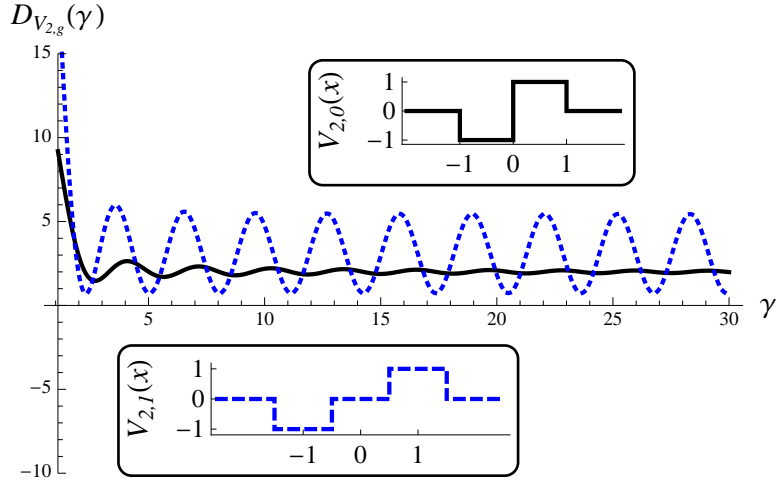
10

Figure 4: The graphs of $D_{V_{2,g}}(\gamma)$ against real γ for the potentials $V_{2,g}(x) := W(x; [-1 - g/2, -g/2, g/2, g/2 + 1]; \{-1, 0, 1\})$ with $g = 0$ (solid black line) and $g = 1$ (dashed blue line).
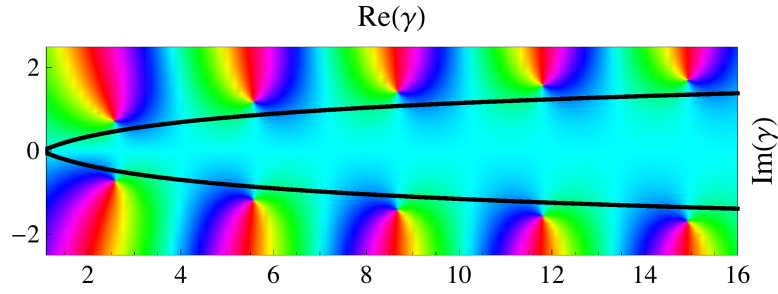
Figure 5: The phase plot of $D_{V_{2,0}}(\gamma)$ for complex γ. The solid black curves (14) illustrate the asymptotic behaviour of the imaginary parts of the eigenvalues.
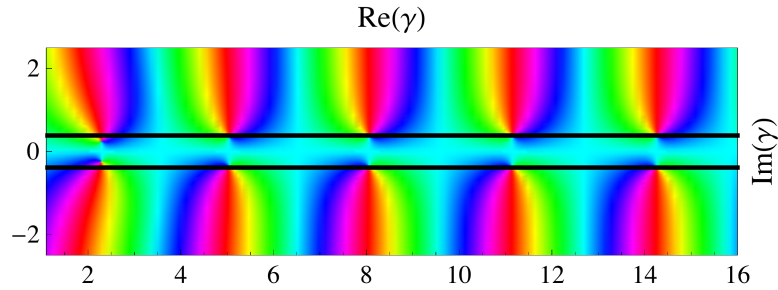
Figure 6: The phase plot of $D_{V_{2,1}}(\gamma)$ for complex γ. The solid black lines (15), $g = 1$, illustrate the asymptotic behaviour of the imaginary parts of the eigenvalues.

Example 2.5 (Illustration of Theorem 1.13). Consider the one-gap potentials $V_{3,g,b}(x) := W(x; [-g - 1, -g, 0, b]; \{-1, 0, 1\})$ parametrised by the gap length $g \geq 0$ and the maximum of the support $b > 0$. For these potentials, $\int_{\mathbb{R}} V_{3,g,b} = b - 1$ and $\|V_{3,g,b}\|_{L^1} = b + 1$. Assume

11

additionally $b \neq 1$. Explicit calculation gives, modulo multiplication by a constant,

$$D_{V_{3,g,b}}(\gamma) = \frac{2}{\widetilde{\gamma}^2} \left[\left((\widetilde{\gamma}^2 + 1) \cos((b-1)\widetilde{\gamma}) - \cos((b+1)\widetilde{\gamma}) + \widetilde{\gamma} \sin((b+1)\widetilde{\gamma}) \right) \cosh(g) \right.$$
$$\left. + \left(\widetilde{\gamma}^2 \cos((b+1)\widetilde{\gamma}) + \widetilde{\gamma} \sin((b+1)\widetilde{\gamma}) \right) \sinh(g) \right].$$

The graphs of $D_{V_{3,g,2}}(\gamma)$ for real γ and $g = 0$ or $g = 1$ are shown in Figure 7.

Figure 7: The graphs of $D_{V_{3,g,2}}(\gamma)$ against real γ for the potentials $V_{3,g,2}(x) := W(x; [-g - 1, -g, 0, 2]; \{-1, 0, 1\})$ with $g = 0$ (solid black line) and $g = 1$ (dashed blue line).

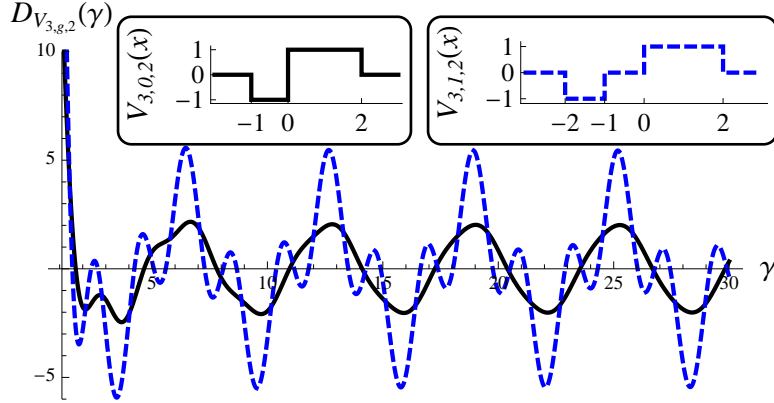We can expect asymptotics of the form

$$\#(\Sigma_{V_{3,g,b}} \cap [0, R]) = C_{g,b} \frac{R}{\pi} + o(R),$$

as $R \to \infty$. For the no-gap potential $V_{3,0,2}$, Theorem 1.9 gives such an asymptotics with $C_{0,2} = 1 = \int_{\mathbb{R}} V_{3,0,2}$. On the hand, $D_{V_{3,1,2}}(\gamma)$ has three times as many real roots as $D_{V_{3,0,2}}(\gamma)$ (for sufficiently large γ). This leads to a constant $C_{1,2} = 3 = \|V_{3,1,2}\|_{L^1}$ in the asymptotics for the one-gap potential $V_{3,1,2}$ as seen in Figure 7; c.f. the discussion in Sections 1.5 and 1.6.

This is just a partial case of a more complicated phenomenon, see Theorem 1.13. Set

$$\alpha = \tanh(g) \quad \text{and} \quad \beta = \left| \frac{b+1}{b-1} \right|.$$

After cancelling some non-zero factors equation (13) with $V = V_{3,g,b}$ takes the asymptotic form

$$\cos((b-1)\widetilde{\gamma}) + \alpha \cos((b+1)\widetilde{\gamma}) + O\left(\widetilde{\gamma}^{-1}\right) = 0 \qquad (16)$$

as $\widetilde{\gamma} \to +\infty$ (where the first and second derivatives of the O-term are also $O\left(\widetilde{\gamma}^{-1}\right)$). Introducing the new variable $x = |b-1|\widetilde{\gamma}$ leads to an equation in the form of (6). The asymptotics for the number of real zeros of (16) can then be obtained from Theorem 1.15. Alternatively, we can use Theorem 1.13 directly; both approaches give

$$C_{g,b} = A(\alpha, \beta) |b-1| = A\left(\tanh(g), \left| \frac{b+1}{b-1} \right|\right) |b-1|,$$

where $A(\alpha, \beta)$ is defined in (5).

Example 2.6 (Illustration of a twin gap effect). The gap dependence illustrated in the previous Example can be made even more dramatic if we consider some special potentials. Introduce the symmetric twin gap potentials

$$V_{4,g}(x) := W\big(x; [-g-2, -g-1, -1, 1, g+1, g+2]; \{-1, 0, 1, 0, -1\}\big)$$

parametrised by the gap length g. Note that $\int_{\mathbb{R}} V_{4,g} = 0$ and $\|V_{4,g}\|_{L^1} = 4$ for any $g \geq 0$. Figure 8 shows the real curves $D_{V_{4,g}}(\gamma)$ for $g = 0.5$ and $g = 1$. One can see that there are only two real eigenvalues for the former, and an infinite number of real eigenvalues for the latter.
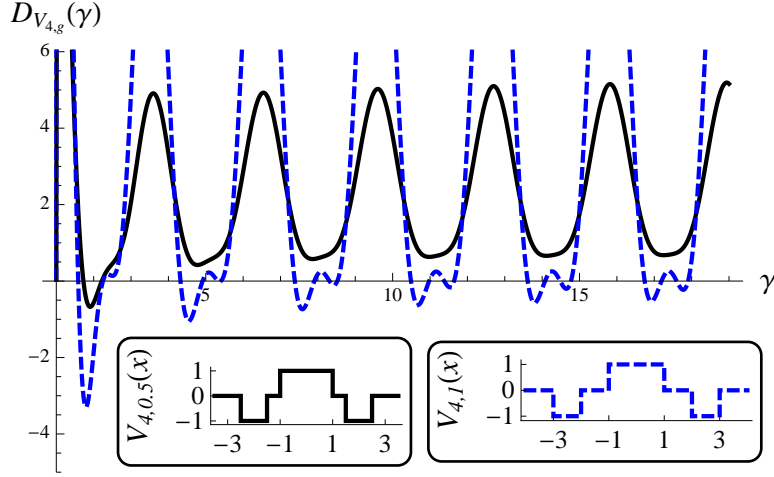

Figure 8: The graphs of $D_{V_{4,g}}(\gamma)$ against real γ for the potentials $V_{4,g}(x) := W(x; [-g-2, -g-1, -1, 1, g+1, g+2]; \{-1, 0, 1, 0, -1\})$ with $g = 0.5$ (solid black line) and $g = 1$ (dashed blue line).

To explain this phenomenon, we once more consider equation (13), now with $V = V_{4,g}$. Although the explicit expression for the determinant $D_{V_{4,g}}$ is rather cumbersome, some simplifications lead to the asymptotic form

$$\left(1 + \sqrt{2}\tanh(g)\cos(2\widetilde{\gamma} - \pi/4)\right)\left(1 + \sqrt{2}\tanh(g)\cos(2\widetilde{\gamma} + \pi/4)\right) + O\left(\widetilde{\gamma}^{-1}\right) = 0 \qquad (17)$$

as $\widetilde{\gamma} \to \infty$ (where the derivative of the O-term is also $O\left(\widetilde{\gamma}^{-1}\right)$). The asymptotics of the number of zeros of (17) reduces to consideration of a pair of elementary equations for cosine. It is then immediate that the asymptotics of $\#(\Sigma_{V_{4,g}} \cap [0, R])$ as $R \to \infty$ changes abruptly between $O(1)$ and $\dfrac{R}{\pi}\|V_{4,g}\|_{L^1} + o(R)$, depending on whether $g < g_0 := \operatorname{arctanh}(1/\sqrt{2}) \approx 0.8814$ or $g > g_0$, respectively. This shows, as already announced in Remark 1.12, that unlike no-gap and one-gap potentials, two-gap potentials with zero integral may produce an infinite number of real eigenvalues.

Example 2.7 (Potential from [HRP]). Using a complicated explicit solution involving special functions, Hartmann, Robinson and Portnoi found that, for the potential $V_{\mathsf{HRP}}(x) = -1/\cosh(x)$ and any $k > 0$, the positive part of the spectrum $\Sigma_{V_{\mathsf{HRP}}}$ coincides with the set $k - \frac{1}{2} + \mathbb{N}$. We treat this potential using the Prüfer method and plot, for real γ, the quantity $\cos(\Delta_{V_{\mathsf{HRP}}}(\gamma))$; by Proposition 3.7, $\gamma \in \Sigma_{V_{\mathsf{HRP}}} \cap \mathbb{R}$ if and only if $\cos(\Delta_{V_{\mathsf{HRP}}}(\gamma)) = 0$ (see Section 3.3 for the definition of $\Delta_{V_{\mathsf{HRP}}}$ and further details). The curves in Figure 9 (drawn for $k = 1$ and $k = 1.5$) illustrate the result of [HRP].
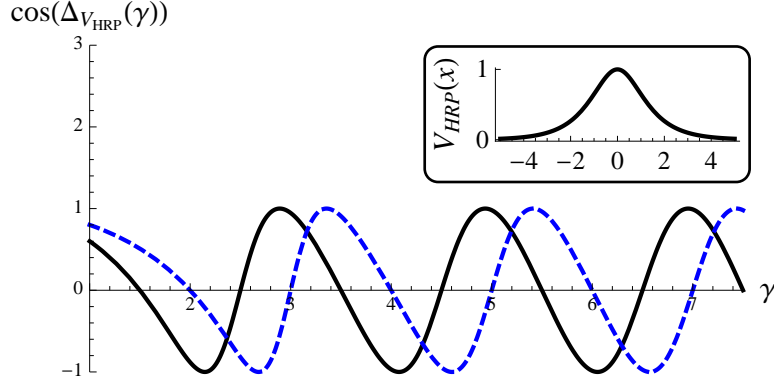
13

Figure 9: The graphs of $\cos(\Delta_{V_{\mathrm{HRP}}}(\gamma))$ for $k = 1$ (solid black line) and $k = 1.5$ (dashed blue line).

3 Main arguments

In this section we give the arguments for the main theorems based on a series of more technical results; the proofs for the latter will be deferred to Section 4.

3.1 General

The unperturbed operator T_0 is an unbounded self-adjoint operator on L^2 whose domain is H^1, the Sobolev space of ($\mathbb{C}^2$ valued) functions on $\mathbb{R}$; that is

$$\mathrm{dom}(\mathrm{T}_0) = H^1 = \{\boldsymbol{\psi} \in L^2 \,:\, \nabla\boldsymbol{\psi} \in L^2\}.$$

In fact it is straightforward to check that $\|\mathrm{T}_0\boldsymbol{\psi}\|^2 = \|\nabla\boldsymbol{\psi}\|^2 + k^2\|\boldsymbol{\psi}\|^2$, so $\|\mathrm{T}_0\boldsymbol{\psi}\|$ is equivalent to $\|\boldsymbol{\psi}\|_{H^1}$. It follows that T_0 defines an isomorphism $H^1 \to L^2$.

Next we consider multiplication by an element of $\mathbb{V}_0$. Firstly note that a norm can be defined on $\mathbb{V}_0$ using the expression

$$\|V\|_{\mathbb{V}_0} = \sup_{x \in \mathbb{R}} \|V\|_{L^2(x-1, x+1)}.$$

This norm makes $\mathbb{V}_0$ a Banach space in which C_0^∞ is a dense subset.

Lemma 3.1. *Multiplication by a fixed $V \in \mathbb{V}_0$ defines a compact map $H^1 \to L^2$.*

Proof. Initially suppose $V \in C_0^\infty$. Choose a bounded interval I with $\mathrm{supp}(V) \subseteq I$. We can view multiplication by V as a composition $\boldsymbol{\psi} \mapsto \boldsymbol{\psi}|_I \mapsto (V\boldsymbol{\psi})|_I \mapsto V\boldsymbol{\psi}$ where we firstly restrict to I, then multiply by V and finally extend by 0. This gives a map $H^1 \to L^2(I) \to L^2(I) \to L^2$, where the last two steps are continuous and the first step is compact (by the Rellich-Kondrachov Theorem; see [Ad], for example).

Since C_0^∞ is dense in $\mathbb{V}_0$ and the set of compact maps is closed, it now suffices to show that multiplication defines a continuous bilinear map $\mathbb{V}_0 \times H^1 \to L^2$. To this end firstly note that the Sobolev Embedding Theorem (*ibid.*) gives

$$\|\boldsymbol{\psi}\|_{L^\infty(x-1, x+1)} \leq C\|\boldsymbol{\psi}\|_{H^1(x-1, x+1)}$$

for some constant C (which is independent of x). Thus

$$\|V\boldsymbol{\psi}\|_{L^2(x-1, x+1)} \leq \|V\|_{L^2(x-1, x+1)} \|\boldsymbol{\psi}\|_{L^\infty(x-1, x+1)} \leq C\,\|V\|_{L^2(x-1, x+1)} \|\boldsymbol{\psi}\|_{H^1(x-1, x+1)}.$$

14

On the other hand

$$\|\boldsymbol{\psi}\|_{L^2}^2 = \frac{1}{2} \int_{\mathbb{R}} \|\boldsymbol{\psi}\|_{L^2(x-1,x+1)}^2 \, \mathrm{d}x,$$

with a similar expression holding for $\|\cdot\|_{H^1}$. Combining the above then gives

$$\begin{aligned}
\|V\boldsymbol{\psi}\|_{L^2}^2 &\le \frac{1}{2}\, C^2 \int_{\mathbb{R}} \|V\|_{L^2(x-1,x+1)}^2 \|\boldsymbol{\psi}\|_{H^1(x-1,x+1)}^2 \, \mathrm{d}x \\
&\le \frac{1}{2}\, C^2 \sup_{x\in\mathbb{R}} \|V\|_{L^2(x-1,x+1)}^2 \int_{\mathbb{R}} \|\boldsymbol{\psi}\|_{H^1(x-1,x+1)}^2 \, \mathrm{d}x = C^2 \, \|V\|_{\mathbb{V}_0}^2 \, \|\boldsymbol{\psi}\|_{H^1}^2,
\end{aligned}$$

for any $V \in \mathbb{V}_0$ and $\boldsymbol{\psi} \in H^1$. $\qquad\square$

Since $H^1 = \mathrm{dom}(\mathrm{T}_0)$, Lemma 3.1 is equivalent to the statement that (multiplication by) $V \in \mathbb{V}_0$ is a relatively compact perturbation of T_0. It follows that the sum $\mathrm{T}_0 + V = \mathrm{T}_V$ defines a self-adjoint operator with $\mathrm{dom}(\mathrm{T}_V) = \mathrm{dom}(\mathrm{T}_0) = H^1$ (see [ReSi]).

Although we're interested in real valued potentials it is helpful to consider some basic results for more general complex valued potentials as well. Let $\widetilde{\mathbb{V}}_0 = \mathbb{V}_0 \otimes_{\mathbb{R}} \mathbb{C}$ denote the complex valued version of $\mathbb{V}_0$ (in other words, $\widetilde{\mathbb{V}}_0$ consists of functions of the form $U + iW$ where $U, W \in \mathbb{V}_0$). Note that $\gamma V \in \widetilde{\mathbb{V}}_0$ for any $V \in \mathbb{V}_0$ and $\gamma \in \mathbb{C}$. Now Lemma 3.1 clearly extends to $\widetilde{\mathbb{V}}_0$, so (multiplication by) some $\mathcal{V} \in \widetilde{\mathbb{V}}_0$ is still a relatively compact perturbation of T_0. Thus the sum $\mathrm{T}_0 + \mathcal{V} = \mathrm{T}_{\mathcal{V}}$ defines a closed operator with $\mathrm{dom}(\mathrm{T}_{\mathcal{V}}) = \mathrm{dom}(\mathrm{T}_0) = H^1$. Although the operator $\mathrm{T}_{\mathcal{V}}$ will not be self-adjoint (unless $\mathcal{V} \in \mathbb{V}_0$), the essential spectrum is still given by (2), while any spectrum of $\mathrm{T}_{\mathcal{V}}$ in $\mathbb{C} \setminus \Lambda_k$ consists of isolated eigenvalues of finite (algebraic) multiplicity.

Proof of Theorem 1.1. We have $\gamma \in \Sigma_V$ iff $(\mathrm{T}_0 + \gamma V)\boldsymbol{\psi} = 0$ for some non-trivial $\boldsymbol{\psi} \in H^1$. In turn this is equivalent to $V\mathrm{T}_0^{-1}\boldsymbol{\varphi} = \mu\boldsymbol{\varphi}$ where $\boldsymbol{\varphi} = \mathrm{T}_0\boldsymbol{\psi} \in L^2$ and $\mu = -1/\gamma$; in other words, $-1/\gamma$ should be an eigenvalue of the operator $V\mathrm{T}_0^{-1}$. However $V\mathrm{T}_0^{-1}$ is a compact operator by Lemma 3.1, so must have discrete spectrum away from 0. The result follows. $\qquad\square$

It is helpful to have a more symmetric version of the idea just used in the proof of Theorem 1.1. Fix $V \in \mathbb{V}_0$ and let J_V be the operator on L^2 given as multiplication by $\mathrm{sgn}(V)$ (where, for definiteness, we can set $\mathrm{sgn}(V)(x) = +1$ whenever $V(x) = 0$). Then $J_V^* = J_V^{-1} = J_V$ while

$$V = \sqrt{|V|}\, J_V \sqrt{|V|}.$$

Define an operator A_V by

$$A_V = \sqrt{|V|}\,\mathrm{T}_0^{-1}\sqrt{|V|}.$$

Using similar ideas to those in the proofs of Theorem 1.1 and Lemma 3.1 we get the following:

Lemma 3.2. *The operator A_V is a compact self-adjoint operator on L^2. Furthermore, we have $\gamma \in \Sigma_V$ iff $-1/\gamma \in \mathrm{spec}(J_V A_V)$.*

If V is single-signed we can choose $J_V = \pm I$ ($+I$ if $V \ge 0$ or $-I$ if $V \le 0$). Then $\gamma \in \Sigma_V$ iff $-1/\gamma$ is in the spectrum of the compact self-adjoint operator $\pm A_V$. This gives a justification of Theorem 1.5, although a more elementary argument is also possible (see Section 3.2).

Remark 3.3. Lemma 3.2 can be viewed as a Birman-Schwinger principle for the Dirac operator T_V. This wide ranging principle has been used to obtain a number of results related to those presented here, both in the single-sign case where the associated Birman-Schwinger operator A_V is self-adjoint (see [Kl, BiLa]), and in the variable sign case where we need to consider the non-self-adjoint operator $J_V A_V$ (see [Sa]). Approaches based on the Birman-Schwinger principle

rely on obtaining spectral information about the operator A_V or $J_V A_V$. Potential sources of information include eigenvalue or singular value estimates (such as [Cw]), or pseudo-differential techniques leading to eigenvalue asymptotics (such as [BiSo]). In Section 3.3 we take a different approach, based on Prüfer techniques; this is convenient in our one-dimensional setting, and allows for slightly less restrictive assumptions on the potential (see Remark 3.16 in particular).

For any $V \in \mathbb{V}_0$ note that $|V|$ is a single-signed potential which is also in $\mathbb{V}_0$. Using the association with compact operators given by Lemma 3.2 we are able to link points in Σ_V and $\Sigma_{|V|}$ with the eigenvalues and singular values of a single operator, and thus estimate the former using the latter (via Weyl's Inequality). Let $0 < \mu_1 < \mu_2 < \dots$ be the positive points in $\Sigma_{|V|}$, ordered by size. Also let $\gamma_1, \gamma_2, \dots$ denote the points in $\Sigma_V \cap \{z \in \mathbb{C} : z \in \mathbb{R}^+ \text{ or } \mathrm{Im}\, z > 0\}$ ordered so as to have non-decreasing modulus (and counted according to algebraic multiplicity).

Lemma 3.4. *For any $N \in \mathbb{N}$ we have*

$$\prod_{n=1}^{N} |\gamma_n| \geq \prod_{n=1}^{N} \mu_n.$$

Proof. By symmetry we know that the points in Σ_V are just $\pm\gamma_n$, $n = 1, 2, \dots$, while the points in $\Sigma_{|V|}$ are just $\pm\mu_n$, $n = 1, 2, \dots$. Lemma 3.2 now implies $\pm 1/\gamma_n$, $n = 1, 2, \dots$ are the non-zero eigenvalues of $J_V A_V$, while $\pm 1/\mu_n$, $n = 1, 2, \dots$ are the non-zero eigenvalues of $J_{|V|} A_{|V|} = A_V$ (note that, we can take $J_{|V|} = I$). However $(J_V A_V)^* J_V A_V = A_V J_V^2 A_V = A_V^2$ so the singular values of $J_V A_V$ are just the eigenvalues of $|A_V|$; that is, $1/\mu_n$, $n = 1, 2, \dots$, where each eigenvalue has multiplicity 2. Given $N \in \mathbb{N}$ we can now use Weyl's Inequality (see [Wey]) to compare the largest $2N$ eigenvalues and singular values of the compact operator $J_V A_V$; this gives

$$\prod_{n=1}^{N} \left(\frac{1}{|\gamma_n|} \right)^2 \leq \prod_{n=1}^{N} \left(\frac{1}{\mu_n} \right)^2.$$

The result follows. $\qquad\square$

3.2 Symmetries

Our unperturbed operator is

$$\mathrm{T}_0 = -i\sigma_2 \nabla + k\sigma_3 = \begin{pmatrix} k & -\nabla \\ \nabla & -k \end{pmatrix}.$$

Since V is real valued we immediately get

$$\overline{(\mathrm{T}_0 + \gamma V)\psi} = (\mathrm{T}_0 + \overline{\gamma} V)\overline{\psi}. \tag{18}$$

It follows that $\gamma \in \Sigma_V$ iff $\overline{\gamma} \in \Sigma_V$. On the other hand, the commutator properties of the Pauli matrices (namely $\sigma_j \sigma_k = -\sigma_k \sigma_j$ if $j \neq k$) give us

$$\sigma_1 (\mathrm{T}_0 + \gamma V) = -(\mathrm{T}_0 - \gamma V)\sigma_1 \quad \text{and} \quad \sigma_2 (\mathrm{T}_0 + \gamma V) = (-i\sigma_2 \nabla - \sigma_3 k + \gamma V)\sigma_2.$$

From the first identity we get $\gamma \in \Sigma_V$ iff $-\gamma \in \Sigma_V$, while the second shows that Σ_V is invariant if we replace k with $-k$ in the definition of T_0.

Remark 3.5. The symmetry corresponding to σ_3 can be used to help study even potentials (compare with our consideration of anti-symmetric potentials below).

Proof of Theorem 1.5. Suppose $T_{\gamma V}\psi = (T_0 + \gamma V)\psi = 0$ for some $\gamma \in \mathbb{C}$ and $\psi \in H^1$. Then $\langle T_0\psi, \psi\rangle = -\gamma\langle V\psi, \psi\rangle$ while $\langle T_0\psi, \psi\rangle, \langle V\psi, \psi\rangle \in \mathbb{R}$. If $\gamma \notin \mathbb{R}$ we must therefore have $\langle V\psi, \psi\rangle = 0$. Since V is single-signed it follows that $V\psi = 0$, leading to $T_0\psi = -\gamma V\psi = 0$ and thus $\psi = 0$ (recall that $T_0 : H^1 \to L^2$ is an isomorphism). $\qquad\square$

Proof of Theorem 1.7. We consider two symmetries of the operator $T_{\gamma V}$; define an anti-linear operator C and a unitary operator S on L^2 by

$$C\psi = \overline{\psi} \quad \text{and} \quad (S\psi)(x) = \sigma_2\psi(-x), \ x \in \mathbb{R}.$$

These operators map H^1 (isometrically) onto H^1, and satisfy $C^2 = I = S^2$ and $CS = -SC$. Furthermore, (18) can be rewritten as $CT_{\gamma V} = T_{\overline{\gamma}V}C$, while $\nabla S = -S\nabla$ and $SV = -VS$ (as V is anti-symmetric) which leads to $ST_{\gamma V} = -T_{\gamma V}S$.

Now suppose $\gamma \in \Sigma_V \cap \mathbb{R}$ and choose $0 \neq \psi \in H^1$ which satisfies $T_{\gamma V}\psi = 0$. Then

$$T_{\gamma V}C\psi = CT_{\gamma V}\psi = 0 \quad \text{and} \quad T_{\gamma V}S\psi = -ST_{\gamma V}\psi = 0.$$

However 0 is a simple eigenvalue of the operator $T_{\gamma V}$ by Lemma 4.1, so we must have $C\psi = \alpha\psi$ and $S\psi = \beta\psi$ for some $\alpha, \beta \in \mathbb{C}$. Then

$$|\alpha|^2\psi = C^2\psi = \psi, \qquad \beta^2\psi = S^2\psi = \psi \qquad \text{and} \qquad \alpha\overline{\beta}\psi = CS\psi = -SC\psi = -\alpha\beta\psi,$$

so $|\alpha|^2 = 1 = \beta^2$ and $\alpha\overline{\beta} = -\alpha\beta$ (recall that $\psi \neq 0$). These equations clearly have no solution, so we must have $\Sigma_V \cap \mathbb{R} = \emptyset$. $\qquad\square$

3.3 General bounds and asymptotics

Suppose $V \in L^1_{\mathrm{loc}}$. We can view our basic equation $T_{\gamma V}\psi = 0$ as the 2×2 system of first order ordinary differential equations on $\mathbb{R}$ given by

$$\begin{aligned}\nabla\psi_1 &= (k - \gamma V)\psi_2, \\ \nabla\psi_2 &= (k + \gamma V)\psi_1.\end{aligned} \tag{19}$$

The basic theory for such equations is well established (see [Ha]; here, and in subsequent references to [Ha], some straightforward modifications of the results are needed in order to cover L^1_{loc} coefficients). In particular, if $x_0 \in \mathbb{R}$ and $\alpha_1, \alpha_2 \in \mathbb{C}$, then there exists a unique absolutely continuous solution to (19) on $\mathbb{R}$ with $\psi_j(x_0) = \alpha_j$ for $j = 1, 2$. Furthermore, for given $x \in \mathbb{R}$, this solution depends continuously on $\alpha_1, \alpha_2, \lambda$ and V (the latter as a function in $L^1(I)$ where I is the interval between x_0 and x). A consequence of the uniqueness of solutions is that for any non-trivial solution ψ of (19) we have $\psi(x) \neq 0$ for all $x \in \mathbb{R}$.

Now suppose $\gamma \in \mathbb{R}$. It follows that all coefficients in (19) are real, so we may restrict our attention to solutions which are also real valued. Since any non-trivial solution ψ is absolutely continuous and satisfies $\psi(x) \neq 0$ for all $x \in \mathbb{R}$ we can define an absolutely continuous function $S : \mathbb{R} \to \mathbb{S}^1$ by

$$S(x) = \frac{\psi_1(x) + \mathrm{i}\psi_2(x)}{(\psi_1^2(x) + \psi_2^2(x))^{1/2}}$$

(here $\mathbb{S}^1$ denotes the unit circle in $\mathbb{C}$). By lifting to $\mathbb{R}$ (the universal cover of $\mathbb{S}^1$) we can define a further absolutely continuous function $\theta : \mathbb{R} \to \mathbb{R}$ so that

$$\mathrm{e}^{\mathrm{i}\theta} = S = \frac{\psi_1 + \mathrm{i}\psi_2}{|\psi|}; \tag{20}$$

this function is the Prüfer argument associated to ψ and is unique up to the addition of a constant in $2\pi\mathbb{Z}$. We note that Prüfer coordinates are a standard tool for problems of this kind (see, for example, [Sch]).

A straightforward calculation gives

$$\mathrm{i}\nabla\theta\,\mathrm{e}^{\mathrm{i}\theta} = \nabla S = \mathrm{i}\,\frac{\psi_1\nabla\psi_2 - \psi_2\nabla\psi_1}{\psi_1^2 + \psi_2^2}\,\mathrm{e}^{\mathrm{i}\theta}.$$

Using (19) it follows that

$$\nabla\theta = \frac{\psi_1\nabla\psi_2 - \psi_2\nabla\psi_1}{\psi_1^2 + \psi_2^2} = \gamma V + k\,\frac{\psi_1^2 - \psi_2^2}{\psi_1^2 + \psi_2^2}.$$

However

$$\frac{\psi_1^2 - \psi_2^2}{\psi_1^2 + \psi_2^2} = \mathrm{Re}\,\mathrm{e}^{2\mathrm{i}\theta} = \cos(2\theta)$$

so θ satisfies the first order non-linear equation

$$\nabla\theta = \gamma V + k\cos(2\theta). \tag{21}$$

Conversely, when $V \in L^1_{\mathrm{loc}}$ the differential equation (21) has a unique absolutely continuous solution valid on $\mathbb{R}$ for a given value of $\theta(x_0)$, $x_0 \in \mathbb{R}$ (see [Ha]). If we have one solution θ of (21) then $\theta + n\pi$ provides a further solution for any $n \in \mathbb{Z}$ (note that, for θ given by (20) the solution $\theta + \pi$ corresponds to taking $-\psi$ as a solution of (19)).

Now suppose $V \in \mathbb{V}_1$ and ψ is a non-trivial solution of (19). Let θ be given by (20). The fact that $\int_{\mathbb{R}}|V(x)|\,\mathrm{d}x < +\infty$ means that, to leading order as $x \to \pm\infty$, ψ behaves like a solution to (19) with $V \equiv 0$. The corresponding asymptotic behaviour of θ can be summarised as follows (see Section 4.1 for more details):

Lemma 3.6. *The quantity $\theta(x)$ has well defined limits as $x \to \pm\infty$ which satisfy*

$$\theta(\pm\infty) := \lim_{x\to\pm\infty}\theta(x) \in \frac{\pi}{4} + \frac{\pi}{2}\,\mathbb{Z}. \tag{22}$$

Furthermore, upon restriction we have $\psi \in L^2(\mathbb{R}^{\pm})$ iff

$$\theta(\pm\infty) \in \mp\frac{\pi}{4} + \pi\mathbb{Z}.$$

For any $\gamma \in \mathbb{R}$ we can uniquely specify two solutions $\theta_{\gamma,+}$ and $\theta_{\gamma,-}$ to (21) by imposing the boundary conditions

$$\theta_{\gamma,\pm}(\pm\infty) = \mp\frac{\pi}{4}. \tag{23}$$

Lemma 3.6 shows that $\theta_{\gamma,\pm}$ correspond to solutions of (19) which are in L^2 on $\mathbb{R}^{\pm}$. We get an L^2 solution on the whole of $\mathbb{R}$ precisely when these solutions 'match' at one (or, equivalently, any) point of $\mathbb{R}$. Choosing 0 as the point at which we check, the matching condition is simply that $\theta_{\gamma,+}(0)$ and $\theta_{\gamma,-}(0)$ must differ by a multiple of π. Define a function $\Delta_V : \mathbb{R} \to \mathbb{R}$ by setting

$$\Delta_V(\gamma) = \big(\theta_{\gamma,+}(+\infty) - \theta_{\gamma,+}(0)\big) + \big(\theta_{\gamma,-}(0) - \theta_{\gamma,-}(-\infty)\big)$$
$$= -\frac{\pi}{2} - \theta_{\gamma,+}(0) + \theta_{\gamma,-}(0). \tag{24}$$

We can thus characterise points in $\Sigma_V \cap \mathbb{R}$ as follows:

Proposition 3.7. *We have $\gamma \in \Sigma_V \cap \mathbb{R}$ iff $\Delta_V(\gamma) \in \dfrac{\pi}{2} + \pi\mathbb{Z}$.*

We have $\Delta_V(0) = 0$ (note that $\theta_{0,\pm}$ are constant functions) while $\Delta_V(\gamma)$ depends continuously on γ (this essentially follows from standard results for the continuous dependence on parameters of solutions of ordinary differential equations; see [Ha]). For large γ we have the following asymptotic behaviour (the proof is given in Section 4.2):

Proposition 3.8. *If $V \in \mathbb{V}_1$ then*

$$\Delta_V(\gamma) = \gamma \int_{\mathbb{R}} V(x) \, \mathrm{d}x + o(\gamma)$$

as $|\gamma| \to \infty$.

Together Propositions 3.7 and 3.8 allow us to establish Theorem 1.4.

Proof of Theorem 1.4. Let $I_R \subset \mathbb{R}$ denote the closed interval with endpoints $\Delta_V(0) = 0$ and $\Delta_V(R)$. By the Intermediate Value Theorem $\Delta_V(\gamma)$ takes each value in I_R at least once for some $\gamma \in [0, R]$. From Proposition 3.7 we then get

$$\#(\Sigma_V \cap [0, R]) = \#\big\{\gamma \in [0, R] : \Delta_V(\gamma) \in (\mathbb{Z} + 1/2)\pi\big\}$$
$$\geq \#(I_R \cap (\mathbb{Z} + 1/2)\pi) \geq \frac{|I_R|}{\pi} - \frac{1}{2}, \tag{25}$$

where $|I_R|$ is the length of I_R. On the other hand, Proposition 3.8 gives

$$|I_R| = |\Delta_V(R)| = R \left| \int_{\mathbb{R}} V(x) \, \mathrm{d}x \right| + o(R) \tag{26}$$

as $R \to \infty$. $\qquad\square$

Remark 3.9. For compactly supported potentials, Theorem 1.4 admits a simpler proof, similarly to [Sch, Theorem 3]. The details of this argument have been written down by Michael Morin (supported by the NSERC Undegraduate Summer Research Award, 2011).

If $V \in \mathbb{V}_1$ is single-signed then Δ_V is monotonic (the proof is given in Section 3.4):

Proposition 3.10. *Suppose $V \in \mathbb{V}_1$ is single-signed and non-trivial. Then Δ_V is strictly increasing if $V \geq 0$ and strictly decreasing if $V \leq 0$.*

Using this monotonicity we can sharpen the argument given for Theorem 1.4 to establish Theorem 1.6.

Proof of Theorem 1.6. If $V \in \mathbb{V}_1$ is single-signed it follows from Proposition 3.10 that there is at most one γ with $\Delta_V(\gamma) = (n + 1/2)\pi$ for any given $n \in \mathbb{Z}$. Thus (25) in the proof of Theorem 1.4 can be modified to give

$$\#(\Sigma_V \cap [0, R]) = \#\big\{\gamma \in [0, R] : \Delta_V(\gamma) \in (\mathbb{Z} + 1/2)\pi\big\}$$
$$= \#(I_R \cap (\mathbb{Z} + 1/2)\pi) \leq \frac{|I_R|}{\pi} + \frac{1}{2}.$$

The required upper asymptotic bound on $\#(\Sigma_V \cap [0, R])$ can then be obtained using (26). $\quad\square$

To give uniform estimates for $\Delta_V(\gamma)$ we firstly introduce an auxiliary function. For any $a \geq 0$ let $\lfloor a \rfloor$ denote the largest integer not exceeding a; that is,

$$\lfloor a \rfloor = \max\{n \in \mathbb{Z} : n \leq a\}.$$

Now define a function $h : [0, +\infty) \to [0, +\infty)$ by

$$h(a) = a + \frac{\pi}{2}\left\lfloor \frac{2}{\pi}a \right\rfloor.$$

In particular, h is strictly increasing, $h(a) = a$ for $a < \pi/2$, $a \leq h(a) \leq 2a$ for all $a \geq 0$, and $h(a) + h(b) \leq h(a+b)$ for all $a, b \geq 0$ (note that $\lfloor a \rfloor + \lfloor b \rfloor \leq \lfloor a+b \rfloor$).

As a general bound on $\Delta_V(\gamma)$ we have the following (the proof is given in Section 4.2):

Proposition 3.11. *If $V \in \mathbb{V}_1$ and $\gamma \in \mathbb{R}$ then $|\Delta_V(\gamma)| \leq h(|\gamma|\,\|V\|_{L^1})$.*

This result leads to uniform lower bounds for points in Σ_V, and in turn helps to justify Theorem 1.3. We firstly deal with the case when V is single-signed (Proposition 3.12), and then consider arbitrary potentials (Proposition 3.13) by using Lemma 3.4 to reduce this to the single-sign case.

Proposition 3.12. *Suppose $V \in \mathbb{V}_1$ is single-signed and non-trivial. Let $0 < \gamma_1 < \gamma_2 < \ldots$ denote the sequence of positive points in Σ_V, arranged in order of increasing size. Then*

$$\gamma_n \geq \frac{\pi}{2\|V\|_{L^1}}\, n, \quad n = 1, 2, \ldots.$$

Note that the constant in the bound is only half the asymptotic value (see Theorem 1.6).

Proof. Suppose $V \geq 0$ (the case $V \leq 0$ can be treated similarly). By Propositions 3.7 and 3.10 we have $\Delta_V(\gamma_n) = (n - 1/2)\pi$ for $n = 1, 2, \ldots$. Now if $a < n\pi/2$ then $\lfloor 2a/\pi \rfloor < n$ so

$$h(a) \leq a + (n-1)\frac{\pi}{2} \leq \left(n - \frac{1}{2}\right)\pi = \Delta_V(\gamma_n) \leq h(\gamma_n\,\|V\|_{L^1})$$

by Proposition 3.11. However h is strictly increasing so we must have $a \leq \gamma_n\,\|V\|_{L^1}$. Taking $a \to (n\pi/2)^-$ now gives the result. $\qquad\square$

Proposition 3.13. *Suppose $V \in \mathbb{V}_1$ is non-trivial and let $\gamma_1, \gamma_2, \ldots$ denote the points in $\Sigma_V \cap \{z \in \mathbb{C} : z \in \mathbb{R}^+ \text{ or } \operatorname{Im} z > 0\}$ ordered so as to have non-decreasing modulus (and counted according to algebraic multiplicity). Then*

$$|\gamma_n| \geq \frac{\pi}{2\mathrm{e}\|V\|_{L^1}}\, n, \quad n = 1, 2, \ldots.$$

Proof. Let $0 < \mu_1 < \mu_2 < \ldots$ denote the positive points in $\Sigma_{|V|}$, arranged in order of increasing size. Noting that $|V| \in \mathbb{V}_1$ is single-signed and non-trivial, Proposition 3.12 gives $\mu_n \geq \nu n$ for $n = 1, 2, \ldots$, where $\nu = \pi/(2\|V\|_{L^1})$. Using Lemma 3.4 and the ordering on the γ_n's we now get

$$|\gamma_N|^N \geq \prod_{n=1}^{N} |\gamma_n| \geq \prod_{n=1}^{N} \mu_n \geq \nu^N N! \geq \nu^N \left(\frac{N}{\mathrm{e}}\right)^N$$

for any $N \in \mathbb{N}$. $\qquad\square$

Proof of Theorem 1.3. If $V = 0$ then $\Sigma_V = \emptyset$ and there is nothing to prove. Now suppose V is non-trivial. Let $R \geq 0$, set $N = \#\big(\Sigma_V \cap \{z \in \mathbb{C} : |z| \leq R\}\big)$ and suppose $N > 0$. Using the symmetry $\Sigma_V = -\Sigma_V$ we know that $N = 2M$ for some $M \in \mathbb{N}$. With $\gamma_1, \gamma_2, \dots$ defined as in Proposition 3.13 it follows that $|\gamma_M| \leq R$. We then obtain

$$N = 2M \leq 2 \frac{2\mathrm{e}}{\pi} \|V\|_{L^1} |\gamma_M| \leq \frac{4\mathrm{e}}{\pi} \|V\|_{L^1} R$$

using this result. $\qquad\square$

Remark 3.14. Various other estimates can obtained from straightforward modifications to the proof of Theorem 1.3 presented above. For example, if $V \in \mathbb{V}_1$ is single-signed we can use Proposition 3.12 in place of Proposition 3.13 to obtain the uniform upper bound

$$\#(\Sigma_V \cap [0, R]) \leq \frac{2}{\pi} \|V\|_{L^1} R$$

for any $R \geq 0$. Alternatively, for any $V \in \mathbb{V}_1$ we can estimate μ_n in the proof of Proposition 3.13 using Theorem 1.6 instead of Proposition 3.12; this leads to the asymptotic bound

$$\#\big(\Sigma_V \cap \{z \in \mathbb{C} : |z| \leq R\}\big) \leq \frac{2\mathrm{e}}{\pi} \|V\|_{L^1} R + o(R)$$

as $R \to \infty$.

As a direct application of the main result in [ElTa] we can extend the general upper bound given by Theorem 1.3 to potentials in $\mathbb{V}_0 \cap L^p$ for $1 < p < \infty$.

Theorem 3.15. *Suppose $V \in \mathbb{V}_0 \cap L^p$ for some $1 < p < \infty$. Then*

$$\#\big(\Sigma_V \cap \{z \in \mathbb{C} : |z| \leq R\}\big) \leq C_p \|V\|_{L^p}^p R^p$$

for any $R \geq 0$, where C_p is a constant depending only on p.

Remark 3.16. The main result in [ElTa] is based on singular value estimates from [Cw]. We can't extend the argument to cover $p = 1$ as it corresponds to an excluded boundary case in [Cw].

3.4 Derivatives

To study the derivatives of $\Delta_V(\gamma)$ we need to obtain more information about the γ dependence of solutions to (21). Firstly consider any closed bounded interval $I = [a, b] \subset \mathbb{R}$ and potential $V \in L^1_{\mathrm{loc}}$ on I. For each $\gamma \in \mathbb{R}$ suppose we have a solution θ_γ of (21) where $\theta_\gamma(b)$ depends twice differentiably on γ. Standard results for ordinary differential equations (see [Ha]) then imply $\theta_\gamma(x)$ is twice differentiable in γ for each $x \in I$ (in fact θ_γ will depend analytically on γ provided $\theta_\gamma(b)$ does), so we can set

$$\omega_\gamma(x) = \frac{\mathrm{d}}{\mathrm{d}\gamma} \theta_\gamma(x) \quad \text{and} \quad \rho_\gamma(x) = \frac{\mathrm{d}^2}{\mathrm{d}\gamma^2} \theta_\gamma(x).$$

From (21) we immediately get

$$\nabla \omega_\gamma = V - 2k \sin(2\theta_\gamma)\, \omega_\gamma \quad \text{and} \quad \nabla \rho_\gamma(x) = -4k \cos(2\theta_\gamma)\, \omega_\gamma^2 - 2k \sin(2\theta_\gamma)\, \rho_\gamma.$$

Then $\nabla(\mathrm{e}^G \omega_\gamma) = \mathrm{e}^G V$, where G is any function satisfying $\nabla G(x) = 2k\sin(2\theta_\gamma(x))$. Thus

$$\omega_\gamma(x) = \mathrm{e}^{G(b)-G(x)}\omega_\gamma(b) - \int_x^b \mathrm{e}^{G(t)-G(x)}V(t)\,\mathrm{d}t$$

$$= \mathrm{e}^{2k\Psi_{[x,b]}}\omega_\gamma(b) - \int_x^b \mathrm{e}^{2k\Psi_{[x,t]}}V(t)\,\mathrm{d}t, \tag{27}$$

where, for any interval $J \subseteq I$,

$$\Psi_J := \int_J \sin(2\theta_\gamma(x))\,\mathrm{d}x. \tag{28}$$

We need to consider (27) with $x = a$ and $b \to +\infty$ when we take $\theta_\gamma = \theta_{\gamma,+}$. Let $\omega_{\gamma,+}$ denote the corresponding derivative in this case. Since $\theta_{\gamma,+}(+\infty) = -\pi/4$ is constant (recall (23)) we would expect $\omega_{\gamma,+}(+\infty) = 0$. Furthermore $\sin(2\theta_{\gamma,+}(t)) \to -1$ as $t \to \infty$, so $\mathrm{e}^{2k\Psi_{[a,t]}} \to 0$ as $t \to \infty$. The precise properties that we require are given in the next result; these can be justified by straightforward if somewhat lengthy arguments.

Proposition 3.17. *The solution $\theta_{\gamma,+}$ to (21) depends on γ differentiably, while the derivative satisfies*

$$\omega_{\gamma,+}(a) := \frac{\mathrm{d}}{\mathrm{d}\gamma}\theta_{\gamma,+}(a) = -\int_a^\infty \mathrm{e}^{2k\Psi_{[a,x]}}V(x)\,\mathrm{d}x$$

for all $a \in \mathbb{R}$.

There is a corresponding result for $\theta_{\gamma,-}$. Proposition 3.10 is now an easy corollary of these results.

Proof of Proposition 3.10. From (24) we get

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\Delta_V(\gamma) = -\omega_{\gamma,+}(0) + \omega_{\gamma,-}(0).$$

Now suppose $V \geq 0$ (the case $V \leq 0$ can be handled similarly). By Proposition 3.17 we have

$$-\omega_{\gamma,+}(0) = \int_0^\infty \mathrm{e}^{2k\Psi_{[0,x]}}V(x)\,\mathrm{d}x.$$

Since $\mathrm{e}^{2k\Psi_J} > 0$ for any (bounded) interval $J \subseteq \mathbb{R}$ the right hand side is non-negative and equal to 0 only if $V = 0$ on $\mathbb{R}^+$ (as an L^1 function). A similar argument shows that $\omega_{\gamma,-}(0)$ is also non-negative and equal to 0 only if $V = 0$ on $\mathbb{R}^-$. The result follows. $\qquad\square$

3.5 No gaps

For a potential $V \in L^1_{\mathrm{loc}}$ and interval $I \subseteq \mathbb{R}$ let $\mathrm{var}_I(V)$ denote the total variation of V on I. We also say that V *has no gaps on I* if

$$\left|I \cap V^{-1}(0)\right| = 0.$$

To work with no-gap potentials we need estimates for the integrals of $\cos(2\theta_\gamma(x))$ and $\sin(2\theta_\gamma(x))$; to complement Ψ_J (see (28)) set

$$\Phi_J = \int_J \cos(2\theta_\gamma(x))\,\mathrm{d}x$$

for any interval $J \subseteq \mathbb{R}$. For large γ equation (21) suggests $\theta_\gamma(x)$ should be changing rapidly wherever $V(x) \neq 0$; it follows that $\cos(2\theta_\gamma(x))$ and $\sin(2\theta_\gamma(x))$ should be rapidly oscillating, leading to cancellation in the integrals defining Φ_J and Ψ_J. This idea lies at the heart of the following result (the proof is given in Section 4.3):

Proposition 3.18. *Let $I \subset \mathbb{R}$ be a closed bounded interval, and suppose a potential V satisfies $\mathrm{var}_I(V) < +\infty$ and has no gaps on I. Also suppose θ_γ satisfies (21) on I. For any sub-interval $J \subseteq I$ we have Φ_J, $\Psi_J = o(1)$ as $\gamma \to \infty$, uniformly in J and possible choices of (the initial condition for) the solution θ_γ.*

These estimates for Φ_J and Ψ_J lead directly to the following asymptotic information about the change in the value of θ_γ across I:

Proposition 3.19. *Let I, V and θ_γ be as in Proposition 3.18. Write $I = [a, b]$. Also suppose $\mathrm{d}^n \theta_\gamma(b) / \mathrm{d}\gamma^n$ exists and is bounded in γ for $n = 1, 2$. Then, for $n = 0, 1, 2$,*

$$\frac{\mathrm{d}^n}{\mathrm{d}\gamma^n} \left(\theta_\gamma(b) - \theta_\gamma(a) - \gamma \int_a^b V(x)\, \mathrm{d}x \right) = o(1) \tag{29}$$

as $\gamma \to \infty$; in particular, $\mathrm{d}^n \theta_\gamma(a) / \mathrm{d}\gamma^n$ also exists and is bounded in γ for $n = 1, 2$.

Proof. Integrating (21) gives

$$\theta_\gamma(b) - \theta_\gamma(a) = \int_a^b \nabla \theta_\gamma(x)\, \mathrm{d}x = \gamma \int_a^b V(x)\, \mathrm{d}x + k\Phi_I,$$

so (29) for $n = 0$ follows directly from Proposition 3.18.

From (27) we get

$$\omega_\gamma(b) - \omega_\gamma(x) - \int_x^b V(t)\, \mathrm{d}t = \left(1 - \mathrm{e}^{2k\Psi_{[x,b]}}\right)\omega_\gamma(b) - \int_x^b \left(1 - \mathrm{e}^{2k\Psi_{[x,t]}}\right) V(t)\, \mathrm{d}t$$

for all $x \in I$. By Proposition 3.18 we have $\Psi_J = o(1)$ and hence $1 - \mathrm{e}^{2k\Psi_J} = o(1)$ as $\gamma \to \infty$, uniformly for all sub-intervals $J \subseteq I$. Setting $W(x) = \omega_\gamma(b) - \int_x^b V(t)\, \mathrm{d}t$ it follows that $W(x)$ is bounded in γ while $W(x) - \omega_\gamma(x) = o(1)$ as $\gamma \to \infty$, both uniformly for $x \in I$. With $x = a$ this becomes (29) for $n = 1$. More generally we have that $\omega_\gamma(x)$ is uniformly bounded for all γ and $x \in I$, while

$$W^2(x) - \omega_\gamma^2(x) = \left(W(x) - \omega_\gamma(x)\right)\left(W(x) + \omega_\gamma(x)\right) = o(1) \quad \text{as } \gamma \to \infty, \tag{30}$$

uniformly for $x \in I$.

Arguing as for (27) we get

$$\rho_\gamma(a) = \mathrm{e}^{2k\Psi_I} \rho_\gamma(b) + 4k \int_a^b \mathrm{e}^{2k\Psi_{[a,x]}} \cos(2\theta_\gamma(x))\, \omega_\gamma^2(x)\, \mathrm{d}x.$$

Thus

$$\rho_\gamma(b) - \rho_\gamma(a) = \left(1 - \mathrm{e}^{2k\Psi_I}\right)\rho_\gamma(b) + 4k \int_a^b \left(1 - \mathrm{e}^{2k\Psi_{[a,x]}}\right) \cos(2\theta_\gamma(x))\, \omega_\gamma^2(x)\, \mathrm{d}x$$

$$+ 4k \int_a^b \cos(2\theta_\gamma(x))\left(W^2(x) - \omega_\gamma^2(x)\right) \mathrm{d}x \; - \; 4k \int_a^b \cos(2\theta_\gamma(x))\, W^2(x)\, \mathrm{d}x.$$

As above $\Psi_J = o(1)$ and hence $1 - \mathrm{e}^{2k\Psi_J} = o(1)$ as $\gamma \to \infty$, uniformly for all sub-intervals $J \subseteq I$. Combined with (30) and the fact that $\cos(2\theta_\gamma)$ and ω_γ are uniformly bounded it follows that the first three terms on the right hand side are $o(1)$ as $\gamma \to \infty$. The following claim deals with the final term and completes the argument.

23

Claim: We have $\int_a^b \cos(2\theta_\gamma(x)) \, W^2(x) \, \mathrm{d}x = o(1)$ *as* $\gamma \to \infty$. Firstly note that

$$\nabla \Phi_{[a,x]} = \cos(2\theta_\gamma(x)) \quad \text{and} \quad \nabla W^2(x) = 2W(x)V(x).$$

Integrating by parts thus gives

$$\int_a^b \cos(2\theta_\gamma(x)) \, W^2(x) \, \mathrm{d}x = \Phi_I W^2(b) - 2 \int_a^b \Phi_{[a,x]} W(x) V(x) \, \mathrm{d}x$$

since $\Phi_{[a,a]} = 0$. By Proposition 3.18 we have $\Phi_J = o(1)$ as $\gamma \to \infty$, uniformly for all sub-intervals $J \subseteq I$. Furthermore $W(x)$ is uniformly bounded for γ and $x \in I$ while $V \in L^1([a,b])$. It follows that $\int_a^b \Phi_{[a,x]} W(x) V(x) \, \mathrm{d}x = o(1)$ as $\gamma \to \infty$, completing the claim. $\qquad \square$

Suppose a potential $V \in L^1_{\mathrm{loc}}$ has support contained in the bounded interval $I = [a,b]$. Clearly constant functions taking values in $\pi/4 + \pi\mathbb{Z}/2$ solve (21) outside I. The condition (23) (together with the uniqueness and continuity of $\theta_{\gamma,\pm}$) then gives us $\theta_{\gamma,-}(x) = \pi/4$ for $x \le a$ and $\theta_{\gamma,+}(x) = -\pi/4$ for $x \ge b$ (this can also be seen from the form of the corresponding solutions to (19)). When defining Δ_V in this case it is convenient to choose the left endpoint of I as the point at which to check whether $\theta_{\gamma,-}$ and $\theta_{\gamma,+}$ can be 'matched'; alternatively, we can keep our current definition of Δ_V if we simply translate our problem so that $a = 0$ (see Remark 1.8). Making such a choice we get

$$\Delta_V(\gamma) = -\frac{\pi}{2} - \theta_{\gamma,+}(a) + \theta_{\gamma,-}(a) = \theta_{\gamma,+}(b) - \theta_{\gamma,+}(a),$$

where $\theta_{\gamma,+}(b) = -\pi/4$ is constant (as a function of γ). We will work with Δ_V in this form for the remainder of the present section.

For compactly supported potentials without gaps we can now rephrase the conclusions of Proposition 3.19 to get the following improved and extended version of Proposition 3.8:

Corollary 3.20. *Suppose $V \in BV_0$ has no gaps. Then, for $n = 0, 1, 2$,*

$$\frac{\mathrm{d}^n}{\mathrm{d}\gamma^n} \left(\Delta_V(\gamma) - \gamma \int_{\mathbb{R}} V(x) \, \mathrm{d}x \right) = o(1)$$

as $\gamma \to \infty$.

When $\int_{\mathbb{R}} V(x) \, \mathrm{d}x \neq 0$ it follows that Δ_V is monotonic for sufficiently large γ. In this case Theorem 1.9 can be proved using an argument very similar to that used for Theorem 1.6. The case $\int_{\mathbb{R}} V(x) \, \mathrm{d}x = 0$ can be treated with a separate observation.

Proof of Theorem 1.9. If $\int_{\mathbb{R}} V(x) \, \mathrm{d}x = 0$ Corollary 3.20 (for $n = 0$) enables us to find $S > 0$ so that $|\Delta_V(\gamma)| < \pi/2$ when $|\gamma| > S$. Then $\Sigma_V \cap \mathbb{R} \subseteq [-S, S]$ by Proposition 3.7. However Σ_V is a discrete subset of $\mathbb{C}$ (Theorem 1.1) so $\Sigma_V \cap [-S, S]$ contains at most finitely many points.

Now suppose $\int_{\mathbb{R}} V(x) \, \mathrm{d}x \neq 0$. By Corollary 3.20 (for $n = 1$) there exists $S > 0$ such that $\Delta_V(\gamma)$ is strictly monotonic for all $\gamma \ge S$. Suppose $R > S$ and let $I_{S,R}$ denote the closed interval with endpoints $\Delta_V(S)$ and $\Delta_V(R)$. Arguing as for Theorem 1.6 we then get

$$\#(\Sigma_V \cap [S, R]) = \#(I_{S,R} \cap (\mathbb{Z} + 1/2)\pi) = \frac{|I_{S,R}|}{\pi} + O(1).$$

However Corollary 3.20 (for $n = 0$) also gives

$$|I_{S,R}| = |\Delta_V(R) - \Delta_V(S)| = |\Delta_V(R)| + O(1) = R \left| \int_{\mathbb{R}} V(x) \, \mathrm{d}x \right| + O(1)$$

as $R \to \infty$ (note that, S is fixed). The fact that $\Sigma_V \cap [0, S]$ contains at most finitely many points (see above) completes the argument. $\qquad \square$

3.6 One gap

Let $V \in BV_0$ be a one-gap potential as considered in Section 1.7. We can use Proposition 3.19 to estimate the change in $\theta_{\gamma,+}$ across the intervals $[a_j, b_j]$ for $j = 1, 2$. Information about the change in $\theta_{\gamma,+}$ across the gap (b_1, a_2) will be obtained from the next result.

Lemma 3.21. *Suppose $\nabla\theta = k\cos(2\theta)$ on some interval $I = (a, b)$ and set $\alpha = \tanh(k(b-a))$. Then*

$$\sin\big(\theta(b) - \theta(a)\big) = \alpha\cos\big(\theta(b) + \theta(a)\big). \tag{31}$$

Proof. Let $L = \dfrac{\pi}{4} + \dfrac{\pi}{2}\mathbb{Z}$, the zero set of $\cos(2\theta)$. Values in L give constant solutions to $\nabla\theta = k\cos(2\theta)$. For such solutions $\theta(b) - \theta(a) = 0$ and $\theta(b) + \theta(a) \in \dfrac{\pi}{2} + \pi\mathbb{Z}$, so both sides of (31) are zero.

Now suppose $\theta(x_0) \notin L$ for some $x_0 \in I$. Using the uniqueness and continuity of solutions to the equation $\nabla\theta = k\cos(2\theta)$ it follows that $\theta(x)$ must remain within the same connected component of $\mathbb{R} \setminus L$ for all $x \in I$. In particular $\cos(2\theta(x)) \neq 0$. Now set

$$F(\theta) = \frac{1 - \sin(2\theta)}{\cos(2\theta)} = \frac{\cos(\theta) - \sin(\theta)}{\cos(\theta) + \sin(\theta)}.$$

Then $F'(\theta) = -2F(\theta)/\cos(2\theta)$ so

$$\nabla\big(e^{2kx}F(\theta(x))\big) = e^{2kx}F(\theta(x))\left(2k - 2\frac{\nabla\theta(x)}{\cos(2\theta(x))}\right) = 0.$$

Hence $e^{2kb}F(\theta(b)) = e^{2ka}F(\theta(a))$. The second expression for $F(\theta)$ then leads to

$$\frac{e^{2kb} - e^{2ka}}{e^{2kb} + e^{2ka}}\cos(\theta(b) + \theta(a)) = \sin(\theta(b) - \theta(a)).$$

The first part of the expression on the left hand side is just $\tanh(k(b - a)) = \alpha$. $\qquad\square$

As in the discussion proceeding Corollary 3.20 we observe that $\theta_{\gamma,-}(x) = \pi/4$ for $x \leq a_1$ and $\theta_{\gamma,+}(x) = -\pi/4$ for $x \geq b_2$ (note that V has support contained in $[a_1, b_2]$). We shall also define Δ_V by choosing a_1 as the point at which to check whether $\theta_{\gamma,-}$ and $\theta_{\gamma,+}$ can be 'matched'. It follows that

$$\Delta_V = \Delta_V(\gamma) = \theta_{\gamma,+}(b_2) - \theta_{\gamma,+}(a_1).$$

Now set $\Delta_j = \Delta_j(\gamma) = \theta_{\gamma,+}(b_j) - \theta_{\gamma,+}(a_j)$ for $j = 1, 2$. Then

$$\theta_{\gamma,+}(a_2) - \theta_{\gamma,+}(b_1) = \Delta_V - \Delta_1 - \Delta_2 \quad \text{and} \quad \theta_{\gamma,+}(a_2) + \theta_{\gamma,+}(b_1) = \Delta_1 - \Delta_2 - \Delta_V - \frac{\pi}{2}.$$

Since $V = 0$ on (b_1, a_2) Lemma 3.21 then gives

$$\begin{aligned}
\sin\big(\Delta_V - \Delta_1 - \Delta_2\big) &= \alpha\sin\big(\Delta_1 - \Delta_2 - \Delta_V\big) \\
\implies \quad \big[\cos(\Delta_1 + \Delta_2) + \alpha\cos(\Delta_1 - \Delta_2)\big]\sin(\Delta_V) & \\
&= \big[\sin(\Delta_1 + \Delta_2) + \alpha\sin(\Delta_1 - \Delta_2)\big]\cos(\Delta_V). \tag{32}
\end{aligned}$$

Lemma 3.22. *We have $\gamma \in \Sigma_V \cap \mathbb{R}$ iff*

$$\cos(\Delta_1 + \Delta_2) + \alpha\cos(\Delta_1 - \Delta_2) = 0. \tag{33}$$

Proof. If $\gamma \in \Sigma_V \cap \mathbb{R}$ we get $\cos(\Delta_V) = 0$ and $\sin(\Delta_V) = \pm 1$ from Proposition 3.7, so (33) follows from (32). Now suppose (33) holds. Then (32) gives either $\cos(\Delta_V) = 0$, in which case $\gamma \in \Sigma_V \cap \mathbb{R}$ by Proposition 3.7, or

$$\sin(\Delta_1 + \Delta_2) + \alpha \sin(\Delta_1 - \Delta_2) = 0. \tag{34}$$

However (33) and (34) imply

$$1 = \cos^2(\Delta_1 + \Delta_2) + \sin^2(\Delta_1 + \Delta_2) = \alpha^2\big(\cos^2(\Delta_1 - \Delta_2) + \sin^2(\Delta_1 - \Delta_2)\big) = \alpha^2,$$

contradicting the fact that $\alpha = \tanh(k(a_2 - b_1)) \in (0, 1)$. $\qquad\square$

Before giving the proofs of the first two results in Section 1.7 we note that if g satisfies (7) and $\beta \in \mathbb{R}$ then it is straightforward to check that $\cos(\beta x + g(x)) - \cos(\beta x)$ also satisfies (7).

Proofs of Theorems 1.11 and 1.13. By Proposition 3.19 applied to V_j on the interval $[a_j, b_j]$, $j = 1, 2$, we have

$$\Delta_1 + \Delta_2 = (v_1 + v_2)\gamma + g_+(\gamma) \quad \text{and} \quad \Delta_1 - \Delta_2 = (v_1 - v_2)\gamma + g_-(\gamma)$$

for some functions g_+ and g_- which satisfy (7). Using the observation proceeding the proof we can then write

$$\cos(\Delta_1 + \Delta_2) + \alpha \cos(\Delta_1 - \Delta_2) = \cos(|v_1 + v_2|\gamma) + \alpha \cos(|v_1 - v_2|\gamma) + \psi(\gamma)$$

for some ψ which also satisfies (7). By Lemma 3.22 we thus have $\gamma \in \Sigma_V \cap \mathbb{R}$ iff

$$\cos(|v_1 + v_2|\gamma) + \alpha \cos(|v_1 - v_2|\gamma) + \psi(\gamma) = 0. \tag{35}$$

Since Σ_V is discrete (Theorem 1.1) the solutions of (35) (in γ) form a discrete subset of $\mathbb{R}$.

If $v_1 + v_2 = 0$ then (35) reduces to $1 + \alpha \cos(2|v_1|\gamma) + \psi(\gamma) = 0$. However $0 < \alpha < 1$ and $\psi(\gamma) = o(1)$ as $\gamma \to \infty$, so we can we can choose S such that $|\psi(\gamma)| < 1 - \alpha$ for all $\gamma \geq S$. Then $\Sigma_V \cap \mathbb{R} \subset [-S, S]$ (recall that Σ_V is symmetric about 0). The discreteness of Σ_V then limits $\Sigma_V \cap \mathbb{R}$ to at most finitely many points, establishing Theorem 1.11.

Now suppose $v_1 + v_2 \neq 0$. Writing $x = |v_1 + v_2|\gamma$, (35) then gives

$$\#(\Sigma_V \cap [0, R]) = \#\big\{x \in [0, |v_1 + v_2|R] : \cos(x) + \alpha \cos(\beta x) + \phi(x) = 0\big\},$$

where $\phi(x) = \psi(x/|v_1 + v_2|)$ which clearly satisfies (7). Theorem 1.13 now follows directly from Theorem 1.15. $\qquad\square$

Proof of Theorem 1.16. Since $A < u$ we can choose w so that $0 < v < A < w \leq u$ and w/v is irrational. Set $v_0 = (u - w)/2 \geq 0$, $v_1 = (v - w)/2 < 0$, $v_2 = (v + w)/2 > 0$ and

$$\beta = \left|\frac{v_1 - v_2}{v_1 + v_2}\right| = \frac{w}{v}.$$

Then $1 < \dfrac{A}{v} < \beta$; thus we can choose $\alpha \in (1/\beta, 1)$ so that $\nu_{\alpha,\beta} = \dfrac{A}{v}$. Now set $g = \dfrac{1}{k}\tanh^{-1}(\alpha) > 0$ and consider the potentials

$$V_1(x) = W\big(x; [-1, 0]; \{v_1\}\big) \quad \text{and} \quad V_2(x) = W\big(x; [g, g+1, g+2]; \{v_2 + v_0, -v_0\}\big)$$

26

(see (9) for notation). Then $V = V_1 + V_2 \in BV_0$ is a one gap potential, with gap from 0 to g. Since $\alpha\beta > 1$ and β is irrational Theorem 1.13 shows that (8) holds with $C = |v_1 + v_2|\,\nu_{\alpha,\beta} = v\dfrac{A}{v} = A$. Furthermore

$$\int_{\mathbb{R}} V(x)\,\mathrm{d}x = \int_{\mathbb{R}} V_1(x)\,\mathrm{d}x + \int_{\mathbb{R}} V_2(x)\,\mathrm{d}x = v_1 + (v_2 + v_0 - v_0) = v$$

while $\|V\|_{L^1} = \|V_1\|_{L^1} + \|V_2\|_{L^1} = -v_1 + (v_2 + 2v_0) = u$. $\qquad\square$

4 Technicalities

4.1 Asymptotic behaviour of ode solutions

We return to viewing our basic equation $\mathrm{T}_{\gamma V}\psi = 0$ as the system of ordinary differential equations (19). When $V \equiv 0$ we have the exponential solutions

$$e^{\pm kx}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \tag{36}$$

(see Lemma 2.1). When $V \in \mathbb{V}_0$ (or, more generally, V belongs to the locally L^1 version of $\mathbb{V}_0$) we can find solutions of (19) with similar asymptotic properties as $x \to \pm\infty$ (see [Ha, Chapter X]). In particular, there are non-trivial solutions $\varphi_{\pm}^{+}$ and $\varphi_{\pm}^{-}$ which satisfy

$$\lim_{x\to+\infty} x^{-1}\log\bigl|\varphi_{\pm}^{+}(x)\bigr| = \pm k \quad\text{and}\quad \lim_{x\to-\infty} x^{-1}\log\bigl|\varphi_{\pm}^{-}(x)\bigr| = \pm k.$$

Since $k \neq 0$ $\varphi_{+}^{+}(x)$ and $\varphi_{-}^{+}(x)$ have different asymptotic behaviour as $x \to +\infty$ (the growth of $\varphi_{\pm}^{+}(x)$ as $x \to +\infty$ is roughly like $e^{\pm kx}$); it follows that these solutions must be linearly independent. A similar discussion applies to φ_{+}^{-} and φ_{-}^{-}.

Lemma 4.1. *If $V \in \mathbb{V}_0$ then either $0 \notin \mathrm{spec}(\mathrm{T}_V)$ or 0 is a simple eigenvalue of T_V.*

Proof. We have $0 \in \mathrm{spec}(\mathrm{T}_V)$ iff 0 is an isolated eigenvalue of T_V (see Remark 1.2). Now suppose ψ is an eigenfunction corresponding to 0. Then ψ satisfies (19) (with $\gamma = 1$). Since φ_{+}^{+} and φ_{-}^{+} are linearly independent solutions of this equation we must have $\psi = \alpha_{+}\varphi_{+}^{+} + \alpha_{-}\varphi_{-}^{+}$ for some constants $\alpha_{\pm}$. Restricting to the interval $\mathbb{R}^{+}$ we have $\varphi_{+}^{+} \notin L^2(\mathbb{R}^{+})$ while $\varphi_{-}^{+} \in L^2(\mathbb{R}^{+})$. It follows that $\alpha_{+} = 0$, and so ψ is a multiple of φ_{-}^{+}. This must also be true for any other eigenfunction corresponding to 0, so any two such eigenfunctions are linearly dependent. $\qquad\square$

Remark 4.2. Using a similar argument we can also get $\psi = \beta_{+}\varphi_{+}^{-}$ for some constant β_{+}, showing that φ_{-}^{+} and φ_{+}^{-} are linearly dependent. In fact this is an alternative characterisation of when 0 is an eigenvalue of T_V.

When $V \in \mathbb{V}_1$ solutions to (19) have well defined leading order asymptotics as $x \to \pm\infty$; these asymptotics are solutions to the same equation with $V \equiv 0$, so must be linear combinations of the exponential functions given in (36) (see [Ha, Chapter X]). In particular, we can choose $\varphi_{\pm}^{+}$ and $\varphi_{\pm}^{-}$ so that

$$\lim_{x\to+\infty}\left|e^{\mp kx}\left(\varphi_{\pm}^{+}(x) - e^{\pm kx}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}\right)\right| = 0 \tag{37a}$$

while

$$\lim_{x\to-\infty}\left|e^{\mp kx}\left(\varphi_{\pm}^{-}(x) - e^{\pm kx}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}\right)\right| = 0. \tag{37b}$$

Furthermore φ_-^+ and φ_+^- are uniquely determined by these asymptotic conditions. (The solutions φ_+^+ and φ_-^- are however only determined up to the addition of a multiple of φ_-^+ and φ_+^- respectively.)

Proof of Lemma 3.6. We consider the case $x \to +\infty$; $x \to -\infty$ can be handled similarly. Since φ_+^+ and φ_-^+ are linearly independent we can write $\psi = \alpha_+\varphi_+^+ + \alpha_-\varphi_-^+$ for some constants $\alpha_\pm$ (which can't both be 0). Now by (37a)

$$\lim_{x \to +\infty} \frac{\psi_1(x)}{\psi_2(x)} = \lim_{x \to +\infty} \frac{(\alpha_+\varphi_+^+(x) + \alpha_-\varphi_-^+(x))_1}{(\alpha_+\varphi_+^+(x) + \alpha_-\varphi_-^+(x))_2} = \begin{cases} 1 & \text{if } \alpha_+ \neq 0 \\ -1 & \text{if } \alpha_+ = 0. \end{cases}$$

However $\tan\theta = \psi_1/\psi_2$ so (22) (for $x \to +\infty$) follows. On the other hand, $\psi \in L^2(\mathbb{R}^+)$ iff $\alpha_+ = 0$, leading to the second part of the result. $\qquad\square$

4.2 Estimates and asymptotics for Δ_V

Propositions 3.11 and 3.8 both follow almost directly from the following estimate:

Lemma 4.3. *Let $V \in \mathbb{V}_1$ and suppose θ solves (21) with $\theta(+\infty) = -\pi/4$. For any $x \in \mathbb{R}$ we then have*

$$\left| \theta(x) + \frac{\pi}{4} \right| \leq h\left(|\gamma| \int_x^\infty |V(t)|\,\mathrm{d}t \right).$$

A similar result holds on $(-\infty, x]$ when $\theta(-\infty) = \pi/4$.

The basic idea is that if $\theta(x)$ increases on an interval $[a, b]$ crossing a range of values where $k\cos(2\theta(x))$ is non-positive then we must have

$$\int_a^b \gamma V(x)\,\mathrm{d}x \geq \theta(b) - \theta(a).$$

We can then add the contributions from each such crossing.

Proof. Let $n \in \mathbb{Z}$ and suppose $\theta(x) \in \left[-(n+3/4)\pi, -(n+1/4)\pi\right]$ for all x in some interval $[a, b] \subset \mathbb{R}$. Then $k\cos(2\theta(x)) \leq 0$ so $\nabla\theta(x) \leq \gamma V(x)$, and hence

$$\theta(b) - \theta(a) \leq \int_a^b \gamma V(t)\,\mathrm{d}t. \qquad (38)$$

This estimate continues to hold with the obvious interpretation when $b = +\infty$.

Now let $x \in \mathbb{R}$ and suppose $\theta(x) < -\pi/4$. Choose $n \in \mathbb{N}_0$ so that $\theta(x) \in \left(-(n+5/4)\pi, -(n+1/4)\pi\right]$. Using the continuity of θ, and the assumption that $\theta(+\infty) = -\pi/4$, we can now choose a sequence of points

$$x \leq a_n < b_n < a_{n-1} < b_{n-1} < \cdots < a_0 < b_0 \leq +\infty$$

such that

(i) For $j = 0, \ldots, n$ and $x \in [a_j, b_j]$ we have $\theta(x) \in \left[-(j+3/4)/\pi, -(j+1/4)\pi\right]$.

(ii) For $j = 0, \ldots, n$ we have $\theta(a_j) = -(j+3/4)\pi$ and $\theta(b_j) = -(j+1/4)\pi$, with the exception that $\theta(a_n) = \theta(x)$ if $\theta(x) \in \left(-(n+3/4)\pi, -(n+1/4)\pi\right]$.

Applying (38) we then get

$$\sum_{j=0}^{n} (\theta(b_j) - \theta(a_j)) \le \sum_{j=0}^{n} \int_{a_j}^{b_j} \gamma V(t) \, \mathrm{d}t \le \int_{x}^{\infty} |\gamma V(t)| \, \mathrm{d}t. \tag{39}$$

If $\theta(x) \in \big(-(n+5/4)\pi, -(n+3/4)\pi\big]$ then each term in the sum on the left of (39) is $\pi/2$ so

$$(n+1)\frac{\pi}{2} \le \int_{x}^{\infty} |\gamma V(t)| \, \mathrm{d}t.$$

It follows that

$$(n+1)\frac{\pi}{2} \le \frac{\pi}{2} \left\lfloor \frac{2}{\pi} \int_{x}^{\infty} |\gamma V(t)| \, \mathrm{d}t \right\rfloor.$$

On the other hand $\theta(x) \ge -(n+5/4)\pi$ so

$$-\theta(x) - \frac{\pi}{4} \le (n+1)\pi \le h\left(\int_{x}^{\infty} |\gamma V(t)| \, \mathrm{d}t\right).$$

Alternatively suppose $\theta(x) \in \big(-(n+3/4)\pi, -(n+1/4)\pi\big]$ so $\theta(a_n) = \theta(x)$. Then the term in the sum on the left of (39) is $\pi/2$ for $j = 0, \ldots, n-1$, so

$$n\frac{\pi}{2} = \sum_{j=0}^{n-1} (\theta(b_j) - \theta(a_j)) \le \int_{x}^{\infty} |\gamma V(t)| \, \mathrm{d}t,$$

leading to

$$n\frac{\pi}{2} \le \frac{\pi}{2} \left\lfloor \frac{2}{\pi} \int_{x}^{\infty} |\gamma V(t)| \, \mathrm{d}t \right\rfloor.$$

Furthermore

$$\sum_{j=0}^{n} (\theta(b_j) - \theta(a_j)) = -\left(n + \frac{1}{4}\right)\pi - \theta(a_n) + n\frac{\pi}{2} = -\theta(x) - \frac{\pi}{4} - n\frac{\pi}{2}.$$

Using (39) again gives

$$-\theta(x) - \frac{\pi}{4} \le \int_{x}^{\infty} |\gamma V(t)| \, \mathrm{d}t + n\frac{\pi}{2} \le h\left(\int_{x}^{\infty} |\gamma V(t)| \, \mathrm{d}t\right).$$

A similar argument can be used to deal with the case $\theta(x) > -\pi/4$ (we need to consider intervals where $\theta(x)$ decreases across the range $\big[(n-1/4)\pi, (n+1/4)\pi\big]$). $\qquad\square$

Proof of Proposition 3.11. By Lemma 4.3 we get

$$\left|\theta_{\gamma,+}(0) + \frac{\pi}{4}\right| \le h\left(|\gamma| \int_{0}^{\infty} |V(t)| \, \mathrm{d}t\right).$$

The equivalent estimate on $(-\infty, 0]$ gives

$$\left|\theta_{\gamma,-}(0) - \frac{\pi}{4}\right| \le h\left(|\gamma| \int_{-\infty}^{0} |V(t)| \, \mathrm{d}t\right).$$

The result now follows from (24) and the fact that $h(a) + h(b) \le h(a+b)$ for any $a, b \ge 0$. $\qquad\square$

Proof of Proposition 3.8. Recalling (24) we can write

$$\Delta_V(\gamma) = \Delta_V^+(\gamma) + \Delta_V^-(\gamma).$$

where $\Delta_V^\pm(\gamma) := -\pi/4 \mp \theta_{\gamma,\pm}(0)$. Now suppose $K \geq 0$. Since $\theta_{\gamma,+}$ satisfies (21) we get

$$\theta_{\gamma,+}(K) - \theta_{\gamma,+}(0) = \gamma \int_0^K V(x)\,\mathrm{d}x + k \int_0^K \cos(2\theta_{\gamma,+}(x))\,\mathrm{d}x$$

so

$$\left| \theta_{\gamma,+}(K) - \theta_{\gamma,+}(0) - \gamma \int_0^\infty V(x)\,\mathrm{d}x \right| \leq |\gamma| \int_K^\infty |V(x)|\,\mathrm{d}x + kK.$$

On the other hand Lemma 4.3 gives us

$$\left| \theta_{\gamma,+}(K) + \frac{\pi}{4} \right| \leq h\left(|\gamma| \int_K^\infty |V(x)|\,\mathrm{d}x \right) \leq 2|\gamma| \int_K^\infty |V(x)|\,\mathrm{d}x.$$

Therefore

$$\left| \Delta_V^+(\gamma) - \gamma \int_0^\infty V(x)\,\mathrm{d}x \right|$$
$$= \left| -\theta_{\gamma,+}(K) - \frac{\pi}{4} + \theta_{\gamma,+}(K) - \theta_{\gamma,+}(0) - \gamma \int_0^\infty V(x)\,\mathrm{d}x \right|$$
$$\leq 3|\gamma| \int_K^\infty |V(x)|\,\mathrm{d}x + kK.$$

Now choose $K = K_\gamma \geq 0$ for each γ so that $K_\gamma \to \infty$ and $|\gamma|^{-1} K_\gamma \to 0$ as $|\gamma| \to \infty$ (we can take $K_\gamma \sim |\gamma|^\mu$ with $0 < \mu < 1$, for example). Then $\int_{K_\gamma}^\infty |V(x)|\,\mathrm{d}x = o(1)$ (since $V \in L^1$) so

$$3|\gamma| \int_{K_\gamma}^\infty |V(x)|\,\mathrm{d}x + kK_\gamma = o(\gamma)$$

as $|\gamma| \to \infty$. A similar estimate can be obtained for $\Delta_V^-(\gamma)$. $\qquad\square$

4.3 Estimates for Φ_J and Ψ_J

Intervals

It is easiest to deal with the potential V in pieces where it is single-signed and bounded away from 0; the next result is the key to identifying these pieces and sets up some of our notation.

Lemma 4.4. *Suppose $V \in BV_0$ has no gaps and set $I = \mathrm{supp}(V)$. For each $\varepsilon > 0$ there exists a finite collection of disjoint closed intervals $I_{\varepsilon,n} \subseteq I$, $n \in \mathcal{N}_\varepsilon$, such that*

(i) *For each $n \in \mathcal{N}_\varepsilon$ V has constant sign on $I_{\varepsilon,n}$ and $|V(x)| \geq \varepsilon$ for all $x \in I_{\varepsilon,n}$.*

(ii) *Setting $E_\varepsilon = I \setminus \bigcup_{n \in \mathcal{N}_\varepsilon} I_{\varepsilon,n}$ we have $|E_\varepsilon| \to 0$ as $\varepsilon \to 0$.*

(iii) *$\#\mathcal{N}_\varepsilon \leq \nu\varepsilon^{-1}$, where $\nu := \mathrm{var}_\mathbb{R}(V)$ is the total variation of V.*

We would ideally like E_ε to be $V^{-1}((-\varepsilon,\varepsilon))$, so the union of the $I_{\varepsilon,n}$'s would be

$$\{x : V(x) \leq -\varepsilon\} \cup \{x : V(x) \geq \varepsilon\}. \tag{40}$$

There are however several technical issues associated with this choice;

- We are not assuming that V is continuous so (40) may not be closed (or even particularly well behaved).

- Even if V is continuous the number of intervals in $V^{-1}((-\varepsilon, \varepsilon))$ may be uncontrollable (even infinite); we need a useful bound on this quantity.

The technicalities in the following argument arise from the need to deal with these issues.

Proof of Lemma 4.4. Set

$$X_\varepsilon^+ = \bigcup_{x \in V^{-1}((-\infty, 2\varepsilon))} (x - \varepsilon, x + \varepsilon) \quad \text{and} \quad X_\varepsilon^- = \bigcup_{x \in V^{-1}((-2\varepsilon, +\infty))} (x - \varepsilon, x + \varepsilon);$$

(so X_ε^+ is the 'ε-neighbourhood' of $\{x : V(x) < 2\varepsilon\}$, while X_ε^- is the 'ε-neighbourhood' of $\{x : V(x) > -2\varepsilon\}$).

Since X_ε^+ is an open subset of $\mathbb{R}$ it consists of a countable union of disjoint open intervals; let $Y_\varepsilon^+ \subseteq X_\varepsilon^+$ be the union of those intervals which intersect $V^{-1}((-\infty, \varepsilon))$. If $J = (\alpha, \beta)$ is a maximal interval in Y_ε^+ it follows that we can find $x \in J$ with $V(x) < \varepsilon$. Furthermore J must also be a maximal interval in X_ε^+ so, if J is bounded, its endpoints satisfy $\alpha, \beta \notin X_\varepsilon^+$. Thus $V(\alpha), V(\beta) \geq 2\varepsilon$ and so $\operatorname{var}_{\overline{J}}(V) \geq |V(x) - V(\alpha)| + |V(\beta) - V(x)| \geq 2\varepsilon$. If J is a half-infinite interval a similar argument shows that $\operatorname{var}_{\overline{J}}(V) \geq \varepsilon$.

Now let M_ε^+ and N_ε^+ denote the number of semi-infinite and bounded maximal intervals in Y_ε^+ (we're presently allowing $N_\varepsilon^+ = +\infty$; if $Y_\varepsilon^+ = \mathbb{R}$ set $M_\varepsilon^+ = 2$, $N_\varepsilon^+ = -1$). Then

$$\nu = \operatorname{var}_{\mathbb{R}}(V) \geq \operatorname{var}_{\overline{Y_\varepsilon^+}}(V) \geq \varepsilon M_\varepsilon^+ + 2\varepsilon N_\varepsilon^+ \quad \Longrightarrow \quad N_\varepsilon^+ \leq \frac{1}{2}(\nu\varepsilon^{-1} - M_\varepsilon^+);$$

in particular, N_ε^+ must be finite. However V has compact support, so Y_ε^+ must be unbounded both above and below. Since Y_ε^+ consists of a finite collection of intervals, it must therefore contain semi-infinite intervals at either end; that is, $M_\varepsilon^+ = 2$. The set $Z_\varepsilon^+ := \mathbb{R} \setminus Y_\varepsilon^+$ is then the union of $N_\varepsilon^+ + 1$ closed bounded intervals; write $Z_\varepsilon^+ = \bigcup_{n \in \mathcal{N}_\varepsilon^+} I_{\varepsilon,n}$ where $\mathcal{N}_\varepsilon^+$ is some indexing set with $\#\mathcal{N}_\varepsilon^+ = N_\varepsilon^+ + 1$. It is straightforward to check that we have $V(x) \geq \varepsilon$ for any $x \in Z_\varepsilon^+$.

We can similarly define Y_ε^-, Z_ε^-, $I_{\varepsilon,n}$ for $n \in \mathcal{N}_\varepsilon^-$ and N_ε^-. Set $\mathcal{N}_\varepsilon = \mathcal{N}_\varepsilon^+ \sqcup \mathcal{N}_\varepsilon^-$. Property (i) is immediate, while (iii) holds since

$$\#\mathcal{N}_\varepsilon = \#\mathcal{N}_\varepsilon^+ + \#\mathcal{N}_\varepsilon^- = N_\varepsilon^+ + N_\varepsilon^- + 2 \leq 2\frac{1}{2}(\nu\varepsilon^{-1} - 2) + 2 = \nu\varepsilon^{-1}.$$

Now $\bigcup_{n \in \mathcal{N}_\varepsilon} I_{\varepsilon,n} = Z_\varepsilon^+ \cup Z_\varepsilon^- =: Z_\varepsilon$. If $0 < \varepsilon_1 < \varepsilon_2$ a straightforward check gives us

$$X_{\varepsilon_1}^\pm \subseteq X_{\varepsilon_2}^\pm \Longrightarrow Y_{\varepsilon_1}^\pm \subseteq Y_{\varepsilon_2}^\pm \Longrightarrow Z_{\varepsilon_1}^\pm \supseteq Z_{\varepsilon_2}^\pm,$$

so $Z_{\varepsilon_1} \supseteq Z_{\varepsilon_2}$ and hence $E_{\varepsilon_1} = I \setminus Z_{\varepsilon_1} \subseteq I \setminus Z_{\varepsilon_2} = E_{\varepsilon_2}$. Since $E_\varepsilon \subseteq I$ for all $\varepsilon > 0$ and $|I| < +\infty$, property (ii) will now follow if we can show $\left| \bigcap_{\varepsilon > 0} E_\varepsilon \right| = 0$ (see [Ru], for example).

For each $\delta > 0$ let
$$\Omega_\delta^\pm = \left\{ x \in \bigcap_{\varepsilon > 0} X_\varepsilon^\pm : \pm V(x) > \delta \right\}.$$

Suppose $x_1, \ldots, x_N$ are distinct points in Ω_δ^+ and let $0 < \varepsilon < \min_{1 \leq i,j \leq N} |x_i - x_j|$. From the definition of Ω_δ^+ we can find $x_{j,\varepsilon}$ with $V(x_{j,\varepsilon}) < 2\varepsilon$ and $|x_{j,\varepsilon} - x_j| < \varepsilon$ for $j = 1, \ldots, N$. It follows that

$$\nu = \operatorname{var}_{\mathbb{R}}(V) \geq \sum_{j=1}^{N} \left| V(x_{j,\varepsilon}) - V(x_j) \right| \geq N(\delta - 2\varepsilon).$$

31

Taking $\varepsilon \to 0^+$ gives $N\delta \le \nu$. Hence Ω_δ^+ is finite (with $\#\Omega_\delta^+ \le \nu\delta^{-1}$). Clearly we also have $\Omega_{\delta_1}^+ \supseteq \Omega_{\delta_2}^+$ if $0 < \delta_1 < \delta_2$.

Similar properties hold for Ω_δ^-. It follows that the set

$$\Omega := \bigcap_{\varepsilon > 0}(X_\varepsilon^+ \cap X_\varepsilon^-) \setminus V^{-1}(0) \subseteq \bigcup_{\delta > 0}(\Omega_\delta^+ \cup \Omega_\delta^-) = \bigcup_{n \in \mathbb{N}}(\Omega_{1/n}^+ \cup \Omega_{1/n}^-)$$

is countable (it is contained in a countable union of finite sets). However

$$E_\varepsilon = I \setminus Z_\varepsilon = I \setminus (Z_\varepsilon^+ \cup Z_\varepsilon^-) \subseteq Y_\varepsilon^+ \cap Y_\varepsilon^- \subseteq X_\varepsilon^+ \cap X_\varepsilon^-$$

so

$$\bigcap_{\varepsilon > 0} E_\varepsilon \subseteq I \cap \bigcap_{\varepsilon > 0}(X_\varepsilon^+ \cap X_\varepsilon^-) \subseteq \left(I \cap V^{-1}(0)\right) \cup \Omega.$$

The no gap condition on V and the countability of Ω now imply $\left|\bigcap_{\varepsilon > 0} E_\varepsilon\right| = 0$. $\qquad\square$

Integrals

To justify Proposition 3.18 we start by considering the cancellations over each period of $\cos$ or $\sin$ (Lemma 4.5) and deal with any incomplete periods (Lemma 4.6). In both cases we work on an interval $J = [a_0, a_1]$ where $\nabla\theta(x) > 0$, so θ is invertible. Making the substitution $u = 2\theta(x)$ we get

$$\mathrm{d}u = 2\nabla\theta(x)\,\mathrm{d}x = 2[\gamma V(x) + k\cos(2\theta(x))]\,\mathrm{d}x,$$

so

$$\int_{a_0}^{a_1} f(2\theta(x))\,\mathrm{d}x = \int_{2\theta(a_0)}^{2\theta(a_1)} \frac{f(u)}{2[\gamma V(\theta^{-1}(u/2)) + k\cos(u)]}\,\mathrm{d}u \tag{41}$$

for any f (we will take either $f = \sin$ or $f = \cos$).

Lemma 4.5. *Suppose θ satisfies (21) on an interval $J = [a_0, a_1]$ where $V(x) \ge \varepsilon$ with $\gamma\varepsilon > k$. If $\theta(a_1) = \theta(a_0) + \pi$ then*

$$|\Phi_J| \le \frac{k|J|}{2(\gamma\varepsilon - k)} + \frac{\mathrm{var}_J(V)}{\varepsilon(\gamma\varepsilon - k)} \quad \text{and} \quad |\Psi_J| \le \frac{\mathrm{var}_J(V)}{\varepsilon(\gamma\varepsilon - k)}.$$

Note that we get a better estimate for the $\sin$ integral; the extra term in the estimate for the $\cos$ integral is needed to cope with the fact that this integral is non-zero even when V is constant.

Proof. Since $\theta(a_1) = \theta(a_0) + \pi$ we can define $\chi : [0, 2\pi) \to [2\theta(a_0), 2\theta(a_1))$ to be the unique bijection with $\chi(u) - u \in 2\pi\mathbb{Z}$ for all $u \in [0, 2\pi)$; χ is piecewise affine with at most one jump (χ will have no jumps iff $\theta(a_0) \in 2\pi\mathbb{Z}$). Using (41) we can then write

$$\Psi_J = \int_{a_0}^{a_1} \sin(2\theta(x))\,\mathrm{d}x = \int_0^{2\pi} \frac{\sin(u)}{2[\gamma V(\theta^{-1}(\chi(u)/2)) + k\cos(u)]}\,\mathrm{d}u = \Psi_+ + \Psi_-$$

where

$$\Psi_+ = \int_0^\pi \frac{\sin(u)}{2[\gamma V(\xi_+(u)) + k\cos(u)]}\,\mathrm{d}u$$

with $\xi_+(u) = \theta^{-1}(\chi(u)/2)$, and

$$\Psi_- = \int_0^\pi \frac{\sin(u + \pi)}{2[\gamma V(\theta^{-1}(\chi(u + \pi)/2)) + k\cos(u + \pi)]}\,\mathrm{d}u = -\int_0^\pi \frac{\sin(u)}{2[\gamma V(\xi_-(u)) - k\cos(u)]}\,\mathrm{d}u$$

32

with $\xi_-(u) = \theta^{-1}(\chi(u+\pi)/2)$; ξ_+ and ξ_- are bounded piecewise continuous functions $[0,\pi] \to \mathbb{R}$ with ranges contained in $J = [a_0, a_1]$ (in fact the range of $\xi_\pm$ is just the set of $x \in J$ where $\pm \sin(2\theta(x)) \geq 0$).

We now seek bounds on $\Psi_\pm$ using (constant) bounds on V. Set $m = \inf\{V(x) : x \in J\}$ and $M = \sup\{V(x) : x \in J\}$ so $\varepsilon \leq m \leq M \leq +\infty$ and $m \leq V(x) \leq M$ for all $x \in J$. Thus

$$0 < \gamma m \pm k \cos(u) \leq \gamma V(\xi_\pm(u)) \pm k \cos(u) \leq \gamma M \pm k \cos(u), \quad u \in [0, \pi],$$

so

$$\int_0^\pi \frac{\sin(u)}{2[\gamma M \pm k \cos(u)]} \, du \leq \pm\Psi_\pm \leq \int_0^\pi \frac{\sin(u)}{2[\gamma m \pm k \cos(u)]} \, du.$$

The integrals appearing in these bounds can be calculated explicitly, leading to

$$\frac{1}{2k} \log\left(\frac{\gamma M + k}{\gamma M - k}\right) \leq \pm\Psi_\pm \leq \frac{1}{2k} \log\left(\frac{\gamma m + k}{\gamma m - k}\right).$$

Hence

$$
\begin{aligned}
|\Psi_J| &= |\Psi_+ - (-\Psi_-)| \\
&\leq \frac{1}{2k}\left[\log\left(\frac{\gamma m + k}{\gamma m - k}\right) - \log\left(\frac{\gamma M + k}{\gamma M - k}\right)\right] \\
&= \frac{1}{2k} \log\left(1 + 2k\frac{\gamma(M - m)}{(\gamma m - k)(\gamma M + k)}\right) \\
&\leq \frac{\gamma(M - m)}{(\gamma m - k)(\gamma M + k)}
\end{aligned}
\tag{42}
$$

(note that, $\log(1 + t) \leq t$ for any $t \geq 0$). Since $(\gamma m - k)(\gamma M + k) \geq (\gamma\varepsilon - k)\gamma\varepsilon > 0$ and $M - m \leq \mathrm{var}_J(V)$ the required estimate for Ψ_J now follows.

In a similar manner we can write

$$\Phi_J = \int_{a_0}^{a_1} \cos(2\theta(x)) \, dx = \Phi_+ + \Phi_-$$

where

$$\Phi_\pm = \pm \int_{-\pi/2}^{\pi/2} \frac{\cos(u)}{2[\gamma V(\eta_\pm(u)) \pm k \cos(u)]} \, du$$

for some bounded piecewise continuous functions $\eta_\pm : [-\pi/2, \pi/2] \to \mathbb{R}$ whose ranges are contained in $J = [a_0, a_1]$ (the range of $\eta_\pm$ is just the set of $x \in J$ where $\pm \cos(2\theta(x)) \geq 0$). For $u \in [-\pi/2, \pi/2]$ we have $0 \leq \cos(u)$ and

$$0 < \gamma m - k \cos(u) \leq \gamma V(\eta_\pm(u)) \pm k \cos(u) \leq \gamma M + k \cos(u),$$

so

$$\int_{-\pi/2}^{\pi/2} \frac{\cos(u)}{2[\gamma M + k \cos(u)]} \, du \leq \pm\Phi_\pm \leq \int_{-\pi/2}^{\pi/2} \frac{\cos(u)}{2[\gamma m - k \cos(u)]} \, du.$$

Now, for any $0 \leq c \leq 1$, $(\gamma m - kc)(\gamma M + kc) \geq (\gamma m - k)(\gamma M + k) > 0$ (recall that $M \geq m$ while $\gamma m \geq \gamma\varepsilon > k$); hence

$$0 \leq \frac{1}{\gamma m - kc} - \frac{1}{\gamma M + kc} = \frac{2kc + \gamma(M - m)}{(\gamma m - kc)(\gamma M + kc)} \leq \frac{2kc + \gamma(M - m)}{(\gamma m - k)(\gamma M + k)}.$$

Putting these estimates together gives

$$
\begin{aligned}
|\Phi_J| &= |\Phi_+ - (-\Phi_-)| \\
&\leq \int_{-\pi/2}^{\pi/2} \frac{\cos(u)}{2[\gamma m - k\cos(u)]} \, \mathrm{d}u - \int_{-\pi/2}^{\pi/2} \frac{\cos(u)}{2[\gamma M + k\cos(u)]} \, \mathrm{d}u \\
&\leq \frac{k}{(\gamma m - k)(\gamma M + k)} \int_{-\pi/2}^{\pi/2} \cos^2(u) \, \mathrm{d}u + \frac{\gamma(M - m)}{2(\gamma m - k)(\gamma M + k)} \int_{-\pi/2}^{\pi/2} \cos(u) \, \mathrm{d}u \\
&= \frac{k\pi}{2(\gamma m - k)(\gamma M + k)} + \frac{\gamma(M - m)}{(\gamma m - k)(\gamma M + k)}.
\end{aligned}
\tag{43}
$$

The second term can be estimated as for (42). On the other hand, (21) implies $|\nabla\theta| \leq \gamma M + k$ on J, so

$$
\pi = \theta(a_1) - \theta(a_0) \leq (\gamma M + k)(a_1 - a_0) = (\gamma M + k)|J|.
$$

The required estimate for the first term in (43) follows. $\qquad\square$

Lemma 4.6. *Suppose θ satisfies* (21) *on an interval $J = [a_0, a_1]$ where $V(x) \geq \varepsilon$ with $\gamma\varepsilon > k$. If $|\theta(a_1) - \theta(a_0)| \leq \pi$ then*

$$
|\Phi_J|, \ |\Psi_J| \leq \frac{2}{\gamma\varepsilon - k}.
$$

Proof. For any $u \in [2\theta(a_0), 2\theta(a_1)]$ we have

$$
\gamma V(\theta^{-1}(u/2)) + k\cos(u) \geq \gamma\varepsilon - k > 0
$$

(note that $\theta^{-1}(u/2) \in J$). For either $f = \sin$ or $f = \cos$ (41) now leads to the estimate

$$
\left| \int_{a_0}^{a_1} f(2\theta(x)) \, \mathrm{d}x \right| \leq \int_{2\theta(a_0)}^{2\theta(a_1)} \frac{|f(u)|}{2(\gamma\varepsilon - k)} \, \mathrm{d}u \leq \frac{1}{2(\gamma\varepsilon - k)} \int_{2\theta(a_0)}^{2\theta(a_0)+2\pi} |f(u)| \, \mathrm{d}u = \frac{2}{\gamma\varepsilon - k},
$$

where the middle step follows since $\theta(a_1) \leq \theta(a_0) + \pi$. $\qquad\square$

The previous Lemmas can be combined in a straightforward way to deal with more general intervals:

Lemma 4.7. *Suppose θ satisfies* (21) *on an interval $J = [a, b]$ where $V(x) \geq \varepsilon$ with $\gamma\varepsilon > k$. Then*

$$
|\Phi_J| \leq \frac{4 + k|J|}{2(\gamma\varepsilon - k)} + \frac{\mathrm{var}_J(V)}{\varepsilon(\gamma\varepsilon - k)} \quad \text{and} \quad |\Psi_J| \leq \frac{2}{\gamma\varepsilon - k} + \frac{\mathrm{var}_J(V)}{\varepsilon(\gamma\varepsilon - k)}.
$$

A similar result holds in the case that $V(x) \leq -\varepsilon$ for all $x \in J$.

Proof. Since θ is absolutely continuous (as is any solution of (21) when $V \in L^1_{\mathrm{loc}}$) and $\nabla\theta(x) \geq \gamma\varepsilon - k > 0$ we can choose $n \geq 0$ and points $a = a_0 < a_1 < \cdots < a_n \leq b$ such that $\theta(a_j) = \theta(a_{j-1}) + \pi$ for $j = 1, \ldots, n$ and $\theta(b) < \theta(a_n) + \pi$. Set $J_j = [a_{j-1}, a_j]$ for $j = 1, \ldots, n$ and $\widetilde{J} = [a_n, b]$ so $J_1, \ldots, J_n, \widetilde{J}$ have disjoint interiors while $J = J_1 \cup \cdots \cup J_n \cup \widetilde{J}$. Now

$$
|\Psi_J| \leq \sum_{j=1}^{n} |\Psi_{J_j}| + |\Psi_{\widetilde{J}}| \leq \sum_{j=1}^{n} \frac{\mathrm{var}_{J_j}(V)}{\varepsilon(\gamma\varepsilon - k)} + \frac{2}{\gamma\varepsilon - k}
$$

using Lemmas 4.5 and 4.6. The Ψ_J estimate now follows from the fact that $\mathrm{var}_{J_1}(V) + \cdots + \mathrm{var}_{J_n}(V) \leq \mathrm{var}_J(V)$. A similar argument leads to the Φ_J estimate. $\qquad\square$

We can now combine the previous result with Lemma 4.4 to deal with arbitrary sub-intervals of $I = \operatorname{supp}(V)$:

Proposition 4.8. *Suppose $V \in BV_0$ has no gaps. Let $\varepsilon > 0$ and consider the notation of Lemma 4.4. If θ satisfies (21) on I and $\gamma\varepsilon > k$ then for any sub-interval $J \subseteq I$ we have*

$$|\Phi_J| \le \frac{6\nu\varepsilon^{-1} + k|J|}{2(\gamma\varepsilon - k)} \quad \text{and} \quad |\Psi_J| \le \frac{3\nu\varepsilon^{-1}}{\gamma\varepsilon - k} + |E_\varepsilon|.$$

Proof. Set $J_{\varepsilon,n} = I_{\varepsilon,n} \cap J$ for $n \in \mathcal{N}_\varepsilon$, and $\mathcal{N}'_\varepsilon = \{n \in \mathcal{N}_\varepsilon : J_{\varepsilon,n} \ne \emptyset\}$. In particular

$$\#\mathcal{N}'_\varepsilon \le \#\mathcal{N}_\varepsilon \le \nu\varepsilon^{-1} \tag{44}$$

by Lemma 4.4(iii), while

$$F_\varepsilon := J \setminus \bigcup_{n \in \mathcal{N}'_\varepsilon} J_{\varepsilon,n} \subseteq I \setminus \bigcup_{n \in \mathcal{N}_\varepsilon} I_{\varepsilon,n} = E_\varepsilon,$$

so

$$|F_\varepsilon| \le |E_\varepsilon|. \tag{45}$$

Now

$$\Psi_J = \sum_{n \in \mathcal{N}'_\varepsilon} \Psi_{J_{\varepsilon,n}} + \int_{F_\varepsilon} \sin(2\theta(x)) \, \mathrm{d}x.$$

Using (45) the modulus of the final term is clearly bounded by $|E_\varepsilon|$. On the other hand, by Lemma 4.4(i) V has constant sign on $J_{\varepsilon,n} \subseteq I_{\varepsilon,n}$ and satisfies $|V(x)| \ge \varepsilon$ for all $x \in J_{\varepsilon,n}$. Lemma 4.7 thus gives

$$|\Psi_{J_{\varepsilon,n}}| \le \frac{2 + \varepsilon^{-1} \operatorname{var}_{J_{\varepsilon,n}}(V)}{\gamma\varepsilon - k}.$$

However the $J_{\varepsilon,n}$ for $n \in \mathcal{N}'_\varepsilon$ are disjoint subintervals of $J \subseteq I$, so

$$\sum_{n \in \mathcal{N}'_\varepsilon} \operatorname{var}_{J_{\varepsilon,n}}(V) \le \operatorname{var}_J(V) \le \operatorname{var}_\mathbb{R}(V) = \nu.$$

Together with (44) we then get

$$\sum_{n \in \mathcal{N}'_\varepsilon} \left(2 + \varepsilon^{-1} \operatorname{var}_{J_{\varepsilon,n}}(V)\right) \le 2\,\#\mathcal{N}'_\varepsilon + \nu\varepsilon^{-1} \le 3\nu\varepsilon^{-1}.$$

The estimate for Ψ_J follows. Since

$$\sum_{n \in \mathcal{N}'_\varepsilon} |J_{\varepsilon,n}| = \left| \bigcup_{n \in \mathcal{N}'_\varepsilon} J_{\varepsilon,n} \right| \le |J|$$

a similar argument leads to the estimate for Φ_J. $\qquad\square$

Proof of Proposition 3.18. Choose $\varepsilon = \varepsilon_\gamma > 0$ for each $\gamma > 0$ so that $\varepsilon_\gamma \to 0$ and $\varepsilon_\gamma^2 \gamma \to \infty$ as $\gamma \to \infty$ (we can take $\varepsilon_\gamma \sim \gamma^{-\mu}$ with $0 < \mu < 1/2$, for example). Letting $\gamma \to \infty$ it follows that $\nu\varepsilon_\gamma^{-1}(\gamma\varepsilon_\gamma - k)^{-1} \to 0$ and $|E_{\varepsilon_\gamma}| \to 0$ (from Lemma 4.4(ii)); both limits are independent of J. On the other hand, $|J| \le |I|$ (since $J \subseteq I$) so we also have $|J|(\gamma\varepsilon_\gamma - k)^{-1} \to 0$ at a rate that can be bounded uniformly in J. The result now follows directly from the estimates in Proposition 4.8. $\qquad\square$

5 Real zeros of a perturbed trigonometric function

This section is devoted to the proof of Theorem 1.15. For notational convenience define a function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \cos(x) + \alpha \cos(\beta x).$$

For any function $\phi : \mathbb{R} \to \mathbb{R}$ we also set $f_\phi = f + \phi$; thus (6) becomes $f_\phi(x) = 0$.

The case $\alpha\beta < 1$ is straightforward; an elementary argument shows that f_ϕ is a "small" perturbation of $\cos(x)$ leading to a one-to-one association between points in the corresponding zero sets (Lemma 5.4).

The case $\alpha\beta > 1$ is dealt with in two main steps; firstly the result is obtained directly when $\phi \equiv 0$ (Section 5.2) and secondly we show that the addition of ϕ cannot change the leading order asymptotic of the number of zeros (Section 5.3). In both steps the arguments for $\beta \in \mathbb{Q}$ and $\beta \notin \mathbb{Q}$ are unified only in the initial stages.

When $\beta \in \mathbb{Q}$ f is periodic and an exact count of the number of zeros can be made. In order to deal with the perturbation ϕ we simply need to avoid cases where f has tangential zeros; this leads to the extra condition $p_\beta + q_\beta \nu_{\alpha,\beta} \notin 4\mathbb{Z}$ in this case (see Lemma 5.13). Since we are then only dealing with the perturbation of transversal zeros we do not require $n = 2$ in condition (7); we will establish the result using the weaker decay condition

$$\phi \in C^1(\mathbb{R}), \quad \phi^{(n)}(x) = o(1) \text{ as } x \to \infty \text{ for } n = 0, 1. \tag{46}$$

When $\beta \notin \mathbb{Q}$ f is no longer periodic but we can appeal to ergodicity to determine the asymptotic distribution of zeros. However dealing with the perturbation ϕ is more subtle in this case as we must consider points where f comes arbitrarily close to having a tangential zero. The number of such points is limited (Corollary 5.23) while condition (7) ensures that the addition of ϕ will alter the number of zeros by at most 2 near each such point (Corollary 5.24).

5.1 Some preliminaries

Let N_ϕ denote the counting function given by

$$\mathsf{N}_\phi(I) = \#\big\{x \in I : f_\phi(x) = 0\big\} \in \mathbb{N} \cup \{0, \infty\}$$

for any interval $I \subseteq \mathbb{R}$; if $R \geq 0$ we'll abuse notation slightly and write $\mathsf{N}_\phi(R)$ for $\mathsf{N}_\phi([0, R])$. We need to determine the limit of $\mathsf{N}_\phi(R)/R$ as $R \to \infty$. To do this it will be convenient to work with sequences of increasing values of R; the next result is straightforward (note that $\mathsf{N}_\phi(R)$ is an non-decreasing function of R).

Lemma 5.1. *The quantity $\mathsf{N}_\phi(R)/R$ has a limit as $R \to \infty$ iff the quantity $\mathsf{N}_\phi(R_n)/R_n$ has a limit as $n \to \infty$ for some (equivalently, any) positive increasing sequence $(R_n)_{n\geq 1}$ with $R_n \to \infty$ and $R_{n-1}/R_n \to 1$ as $n \to \infty$. When the limits exist they are equal.*

Perturbation of zeros

For any function $h \in C^1(\mathbb{R})$ set $E_h = (h)^2 + (h')^2$.

Lemma 5.2. *Suppose $g, \psi \in C^1(\mathbb{R})$ satisfy $E_\psi < E_g$ on an interval I where g' doesn't change sign (that is, either $g' \geq 0$ or $g' \leq 0$). Then $g + \psi$ can have at most one zero on I.*

Proof. Assume $g' \geq 0$ (the case $g' \leq 0$ can be handled similarly). Now suppose $g + \psi$ has (at least) two zeros on I. Then (at least) one of these zeros, say x_0, satisfies $(g + \psi)'(x_0) \leq 0$, so

$$\psi(x_0) = -g(x_0) \quad \text{and} \quad \psi'(x_0) \leq -g'(x_0) \leq 0.$$

This leads to the contradiction $E_\psi(x_0) \geq E_g(x_0)$. $\qquad\qquad\square$

Lemma 5.3. *Let $g, \psi \in C^1(\mathbb{R})$ and suppose $I = [s,t]$ is a closed interval with $g'(s) = 0 = g'(t)$ and $g'(x) \neq 0$ for $x \in (s,t)$. Also suppose $E_\psi < E_g$ on I. Then $g + \psi$ and g have the same number of zeros on I. Furthermore the endpoints of I can't be zeros of either function.*

Proof. Since $g'(x) \neq 0$ for $x \in (s,t)$, g can have at most one zero on I. Also $g' \geq 0$ or $g' \leq 0$ on I (by continuity), so $g + \psi$ has at most one zero on I by Lemma 5.2.

We have $0 \leq \psi^2(s) \leq E_\psi(s) < E_g(s) = g^2(s)$. Thus g and $g + \psi$ are non-zero and have the same sign at s. A similar result applies at t.

If $g(s)g(t) < 0$ then the same is true for $g + \psi$. In this case both functions have at least one, and hence exactly one, zero on I.

If $g(s)g(t) > 0$ (equivalently $(g+\psi)(s)\,(g+\psi)(t) > 0$) then g (respectively $g+\psi$) either has no zeros on I or a single zero $x_0 \in I$ which is also a turning point. It follows that $E_g(x_0) = 0$ (respectively $\psi(x_0) = -g(x_0)$, $\psi'(x_0) = -g'(x_0)$) which leads to a contradiction since $E_g > 0$ on I (respectively $E_\psi(x_0) \neq E_g(x_0)$). Hence neither g nor $g + \psi$ have any zeros on I when $g(s)g(t) > 0$. $\qquad\qquad\square$

The case $\alpha\beta < 1$

The case $\alpha\beta < 1$ in Theorem 1.15 follows easily from the next result.

Lemma 5.4. *Suppose $0 \leq \alpha, \alpha\beta < 1$ and ϕ satisfies (46). Then there exists $N \in \mathbb{N}$ such that f_ϕ has exactly one zero in $[n\pi, (n+1)\pi]$ for any $n \geq N$; furthermore this zero cannot occur at an endpoint of the interval.*

Proof. Set $\psi(x) = \alpha\cos(\beta x) + \phi(x)$ and $\varepsilon = \frac{1}{2}\min\{1 - \alpha^2, 1 - \alpha^2\beta^2\} > 0$. Now

$$E_\psi(x) = \alpha^2\cos^2(\beta x) + \alpha^2\beta^2\sin^2(\beta x) + 2\alpha\cos(\beta x)\phi(x) + 2\alpha\beta\sin(\beta x)\phi'(x) + E_\phi(x)$$
$$\leq \alpha^2\max\{1, \beta^2\} + 2\alpha|\phi(x)| + 2\alpha\beta|\phi'(x)| + E_\phi(x)$$
$$= 1 - 2\varepsilon + 2\alpha|\phi(x)| + 2\alpha\beta|\phi'(x)| + E_\phi(x).$$

Our assumptions on ϕ then allow us to find $N \in \mathbb{N}$ such that $E_\psi(x) \leq 1 - \varepsilon$ for all $x \geq N\pi$. If $n \geq N$ then Lemma 5.3 (with $g = \cos$ so $E_g = 1$) shows that $\cos$ and $\cos +\psi = f_\phi$ have the same number of zeros in $[n\pi, (n+1)\pi]$ and any zeros lie in the interior. The result follows. $\quad\square$

5.2 The unperturbed function

Throughout this section we shall assume $\alpha\beta > 1$ (although several of the results can be extended to cover other cases). Since $\alpha < 1$ it follows that $\beta > 1$. Define $\xi, \eta \in (0, \pi/2)$ by

$$\xi = \arcsin\frac{\sqrt{\alpha^2\beta^2 - 1}}{\sqrt{\beta^2 - 1}} \quad \text{and} \quad \eta = \arcsin\frac{\sqrt{1 - \alpha^2}}{\alpha\sqrt{\beta^2 - 1}}. \qquad (47)$$

The complementary angles satisfy

$$\xi' = \frac{\pi}{2} - \xi = \arcsin\frac{\beta\sqrt{1 - \alpha^2}}{\sqrt{\beta^2 - 1}} \quad \text{and} \quad \eta' = \frac{\pi}{2} - \eta = \arcsin\frac{\sqrt{\alpha^2\beta^2 - 1}}{a\sqrt{\beta^2 - 1}}. \qquad (48)$$

If we fix $\beta > 1$ and vary α from $1/\beta$ to 1 it is easy to check that ξ increases from 0 to $\pi/2$ and η decreases from $\pi/2$ to 0. Also note that $\nu_{\alpha,\beta} = \dfrac{2}{\pi}(\beta\xi + \eta)$ (recall (4)).

Zeros of f

The aim of this section is to establish Theorem 1.15 in the case that $\phi \equiv 0$. For any $n \in \mathbb{Z}$ set $\mathsf{n}_n = \mathsf{N}_0(2\pi n + [0, 2\pi))$; in particular

$$\mathsf{N}_0(2\pi N) = \sum_{n=0}^{N-1} \mathsf{n}_n \qquad (49)$$

for any $N \in \mathbb{N}$. For any $t \in \mathbb{R}$ let

$$\mathsf{m}(t) = \#\big\{x \in [0, \pi) : \cos(x) + \alpha \cos(\beta x + t) = 0\big\}.$$

Clearly $\mathsf{m}(t)$ is 2π-periodic in t while

$$\cos 0 + \alpha \cos(\beta.0 + t) \geq 1 - \alpha > 0 \quad \text{and} \quad \cos \pi + \alpha \cos(\beta\pi + t) \leq -1 + \alpha < 0$$

so the inclusion/exclusion of the possibility $x = 0$ or $x = \pi$ does not alter the definition of m.

Lemma 5.5. *For all $n \in \mathbb{Z}$ we have*

$$\mathsf{n}_n = \mathsf{m}(2\pi n\beta) + \mathsf{m}(-2\pi(n+1)\beta).$$

Proof. Let $x_1 = x - 2\pi n$, $x_2 = 2\pi - x_1$ (so $\cos(x) = \cos(x_1) = \cos(x_2)$). Then

$$f(x) = 0 \quad \Longleftrightarrow \quad \cos(x_1) + \alpha \cos(\beta x_1 + 2\pi n\beta) = 0 \qquad (50)$$
$$\Longleftrightarrow \quad \cos(x_2) + \alpha \cos(\beta x_2 - 2\pi(n+1)\beta) = 0. \qquad (51)$$

Clearly $x_2 \in (0, \pi)$ iff $x_1 \in (\pi, 2\pi)$. Thus

$$\mathsf{n}_n = \#\big\{x_1 \in (0, \pi) : (50) \text{ holds}\big\} + \#\big\{x_2 \in (0, \pi) : (51) \text{ holds}\big\}$$
$$= \mathsf{m}(2\pi n\beta) + \mathsf{m}(-2\pi(n+1)\beta)$$

(from the definition of m). $\qquad \square$

Set $\mu = \beta\xi - \eta'$ (recall (47) and (48)) and define an open interval by

$$J = -\beta\frac{\pi}{2} + \frac{3\pi}{2} + (-\mu, \mu);$$

in particular

$$\nu_{\alpha,\beta} = 1 + \frac{2}{\pi}\mu \quad \text{and} \quad |J| = 2\mu. \qquad (52)$$

For any open interval I set

$$\widetilde{\chi}_I = \frac{1}{2}\big(\chi_I + \chi_{\overline{I}}\big)$$

(so $\widetilde{\chi}_I$ is 1 on I, takes the value $1/2$ at the end points of I, and is 0 elsewhere).

Lemma 5.6. *For any t we have*

$$\mathsf{m}(t) = 1 + 2\sum_{n \in \mathbb{Z}} \widetilde{\chi}_J(t - 2\pi n).$$

Remark 5.7. If $\alpha\beta < 1$ a simplified version of the following argument can be used to show $\mathsf{m}(t) = 1$ for all t.

Proof of Lemma 5.6. Set $\zeta = \arcsin(\alpha) \in \left(0, \frac{\pi}{2}\right)$ and $K = \frac{\pi}{2} + (-\zeta, \zeta)$. Now suppose

$$\cos(\beta x + t) = -\frac{1}{\alpha} \cos(x). \tag{53}$$

for some $x \in [0, \pi)$ and $t \in \mathbb{R}$. Firstly observe that $|\cos(x)| \le \alpha$ so $x \in \overline{K}$. Define $\theta_- : \overline{K} \to \mathbb{R}$ and $\theta_+ : K \to \mathbb{R}$ by

$$\theta_-(x) = \arccos\left(-\frac{1}{\alpha} \cos(x)\right) \quad \text{and} \quad \theta_+(x) = 2\pi - \theta_-(x).$$

[These functions correspond to the two basic branches of the inverse of cos; the remaining branches can be obtained by adding multiples of 2π.] Also define $\omega_- : \overline{K} \to \mathbb{R}$ and $\omega_+ : K \to \mathbb{R}$ by $\omega_s(x) = -\beta x + \theta_s(x)$ for $s \in \{-, +\}$.

Now (53) is equivalent to the existence of unique $s \in \{-, +\}$ and $n \in \mathbb{Z}$ such that $\beta x + t = \theta_s(x) + 2\pi n$. In turn, this is equivalent to

$$\text{there exists unique } s \in \{-, +\}, \ n \in \mathbb{Z} \text{ with } t - 2\pi n = \omega_s(x) \tag{54}$$

(note that, if $x = \pi/2 \pm \zeta$ the we must take $s = -$). We can determine m(t) using (54) if we know the ranges of ω_- and ω_+, together with the multiplicity of covering.

Range of ω_-: The function θ_- is monotonically decreasing while $\beta > 0$ so ω_- is also monotonically decreasing on $\overline{K}$. Thus $\mathrm{Ran}\,\omega_- = I_-$ where

$$I_- = \left[\omega_-\left(\frac{\pi}{2} + \zeta\right), \omega_-\left(\frac{\pi}{2} - \zeta\right)\right] = \left[-\beta\left(\frac{\pi}{2} + \zeta\right), -\beta\left(\frac{\pi}{2} - \zeta\right) + \pi\right].$$

The multiplicity of covering is 1.

Range of ω_+: The turning points of ω_+ (on K) satisfy

$$\theta_+'(x) = \beta \quad \Longleftrightarrow \quad \frac{\sin x}{\sqrt{\alpha^2 - \cos^2(x)}} = \beta \quad \Longleftrightarrow \quad \cos^2(x) = \frac{\alpha^2 \beta^2 - 1}{\beta^2 - 1}.$$

This gives precisely two turning points, at $x_\pm = \frac{\pi}{2} \pm \xi$. Furthermore ω_+ is monotonically increasing on $\left(\frac{\pi}{2} - \zeta, x_-\right)$ and $\left(x_+, \frac{\pi}{2} + \zeta\right)$, and monotonically decreasing on $[x_-, x_+]$. Now

$$\theta_+(x_\pm) = 2\pi - \arccos\left(-\frac{1}{\alpha} \cos(x_\pm)\right) = 2\pi - \frac{\pi}{2} + \arcsin\left(\pm\frac{\sqrt{\alpha^2\beta^2 - 1}}{\alpha\sqrt{\beta^2 - 1}}\right) = \frac{3\pi}{2} \pm \eta',$$

so

$$\omega_+(x_\pm) = -\beta\frac{\pi}{2} \mp \beta\xi + \frac{3\pi}{2} \pm \eta' = -\beta\frac{\pi}{2} + \frac{3\pi}{2} \mp \mu.$$

Hence

$$\mathrm{Ran}\,\omega_+ = \left(\omega_+\left(\frac{\pi}{2} - \zeta\right), \omega_+(x_-)\right) \cup \left(\omega_+(x_+), \omega_+\left(\frac{\pi}{2} + \zeta\right)\right) \cup [\omega_+(x_+), \omega_+(x_-)]$$
$$= I_+^- \cup I_+^+ \cup \overline{J},$$

where

$$I_+^- = \left(-\beta\left(\frac{\pi}{2} - \zeta\right) + \pi, -\beta\frac{\pi}{2} + \frac{3\pi}{2} + \mu\right) \quad \text{and} \quad I_+^+ = \left(-\beta\frac{\pi}{2} + \frac{3\pi}{2} - \mu, -\beta\left(\frac{\pi}{2} + \zeta\right) + 2\pi\right).$$

39

Each interval has a multiplicity of 1.

To complete the proof note that the definition of m and (54) give

$$m(t) = \sum_{n \in \mathbb{Z}} \left(\chi_{I_-} + \chi_{I_+^-} + \chi_{I_+^+} + \chi_{\overline{J}} \right)(t - 2\pi n).$$

Now I_- and I_+^- are disjoint sets with union

$$\widetilde{I} = \left[-\beta \left(\frac{\pi}{2} + \zeta \right), \ -\beta \frac{\pi}{2} + \frac{3\pi}{2} + \mu \right).$$

It follows that

$$\chi_{I_-} + \chi_{I_+^-} + \chi_{I_+^+} = \chi_{\widetilde{I}} + \chi_{I_+^+} = \chi_{-\beta(\pi/2 + \zeta) + [0, 2\pi)} + \chi_J,$$

where the last step uses the identity $\chi_{[p,r)} + \chi_{(q,s)} = \chi_{[p,s)} + \chi_{(q,r)}$ which holds whenever $p, q < r, s$. The result follows. $\qquad \square$

Ergodicity can now be used in the case $\beta \notin \mathbb{Q}$. A convenient ergodic theorem gives us

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} g(2\pi n \beta) = \frac{1}{2\pi} \int_0^{2\pi} g(x) \, \mathrm{d}x \tag{55}$$

whenever $\beta \notin \mathbb{Q}$ and $g : \mathbb{R} \to \mathbb{R}$ is a 2π-periodic Riemann integrable function (this is a version of Weyl equidistribution; see [StSh] for example).

Proof of Theorem 1.15 when $\phi \equiv 0$, $\alpha\beta > 1$ and $\beta \notin \mathbb{Q}$. Combine (49) with Lemmas 5.5 and 5.1 to get

$$2\pi \lim_{R \to \infty} \frac{\mathsf{N}_0(R)}{R} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathsf{n}_n = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \left(m(2\pi n\beta) + m(-2\pi(n+1)\beta) \right).$$

Since $\beta \notin \mathbb{Q}$ and m is a 2π-periodic piecewise constant function (55) and Lemma 5.6 then give

$$2\pi \lim_{R \to \infty} \frac{\mathsf{N}_0(R)}{R} = \frac{1}{2\pi} \int_0^{2\pi} \left(m(x) + m(-x - 2\pi\beta) \right) \mathrm{d}x = \frac{1}{\pi} \int_0^{2\pi} m(x) \, \mathrm{d}x$$

$$= \frac{1}{\pi} \left(2\pi + 2 \sum_{n \in \mathbb{Z}} \int_0^{2\pi} \widetilde{\chi}_J(x - 2\pi n) \, \mathrm{d}x \right) = \frac{1}{\pi} \left(2\pi + 2 \int_{\mathbb{R}} \widetilde{\chi}_J \, \mathrm{d}x \right) = \frac{1}{\pi}(2\pi + 2|J|).$$

The result now follows from (52). $\qquad \square$

Now suppose $\beta \in \mathbb{Q}$. Write $\beta = p/q$ where $p, q \in \mathbb{N}$ are coprime.

Lemma 5.8. *We have*

$$\lim_{R \to \infty} \frac{\mathsf{N}_0(R)}{R} = \frac{1}{\pi} \left(1 + \frac{2}{q} \sum_{n \in \mathbb{Z}} \widetilde{\chi}_J(2\pi n/q) \right).$$

Proof. For $N \in \mathbb{N}$ Lemma 5.5 gives

$$\sum_{n=0}^{qN-1} \mathsf{n}_n = \sum_{j=0}^{N-1}\sum_{k=0}^{q-1}\left[\mathsf{m}\left(2\pi(k+qj)\,\frac{p}{q}\right) + \mathsf{m}\left(-2\pi(k+qj+1)\,\frac{p}{q}\right)\right]$$

$$= N\sum_{k=0}^{q-1}\left[\mathsf{m}\left(2\pi k\,\frac{p}{q}\right) + \mathsf{m}\left(-2\pi(k+1)\,\frac{p}{q}\right)\right]$$

since m is 2π-periodic. Now the mappings

$$k \mapsto kp \mod q \quad \text{and} \quad k \mapsto -(k+1)p \mod q$$

give bijections on $\{0, 1, \dots, q-1\}$ (since p and q are coprime). Together with Lemma 5.1 and (49) we then get

$$\lim_{R\to\infty}\frac{\mathsf{N}_0(R)}{R} = \frac{1}{2\pi}\lim_{N\to\infty}\frac{1}{qN}\sum_{n=0}^{qN-1}\mathsf{n}_n = \frac{1}{\pi q}\sum_{k=0}^{q-1}\mathsf{m}(2\pi k/q).$$

On the other hand, Lemma 5.6 gives

$$\sum_{k=0}^{q-1}\mathsf{m}(2\pi k/q) = \sum_{k=0}^{q-1}\left(1 + 2\sum_{l\in\mathbb{Z}}\widetilde{\chi}_J\left(2\pi(k-lq)\,\frac{1}{q}\right)\right) = q + 2\sum_{n\in\mathbb{Z}}\widetilde{\chi}_J(2\pi n/q),$$

with $n = k - lq$. $\qquad\square$

Proof of Theorem 1.15 when $\phi \equiv 0$, $\alpha\beta > 1$ and $\beta \in \mathbb{Q}$. Let $n \in \mathbb{Z}$. Then

$$\widetilde{\chi}_J(2\pi n/q) = 1 \quad\Longleftrightarrow\quad \frac{2\pi n}{q} \in J$$

$$\Longleftrightarrow\quad -\frac{\beta}{4} + \frac{3}{4} - \frac{1}{4}\frac{2}{\pi}\mu < \frac{n}{q} < -\frac{\beta}{4} + \frac{3}{4} + \frac{1}{4}\frac{2}{\pi}\mu$$

$$\Longleftrightarrow\quad -\frac{1}{4}(p + q\nu_{\alpha,\beta}) + q < n < \frac{1}{4}(p + q\nu_{\alpha,\beta}) - \frac{p+q}{2} + q. \qquad (56)$$

We get $\widetilde{\chi}_J(2\pi n/q) = 1/2$ iff n is equal to one of the endpoints, and $\widetilde{\chi}_J(2\pi n/q) = 0$ iff n lies beyond the given range. To ensure we miss the endpoints we require $p + q\nu_{\alpha,\beta} \notin 4\mathbb{Z}$ (left endpoint) and $(p + q\nu_{\alpha,\beta}) - 2(p+q) \notin 4\mathbb{Z}$ (right endpoint). If $p+q$ is even these conditions are equivalent; otherwise they combine as the requirement $p + q\nu_{\alpha,\beta} \notin 2\mathbb{Z}$. We will now assume this condition is satisfied. From (56) we then get

$$N := \sum_{n\in\mathbb{Z}}\widetilde{\chi}_J(2\pi n/q) = \left\lfloor\frac{1}{4}(p + q\nu_{\alpha,\beta}) - \frac{p+q}{2} + q\right\rfloor - \left\lfloor-\frac{1}{4}(p + q\nu_{\alpha,\beta}) + q\right\rfloor$$

$$= \left\lfloor\frac{1}{4}(p + q\nu_{\alpha,\beta}) - \frac{p+q}{2}\right\rfloor + \left\lfloor\frac{1}{4}(p + q\nu_{\alpha,\beta})\right\rfloor + 1$$

(since $-\lfloor -x\rfloor = \lfloor x\rfloor + 1$ for any $x \notin \mathbb{Z}$).

Case p, q are both odd. Then $(p+q)/2 \in \mathbb{Z}$ so

$$N = 2\left\lfloor\frac{1}{4}(p + q\nu_{\alpha,\beta})\right\rfloor - \frac{p+q}{2} + 1.$$

41

Lemma 5.8 now gives

$$\lim_{R\to\infty}\frac{\mathsf{N}_0(R)}{R}=\frac{1}{\pi}\left(1+\frac{2}{q}\,N\right)=\frac{1}{\pi}\left(\frac{4}{q}\left\lfloor\frac{1}{4}(p+q\nu_{\alpha,\beta})\right\rfloor-\frac{p}{q}+\frac{2}{q}\right);$$

the right hand side is just $A(\alpha,\beta)/\pi$ since $p_\beta=p$, $q_\beta=q$ in this case.

Case p, q have opposite parity. Then $(p+q-1)/2\in\mathbb{Z}$ so

$$N=\left\lfloor\frac{1}{4}(p+q\nu_{\alpha,\beta})-\frac{1}{2}\right\rfloor-\frac{p+q-1}{2}+\left\lfloor\frac{1}{4}(p+q\nu_{\alpha,\beta})\right\rfloor+1=\left\lfloor\frac{1}{2}(p+q\nu_{\alpha,\beta})\right\rfloor-\frac{p+q}{2}+\frac{1}{2}$$

(since $\lfloor x-1/2\rfloor+\lfloor x\rfloor=\lfloor 2x\rfloor-1$ whenever $2x\notin\mathbb{Z}$). Lemma 5.8 now gives

$$\lim_{R\to\infty}\frac{\mathsf{N}_0(R)}{R}=\frac{1}{\pi}\left(1+\frac{2}{q}\,N\right)=\frac{1}{\pi}\left(\frac{2}{q}\left\lfloor\frac{1}{2}(p+q\nu_{\alpha,\beta})\right\rfloor-\frac{p}{q}+\frac{1}{q}\right);$$

since $p_\beta=2p$, $q_\beta=2q$ in this case, the right hand side becomes $A(\alpha,\beta)/\pi$ (note that, the condition $p+q\nu_{\alpha,\beta}\notin 2\mathbb{Z}$ becomes $p_\beta+q_\beta\nu_{\alpha,\beta}\notin 4\mathbb{Z}$). $\qquad\square$

Turning points

Consider the set of non-negative turning points of f,

$$\mathcal{T}=\{x\geq 0:f'(x)=0\}.$$

Since f is an analytic function $\mathcal{T}$ is a discrete subset of $\mathbb{R}$. List the points of $\mathcal{T}$ in increasing order as $t_0=0,t_1,t_2,\dots$ (note that $f'(0)=0$). For each $n\geq 1$ set $J_n=[t_{n-1},t_n]$.

Lemma 5.9. *We have $t_n\to\infty$ as $n\to\infty$ and $t_n-t_{n-1}\leq 2\pi/\beta$ for all $n\in\mathbb{N}$. It follows that $t_n/t_{n-1}\to 1$ as $n\to\infty$.*

Remark 5.10. When $\alpha\beta\leq 1$ the same result holds with 2π in place of $2\pi/\beta$.

Proof of Lemma 5.9. Since $f'(x)=-\sin(x)-\alpha\beta\sin(\beta x)$ we have $(-1)^n f'(x)<0$ when $x=(n+1/2)\pi/\beta$. Thus $\mathcal{T}$ contains at least one point in the interval $\big((n-1/2)\pi/\beta,\ (n+1/2)\pi/\beta\big)$ for any $n\in\mathbb{Z}$; this forces $t_n\to\infty$ and $t_n-t_{n-1}\leq 2\pi/\beta$. $\qquad\square$

Bound on f''

Let $x\in\mathbb{R}$ and set $u_n=f^{(n)}(x)$ for $n=0,1,2$. Then $|u_n|\leq 1+\alpha\beta^n$ while

$$\cos(x)=-\alpha\cos(\beta x)+u_0,$$
$$\sin(x)=-\alpha\beta\sin(\beta x)+u_1,$$
$$\cos(x)=-\alpha\beta^2\cos(\beta x)+u_2.$$

Squaring and rearranging each equation leads to

$$\sin^2(x)-\alpha^2\sin^2(\beta x)=1-\alpha^2+v_0,\qquad v_0=2u_0\alpha\cos(\beta x)-u_0^2,\qquad(57\text{a})$$
$$\sin^2(x)-\alpha^2\beta^2\sin^2(\beta x)=v_1,\qquad v_1=-2u_1\alpha\beta\sin(\beta x)+u_1^2,\qquad(57\text{b})$$
$$\sin^2(x)-\alpha^2\beta^4\sin^2(\beta x)=1-\alpha^2\beta^4+v_2,\qquad v_2=2u_2\alpha\beta^2\cos(\beta x)-u_2^2.\qquad(57\text{c})$$

In particular,

$$|v_n|\leq 2\alpha\beta^n|u_n|+|u_n|^2\leq(1+3\alpha\beta^n)|u_n|.\qquad(58)$$

Solving (57a) and (57b) as linear equations for $\sin^2(x)$ and $\sin^2(\beta x)$ leads to

$$\begin{pmatrix} \sin^2(x) \\ \sin^2(\beta x) \end{pmatrix} = \frac{1-\alpha^2}{\alpha^2(\beta^2-1)} \begin{pmatrix} \alpha^2\beta^2 \\ 1 \end{pmatrix} + \frac{1}{\alpha^2(\beta^2-1)} \begin{pmatrix} \alpha^2\beta^2 v_0 + \alpha^2 v_1 \\ -\alpha^2 v_0 - v_1 \end{pmatrix} \tag{59}$$

(recall that $\beta > 1$). Using (57c) then gives

$$v_2 = (1+\beta^2)(\alpha^2\beta^2 - 1) + \frac{\beta^2(1+\alpha^2\beta^2)}{\beta^2-1} v_0 + \frac{1+\beta^4}{\beta^2-1} v_1.$$

The estimates (58) now imply $|u_2| \geq c_0 - c_{1,0}|u_0| - c_{1,1}|u_1|$ where

$$c_0 = \frac{(1+\beta^2)(\alpha^2\beta^2-1)}{1+3\alpha^2\beta^4}, \quad c_{1,0} = \frac{\beta^2(1+\alpha^2\beta^2)(1+3\alpha^2)}{(\beta^2-1)(1+3\alpha^2\beta^4)} \quad \text{and} \quad c_{1,1} = \frac{(1+\beta^4)(1+3\alpha^2\beta^2)}{(\beta^2-1)(1+3\alpha^2\beta^4)}$$

are positive constants. Taking $c_1 = \sqrt{2}\max\{c_{1,0}, c_{1,1}\}$ we have thus established the following.

Lemma 5.11. *There exist positive constants c_0 and c_1 (depending only on α and β) such that $|f''(x)| \geq c_0 - c_1\sqrt{E_f(x)}$ for all $x \in \mathbb{R}$.*

It follows that we can find $\kappa > 0$ so that

$$\text{if } E_f(x) \leq \kappa^2 \text{ for some } x \in \mathbb{R} \text{ then } |f''(x)| \geq \kappa \tag{60}$$

(we can choose κ to be anything in $(0, c_0/(1+c_1)]$). Now set

$$U_\kappa = \{x \in \mathbb{R}^+ : E_f(x) < \kappa^2\}.$$

Then U_κ is open (as E_f is continuous) and $|f''(x)| \geq \kappa$ for all $x \in U_\kappa$ (by (60)). Further useful properties are as follows.

Lemma 5.12. *Let I be a maximal connected component of U_κ.*

(i) If $E_f'(x) = 0$ for some $x \in U_\kappa$ then $x \in \mathcal{T}$.

(ii) I contains a unique point of $\mathcal{T}$.

(iii) If $\phi \in C^2(\mathbb{R})$ satisfies $|\phi''(x)| < \kappa$ on I then f_ϕ can have at most two zeros on I.

Proof. For part (i) let $x \in U_\kappa$ and suppose $0 = E_f'(x) = 2f'(x)(f(x) + f''(x))$. However $|f(x)|^2 \leq E_f(x) < \kappa^2$ so $f(x) + f''(x) \neq 0$. Thus $f'(x) = 0$.

For part (ii) write $I = (s, t)$. Then $E_f(s) = \kappa^2 = E_f(t)$ so $E_f'(x) = 0$ for some $x \in I$ and hence $x \in \mathcal{T}$ by part (i) (note that, $E_f(0) = (1+a)^2 > 1 > c_0^2 > \kappa^2$ so we can't have $s = 0$). If there were distinct points $x_1, x_2 \in I \cap \mathcal{T}$ then we could find x_0 between x_1 and x_2 (and hence in I) with $f''(x_0) = 0$, contradicting the fact that $|f''(x_0)| \geq \kappa$.

For part (iii) note that we have $|f_\phi''(x)| \geq |f''(x)| - |\phi''(x)| > \kappa - \kappa = 0$ for all $x \in I$. $\square$

Tangential zeros

In this section we use the notation of Theorem 1.15.

Lemma 5.13. *Suppose $\beta \in \mathbb{Q}$. If $f(x) = 0 = f'(x)$ for some x then $p_\beta + q_\beta \nu_{\alpha,\beta} \in 4\mathbb{Z}$.*

Proof. Suppose $f(x) = 0 = f'(x)$ so $\cos(x) = -\alpha\cos(\beta x)$ and $\sin(x) = -\alpha\beta\sin(\beta x)$. In particular $\cos(x)$ and $\cos(\beta x)$ have opposite signs, as do $\sin(x)$ and $\sin(\beta x)$. It follows that x and βx lie in diametrically opposite quadrants.

From (59) (with $v_0 = 0 = v_1$) we get

$$\sin^2(x) = \frac{\beta^2(1-\alpha^2)}{\beta^2-1} \quad \text{and} \quad \sin^2(\beta x) = \frac{1-\alpha^2}{\alpha^2(\beta^2-1)}.$$

Hence

$$x \in (\xi' + \pi\mathbb{Z}) \cup (-\xi' + \pi\mathbb{Z}) \quad \text{and} \quad \beta x \in (\eta + \pi\mathbb{Z}) \cup (-\eta + \pi\mathbb{Z}).$$

Since $\xi', \eta \in (0, \pi/2)$ this can be combined with the earlier observation to give one of four possibilities;

$$\begin{cases} x \in \xi' + \pi 2\mathbb{Z}, & \beta x \in \eta + \pi(2\mathbb{Z}+1), \\ x \in \xi' + \pi(2\mathbb{Z}+1), & \beta x \in \eta + \pi 2\mathbb{Z}, \\ x \in -\xi' + \pi 2\mathbb{Z}, & \beta x \in -\eta + \pi(2\mathbb{Z}+1), \\ x \in -\xi' + \pi(2\mathbb{Z}+1), & \beta x \in -\eta + \pi 2\mathbb{Z}. \end{cases}$$

Comparing the expressions for x and βx we can thus find integers $m, n \in \mathbb{Z}$ of opposite parity such that $\beta(\xi' + m\pi) = \eta + n\pi$. Then

$$\beta + \nu_{\alpha,\beta} = \beta + \frac{2}{\pi}(\beta\xi + \eta) = 2\beta + \frac{2}{\pi}(-\beta\xi' + \eta) = 2\big((1+m)\beta - n\big).$$

However $\beta = p_\beta/q_\beta$ so

$$p_\beta + q_\beta\nu_{\alpha,\beta} = q_\beta(\beta + \nu_{\alpha,\beta}) = 2\big((1+m)p_\beta - nq_\beta\big).$$

Now p_β and q_β have the same parity, as do $1+m$ and n. Therefore $(1+m)p_\beta - nq_\beta \in 2\mathbb{Z}$ and hence $p_\beta + q_\beta\nu_{\alpha,\beta} \in 4\mathbb{Z}$. $\qquad\square$

If $\beta \in \mathbb{Q}$ then E_f is smooth, non-negative and periodic, so E_f can be uniformly bounded away from 0 if it is nowhere zero. Lemma 5.13 thus leads to the following.

Corollary 5.14. *Suppose $\beta \in \mathbb{Q}$. If $p_\beta + q_\beta\nu_{\alpha,\beta} \notin 4\mathbb{Z}$ then there exists $\delta > 0$ such that $E_f(x) \geq \delta$ for all $x \in \mathbb{R}$.*

5.3 Perturbations

Suppose ϕ satisfies (46). Choose a decreasing function $\sigma : \mathbb{R} \to \mathbb{R}^+$ so that $E_\phi(x) < \sigma^2(x)$ for all x and $\sigma(x) \to 0$ as $x \to \infty$. Set

$$\widetilde{\mathcal{T}} = \big\{t_n : n \geq 1, \; E_f(t_n) < \sigma^2(t_{n-1})\big\};$$

these are the turning points of f which are 'small' in some sense (relative to ϕ) and can cause changes in the number of zeros when ϕ is added to f.

Choose $m_0 > 1$ so that $\sigma(t_{m_0-2}) < \kappa$ (which is possible since $\sigma(x) \to 0$ as $x \to \infty$). Now suppose $t_n \in \widetilde{\mathcal{T}}$ for some $n \geq m_0 - 1$. Then

$$E_f(t_n) < \sigma^2(t_{n-1}) \leq \sigma^2(t_{m_0-2}) < \kappa$$

so $t_n \in U_\kappa$. Let K_n denote the maximal connected component of U_κ which contains t_n. If $t_n \notin \widetilde{\mathcal{T}}$ for some $n \geq m_0 - 1$ set $K_n = \emptyset$.

For any $n \geq m_0$ let $\widetilde{J}_n = J_n \setminus (K_{n-1} \cup K_n)$. Firstly note that $\widetilde{J}_n$ is non-empty (as otherwise we would have $J_n \subseteq K_{n-1} \cup K_n \subseteq U_\kappa$, implying that U_κ has a connected component containing the distinct elements t_{n-1}, t_n of $\mathcal{T}$). Also $\widetilde{J}_n$ is an interval (removal of K_{n-1} and K_n could only split the interval J_n if either $K_{n-1} \subset J_n$ or $K_n \subset J_n$ which, in turn, is only possible if $K_{n-1} = \emptyset$ or $K_n = \emptyset$).

Lemma 5.15. *Let $n \geq m_0$ and suppose I is a closed and bounded sub-interval of J_n with $E_f(x) < \kappa$ for some $x \in I$. Then the minimum of E_f on I occurs at an endpoint.*

Proof. Suppose the minimum of E_f on I occurs at x_0 which is an interior point of I. Then x_0 is also in the interior of J_n and hence $x_0 \notin \mathcal{T}$. On the other hand, we must have $E_f'(x_0) = 0$ and $E_f(x_0) < \kappa$, so $x_0 \in \mathcal{T}$ by Lemma 5.12(i). $\qquad\square$

Lemma 5.16. *For any $n \geq m_0$ we have $E_f(x) \geq \sigma^2(x)$ for all $x \in \widetilde{J}_n$.*

Proof. Set $\widetilde{J}_n = [s, t]$ and suppose $E_f(x) < \sigma^2(x)$ for some $x \in \widetilde{J}_n$. Since $x \geq t_{n-1} > t_{m_0-2}$ we get $E_f(x) < \sigma^2(t_{m_0-2}) < \kappa$. Lemma 5.15 then shows that the minimum of E_f on $\widetilde{J}_n$ must occur at either s or t. Now $E_f(s) = \kappa$ if $t_{n-1} \in \widetilde{\mathcal{T}}$ while $s = t_{n-1}$ if $t_{n-1} \notin \widetilde{\mathcal{T}}$. However κ is not the minimum value of E_f on $\widetilde{J}_n$, while having $E_f(t_{n-1})$ as the minimum value would imply

$$E_f(t_{n-1}) \leq E_f(x) < \sigma^2(x) \leq \sigma^2(t_{n-1}) \leq \sigma^2(t_{n-2}),$$

leading to the contradiction $t_{n-1} \in \widetilde{\mathcal{T}}$. A similar argument shows that the minimum value of E_f on $\widetilde{J}_n$ can't occur at t. $\qquad\square$

If $n \geq 1$ then $f'(x) \neq 0$ for any $x \in (t_{n-1}, t_n)$ so f can have at most 1 zero on J_n; that is, $\mathsf{N}_0(J_n) \leq 1$. Since $E_\phi(x) < \sigma^2(x)$ for all x we immediately get the following corollary of Lemmas 5.2, 5.3 and 5.16.

Corollary 5.17. *Suppose $n \geq m_0$. Then $\mathsf{N}_\phi(\widetilde{J}_n) \leq 1$. Furthermore, if $\widetilde{J}_n = J_n$ then $\mathsf{N}_\phi(J_n) = \mathsf{N}_0(J_n)$; in this case neither f nor f_ϕ can have a zero at an endpoint of J_n.*

Note that, the requirement $\widetilde{J}_n = J_n$ is equivalent to $K_{n-1} = \emptyset = K_n$, or $J_n \cap \widetilde{\mathcal{T}} = \emptyset$.

Proof of Theorem 1.15 when $\alpha\beta > 1$, $\beta \in \mathbb{Q}$

Suppose $\alpha\beta > 1$ and $\beta \in \mathbb{Q}$. Theorem 1.15 for general ϕ then follows from the case $\phi \equiv 0$ and the following result.

Proposition 5.18. *Suppose ϕ satisfies (46) and $f_\phi^{-1}(0)$ is a discrete subset of $\mathbb{R}$. Then*

$$\lim_{R \to \infty} \frac{|\mathsf{N}_\phi(R) - \mathsf{N}_0(R)|}{R} = 0.$$

Proof. Using Corollary 5.14 we can choose $m \geq m_0$ so that $E_f(x) \geq \sigma^2(t_{m-1})$ for all x. Then $\widetilde{\mathcal{T}} \cap [t_m, \infty) = \emptyset$ so $\widetilde{J}_n = J_n$ for any $n > m$. Corollary 5.17 then gives $\mathsf{N}_\phi([t_m, t_M]) = \mathsf{N}_0([t_m, t_M])$ for any $M > m$. On the other hand, $f_0^{-1}(0)$ and $f_\phi^{-1}(0)$ are discrete subsets of $\mathbb{R}$ so $\mathsf{N}_0([0, t_m))$ and $\mathsf{N}_\phi([0, t_m))$ are both finite. Thus $\left|\mathsf{N}_\phi(t_M) - \mathsf{N}_0(t_M)\right| = O(1)$ as $M \to \infty$. Then

$$\lim_{R \to \infty} \frac{|\mathsf{N}_\phi(R) - \mathsf{N}_0(R)|}{R} = \lim_{M \to \infty} \frac{|\mathsf{N}_\phi(t_M) - \mathsf{N}_0(t_M)|}{t_M} = 0$$

with the help of Lemmas 5.1 and 5.9. $\qquad\square$

Distribution of points in $\widetilde{\mathcal{T}}$

Lemma 5.19. *There exists a constant $c > 0$ such that $|t - s| \geq c$ for all distinct $s, t \in \widetilde{\mathcal{T}}$.*

Proof. Let $s, t \in \widetilde{\mathcal{T}}$ with $s < t$. Now $s, t \in U_\kappa$ so $|f''(s)|, |f''(t)| \geq \kappa$. Also $f'(s) = 0 = f'(t)$ so we can find $x_0 \in (s, t)$ with $f''(x_0) = 0$. Then the total variation of f'' between s and t is at least 2κ. However $|f'''(x)| \leq 1 + \alpha\beta^3$ for all x; thus $t - s \geq 2\kappa/(1 + \alpha\beta^3)$. $\square$

Split $\widetilde{\mathcal{T}}$ into a pair of increasing sequences of distinct points $(s_n^+)_{n \geq 1}$ and $(s_n^-)_{n \geq 1}$ so that $s_n^+ \in [0, \pi/2) + \pi\mathbb{Z}$ and $s_n^- \in [\pi/2, \pi) + \pi\mathbb{Z}$ for all n.

Lemma 5.20. *Suppose $(s_n^+)_{n \geq 1}$ is an infinite sequence. Then $s_n^+ - s_{n-1}^+$, $b(s_n^+ - s_{n-1}^+) \longrightarrow \pi\mathbb{N}$ as $n \to \infty$. A similar result hold for $(s_n^-)_{n \geq 1}$.*

If $(x_n)_{n \geq 1}$ is a sequence and $X \subseteq \mathbb{R}$ then $x_n \to X$ simply means $\mathrm{dist}(x_n, X) \to 0$.

Proof of Lemma 5.20. We have $f'(s_n^+) = 0$ for all n and $|f(s_n^+)|^2 = E_f(s_n^+) \leq \sigma^2(s_{n-1}^+) \to 0$ as $n \to \infty$. From (59) it follows that

$$\sin^2(s_n^+) \longrightarrow \frac{\beta^2(1 - \alpha^2)}{\beta^2 - 1} \quad \text{and} \quad \sin^2(\beta s_n^+) \longrightarrow \frac{1 - \alpha^2}{\alpha^2(\beta^2 - 1)}$$

as $n \to \infty$. Hence

$$s_n^+ \longrightarrow (\xi' + \pi\mathbb{Z}) \cup (-\xi' + \pi\mathbb{Z}) \quad \text{and} \quad \beta s_n^+ \longrightarrow (\eta + \pi\mathbb{Z}) \cup (-\eta + \pi\mathbb{Z})$$

as $n \to \infty$ (recall (47) and (48) for the definitions of ξ' and η). However, $s_n^+ \in [0, \pi/2) + \pi\mathbb{Z}$ for all n, while $\cos(s_n^+) = -\alpha\cos(\beta s_n^+) + f(s_n^+)$ and $\sin(s_n^+) = -\alpha\beta\sin(\beta s_n^+)$ so (modulo 2π) βs_n^+ must tend to a quadrant diametrically opposite s_n^+. Hence $s_n^+, \beta s_n^+ \longrightarrow [0, \pi/2] + \pi\mathbb{Z}$ as $n \to \infty$. Comparing with the previous expressions we then get

$$s_n^+ \longrightarrow \xi' + \pi\mathbb{Z} \quad \text{and} \quad \beta s_n^+ \longrightarrow \eta + \pi\mathbb{Z} \quad \Longrightarrow \quad s_n^+ - s_{n-1}^+, \ \beta(s_n^+ - s_{n-1}^+) \longrightarrow \pi\mathbb{Z}$$

as $n \to \infty$ (note that $\xi', \eta \in (0, \pi/2)$). Finally note that Lemma 5.19 gives $s_n^+ - s_{n-1}^+ \geq c > 0$ for all n, so we can replace the final $\pi\mathbb{Z}$ with $\pi\mathbb{N}$. $\square$

Lemma 5.21. *If $\beta \notin \mathbb{Q}$ then $\#\{n : s_n^+ - s_{n-1}^+ \leq C\} < \infty$ for all constants C. A similar result holds for $(s_n^-)_{n \geq 1}$.*

Proof. Suppose $\#\{n : s_n^+ - s_{n-1}^+ \leq C\} = \infty$ for some C. Then we can find $x \in [c, C]$ and a sub-sequence $s_{n_k}^+$ such that $s_{n_k}^+ - s_{n_k-1}^+ \to x$ as $k \to \infty$. Lemma 5.20 then gives $x = q\pi$ and $\beta x = p\pi$ for some $p, q \in \mathbb{N}$, so $\beta = p/q \in \mathbb{Q}$. $\square$

Lemma 5.22. *If $(x_n)_{n \geq 1}$ is an increasing positive sequence with $\#\{n : x_n - x_{n-1} \leq C\} < \infty$ for all C then $\#\{n : x_n \leq R\} = o(R)$ as $R \to \infty$.*

Proof. Let $\varepsilon > 0$ and choose N so that $x_n - x_{n-1} \leq 1/\varepsilon$ for all $n > N$. Then, for all $m \geq 1$, $x_{N+m} \geq x_N + m/\varepsilon$. Hence

$$\#\{n \geq 1 : x_n \leq R\} \leq N + \#\{m \geq 1 : x_{N+m} \leq R\}$$
$$\leq N + \#\{m \geq 1 : x_N + m/\varepsilon \leq R\} \leq N + \varepsilon(R - x_N)$$

whenever $R \geq x_N$. Thus

$$\limsup_{R \to \infty} \frac{1}{R} \#\{n \geq 1 : x_n \leq R\} \leq \limsup_{R \to \infty} \left(\varepsilon + \frac{N - \varepsilon x_N}{R} \right) = \varepsilon.$$

Taking $\varepsilon \to 0^+$ completes the result. $\square$

Since $\#(\widetilde{\mathcal{T}} \cap [0,R]) = \#\{n : s_n^+ \leq R\} + \#\{n : s_n^- \leq R\}$ Lemmas 5.21 and 5.22 immediately lead to the following.

Corollary 5.23. *If $\beta \notin \mathbb{Q}$ then $\#(\widetilde{\mathcal{T}} \cap [0,R]) = o(R)$ as $R \to \infty$.*

Proof of Theorem 1.15 when $\alpha\beta > 1$, $\beta \notin \mathbb{Q}$

If ϕ satisfies (7) we can choose $m \geq m_0$ so that $|\phi''(x)| < \kappa$ for all $x \geq t_{m-1}$. Lemma 5.12(iii) immediately gives the following (note that, if $K_n = \emptyset$ the result is trivial).

Corollary 5.24. *If $n \geq m$ then $\mathsf{N}_\phi(K_n) \leq 2$.*

Now suppose $\alpha\beta > 1$ and $\beta \notin \mathbb{Q}$. Theorem 1.15 for general ϕ then follows from the case $\phi \equiv 0$ and the following result.

Proposition 5.25. *Suppose ϕ satisfies (7) and $f_\phi^{-1}(0)$ is a discrete subset of $\mathbb{R}$. Then*

$$\lim_{R \to \infty} \frac{|\mathsf{N}_\phi(R) - \mathsf{N}_0(R)|}{R} = 0.$$

Proof. If $M > m$ then

$$[t_m, t_M] = \bigcup_{m < n \leq M} J_n' = \bigcup_{m < n \leq M} \widetilde{J}_n \cup \bigcup_{m < n < M} K_n \cup K_m' \cup K_M',$$

where $J_n' = J_n \setminus \{t_n\} = [t_{n-1}, t_n)$ for $m < n < M$, $J_M' = J_M$, $K_m' = K_m \cap [t_m, \infty)$ and $K_M' = K_M \cap [0, t_M]$. Furthermore the intervals in the first covering are disjoint while those in the second covering can only overlap at points $t_n \in \mathcal{T} \setminus \widetilde{\mathcal{T}}$; for such points

$$f^2(t_n) = E_f(t_n) \geq \sigma^2(t_{n-1}) \geq \sigma^2(t_n) > E_\phi(t_n) \geq \phi^2(t_n) \geq 0$$

so t_n is not zero of f_ϕ. Therefore

$$\mathsf{N}_0([t_m, t_M]) = \sum_{m < n \leq M} \mathsf{N}_0(J_n')$$

and

$$\mathsf{N}_\phi([t_m, t_M]) = \sum_{m < n \leq M} \mathsf{N}_0(\widetilde{J}_n) + \sum_{m < n < M} \mathsf{N}_\phi(K_n) + \mathsf{N}_\phi(K_m') + \mathsf{N}_\phi(K_M').$$

Now set

$$\mathsf{k}_{m,M} = \#\{m \leq n \leq M : K_n \neq \emptyset\} = \#(\widetilde{\mathcal{T}} \cap [t_m, t_M]).$$

By Corollary 5.24

$$\sum_{m < n < M} \mathsf{N}_\phi(K_n) + \mathsf{N}_\phi(K_m') + \mathsf{N}_\phi(K_M') \leq \sum_{m \leq n \leq M} \mathsf{N}_\phi(K_n) \leq 2\mathsf{k}_{m,M}.$$

By Corollary 5.17 (and the discussion proceeding it) $\mathsf{N}_\phi(\widetilde{J}_n) - \mathsf{N}_\phi(J_n') = 0$ if $\widetilde{J}_n = J_n$, while $|\mathsf{N}_\phi(\widetilde{J}_n) - \mathsf{N}_\phi(J_n')| \leq 1$ in general. Since $\#\{m < n \leq M : \widetilde{J}_n \neq J_n\} \leq 2\mathsf{k}_{m,M}$ we then get

$$\left| \sum_{m < n \leq M} \mathsf{N}_0(\widetilde{J}_n) - \sum_{m < n \leq M} \mathsf{N}_0(J_n') \right| \leq 2\mathsf{k}_{m,M}.$$

Combining the above estimates now gives

$$\big|\mathsf{N}_\phi([t_m, t_M]) - \mathsf{N}_0([t_m, t_M])\big| \le 4\mathsf{k}_{m,M}.$$

However $\mathsf{k}_{m,M} \le \#(\widetilde{\mathcal{T}} \cap [0, t_M])$ so Corollary 5.23 (together with Lemma 5.9) implies

$$\big|\mathsf{N}_\phi([t_m, t_M]) - \mathsf{N}_0([t_m, t_M])\big| = o(t_M)$$

as $M \to \infty$. On the other hand, $f_0^{-1}(0)$ and $f_\phi^{-1}(0)$ are discrete subsets of $\mathbb{R}$ so $\mathsf{N}_0([0, t_m))$ and $\mathsf{N}_\phi([0, t_m))$ are both finite. Thus $\big|\mathsf{N}_\phi(t_M) - \mathsf{N}_0(t_M)\big| = o(t_M)$ as $M \to \infty$ and so

$$\lim_{R \to \infty} \frac{|\mathsf{N}_\phi(R) - \mathsf{N}_0(R)|}{R} = \lim_{M \to \infty} \frac{|\mathsf{N}_\phi(t_M) - \mathsf{N}_0(t_M)|}{t_M} = 0$$

with the help of Lemmas 5.1 and 5.9. $\qquad\square$

Remark 5.26. It is instructive to look at the limiting cases in Theorem 1.15. When $\alpha = 1$ we have
$$f(x) = \cos(x) + \cos(\beta x) = 2\cos\Big(\frac{\beta + 1}{2}\, x\Big) \cos\Big(\frac{\beta - 1}{2}\, x\Big).$$

The zeros of this function occur precisely when $(\beta \pm 1)x \in (2\mathbb{Z} + 1)\pi$. If $\beta \ne p/q$ for some $p, q \in \mathbb{N}$ with opposite parity then all zeros of f are simple and

$$\lim_{R \to \infty} \frac{\mathsf{N}_0(R)}{R} = \frac{1}{\pi}\Big(\frac{\beta + 1}{2} + \frac{|\beta - 1|}{2}\Big) = \frac{1}{\pi}\begin{cases} 1 & \text{if } \beta < 1, \\ \beta & \text{if } \beta > 1. \end{cases}$$

The same formula holds for arbitrary β if we count zeros with multiplicity. This agrees with Theorem 1.15 and the limiting behaviour of $\nu_{\alpha,\beta}$ as $\alpha \to 1^-$.

At the same time, Theorem 1.15 does not extend in a straightforward manner to the case $\alpha\beta = 1$. This can be seen by taking $\alpha = 1/3$, $\beta = 3$; then

$$f(x) = \cos(x) + \frac{1}{3}\cos(3x) = \frac{4}{3}\cos^3(x).$$

Although the zeros of this function are precisely the points $\big(n + \frac{1}{2}\big)\pi$ for $n \in \mathbb{Z}$, we also have

$$f'\Big(\Big(n + \frac{1}{2}\Big)\pi\Big) = f''\Big(\Big(n + \frac{1}{2}\Big)\pi\Big) = 0.$$

It is then straightforward to construct a perturbation ϕ satisfying (7) so that $f_\phi = f + \phi$ has arbitrarily many zeros close to $\big(n + \frac{1}{2}\big)\pi$ for each $n \in \mathbb{Z}$.

References

[Ad]   R. A. Adams, *Sobolev Spaces*, Pure and Applied Mathematics **65**, Academic Press, New York-London, 1975.

[AMS]   J. Avron, P. H. M. v. Mouche, B. Simon, *On the measure of the Spectrum for the almost Mathieu operator*, Commun. Math. Phys. **132**, 103–118 (1990).

[BaTiBr] J. H. Bardarson, M. Titov, P. W. Brouwer, *Electrostatic Confinement of Electrons in an Integrable Graphene Quantum Dot*, Phys. Rev. Lett. **102**, 226803 (2009).

[BiLa]   M. S. Birman, A. Laptev, *Discrete spectrum of the perturbed Dirac operator*, Ark. Mat. **32**, 13–32 (1994).

[BiSo]   M. S. Birman, M. Z. Solomyak, *Spectral asymptotics of pseudodifferential operators with anisotropic homogeneous symbols*, I, Vestnik Leningrad. Univ. Mat. Mekh. Astronom. **13**, 13–21 (1977); II, Vestnik Leningrad. Univ. Mat. Mekh. Astronom. **13**, 5–10 (1979) (Russian).

[BrFr]   L. Brey, H. A. Freitig, *Emerging Zero Modes for Graphene in a Periodic Potential*, Phys. Rev. Lett. **103**, 046809 (2009).

[Cw]   M. Cwikel, *Weak type estimates for singular values and the number of bound states of Schrödinger operators*, Ann. of Math. **106**, 93–100 (1977).

[GGL]   K. Golden, S. Goldstein, J.L. Lebowitz, *Classical transport in modulated structures*, Phys. Rev. Lett. **55**, no. 24, 2629–2632 (1985).

[ElTa]   D. M. Elton, N. T. Ta, *Eigenvalue Counting Estimates for a Class of Linear Spectral Pencils with Applications to Zero Modes*, J. Math. Anal. Appl. **391**, 613–618 (2012) .

[GGHKSSV]   F. Gesztesy, D. Gurarie, H. Holden, M. Klaus, L. Sadun, B. Simon, P. Vogl, *Trapping and cascading eigenvalues in the large coupling limit*, Commun. Math. Phys. **118**, 597–634 (1988).

[Ha]   P. Hartman, *Ordinary Differential Equations*, John Wiley and Sons, New York, 1964.

[HRP]   R. R. Hartmann, N. J. Robinson, M. E. Portnoi, *Smooth electron waveguides in graphene*, Phys. Rev. B **81**, 245431 (2010) .

[JM]   S. Jitomirskaya, C. A. Marx, *Analytic quasi-periodic Schrdinger operators and rational frequency approximants*, Geom. Funct. Anal. **22**, no. 5, 1407–1443 (2012).

[Kac]   M. Kac, *On the distribution of values of trigonometric sums with linearly independent frequencies*, Amer. J. Math. **65**, no. 4, 609–615 (1943).

[KKW]   M. Kac, E. R. van Kampen, A. Wintner, *On the distribution of the values of real almost periodic functions*, Amer. J. Math. **61**, 985–991 (1939).

[Kl]   M. Klaus, *On the point spectrum of Dirac operators*, Helv. Phys. Acta **53**, 453–462 (1980).

[ReSi]   M. Reed, B. Simon, *Methods of Modern Mathematical Physics IV: Analysis of Operators*, Academic Press, San Diego, 1978.

[Ru]   W. Rudin, *Real and Complex Analysis*, 3rd Edition, McGraw-Hill, Singapore, 1987.

[Sa]   O. Safronov, *The discrete spectrum of selfadjoint perturbations of variable sign*, Commun. PDE **26**, no. 3-4, 629–649 (2001).

[Sch]   K. M. Schmidt, *Spectral properties of rotationally symmetric massless Dirac operators*, Lett. Math. Phys. **92**, 231–241 (2010).

[StSh]   E. M. Stein, R. Shakarchi, *Fourier Analysis: An Introduction*, Princeton University Press, Princeton, 2003.

[St]   P. Stein, *On the Real Zeros of a certain Trigonometric Function*, Math. Proc. Cambridge Philos. Soc. **31**, 455–467 (1935).

[StDoPo]   D. A. Stone, C. A. Downing, M. E. Portnoi, *Searching for confined modes in graphene channels: The variable phase method*, Phys. Rev. B **86**, 075464 (2012).

[WeSe]   E. Wegert, G. Semmler, *Phase plots of complex functions: a journey in illustration*, Notices AMS **58**, 768–780 (2010).

[Wey]   H. Weyl, *Inequalities between two kinds of eigenvalues of a linear transformation*, Proc. Natl. Acad. Sci. USA **35**, 408–411 (1949).