Acta Math., 208 (2012), 151–209
DOI: 10.1007/s11511-012-0076-4


# Spectral gaps for sets and measures

by

Alexei Poltoratski

*Texas A&M University*
*College Station, TX, U.S.A.*

## 1. Introduction

Let $\mu$ be a non-zero finite complex measure on the real line. By $\hat{\mu}$ we denote its Fourier transform

$$\hat{\mu}(z) = \int_{\mathbb{R}} e^{-izt}\, d\mu(t).$$

Various properties of the Fourier transform of a measure have been studied by harmonic analysts for more than a century. One of the reasons for such a prolonged interest is the natural physical sense of the quantity $\hat{\mu}(t)$. For instance, in mathematical models of quantum mechanics, if $\mu$ is a spectral measure of a Hamiltonian, then $|\hat{\mu}(t)|^2$ represents the so-called survival probability of the particle, i.e. the probability to find the particle in its initial state, at the moment $t$. The problems considered in the present paper belong to the area of the uncertainty principle in harmonic analysis, whose name itself suggests relations and similarities with physics.

The uncertainty principle in harmonic analysis, as formulated in [13], says that a measure (function, distribution) and its Fourier transform cannot be simultaneously small. This broad statement gives rise to a multitude of exciting mathematical problems, each corresponding to a particular sense of "smallness".

One such problem is the well-known gap problem. Here the smallness of $\mu$ and $\hat{\mu}$ is understood in the sense of porosity of their supports. The statement that one hopes to obtain is that if the support of $\hat{\mu}$ has a large gap, then the support of $\mu$ cannot be too "rare". As usual, the ultimate challenge is to obtain quantitative estimates relating the two supports, something that we will attempt to do in this paper.

Beurling's gap theorem says that if the sequence of gaps in the support of $\mu$ is long, in the sense given by (4.3), then the support of $\hat{\mu}$ cannot have *any* gaps, unless $\mu$ is trivial,

---

The author was supported by N.S.F. Grant No. 0800300.

see [3] or [15, Volume 1, p. 237]. We discuss this classical result in §5. Beurling's proof used some of the methods of an earlier gap theorem by Levinson [19]. In [7, Theorem 66, p. 271] de Branges proved that the existence of a measure with a given spectral gap is equivalent to the existence of a certain entire function of exponential type. In §8 we look at this result from the point of view of Toeplitz kernels and formulate its extension. Further results and references concerning the gap problem can be found in [2], [3], [6], [10], [15], [19] and [32].

To state the gap problem more precisely let us give the following definition. If $X$ is a closed subset of the real line, denote by $G_X$ its *gap characteristic*, i.e. the supremum of the size of the gap in the support of $\hat{\mu}$, taken over all non-trivial finite complex measures $\mu$ supported on $X$, see §2. The gap problem is the problem of finding $G_X$ in terms of $X$. To formulate a solution, we introduce a new metric characteristic of a closed set, $\mathcal{C}_X$, see §4. Our main result is Theorem 4.7, which says that $G_X$ is equal to $2\pi\mathcal{C}_X$.

The definition of $\mathcal{C}_X$ contains two conditions that, for the purposes of this paper, we call the density condition and the energy condition. The density condition is similar to some of the definitions of densities used in the area of the uncertainty principle, see §4. The physical flavor of the energy condition seems to suggest new connections for the gap problem that are yet to be fully understood.

As discussed in §2, the gap problem can be equivalently reformulated as follows. Let $\mu$ be a finite complex measure on $\mathbb{R}$. Find the supremum $G_\mu$ of the size of the spectral gap of the measure $f\mu$, taken over all non-trivial $f\in L^1(|\mu|)$. In this reformulation $G_\mu=G_X$ for $X=\operatorname{supp}\mu$, see Proposition 2.1.

A close relative of the gap problem is another classical question of harmonic analysis, the so-called type problem. It can be stated in several equivalent ways. For instance, if one replaces $L^1$ with $L^2$ in the above definition of $G_\mu$ one can introduce a similar quantity $G^2_\mu$, the supremum of the size of the gap in the support of $\widehat{f\mu}$, taken over all non-zero $f\in L^2(|\mu|)$. Via duality, $G^2_\mu$ can also be defined as the infimum of $a$ such that the family of exponential functions

$$\mathcal{E}_a=\{e^{i\lambda x}:\lambda\in[0,a]\}$$

spans $L^2(\mu)$. The type problem asks for finding $G^2_\mu$ in terms of $\mu$.

The type problem dates back to the works of Wiener, Kolmogorov and Kreĭn on stationary Gaussian processes. In that context the property that the family of exponentials $\mathcal{E}_a$ is complete in $L^2(\mu)$, where $\mu$ is the spectral measure of the process, is equivalent to the property that the process at any time can be predicted from the data for the time period from 0 to $a$. As any positive even measure is a spectral measure of a stationary Gaussian process and vice versa, this reformulation is practically equivalent, see [11], [16] and [18].

Since for finite measures

$$G_\mu^2 \leqslant G_\mu = G_{\operatorname{supp}\mu},$$

Theorem 4.7 gives an upper estimate for $G_\mu^2$. The methods we develop in this paper can be applied to give further results for the type problem, see [28].

In addition to $G_\mu^1 = G_\mu$ and $G_\mu^2$, one can define and study similar quantities for other $p$, see for instance the book by Koosis [15] for the case $p=\infty$, or [28].

Such problems can also be restated in terms of the Bernstein uniform weighted approximation, see for example [15]. From that point of view, $G_X$ is the minimal size of the interval such that continuous functions on $X$ admit weighted approximation by trigonometric polynomials with frequencies from that interval. We give a short example of a similar connection at the end of §3. Important relations with spectral theory of second order differential operators were studied by Kreĭn [17], [18], Gelfand and Levitan [12], and Borichev and Sodin [6].

Our methods are based on the approach developed by N. Makarov and the author in [20] and [21]. We utilize close connections between most problems from this area of harmonic analysis and the problem of injectivity of Toeplitz operators. In the case of the gap problem, this connection is expressed by Theorem 3.2 below. The Toeplitz approach for similar problems was first suggested by Nikol′skii in [24], see also [25]. Our main proof utilizes several important ideas of the Beurling–Malliavin theory [4], [5], [21], including its famous multiplier theorem.

One of the advantages of the Toeplitz approach is that it reveals hidden connections between various problems of analysis and mathematical physics, see [20]. The relations between the gap problem and the Beurling–Malliavin theory on the completeness of exponentials in $L^2$ on an interval have been known to experts, at a rather intuitive level, for several decades. Now we can see this connection formulated in precise mathematical terms. Namely, the Beurling–Malliavin problem is equivalent to the problem of triviality of the kernel of a Toeplitz operator with the symbol

$$\phi = e^{-iax}\theta,$$

for a suitable meromorphic inner function $\theta$, while the gap problem reduces to the triviality of the kernel of the Toeplitz operator with the symbol

$$\bar{\phi} = e^{iax}\bar{\theta},$$

see [20, §4.6] and Theorem 3.2 below.

The paper is organized as follows:

- In §2 we discuss an alternative formulation of the gap problem and show that the supremum of the size of the spectral gap for a fixed measure, taken over all possible densities, is determined by the support of the measure.

- In §3 we restate the gap problem in terms of kernels of Toeplitz operators and introduce the approach that will be used in the main proof. We point out a connection with problems on uniform approximation.
- §4 contains the main definition and the main result of the paper. For a closed real set $X$ we define a metric characteristic $\mathcal{C}_X$ that determines the maximal size of the gap over all non-zero complex measures supported on $X$.
- In §5 we discuss examples pertinent to the main theorem as well as its relations with some of the known results.
- §6 contains the main proof.
- In §7 we prove several technical lemmas and corollaries used in §4 and §6.
- §8 can be viewed as an appendix. It contains a Toeplitz version of the statement and proof of de Branges' theorem [7, Theorem 66].

*Acknowledgments.* I would like to thank Nikolai Makarov whose mathematical intuition led to the development of the approach used in this paper. I am also grateful to Misha Sodin for getting me interested in the gap and type problems and for numerous invaluable discussions.

## 2. Spectral gap as a property of the support

Let $M$ be a set of all finite Borel complex measures on the real line. If $X$ is a closed subset of the real line we set

$$G_X=\sup\{a: \text{there is } \mu\in M, \text{ with } \mu\not\equiv 0 \text{ and } \operatorname{supp}\mu\subset X, \text{ such that } \hat{\mu}=0 \text{ on } [0,a]\}.$$

Now let $\mu\in M$. Denote

$$G_\mu=\sup\{a: \text{there is } f\in L^1(|\mu|) \text{ such that } \widehat{f\mu}=0 \text{ on } [0,a]\}.$$

PROPOSITION 2.1. $G_\mu=G_{\operatorname{supp}\mu}$.

*Proof.* Obviously, $G_{\operatorname{supp}\mu}\geqslant G_\mu$. To prove the opposite inequality, notice that by Lemma 8.2 below there exists a finite discrete measure

$$\nu=\sum_n \alpha_n\delta_{x_n}, \quad \{x_n\}_n\subset\operatorname{supp}\mu,$$

such that $\hat{\mu}$ has a gap of size greater than $G_{\operatorname{supp}\mu}-\varepsilon$. Around each $x_n$ choose a small neighborhood $V_n=(a_n,b_n)$ such that for any sequence of points

$$Y=\{y_n\}_n, \quad y_n\in V_n,$$

there exists a non-trivial measure $\eta_Y=\sum_n \beta_n\delta_{y_n}$ such that $\hat{\eta}$ has a gap of size greater than $G_{\operatorname{supp}\eta}-\varepsilon$. The existence of such a collection of neighborhoods follows from the results of [2] (for some sequences) and [6], as well as from Theorem 4.7 below.

Now one can choose a family of finite measures $\eta_\tau$, $\tau\in[0,1]$, with the following properties:

- For each $\tau$,

$$\eta_\tau=\sum_n \beta_n^\tau\delta_{y_n^\tau}$$

where $y_n^\tau\in V_n$ and $\hat{\eta}_\tau$ has a gap of size greater than $G_{\operatorname{supp}\eta}-\varepsilon$ centered at 0;

- The measure

$$\gamma=\int_0^1 \eta_\tau\,d\tau$$

is non-trivial and absolutely continuous with respect to $\mu$.

It remains to observe that the support of $\hat{\gamma}$ has a gap of size at least $G_{\operatorname{supp}\eta}-\varepsilon$. □

## 3. Clark measures, Toeplitz kernels and uniform approximation

By $H^2$ we denote the Hardy space in the upper half-plane $\mathbb{C}_+$. We say that an inner function $\theta(z)$ in $\mathbb{C}_+$ is *meromorphic* if it has a meromorphic extension to the whole complex plane. The meromorphic extension to the lower half-plane $\mathbb{C}_-$ is given by

$$\theta(z)=\frac{1}{\bar{\theta}(\bar{z})}.$$

Each inner function $\theta(z)$ determines a model subspace

$$K_\theta=H^2\ominus\theta H^2$$

of the Hardy space $H^2(\mathbb{C}_+)$. These subspaces play an important role in complex and harmonic analysis, as well as in operator theory, see [25].

Each inner function $\theta(z)$ determines a positive harmonic function

$$\operatorname{Re}\frac{1+\theta(z)}{1-\theta(z)}$$

and, by the Herglotz representation, a positive measure $\sigma$ such that

$$\operatorname{Re}\frac{1+\theta(z)}{1-\theta(z)}=py+\frac{1}{\pi}\int_{\mathbb{R}}\frac{y\,d\sigma(t)}{(x-t)^2+y^2},\quad z=x+iy, \tag{3.1}$$

for some $p\geqslant 0$. The number $p$ can be viewed as a point mass at infinity. The measure $\sigma$ is singular, supported on the set where non-tangential limits of $\theta$ are equal to 1 and satisfies

$$\int_{\mathbb{R}} \frac{d\sigma(t)}{1+t^2} < \infty. \tag{3.2}$$

The measure $\sigma+p\delta_\infty$ on $\widehat{\mathbb{R}}$ is called the *Clark measure* for $\theta(z)$. (Following a standard notation, we will sometimes denote the Clark measure defined in (3.1) by $\sigma_1$.)

Conversely, for every positive singular measure $\sigma$ satisfying (3.2) and a number $p\geqslant 0$, there exists an inner function $\theta(z)$ satisfying (3.1).

Every function $f\in K_\theta$ has non-tangential boundary values $\sigma$-a.e. and can be recovered from these values via the formula

$$f(z) = \frac{p}{2\pi i}(1-\theta(z)) \int_{\mathbb{R}} f(t)\overline{(1-\theta(t))}\, dt + \frac{1-\theta(z)}{2\pi i} \int_{\mathbb{R}} \frac{f(t)}{t-z}\, d\sigma(t), \tag{3.3}$$

see [26]. If the Clark measure does not have a point mass at infinity, the formula is simplified to

$$f(z) = \frac{1}{2\pi i}(1-\theta(z))Kf\sigma,$$

where $Kf\sigma$ stands for the Cauchy integral

$$Kf\sigma(z) = \int_{\mathbb{R}} \frac{f(t)}{t-z}\, d\sigma(t).$$

This gives an isometry of $L^2(\sigma)$ onto $K_\theta$. In the case of meromorphic $\theta(z)$, every function $f\in K_\theta$ also has a meromorphic extension in $\mathbb{C}$, which is given by the formula (3.3). The corresponding Clark measure is discrete with masses at the points of $\{z:\theta(z)=1\}$ given by

$$\sigma(\{x\}) = \frac{2\pi}{|\theta'(x)|}.$$

For more details on Clark measures the reader may consult [29] or the references therein.

Each meromorphic inner function $\theta(z)$ can be written as $\theta(t)=e^{i\phi(t)}$ on $\mathbb{R}$, where $\phi(t)$ is a real-analytic and strictly increasing function. The function $\phi(t)=\arg\theta(t)$ is the continuous argument of $\theta(z)$.

Recall that the Toeplitz operator $T_U$ with a symbol $U\in L^\infty(\mathbb{R})$ is the map

$$\begin{aligned} T_U\colon H^2 &\longrightarrow H^2, \\ F &\longmapsto P_+(UF), \end{aligned}$$

where $P_+$ is the orthogonal projection in $L^2(\mathbb{R})$ onto the Hardy space $H^2=H^2(\mathbb{C}_+)$.

We will use the following notation for kernels of Toeplitz operators (or *Toeplitz kernels* in $H^2$):

$$N[U]=\ker T_U.$$

For example, $N[\bar\theta]=K_\theta$ if $\theta$ is an inner function. Along with $H^2$-kernels, one may consider Toeplitz kernels $N^p[U]$ in other Hardy classes $H^p$, the kernel $N^{1,\infty}[U]$ in the "weak" space $H^{1,\infty}=H^p\cap L^{1,\infty}$, $0<p<1$, or the kernel in the Smirnov class $\mathcal{N}^+(\mathbb{C}_+)$:

$$N^+[U]=\{f\in\mathcal{N}^+\cap L^1_{\mathrm{loc}}(\mathbb{R}):\overline{U}\bar f\in\mathcal{N}^+\}.$$

If $\theta$ is a meromorphic inner function, $K^+_\theta=N^+[\bar\theta]$ can also be considered. For more on such kernels see [20] and [21].

For any inner function $\theta$ in the upper half-plane we denote by $\mathrm{spec}_\theta$ the closure of the set $\{z:\theta(z)=1\}$, the set of points on the line where the non-tangential limit of $\theta$ is equal to 1, plus the infinite point if the corresponding Clark measure has a point mass at infinity, i.e. if $p$ in (3.1) is positive. If $\mathrm{spec}_\theta\subset\mathbb{R}$, as in the next definition, then $p$ in (3.1) is 0. Throughout the paper, $S$ stands for the exponential inner function $S(z)=e^{iz}$.

We call a sequence of real points *discrete* if it has no finite accumulation points. Note that $\{z:\theta(z)=1\}$ is discrete if and only if $\theta$ is meromorphic.

*Definition* 3.1. If $X\subset\mathbb{R}$ is a closed set, we define

$$T_X=\sup\{a:N[\bar\theta S^a]\neq 0 \text{ for some meromorphic inner } \theta \text{ with } \mathrm{spec}_\theta\subset X\}.$$

The following theorem (see [22, §2.1]) shows the connection between the gap problem and the problem of triviality of Toeplitz kernels. Such connections will be used throughout the paper.

THEOREM 3.2. ([22]) $G_X=T_X$.

The gap problem is closely related to problems of uniform approximation of continuous functions by trigonometric polynomials, see [6] for a more detailed discussion and further references. To give a simple example of such a connection we consider the following version of the problem.

Let again $X$ be a closed subset of the line. Denote by $C_0(X)$ the space of all continuous functions on $X$ tending to 0 at infinity, with the usual sup-norm. It is not possible to discuss approximation by trigonometric polynomials in this particular space directly, since finite linear combinations of exponential functions do not belong to $C_0(X)$. The standard solution is to consider "generalized" linear combinations of exponentials, i.e. the Paley–Wiener space

$$\mathrm{PW}_a=\{\hat f:f\in L^2([-a,a])\}.$$

Let us define

$$A_X = \inf\{a>0 : \mathrm{PW}_a \text{ is dense in } C_0(X)\}$$

or $A_X=\infty$ if the set is empty. The following statement is a product of the standard duality argument.

Proposition 3.3. $G_X = 2A_X$.

Together with Theorem 4.7 this statement gives a formula for $A_X$.

## 4. The main theorem

Before stating our main result we need to give several definitions.

Let $\Lambda=\{\lambda_1, ..., \lambda_n\}$ be a finite set of points on $\mathbb{R}$. Consider the quantity

$$E(\Lambda) = \sum_{\substack{\lambda_k,\lambda_l\in\Lambda \\ \lambda_k\neq\lambda_l}} \log|\lambda_k-\lambda_l|. \tag{4.1}$$

### Physical interpretation

According to the 2-dimensional Coulomb law, $E(\Lambda)$ is the energy of a system of "flat" electrons placed at the points of $\Lambda$. The 2-dimensional Coulomb-gas formalism corresponds to the planar potential theory with logarithmic potential and assumes the potential energy at infinity to be equal to $-\infty$, see for example [9], [23] and [30].

Physically, the 2-dimensional Coulomb law can be derived from the standard 3-dimensional law via a method of "reduction". According to this method, one replaces each electron in the plane with a uniformly charged string orthogonal to the plane. After that one applies the 3-dimensional law and a renormalization procedure.

### Key example

*Let $I\subset\mathbb{R}$ be an interval, $C>0$ and let $\Lambda$ be a set of $k$ points uniformly distributed on $I$:*

$$\Lambda = I\cap C\mathbb{Z} = \{(n+1)C, (n+2)C, ..., (n+k)C\}.$$

*Then*

$$E(\Lambda) = \sum_{m=1}^{k} \log[C^{k-1}(m-1)!(k-m)!] = k^2\log|I| + O(|I|^2) \tag{4.2}$$

*as follows from Stirling's formula. Here $|I|$ stands for the length of $I$ and the notation $O(|I|^2)$ corresponds to the direction $|I|\to\infty$ (with $C$ remaining fixed).*

*Remark* 4.1. The uniform distribution of points on the interval does not maximize the energy $E(\Lambda)$ but comes within $O(|I|^2)$ from the maximum, which is negligible for our purposes, see the main definition and its discussion below. It is interesting to observe that the maximal energy for $k$ points is achieved when the points are placed at the endpoints of $I$ and at the zeros of the Jacobi $(1,1)$-polynomial of degree $k-2$, see for example [14].

We call a sequence of disjoint intervals $\{I_n\}_n$ on the real line *long* (in the sense of Beurling and Malliavin) if

$$\sum_n \frac{|I_n|^2}{1+\operatorname{dist}^2(0,I_n)}=\infty. \tag{4.3}$$

If the sum is finite, we call $\{I_n\}_n$ *short*.

Let

$$\ldots<a_{-2}<a_{-1}<a_0=0<a_1<a_2<\ldots$$

be a two-sided sequence of real points. We say that the intervals $I_n=(a_n,a_{n+1}]$ form a *short partition* of $\mathbb{R}$ if $|I_n|\to\infty$ as $n\to\pm\infty$ and the sequence $\{I_n\}_n$ is short.

**Main definition**

Let $\Lambda=\{\lambda_n\}_n$ be a sequence of distinct real points. We write $\mathcal{C}_\Lambda\geqslant a$ if there exists a short partition $\{I_n\}_n$ such that

$$\Delta_n\geqslant a|I_n| \quad \text{for all } n \quad \text{(density condition)} \tag{4.4}$$

and

$$\sum_n \frac{\Delta_n^2\log|I_n|-E_n}{1+\operatorname{dist}^2(0,I_n)}<\infty \quad \text{(energy condition)}, \tag{4.5}$$

where

$$\Delta_n=\#(\Lambda\cap I_n)$$

and

$$E_n=E(\Lambda\cap I_n)=\sum_{\substack{\lambda_k,\lambda_l\in I_n\\ \lambda_k\neq\lambda_l}}\log|\lambda_k-\lambda_l|.$$

If $X$ is a closed subset of $\mathbb{R}$, we put

$$\mathcal{C}_X=\sup\{a: \text{ there is a sequence } \Lambda\subset X \text{ such that } \mathcal{C}_\Lambda\geqslant a\}.$$

*Remark* 4.2. Notice that the series in the energy condition is positive. Indeed, every term in the sum defining $E_n$ is at most $\log|I_n|$ and there are less than $\Delta_n^2$ terms.

The example before the definition shows that the numerator in the energy condition is (up to lower order terms) the difference between the energy of the optimal configuration, when the points are spread uniformly on $I_n$, and the energy of $\Lambda\cap I_n$.

Thus the energy condition is a requirement that the placement of the points of $\Lambda$ is close to uniform, in the sense that the work needed to spread the points of $\Lambda$ uniformly on each interval is summable with respect to the Poisson weight.

*Remark* 4.3. The inequality $\log|I_n|-\log|\lambda_k-\lambda_l|>0$, which holds for any $\lambda_k,\lambda_l\in I_n$, also implies that if a sequence $\Lambda$ satisfies the energy condition (4.5), then any subsequence of $\Lambda$ also satisfies (4.5) on $\{I_n\}_n$. A deletion of points from $\Lambda$ will eliminate some positive terms from the numerator in (4.5) which can only make the sum smaller.

*Remark* 4.4. We say that a partition $\{I_n\}_n$ is *monotone* if $|I_n|\leqslant|I_{n+1}|$ for $n\geqslant 0$ and $|I_{n+1}|\leqslant|I_n|$ for $n<0$. Corollary 7.7 shows that in the above definition the words "short partition" can be replaced by "short monotone partition". Since monotone partitions are easier to work with, this modified definition will be used in the proof of Theorem 4.7.

*Remark* 4.5. The requirement that the partition $I_n=(a_n,a_{n+1}]$ satisfied $|I_n|\to\infty$ is not essential and can be omitted if one slightly changes the definitions of $\Delta_n$ and $E_n$ in (4.5). One could, for instance, use

$$\Delta_n=\#(\Lambda\cap(a_n-1,a_{n+1}+1])$$

and

$$E_n=\sum_{\substack{\lambda_k,\lambda_l\in(a_n-1,a_{n+1}+1]\\ \lambda_k\neq\lambda_l}}\log|\lambda_k-\lambda_l|.$$

*Remark* 4.6. The density condition says simply that the lower (interior) density of the sequence in the sense of Beurling and Malliavin, is at least $a$. Such a density can be defined in several different ways:

- If $\Lambda$ is a real sequence define $d_1(\Lambda)$ to be the supremum of all $a$ such that there exists a short monotone partition $\{I_n\}_n$ satisfying (4.4).
- Denote by $d_2(\Lambda)$ the supremum of all $a$ such that there exists a short (not necessarily monotone) partition $\{I_n\}_n$ satisfying (4.4).
- Define $d_3(\Lambda)$ to be the supremum of all $a$ such that there exists a subsequence of $\Lambda$ whose counting function $n(x)$ satisfies

$$\int_{\mathbb{R}}\frac{|n(x)-ax|}{1+x^2}\,dx<\infty.$$

This definition was used in [7].

• Finally, define $d_4(\Lambda)$ to be the infimum of all $a$ such that there exists a long sequence of disjoint intervals $I_n$ satisfying

$$\#(\Lambda\cap I_n) < a|I_n|.$$

This definition was used in [22].

One can easily show that all these definitions are equivalent, i.e.

$$d_1(\Lambda) = d_2(\Lambda) = d_3(\Lambda) = d_4(\Lambda).$$

As was mentioned above, such a density $d$ was introduced in [5], where it was called *interior* density. A closely related notion of *exterior* density appears in the Beurling–Malliavin theorem on the completeness of exponential functions in $L^2$ on an interval, see [5] or [22]. We give a definition of exterior density in §7.

Now we are ready to return to the gap problem and use our newly defined metric characteristic $\mathcal{C}_X$ of a closed set $X\subset\mathbb{R}$. Our main result is the following theorem.

THEOREM 4.7. *$G_X = 2\pi\mathcal{C}_X$.*

It will be proved in §6.

## 5. Examples and applications

In this section we discuss examples related to Theorem 4.7, including some of its relations with existing results.

*Example* 5.1. As discussed above, if the points of the sequence are spread uniformly over the interval, then

$$E_n = \sum_{\lambda_k,\lambda_l\in I_n} \log|\lambda_k-\lambda_l|$$

is roughly (up to $O(|I_n|^2)$, which is small for short sequences $\{I_n\}_n$) equal to $\Delta_n^2\log|I_n|$ as follows from Stirling's formula. This happens for instance when the sequence $\Lambda$ is separated, i.e. satisfies $|\lambda_n-\lambda_{n+1}|>\delta>0$ for all $n$. Thus for separated sequences $\Lambda$ the energy condition disappears and

$$G_\Lambda = 2\pi d_j(\Lambda),$$

where $d_j$, $j=1,2,3,4$, is any of the equivalent densities defined in the last remark, i.e. the interior density of $\Lambda$. This is one of the results of [22].

For example, as follows from Proposition 2.1, if the support of a measure $\mu$ contains a separated sequence of interior density $D$, then for any $\varepsilon>0$ there exists $f\in L^1(|\mu|)$ such that $\widehat{f\mu}=0$ on $[0,2\pi D-\varepsilon]$.

*Example* 5.2. Let $\Lambda$ be a real sequence such that the density condition (4.4) holds for some $a>0$ and some partition $\{I_n\}_n$ satisfying the stronger shortness condition

$$\sum_n \frac{|I_n|^2 \log |I_n|}{1+\operatorname{dist}^2(0, I_n)} < \infty.$$

Then we will automatically have that

$$\sum_n \frac{\Delta_n^2 \log |I_n| - \sum_{\lambda_k, \lambda_l \in I_n} \log_+ |\lambda_k - \lambda_l|}{1+\operatorname{dist}^2(0, I_n)} < \infty.$$

Accordingly, condition (4.5) will be significantly simplified and one will only need to check that

$$\sum_{\substack{\lambda_k, \lambda_l \in \Lambda \\ \lambda_k \neq \lambda_l}} \frac{\log_- |\lambda_k - \lambda_l|}{1+\lambda_k^2} < \infty$$

to conclude that $G_\Lambda = 2\pi \mathcal{C}_\Lambda \geqslant 2\pi a$.

Consider, for instance, the log-short partition

$$I_0 = (-1, 1], \quad I_n = (n^\alpha, (n+1)^\alpha], \quad I_{-n} = (-(n+1)^\alpha, -n^\alpha], \quad n = 1, 2, ...,$$

for some $\alpha > 1$. Let an increasing discrete sequence $\Lambda = \{\lambda_n\}_n$ be such that

$$\alpha |n|^{\alpha-1} \leqslant \#(\Lambda \cap I_n) \leqslant \alpha |n|^{\alpha-1} + 1 \tag{5.1}$$

for all $n$ and

$$\lambda_{k+1} - \lambda_k \geqslant \operatorname{const} e^{-|k|/\log^2 |k|} \tag{5.2}$$

for all $k$, $|k|>1$. Then by the previous discussion $G_\Lambda = 2\pi$.

Similarly to the last example, if $\mu$ is a finite measure whose support contains $\Lambda$, then for any $\varepsilon > 0$ there exists $f \in L^1(|\mu|)$ such that $\widehat{f\mu} = 0$ on $[0, 2\pi - \varepsilon]$.

On the other hand, if (5.1) holds, but instead of (5.2) we have that on each $J_n$,

$$\lambda_{k+1} - \lambda_k \leqslant \operatorname{const} e^{-|k|/\log |k|}$$

for any $\lambda_k, \lambda_{k+1} \in I_n$, $|k|>1$, then the interior density of $\Lambda$ is still 1, but the energy condition is not satisfied by any subsequence of $\Lambda$ of positive interior density on any short partition. Thus $G_\Lambda = 0$.

As was mentioned in the introduction, one of the classical results on the gap problem is the following theorem by Beurling.

Theorem 5.3. ([3], [15]) *Let $\mu$ be a finite complex measure on $\mathbb{R}$ such that the complement of its support contains a long sequence of intervals* (*in the sense of* (4.3)). *Suppose that $\hat{\mu}=0$ on an interval of positive length. Then $\mu\equiv 0$.*

To obtain this statement from Theorem 4.7, notice that if $\mu$ is a measure as in the statement, then for any short partition of $\mathbb{R}$ one of the intervals of the partition will be contained in a gap of $X=\operatorname{supp}\mu$. Hence $X$ does not contain a sequence $\Lambda$ that satisfies the density condition (4.4) on a short partition with $a>0$. Therefore $G_X=0$.

To see another application of Theorem 4.7 in the "positive" direction, let us formulate the following result proved by Benedicks in [2]. This statement provides one of the very few non-trivial examples of sets with positive gap characteristic that exist in the literature.

Theorem 5.4. ([2]) *Let $\ldots<a_{-1}<a_0<a_1<a_2<\ldots$ be a discrete sequence of points and let $I_n=(a_n,a_{n+1}]$ be the corresponding partition of $\mathbb{R}$. Suppose that there exist positive constants $C_1$, $C_2$ and $C_3$ such that*

(1) *if $C_1^{-1}a_{2n+1}<a_{2k+1}<C_1a_{2n+1}$, then*

$$C_2^{-1}|I_{2n+1}|<|I_{2k+1}|<C_2|I_{2n+1}|;$$

(2) *for all $n$,*

$$C_1^{-1}|a_{2n+1}|<|a_{2n-1}|<C_1|a_{2n+1}|;$$

(3) *for all $n$,*

$$|I_{2n+1}|>C_3\max\{|I_{2n}|,1\};$$

(4)

$$\sum_n \frac{|I_{2n+1}|^2}{1+a_{2n+1}^2}\left(\log_+\frac{|I_{2n+1}|}{|I_{2n}|}+1\right)<\infty.$$

*Then for any real number $A>0$ and $1\leqslant p<\infty$ there exists a non-zero function*

$$f\in L^1(\mathbb{R})\cap L^p(\mathbb{R})\cap C^\infty(\mathbb{R}),\quad \operatorname{supp} f\subset\bigcup_n I_{2n},$$

*such that $\hat{f}=0$ on $[0,A]$.*

One can show that the part pertaining to the gap problem, i.e. the existence of $f\in L^1\left(\bigcup_n I_{2n}\right)$ with a spectral gap of size $A$, follows from Theorem 4.7.

For simplicity, let us consider the case when $|I_n|\to\infty$. In this case conditions (1) and (2) of the theorem prove to be redundant. Indeed, put $X=\operatorname{clos}\left(\bigcup_n I_{2n}\right)$ and let $\mu$ be the restriction of the Poisson measure $dx/(1+x^2)$ to $X$. By Proposition 2.1, we need to show that $G_\mu>A$.

Let

$$J_n=(a_{2n},a_{2n+2}]=I_{2n}\cup I_{2n+1}.$$

Then, by conditions (3) and (4) of the theorem, $J_n$ is a short partition.

Consider the sequence $\Lambda\subset\bigcup_n I_{2n}$ such that

$$2A|J_n|\leqslant\#(\Lambda\cap I_{2n})\leqslant 2A|J_n|+1$$

for each $n$ and the points of $\Lambda$ are spread uniformly on each $I_{2n}$. Then $\Lambda$ satisfies the density condition (4.4) on $\{J_n\}_n$ with $a=2A$. As for the energy condition, the numerator in (4.5) is at most

$$\mathrm{const}(|J_n|^2\log_+|J_n|-|J_n|^2\log_+|I_{2n}|)+O(|J_n|^2)$$

by (4.2), and the energy condition follows immediately from conditions (3) and (4) of the theorem. Hence $G_\mu=G_X\geqslant 2A$.

Similar examples with $1\leqslant p\leqslant\infty$ are discussed in [28].

## 6. The proof of the main theorem

Before starting the proof, let us introduce the following notation. If $f$ is a function on $\mathbb{R}$ and $I\subset\mathbb{R}$ we denote by $f|_I$ the function that is equal to $f$ on $I$ and to 0 on $\mathbb{R}\setminus I$.

In our estimates we write $a(n)\lesssim b(n)$ if $a(n)<Cb(n)$ for some positive constant $C$, not depending on $n$, and large enough $|n|$. We write $a(n)\asymp b(n)$ if $ca(n)<b(n)<Ca(n)$ for some $C>c>0$. Some formulas will have other parameters in place of $n$ or no parameters at all. For instance, $\asymp$ may be put between two improper integrals to indicate that they either both converge or both diverge.

By $\Pi$ we denote the Poisson measure $dx/(1+x^2)$ on the real line. In particular, $L^p_\Pi=L^p(\mathbb{R},dx/(1+x^2))$.

We will denote by $\mathcal{D}(\mathbb{R})$ the standard Dirichlet space on $\mathbb{R}$ (in $\mathbb{C}_+$). Recall that the Hilbert space $\mathcal{D}=\mathcal{D}(\mathbb{R})$ consists of functions $h\in L^1_\Pi$ such that the harmonic extension $u=u(z)$ of $h$ to $\mathbb{C}_+$ has a finite gradient norm,

$$\|h\|^2_{\mathcal{D}}\equiv\|u\|^2_\nabla\stackrel{\text{def}}{=}\int_{\mathbb{C}_+}|\nabla u|^2\,dA<\infty,$$

where $dA$ is the area measure. If $h\in\mathcal{D}(\mathbb{R})$ is a smooth function, then we also have

$$\|h\|^2_{\mathcal{D}}=\int_{\mathbb{R}}\bar h\tilde h'\,dx,$$

where $\tilde h$ denotes a harmonic conjugate function.

**6.1. Proof of the main theorem, part I**

First suppose that $\mathcal{C}_X > 1/2\pi$. We will show that $G_X \geqslant 1$.

Choose $\varepsilon > 0$. As $\mathcal{C}_X > 1/2\pi$, there exists a sequence $\Lambda = \{\lambda_n\}_n \subset X$, $\mathcal{C}_\Lambda > 1/2\pi$. Let

$$I_n = (a_n, a_{n+1}]$$

be the corresponding short monotone partition, see Remark 4.4. Without loss of generality, we may assume that

$$\frac{1}{2\pi}|I_n| < \#(\Lambda \cap I_n) \leqslant \frac{1}{2\pi}|I_n| + 1$$

(otherwise just delete some of the points from $\Lambda$, see Remark 4.3). We will assume that $|I_n| \gg 1/\varepsilon \gg 1$ for all $n$.

By Lemma 7.1 and Corollary 7.5, we may assume that the lengths of the intervals $(\lambda_n, \lambda_{n+1})$ are bounded from above. It will be convenient for us to assume that the endpoints of $I_n$ belong to $\Lambda$, i.e. that $I_n = (\lambda_{k_n}, \lambda_{k_{n+1}}]$ for some $\lambda_{k_n}, \lambda_{k_{n+1}} \in \Lambda$. We will also include the endpoints of the intervals into the energy condition, by defining $E_n$ as

$$E_n = \sum_{\substack{\lambda_{k_n} \leqslant \lambda_k, \lambda_l \leqslant \lambda_{k_{n+1}} \\ \lambda_k \neq \lambda_l}} \log|\lambda_k - \lambda_l|, \tag{6.1}$$

and assuming that (4.5) is satisfied with these $E_n$. Such an assumption can be made because if the sum in (4.5) becomes infinite with $E_n$ defined by (6.1), one can, for instance, delete the first point $\lambda_{n_k+1}$ from $\Lambda$ on all $I_n$ for large $n$. After the addition of $\lambda_{k_n}$ and deletion of $\lambda_{n_k+1}$ in the sum defining $E_n$, each term in (4.5) will become smaller and the sum will remain finite. At the same time, since

$$|I_n| \asymp \#(\Lambda \cap I_n) \to \infty,$$

the subsequence will still have more than $|I_n|$ points on each $I_n$ and will satisfy the density condition.

Our goal is to show that $G_\Lambda \geqslant 1$ by producing a measure on $\Lambda$ with spectral gap of size arbitrarily close to 1. Due to connections discussed in §3, the existence of such a measure will follow from the non-triviality of a certain Toeplitz kernel.

Since the lengths of $(\lambda_n, \lambda_{n+1})$ are bounded from above, we can apply Lemma 7.8. Denote by $\theta$ the corresponding meromorphic inner function with $\operatorname{spec}_\theta = \Lambda$.

Let $u = \arg(\theta \bar{S}) = \arg\theta - x$. First, we choose a larger partition $J_n = (b_n, b_{n+1})$ and a small "correction" function $v$ such that $u - v$ becomes an atom on each $J_n$.

Claim 6.1. *There exists a subsequence* $\{b_n\}_n$ *of the sequence* $\{a_n\}_n$ *and smooth functions* $v_1$ *and* $v_2$ *such that*

(1) $|v_1'|<\frac{1}{2}\varepsilon$ *and* $u-v_1=0$ *at all* $a_n$;

(2) $J_n=(b_n,b_{n+1})$ *is a short monotone partition*;

(3) $|v_2'|<\frac{1}{2}\varepsilon$ *and* $u-v=u-(v_1+v_2)=0$ *at all* $b_n$;

(4) $\int_{J_n}(u-v)\,dx=0$ *for all* $n$;

(5) $\tilde{u}-\tilde{v}\in L^1_\Pi$.

*Proof.* First, choose a smooth function $v_1$ satisfying (1). Such a function exists because

$$\big|2\pi\Delta_n-|I_n|\big|\leqslant 2\pi\ll\tfrac{1}{2}\varepsilon|I_n|.$$

Notice that because the sequence $I_n$ is short and

$$(u-v_1)'>-1-\tfrac{1}{2}\varepsilon,$$

condition (1) implies that

$$u-v_1\in L^1_\Pi. \tag{6.2}$$

Choose $b_0=a_0=0$. Choose $b_1=a_{n_1}>b_0$ to be the smallest element of $\{a_k\}_k$ satisfying

$$\left|\int_{b_0}^{a_{n_1}}(u-v_1)\,dx\right|<\frac{\varepsilon}{8}(a_{n_1}-b_0)^2.$$

Notice that, because of (6.2), such an $a_{n_1}$ will always exist. After that proceed choosing $b_2,b_3,\dots$ in the following way: If $b_j$ is chosen, choose $b_{j+1}=a_{n_{j+1}}$ to be the smallest element of $\{a_k\}_k$ satisfying $a_{n_{j+1}}>b_j$,

$$\left|\int_{b_j}^{a_{n_{j+1}}}(u-v_1)\,dx\right|<\frac{\varepsilon}{8}(a_{n_{j+1}}-b_j)^2 \tag{6.3}$$

and

$$a_{n_{j+1}}-b_j\geqslant b_j-b_{j-1}.$$

Choose $b_k$, $k<0$, in the same way.

We claim that the resulting sequence $J_k=(b_{k-1},b_k)$ forms a short monotone partition.

Let $k$ be positive. By our construction, $I_{n_k}$ is the last (rightmost) among the intervals $I_n$ contained in $J_k$. Notice that, because of monotonicity, $I_{n_k}$ is the largest interval among the intervals $I_n$ contained in $J_k$. We will show that for each $k$,

$$|J_k|<\left(\left\lfloor\frac{10}{\varepsilon}\right\rfloor+1\right)|I_{n_k}|, \tag{6.4}$$

where $\lfloor\cdot\rfloor$ stands for the integer part (floor function) of a real number.

This can be proved by induction. The basic step: By our construction $b_1=a_{n_1}$ and

$$\left|\int_{b_0}^{a_{n_1-1}}(u-v_1)\,dx\right|\geqslant\frac{\varepsilon}{8}(a_{n_1-1}-b_0)^2.$$

Since $(u-v_1)'>-1-\varepsilon$ and $u-v_1=0$ at all $a_n$, $|u-v_1|\leqslant(1+\varepsilon)|I_{n_1-1}|$ on $(b_0,a_{n_1-1})$. Hence

$$(1+\varepsilon)|I_{n_1-1}|(a_{n_1-1}-b_0)\geqslant\left|\int_{b_0}^{a_{n_1-1}}(u-v_1)\,dx\right|\geqslant\frac{\varepsilon}{8}(a_{n_1-1}-b_0)^2$$

and

$$a_{n_1-1}-b_0\leqslant 8\frac{1+\varepsilon}{\varepsilon}|I_{n_1-1}|.$$

It follows that

$$|J_1|=(a_{n_1-1}-b_0)+|I_{n_1}|\leqslant\frac{9}{\varepsilon}|I_{n_1-1}|+|I_{n_1-1}|\leqslant\frac{10}{\varepsilon}|I_{n_1-1}| \tag{6.5}$$

(if $\varepsilon$ is small enough). For the inductional step, assume that (6.4) holds for $k=l-1$. For $J_l=(b_{l-1},b_l)$, $b_l=a_{n_l}$, there are two possibilities:

$$\left|\int_{b_{l-1}}^{a_{n_l-1}}(u-v_1)\,dx\right|\geqslant\frac{\varepsilon}{8}(a_{n_l-1}-b_{l-1})^2$$

or

$$a_{n_l-1}-b_{l-1}<b_{l-1}-b_{l-2}.$$

In the first case we prove (6.5) in the same way as in the basic step. In the second case we note that, by the monotonicity of $\{I_n\}_n$, the number of intervals $I_n$ inside $(b_{l-1},a_{n_l-1})$ is at most $(a_{n_l-1}-b_{l-1})/|I_{n_{l-1}}|$, which is strictly less than $|J_{l-1}|/|I_{n_{l-1}}|\leqslant\lfloor 10/\varepsilon\rfloor+1$. Accordingly the number of intervals in $(b_{l-1},a_{n_l-1})$ is at most $\lfloor 10/\varepsilon\rfloor$. Therefore the number of intervals in $J_l=(b_{l-1},b_l)$ is at most $\lfloor 10/\varepsilon\rfloor+1$. Now, since $I_{n_l}$ is the largest interval in $J_l$, we again get (6.4), which implies the shortness of $\{J_n\}_n$. The monotonicity follows from our construction.

Now define the function $v_2$ on each $J_k$ in the following way. First consider the tent function $T_k$ defined on $\mathbb{R}$ as

$$T_k(x)=\tfrac{1}{4}\varepsilon\operatorname{dist}(x,\mathbb{R}\setminus J_k).$$

Notice that, because of (6.3), for each $k$ there exists a constant $C_k$, $|C_k|\leqslant 1$, such that

$$\int_{J_k}[(u-v_1)-C_kT_k]\,dx=0.$$

Now define $v_2$ as a smoothed-out sum $\sum_k C_k T_k$ that satisfies $|v_2'|<\frac{1}{2}\varepsilon$ and still has the properties that $v_2(b_k)=0$ and

$$\int_{J_k} [(u-v_1)-v_2]\,dx=0$$

for each $k$. Finally, let $v=v_1+v_2$. The last condition of the claim will be satisfied because the restrictions $(u-v)|_{J_k}$ form a collection of atoms with a finite sum of $L^1_\Pi$-norms:

$$\|(u-v)|_{J_k}\|_{L^1_\Pi} \lesssim \frac{|J_k|^2}{1+\operatorname{dist}^2(0,J_k)}$$

(for more on atomic decompositions see [8]). □

The function $v$ from the last claim is a smooth function satisfying $|v'|\leqslant\varepsilon$. Therefore it can be represented as $v=v_+-v_-$, where $v_\pm$ are smooth growing functions, $\varepsilon\leqslant v_\pm'\leqslant 2\varepsilon$. Hence one can choose two meromorphic inner functions $I_\pm$ satisfying

$$\{x:\arg I_\pm(x)=k\pi\}=\{x:v_\pm(x)=k\pi\}$$

and

$$|I_\pm'|\lesssim\varepsilon.$$

The existence of such $I_\pm$ follows from Lemma 7.8 below.

Note that then, automatically, $|\arg(\bar{I}_+I_-)-v|<2\pi$. The function $\arg(\theta\bar{S}I_+\bar{I}_-)$, as well as its harmonic conjugate, still belongs to $L^1_\Pi$.

Without loss of generality, we may assume that $\arg(\theta\bar{S}I_+\bar{I}_-)=0$ at 0.

Claim 6.2. *The function*

$$\frac{\arg(\theta(x)\bar{S}(x)I_+(x)\bar{I}_-(x))}{x}$$

*belongs to the Dirichlet class* $\mathcal{D}(\mathbb{R})$.

*Proof.* We will actually prove that the function $w/x$, $w=\arg\theta-x-v$, belongs to $\mathcal{D}(\mathbb{R})$ instead (again, without loss of generality, we may assume that $w(0)=0$ with large multiplicity). The function

$$\frac{v-(\arg I_+-\arg I_-)}{x}$$

is a bounded function with bounded derivative which obviously belongs to $\mathcal{D}(\mathbb{R})$.

Let $q$ be the harmonic extension of $w/x$ in the upper half-plane. We need to show that the gradient norm of $q+i\tilde{q}$ in $\mathbb{C}_+$ is finite, i.e. that

$$\|q+i\tilde{q}\|^2_\nabla=\lim_{r\to\infty}\int_{\partial D(r)} q\,d\tilde{q}=-\lim_{r\to\infty}\int_{\partial D(r)}\tilde{q}\,dq<\infty,$$

where $D(r)$ is the semidisk $\{z\in\mathbb{C}_+:|z|<r\}$.

We first prove that the integrals over $\partial D(r)\cap\mathbb{R}$ are uniformly bounded from above, i.e. that

$$-\int_{\mathbb{R}}\tilde{q}\,dq<\infty.$$

First, notice that the harmonic conjugate of $w/x=q$ is $\widetilde{w}/x=\tilde{q}$ (we may assume that $\widetilde{w}(0)=0$) and $(w/x)'=w'/x-w/x^2$. Recall that, by our construction (see Claim 6.1), $w'$ is bounded from below and is zero at the endpoints of every $J_n$. Hence $|w|\lesssim|J_n|$ on every $J_n$. Since the partition $\{J_n\}_n$ is short, it follows that $|w(x)|=o(|x|)$ and that $w/x^2$ is a bounded function. Therefore,

$$-\int_{\mathbb{R}}\tilde{q}\,dq\asymp-\int_{\mathbb{R}}w'\widetilde{w}\,\frac{dx}{x^2}$$

and we can estimate the last integral instead.

If $I$ is an interval, then $2I$ denotes the interval with the same center as $I$ satisfying $|2I|=2|I|$.

Put $w_n=w|_{J_n}$. Then

$$\int_{\mathbb{R}}w'\widetilde{w}\,\frac{dx}{x^2}=\sum_n\sum_k\int_{J_n}w'\widetilde{w}_k\,\frac{dx}{x^2}. \tag{6.6}$$

To estimate the last integral, let us first consider the case when the intervals $J_n$ and $J_k$ are far from each other:

$$\max\{|J_n|,|J_k|\}\leqslant\operatorname{dist}(J_n,J_k).$$

In this case

$$\left|\int_{J_n}w'\widetilde{w}_k\,\frac{dx}{x^2}\right|\lesssim\int_{J_n}|w'|\frac{|J_k|^3}{\operatorname{dist}^2(J_k,x)}\,\frac{dx}{x^2}\lesssim\frac{|J_k|^3}{1+\operatorname{dist}^2(J_n,0)}\int_{J_n}\frac{dx}{\operatorname{dist}^2(J_k,x)}. \tag{6.7}$$

Here we used the property that each $w_k$ is an atom supported on $J_k$ whose $L^1$-norm is $\lesssim|J_k|^2$ and employed the standard estimates from the theory of atomic decompositions, see [8]. In the last inequality we used the property

$$\int_{J_n}|w'(x)|\,dx\lesssim|J_n|. \tag{6.8}$$

Now let us consider the "mid-range" case when

$$\min\{|J_n|,|J_k|\}\leqslant\operatorname{dist}(J_n,J_k)<\max\{|J_n|,|J_k|\}.$$

Assume that $0<k<n$ (other cases are analogous). Then, by monotonicity $|J_k|\leqslant|J_n|$ and

$$\begin{aligned}\left|\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2}\right| &\lesssim \frac{1}{1+\operatorname{dist}^2(J_k,0)}\int_{J_n}|w'|\frac{|J_k|^3}{\operatorname{dist}^2(J_k,x)}\,dx \\ &\leqslant \frac{|J_k|}{1+\operatorname{dist}^2(J_k,0)}\,\frac{|J_k|^2}{\operatorname{dist}^2(J_k,J_n)}\int_{J_n}|w'|\,dx \lesssim \frac{|J_k|\,|J_n|}{1+\operatorname{dist}^2(J_k,0)}.\end{aligned}\tag{6.9}$$

Finally, the last case is

$$\operatorname{dist}(J_n,J_k)<\min\{|J_n|,|J_k|\}.\tag{6.10}$$

Again we assume that $n\geqslant k>0$. Then, by monotonicity, either $n=k$ or $|n-k|=1$, i.e. the intervals are either the same or adjacent. The estimates in this case are more complicated and will be done differently. First, integrating by parts we get

$$-\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2}=\int_{J_n} w'\left(\int_{J_k}\frac{w(t)\,dt}{t-x}\right)\frac{dx}{x^2}=-\int_{J_n} w'\left(\int_{J_k}\log|t-x|w'(t)\,dt\right)\frac{dx}{x^2}.$$

By the first inequality of (7.5) in Lemma 7.10, applied to $h=w'$, $f=(\arg\theta)'$ and $g=(x+v)'$, for $1\ll k\leqslant n$ we have

$$\begin{aligned}&-\int_{J_n} w'\left(\int_{J_k}\log|t-x|w'(t)\,dt\right)\frac{dx}{x^2}\\ &\qquad\lesssim -\frac{1}{1+\operatorname{dist}^2(J_n,0)}\left(\iint_{J_n\times J_k}\log|t-x|w'(x)w'(t)\,dx\,dt+C|J_n|^2\right)\end{aligned}\tag{6.11}$$

and we can work with the latter integral instead of the former. To verify the conditions of Lemma 7.10, note that if $E=J_n$ and $I=J_k$, then, for large enough $k$ and $n$, we will have $E\cup I\subset[d,2d]$ as a consequence of the shortness condition and (6.10). The relation (7.4) will be satisfied because $w=0$ at the endpoints of $J_k$, see condition (3) of Claim 6.1. The constant $D_1$ satisfies

$$D_1\leqslant\int_{J_n}(\arg\theta)'\,dx+\int_{J_n}(x+v)'\,dx\lesssim|J_n|.$$

Finally,

$$\frac{D_2}{d^2}\lesssim\left\|\int_{J_k}\log|t-x|w'(t)\,dt\right\|_{L^1_\Pi}=\|\widetilde{w}_k\|_{L^1_\Pi}\lesssim\|w_k\|_{L^1_\Pi}\lesssim\frac{|J_k|^2}{d^2},$$

because $w_k$ is an atom.

To estimate the integral in the right-hand side of (6.11), let

$$p=\arg\theta-x-v_1=w+v_2,$$

where the functions $v_1$ and $v_2$ are from Claim 6.1. Also let $p_n=p|_{J_n}$ and $v_2^n=v_2|_{J_n}$. The key properties of $v_1$ that we will use are that $\arg\theta-x-v_1=0$ at the endpoints of all $I_n$, $v_2=0$ at the endpoints of $J_n$ and $|v_1'|,|v_2'|<\varepsilon$. Then

$$-\iint_{J_n\times J_k}\log|t-x|w'(x)w'(t)\,dx\,dt=-\iint_{J_n\times J_k}\log|t-x|p'(x)p'(t)\,dx\,dt \\ -\int_{J_k}(\tilde p_n v_2'+\tilde v_2^n p'+\tilde v_2^n v_2')\,dx.$$

Notice that

$$\left|\int_{J_k}\tilde p_n v_2'\,dx\right|\leqslant\varepsilon\|\tilde p_n\|_2\sqrt{|J_k|}\leqslant\varepsilon\|p_n\|_2\sqrt{|J_k|}\lesssim|J_n|^2,$$

because $|p_n|\lesssim|J_n|$ on $J_n$ and $p_n=0$ outside. Also,

$$\left|\int_{J_k}p'\tilde v_2^n\,dx\right|=|\langle p_k,v_2^n\rangle_{\mathcal D}|=\left|\int_{J_n}\tilde p_k v_2'\,dx\right|\lesssim|J_n|^2,$$

by the same estimate. Similarly, since $|v_2^n|\lesssim|J_n|$ on $J_n$ and equals zero outside,

$$\left|\int_{J_k}\tilde v_2^n v_2'\,dx\right|\lesssim|J_n|^2.$$

Hence

$$-\iint_{J_n\times J_k}\log|t-x|w'(x)w'(t)\,dx\,dt=-\iint_{J_n\times J_k}\log|t-x|p'(x)p'(t)\,dx\,dt+O(|J_n|^2). \tag{6.12}$$

For the last integral we have

$$-\iint_{J_n\times J_k}\log|t-x|p'(x)p'(t)\,dx\,dt=-\sum_{I_j\subset J_k}\sum_{I_{j'}\subset J_n}\iint_{I_j\times I_{j'}}\log|t-x|p'(x)p'(t)\,dx\,dt. \tag{6.13}$$

To estimate

$$-\iint_{I_j\times I_{j'}}\log|t-x|p'(x)p'(t)\,dx\,dt=\iint_{I_j\times I_{j'}}\log_-|t-x|p'(x)p'(t)\,dx\,dt \\ -\iint_{I_j\times I_{j'}}\log_+|t-x|p'(x)p'(t)\,dx\,dt \tag{6.14}$$

we consider three cases. First, to estimate the integral in the case when $j=j'$, notice that, since $1+v_1'$ is bounded,

$$\int_{I_j}\log_-|x-t|(1+v_1'(x))\,dx<\text{const}$$

for any $t\in I_j$. Once again, the positive functions $(\arg\theta)'$ and $v_1'+1$ satisfy

$$\int_{I_l}(\arg\theta)'\,dx=\int_{I_l}(v_1'+1)\,dx=2\pi\Delta_l+O(1)=|I_l|+O(1). \tag{6.15}$$

Hence

$$\begin{aligned}\iint_{I_j\times I_j} & \log_-|t-x|p'(x)p'(t)\,dx\,dt\\ &=\iint_{I_j\times I_j}\log_-|t-x|(\arg\theta)'(x)(\arg\theta)'(t)\,dx\,dt\\ &\quad -2\iint_{I_j\times I_j}\log_-|t-x|(1+v_1'(x))(\arg\theta)'(t)\,dx\,dt\\ &\quad +\iint_{I_j\times I_j}\log_-|t-x|(1+v_1'(x))(1+v_1'(t))\,dx\,dt\\ &=\iint_{I_j\times I_j}\log_-|t-x|(\arg\theta)'(x)(\arg\theta)'(t)\,dx\,dt+O(|I_j|).\end{aligned}$$

For the last integral we have

$$\begin{aligned}\iint_{I_j\times I_j} & \log_-|t-x|(\arg\theta)'(x)(\arg\theta)'(t)\,dx\,dt\\ &=\sum_{(\lambda_l,\lambda_{l+1})\subset I_j}\ \sum_{(\lambda_m,\lambda_{m+1})\subset I_j}\int_{\lambda_l}^{\lambda_{l+1}}\int_{\lambda_m}^{\lambda_{m+1}}(\arg\theta)'(x)(\arg\theta)'(t)\,dx\,dt.\end{aligned}$$

Using that

$$\int_{\lambda_s}^{\lambda_{s+1}}(\arg\theta)'\,dx=2\pi$$

and that

$$(\arg\theta)'\lesssim\frac{1}{\min\{|I_{s-1}|,|I_s|,|I_{s+1}|\}}+\frac{1}{|I_s|^2}\quad\text{on } (\lambda_s,\lambda_{s+1}),$$

for all $s$ by Lemma 7.8, we can apply Lemma 7.9 (1)–(3). Assuming that $\lambda_l\leqslant\lambda_m$, we conclude that

$$\begin{aligned}\int_{\lambda_l}^{\lambda_{l+1}}\int_{\lambda_m}^{\lambda_{m+1}} & \log_-|t-x|(\arg\theta)'(x)(\arg\theta)'(t)\,dx\,dt\\ &\lesssim\begin{cases}\log_-(\lambda_m-\lambda_{l+1}), & \text{if } \lambda_m>\lambda_{l+1},\\ \max\{\log_-(\lambda_l-\lambda_{l-1}),\log_-(\lambda_{l+1}-\lambda_l),\log_-(\lambda_{l+2}-\lambda_{l+1}),\log_-(\lambda_{l+3}-\lambda_{l+2})\}+1, & \text{if } \lambda_m=\lambda_{l+1},\\ \max\{\log_-(\lambda_l-\lambda_{l-1}),\log_-(\lambda_{l+1}-\lambda_l),\log_-(\lambda_{l+2}-\lambda_{l+1})\}+1, & \text{if } \lambda_m=\lambda_l,\end{cases}\end{aligned}$$

which implies that

$$\iint_{I_j\times I_j}\log_-|t-x|p'(x)p'(t)\,dx\,dt\lesssim\sum_{\substack{a_j\leqslant\lambda_k,\lambda_l\leqslant a_{j+1}\\ \lambda_k\neq\lambda_l}}\log_-|\lambda_k-\lambda_l|+|I_j|. \tag{6.16}$$

To estimate the integral of $\log_+$, first notice that, by Lemma 7.9 (5) and (6.15),

$$\int_{I_j}\log_+|x-t|(1+v_1'(x))\,dx=|I_j|\log_+|I_j|+O(|I_j|)$$

for any $t\in I_j$.

Together with Lemma 7.9 (4) and (6.15), we get

$$\begin{aligned}
&\iint_{I_j\times I_j}\log_+|t-x|p'(x)p'(t)\,dx\,dt\\
&\quad=\iint_{I_j\times I_j}\log_+|t-x|(\arg\theta)'(x)(\arg\theta)'(t)\,dx\,dt\\
&\qquad-2\iint_{I_j\times I_j}\log_+|t-x|(v_1'(x)+1)(\arg\theta)'(t)\,dx\,dt\\
&\qquad+\iint_{I_j\times I_j}\log_+|t-x|(v_1'(x)+1)(v_1'(t)+1)\,dx\,dt\\
&\quad=\sum_{a_j\leqslant\lambda_k,\lambda_l<a_{j+1}}\int_{\lambda_k}^{\lambda_{k+1}}\int_{\lambda_l}^{\lambda_{l+1}}\log_+|t-x|(\arg\theta)'(x)(\arg\theta)'(t)\,dx\,dt\\
&\qquad\qquad-|I_j|^2\log|I_j|+O(|I_j|^2)\\
&\quad\geqslant4\pi^2\sum_{a_j\leqslant\lambda_k,\lambda_l<a_{j+1}}\log_+|\lambda_k-\lambda_l|-|I_j|^2\log|I_j|+O(|I_j|^2).
\end{aligned}\tag{6.17}$$

Next, let us consider the case when $j\neq j'$ and the intervals $I_j$ and $I_{j'}$ are not adjacent. This estimate is similar to (6.9), but we will treat it using a different technique. Assume for example that $j'>j+1$.

For $\log_-$, recalling that $|I_k|>1$ for all $k$, we get

$$-\iint_{I_j\times I_{j'}}\log_-|t-x|p'(x)p'(t)\,dx\,dt=0. \tag{6.18}$$

For $\log_+$ we have

$$\begin{aligned}
&-\iint_{I_j\times I_{j'}} \log_+|t-x|p'(x)p'(t)\,dx\,dt \\
&\quad = -\int_{a_j}^{a_{j+1}}\int_{a_{j'}}^{a_{j'+1}} \log_+|t-x|p'(x)p'(t)\,dx\,dt \\
&\quad = -\int_{a_j}^{a_{j+1}}\int_{a_{j'}}^{a_{j'+1}} \log_+|t-x|(\arg\theta-x-v_1)'(x)(\arg\theta-x-v_1)'(t)\,dx\,dt \\
&\quad \leqslant -\int_{I_{j'}}\biggl(\log|a_{j+1}-t|\int_{I_j}(\arg\theta)'(x)\,dx-\log|a_j-t|\int_{I_j}(v_1'+1)(x)\,dx\biggr)(\arg\theta)'(t)\,dt \\
&\qquad\quad +\int_{I_{j'}}\biggl(\log|a_j-t|\int_{I_j}(\arg\theta)'(x)\,dx-\log|a_{j+1}-t|\int_{I_j}(v_1'+1)(x)\,dx\biggr)(v_1'+1)(t)\,dt \\
&\quad \leqslant 2|I_j|\,|I_{j'}|. \qquad (6.19)
\end{aligned}$$

Here we used that $\operatorname{dist}(I_j,I_{j'})\geqslant|I_j|$ by monotonicity and (6.15).

In the case when $I_j$ and $I_{j'}$ are adjacent, i.e. $j'=j+1$, the estimate can be done differently. Note that $p_n=p|_{I_n}$ is a compactly supported function with bounded derivative (the bound depends on $n$). Therefore it belongs to the Dirichlet space $\mathcal{D}(\mathbb{R})$. The estimates (6.16) and (6.17) yield

$$\|p_n\|^2_{\mathcal{D}} \lesssim \frac{1}{4\pi^2}|I_n|^2\log|I_n|-E_n+|I_n|^2. \qquad (6.20)$$

Hence

$$\begin{aligned}
&\iint_{I_j\times I_{j+1}} \log|t-x|p'(x)p'(t)\,dx\,dt=\langle p_j,p_{j+1}\rangle_{\mathcal{D}}\leqslant\|p_j\|^2+\|p_{j+1}\|^2 \\
&\quad\lesssim\biggl(\frac{1}{4\pi^2}|I_j|^2\log|I_j|-E_j\biggr)+\biggl(\frac{1}{4\pi^2}|I_{j+1}|^2\log|I_{j+1}|-E_{j+1}\biggr)+|I_j|^2+|I_{j+1}|^2.
\end{aligned} \qquad (6.21)$$

Now we can return to estimating

$$-\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2}$$

in the case when $|k-n|\leqslant 1$. Using the estimates (6.11)–(6.13), we obtain

$$\begin{aligned}
&-\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2} \\
&\quad = -\frac{1}{1+\operatorname{dist}^2(0,J_n)}\biggl(\sum_{I_j\subset J_k}\sum_{I_{j'}\subset J_n}\iint_{I_j\times I_{j'}}\log|t-x|p'(x)p'(t)\,dx\,dt+O(|J_n|^2)\biggr).
\end{aligned}$$

The estimates (6.16)–(6.21) yield

$$
\begin{aligned}
-\sum_{I_j\subset J_n}\sum_{I_{j'}\subset J_k}\iint_{I_j\times I_{j'}} &\log|t-x|p'(x)p'(t)\,dx\,dt\\
&\lesssim \sum_{I_j\subset J_k\cup J_n}\left(\frac{1}{4\pi^2}|I_j|^2\log|I_j|-E_j+|I_j|^2\right)+\sum_{I_j,I_{j'}\subset J_k\cup J_n}|I_j|\,|I_{j'}|\\
&\leqslant \sum_{I_j\subset J_k\cup J_n}\left(\frac{1}{4\pi^2}|I_j|^2\log|I_j|-E_j\right)+|J_n|^2+|J_k|^2.
\end{aligned}
$$

All in all, in the case $|n-k|\leqslant 1$, we have

$$
-\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2}\lesssim\frac{1}{1+\mathrm{dist}^2(0,J_n)}\left(\sum_{I_j\subset J_k\cup J_n}\left(\frac{1}{4\pi^2}|I_j|^2\log|I_j|-E_j\right)+|J_n|^2+|J_k|^2\right). \tag{6.22}
$$

To continue the estimate of the right-hand side of (6.6), notice that

$$
\begin{aligned}
\sum_n\sum_k&\left(-\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2}\right)\\
&=\sum_n\sum_{\substack{k\\ \max\{|J_k|,|J_n|\}\leqslant\mathrm{dist}(J_k,J_n)}}\left(-\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2}\right)\\
&\quad+\sum_n\sum_{\substack{k\\ \min\{|J_k|,|J_n|\}\leqslant\mathrm{dist}(J_k,J_n)<\max\{|J_k|,|J_n|\}}}\left(-\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2}\right)\\
&\quad+\sum_n\sum_{\substack{k\\ \mathrm{dist}(J_k,J_n)<\min\{|J_k|,|J_n|\}}}\left(-\int_{J_n} w'\widetilde{w}_k\,\frac{dx}{x^2}\right)\\
&=:\mathrm{I}+\mathrm{II}+\mathrm{III}.
\end{aligned}
$$

For the first sum, by (6.7), we get

$$
\begin{aligned}
\mathrm{I}&\lesssim\sum_k\sum_{\substack{n>k\\ \mathrm{dist}(J_k,J_n)>|J_k|}}\frac{|J_k|^3}{1+\mathrm{dist}^2(J_n,0)}\int_{J_n}\frac{dx}{\mathrm{dist}^2(J_k,x)}\\
&\leqslant\sum_k\frac{|J_k|^3}{1+\mathrm{dist}^2(J_k,0)}\,\frac{1}{|J_k|}=\sum_k\frac{|J_k|^2}{1+\mathrm{dist}^2(J_k,0)}<\infty.
\end{aligned} \tag{6.23}
$$

For the second sum, by (6.9),

$$
\mathrm{II}\lesssim\sum_n\sum_{\substack{k\neq n\\ \mathrm{dist}(J_k,J_n)<\max\{|J_k|,|J_n|\}}}\frac{|J_k|\,|J_n|}{1+\mathrm{dist}^2(J_n,0)}.
$$

Recall that, by our assumption, $|J_n|\leqslant\operatorname{dist}(J_n,0)$ for all $n\neq 0,-1$. We may also assume that $|J_{-1}|=|J_0|$. Then in each term in the last sum $k$ and $n$ have the same sign. Let us estimate the part of the sum with non-negative $k$ and $n$.

$$\sum_{n\geqslant 0}\sum_{\substack{0\leqslant k\neq n\\ \operatorname{dist}(J_k,J_n)<\max\{|J_k|,|J_n|\}}}\frac{|J_k|\,|J_n|}{1+\operatorname{dist}^2(J_n,0)}=2\sum_{n\geqslant 0}\sum_{\substack{0\leqslant k<n\\ \operatorname{dist}(J_k,J_n)<|J_n|}}\frac{|J_k|\,|J_n|}{1+\operatorname{dist}^2(J_n,0)}$$
$$\leqslant 4\sum_{n\geqslant 0}\frac{|J_n|^2}{1+\operatorname{dist}^2(J_n,0)}<\infty.$$

Terms with negative indices $k$ and $n$ can be estimated similarly to conclude that

$$\mathrm{II}\lesssim\sum_n\frac{|J_n|^2}{1+\operatorname{dist}^2(J_n,0)}<\infty.$$

Finally, for the third sum, by (6.22),

$$\mathrm{III}\lesssim\sum_n\sum_{\substack{k\\ |k-n|\leqslant 1}}\frac{1}{1+\operatorname{dist}^2(0,J_n)}\biggl(\sum_{I_j\subset J_k\cup J_n}\biggl(\frac{1}{4\pi^2}|I_j|^2\log|I_j|-E_j\biggr)+|J_n|^2+|J_k|^2\biggr)$$
$$\lesssim\sum_n\frac{1}{1+\operatorname{dist}^2(0,J_n)}\biggl(\sum_{I_j\subset J_n}\biggl(\frac{1}{4\pi^2}|I_j|^2\log|I_j|-E_j\biggr)+|J_n|^2\biggr)<\infty,$$

because $\Lambda$ satisfies the energy condition on $I_n$. Altogether these estimates give us

$$-\int_{\mathbb{R}}\widetilde{w}w'\,\frac{dx}{x^2}<\infty.$$

The integrals over the circular part of $\partial D(r)$ can be estimated like in [21, §5.3]. We need to show that the integrals

$$-\int_{\partial D(r)\setminus\mathbb{R}}\tilde{q}\,dq=rI'(r)\quad\text{and}\quad I(r):=\frac{1}{2}\int_0^{\pi}\tilde{q}^2(re^{i\phi})\,d\phi$$

do not tend to $\infty$ as $r\to\infty$. In fact, it is enough to show that

$$I(r)\not\to\infty,$$

because if $rI'(r)\to\infty$, then $I'(r)\geqslant 1/r$ for all $r\gg 1$, and we have $I(r)\to\infty$.

As we will see shortly, $I(r)$ does not tend to $\infty$ for *any* $h\in L^1(1+|x|^{-1})$ in place of $\tilde{q}$ (recall that $\tilde{q}=\widetilde{w}/x$ and $\widetilde{w}\in L^1_\Pi$). It will be more convenient for us to prove an equivalent statement in the unit disk $\mathbb{D}$.

Let $h+i\tilde{h}$ be an analytic function in $\mathbb{D}$ such that

$$\frac{h(\zeta)}{1-|\zeta|}\in L^1(\mathbb{T}),$$

where $\mathbb{T}=\partial\mathbb{D}$. Define

$$l(z)=\frac{1+z}{1-z}(h(z)+i\tilde{h}(z)),\quad z\in\mathbb{D},$$

and denote by $l^M(\zeta)$, $\zeta\in\mathbb{T}$, the angular maximal function. Then $\operatorname{Im} l\in L^1(\mathbb{T})$ and, by the Hardy–Littlewood maximal theorem,

$$l^M\in L^{1,\infty}(\mathbb{T}). \tag{6.24}$$

Let us show that, as $\varepsilon\to 0$,

$$\frac{1}{\varepsilon}\int_{C_\varepsilon}|h+i\tilde{h}|^2\,|dz|\not\to\infty,\quad C_\varepsilon:=\{z\in\mathbb{D}:|1-z|=\varepsilon\}.$$

We have

$$\frac{1}{\varepsilon}\int_{C_\varepsilon}|h+i\tilde{h}|^2\,|dz|\leqslant\varepsilon\int_{C_\varepsilon}|l|^2\,|dz|\lesssim(\varepsilon l^M(\zeta))^2+(\varepsilon l^M(\bar{\zeta}))^2,$$

where $\zeta\in\mathbb{T}$, $|1-\zeta|=\varepsilon$. The right-hand side cannot tend to infinity because otherwise, for all small $\varepsilon$, we would have

$$l^M(\zeta)+l^M(\bar{\zeta})\gg\frac{1}{\varepsilon}$$

on an interval of length $\varepsilon$, which would contradict (6.24). $\square$

Let

$$\phi=\tfrac{1}{2}\arg(\theta\bar{S}I_+\bar{I}_-).$$

Recall that $\phi,\tilde{\phi}\in L^1_\Pi$. By the last claim $\phi/x$ belongs to the Dirichlet class. Since $\tilde{\phi}/x$ is the conjugate of $\phi/x$, $\tilde{\phi}/x$ belongs to $\mathcal{D}(\mathbb{R})$ as well. Hence, by the Beurling–Malliavin multiplier theorem, see for instance [21, §5.1], there exists a smooth function $m$ on $\mathbb{R}$ satisfying

$$m'<\varepsilon,\quad \widetilde{m}\in L^1_\Pi\quad\text{and}\quad\widetilde{m}\geqslant\max\{0,-\tilde{\phi}\}.$$

In other words, if $\Phi$ and $M$ are outer functions,

$$\Phi=e^{i\phi-\tilde{\phi}}\quad\text{and}\quad M=e^{im-\widetilde{m}},$$

then $\Phi M$ is bounded in $\mathbb{C}_+$.

Since $m'<\varepsilon$, $\varepsilon x-m$ is an increasing function. There exists a meromorphic inner function $J$ such that

$$\{x:J(x)=\pm 1\}=\{x:2(\varepsilon x-m)=k\pi\}.$$

Let

$$d_1 = 2(\varepsilon x - m) \quad \text{and} \quad d_2 = \arg J.$$

Then the difference

$$d = d_1 - d_2 = 2(\varepsilon x - m) - \arg J$$

satisfies $|d| < \pi$.

Put

$$l(x) = \varepsilon x - \tfrac{1}{2} \arg J.$$

Notice that $\tilde{l} \in L^1_\Pi$ because $2l = d + 2m$, where $d$ is bounded and $\widetilde{m} \in L^1_\Pi$.

Consider an outer function $\Psi = e^{il - \tilde{l}}$. Then

$$\bar{S}^{2\varepsilon} \Psi = \bar{J} \bar{\Psi}, \quad \text{or equivalently} \quad \bar{S}^{2\varepsilon} J \Psi = \bar{\Psi}$$

on $\mathbb{R}$. Thus $\Psi \in N^+[\bar{S}^{2\varepsilon} J]$.

Moreover, the ratio $\Psi/M$ is equal to $e^{id/2 - \tilde{d}/2}$. Since $|d| < \pi$, $\Psi/M$ belongs to every $L^p_\Pi$, $p < 1$. Our next goal is to construct another "small" outer multiplier function $k$ so that $k\Psi/M \in L^2_\Pi$.

Consider the step function

$$\alpha(x) = \frac{\pi}{5} \left\lfloor \frac{5}{\pi} d_1 \right\rfloor - \frac{\pi}{5} \left\lfloor \frac{5}{\pi} d_2 \right\rfloor,$$

where again $\lfloor \cdot \rfloor$ denotes the integer part of a real number. Then

$$|d - \alpha| < \tfrac{2}{5} \pi. \tag{6.25}$$

Since $d_1 = d_2 = \pi m$ at the points $\{c_m\}_m = \{x : J(x) = \pm 1\}$, the function $\alpha$ only takes values $\frac{1}{5} k\pi$, $k = -4, ..., 4$. Therefore $\alpha$ can be represented as

$$\alpha = \frac{\pi}{5} \left( \sum_{n=1}^{4} \beta_n - \sum_{n=5}^{8} \beta_n \right),$$

where $\beta_n$ are elementary step functions, each taking only two values, 0 and 1, and making at most one positive and one negative jump on each interval $[c_m, c_{m+1}]$. For each $n = 1, ..., 8$ one can choose an inner function $Q_n$ such that

$$\frac{1 - Q_n}{1 + Q_n} = \text{const} e^{K\beta_n}.$$

Notice that then

$$e^{\tilde{\alpha} - i\alpha} = \text{const} \left( \prod_{n=1}^{4} \frac{1 + Q_n}{1 - Q_n} \prod_{n=5}^{8} \frac{1 - Q_n}{1 + Q_n} \right)^{1/5}.$$

Because of (6.25), we have

$$\left|\frac{\Psi}{M}\prod_{n=1}^{4}(1+Q_n)\prod_{n=5}^{8}(1-Q_n)\right|\lesssim\left|\frac{\Psi}{M}\right|\left|\prod_{n=1}^{4}\frac{1+Q_n}{1-Q_n}\prod_{n=5}^{8}\frac{1-Q_n}{1+Q_n}\right|^{1/10}$$
$$=\operatorname{const}e^{\tilde{\alpha}/2-\tilde{d}/2}\in L^2_\Pi(\mathbb{R})$$

and, since the function $M\Phi$ is bounded,

$$\Psi\Phi\prod_{n=1}^{4}(1+Q_n)\prod_{n=5}^{8}(1-Q_n)=\frac{\Psi}{M}M\Phi\prod_{n=1}^{4}(1+Q_n)\prod_{n=5}^{8}(1-Q_n)\in L^2_\Pi(\mathbb{R}).$$

Now notice that since $N^+[\bar{S}^{2\varepsilon}J]\neq 0$, the set $\{x:J(x)=1\}$ has Beurling–Malliavin density at most $2\varepsilon$, see [20, §4.6]. By our construction, the Beurling–Malliavin density of each of the sets $\{x:Q_n(x)=1\}$ is the same as that of $\{x:J(x)=1\}$, i.e. at most $2\varepsilon$. Consequently, the kernel

$$N^\infty\left[\bar{S}^{17\varepsilon}\prod_{n=1}^{8}Q_n\right]$$

contains a non-zero function $\tau$, see [20, §4.2 and §4.6].

Similarly, since the Beurling–Malliavin density of $\{x:I_+(x)=1\}$ is less than $\varepsilon$, the kernel $N^\infty[\bar{S}^\varepsilon I_+]$ is infinite-dimensional. Hence it contains a non-trivial function $\eta$ with at least one zero $a$ in $\mathbb{C}_+$. Then the function $\varkappa=\eta/(z-a)$ also belongs to $N^\infty[\bar{S}^\varepsilon I_+]$ and satisfies $|\varkappa|\lesssim(1+|x|)^{-1}$ on $\mathbb{R}$.

Therefore

$$\bar{\theta}S^{1-20\varepsilon}\varkappa\tau\prod_{n=1}^{4}(1+Q_n)\prod_{n=5}^{8}(1-Q_n)\Psi\Phi$$
$$=(\bar{S}^\varepsilon I_+\varkappa)\left(\bar{S}^{17\varepsilon}\prod_{n=1}^{4}(1+Q_n)\prod_{n=5}^{8}(1-Q_n)\tau\right)(\bar{S}^{2\varepsilon}\Psi)(\bar{\theta}S^1\bar{I}_+\Phi)\in\overline{H}^2.$$

Accordingly, the space $K_\theta$ contains the function

$$f=S^{1-20\varepsilon}\varkappa\tau\prod_{n=1}^{4}(1+Q_n)\prod_{n=5}^{8}(1-Q_n)\Psi\Phi.$$

Now we could simply refer to Theorem 3.2 to conclude this part of the proof. For the sake of completeness we also present a direct argument.

By the Clark representation formula

$$f=(1+\theta)Kf\sigma_1,$$

where $\sigma_1$ is the Clark measure corresponding to $\theta$ concentrated on $\{x:\theta(x)=1\}=\Lambda$. Since $1+\theta$ is bounded in the upper half-plane and $f$ decreases faster than $e^{-(1-21\varepsilon)y}$ along the positive $y$-axis, so does $Kf\mu$. Hence $f\mu$ is the measure concentrated on $\Lambda$ with the spectral gap at least $1-21\varepsilon$.

**Proof of the main theorem, part II**

Now suppose that $G_X > 1$ but $\mathcal{C}_X < 1/2\pi$.

By Lemma 8.3 there exists a discrete increasing sequence $\Lambda = \{\lambda_n\}_{n\in\mathbb{Z}} \subset X$ and a measure $\nu$, with $\operatorname{supp}\nu = \Lambda$, such that $\nu$ has a spectral gap of size 1 and $K\nu$ does not have any zeros in $\mathbb{C}$.

Similarly to the previous part, we assume that $\sup_n(\lambda_n - \lambda_{n-1}) < \infty$. The general case is discussed at the end of the proof. If $\sup_n(\lambda_n - \lambda_{n-1}) < \infty$, we can apply Lemma 7.8 and consider the inner function $\theta$ corresponding to $\Lambda$. A function $f \in N[\phi]$ is called *purely outer* if $f$ is outer in the upper half-plane and $\phi f = \bar{g}$ is outer in the lower half-plane. Since $K\nu$ is divisible by $S$, the function

$$f = S^{-1}(1-\theta)K\nu \in K_\theta^{1,\infty}$$

is a purely outer element of $N^{1,\infty}[\bar{\theta}S]$, see §8. Note that $f = e^{i\phi - \tilde{\phi}}$ in $\mathbb{C}_+$, where

$$2\phi = \arg\theta - x.$$

Denote by $\Gamma_n$ the middle third of the interval $(\lambda_n, \lambda_{n+1})$. Our plan is to calculate the integral

$$\int_{\bigcup_n \Gamma_n} \phi' \tilde{\phi} \frac{dx}{x^2} \tag{6.26}$$

in two different ways and arrive at a contradiction by obtaining two different answers.

First let us choose a short monotone partition $\{I_n\}_n$ of $\mathbb{R}$ such that $\Lambda$ satisfies the density condition (4.4) with $a = 1/2\pi$ on that partition.

Put $a_0 = 0$. Choose $a_1 > a_0$ to be the smallest point such that

$$\#(\Lambda \cap (a_0, a_1]) \geqslant \frac{a_1 - a_0}{2\pi}.$$

Note that such a point always exists because $\Lambda$ supports a measure with a spectral gap greater than 1: otherwise we would be able to choose a long sequence of intervals satisfying (7.1) in Lemma 7.2 with $a = 1/2\pi$ and arrive at a contradiction. After $a_j$, $j \geqslant 1$, is chosen, choose $a_{j+1} > a_j$ as the smallest point such that

$$\#(\Lambda \cap (a_j, a_{j+1}]) \geqslant \frac{a_{j+1} - a_j}{2\pi} \quad \text{and} \quad a_{j+1} - a_j \geqslant a_j - a_{j-1}.$$

Choose $a_j$, $j < 0$, in a similar way. Put $I_n = (a_n, a_{n+1}]$. Again, by Lemma 7.2, $\{I_n\}_n$ has to be short.

In what follows we will assume, without loss of generality, that $|I_n|/2\pi = \#(\Lambda \cap I_n)$.

Note that, since $\mathcal{C}_X < 1$, the sum in the energy condition (4.5) has to be infinite. At the same time, a part of that sum has to be finite, as we state in the next claim.

CLAIM 6.3.

$$\sum_n \frac{\log_-(\lambda_{n+1}-\lambda_n)}{\lambda_n^2+1}<\infty.$$

*Proof.* Suppose that the sum is infinite. Put $\mu=|\nu|$ and let $\Phi$ be the inner function such that $\mu$ is its Clark measure. Let $\psi=\arg\Phi-x$.

Define the intervals $J_n$ and the function $v$ as in Claim 6.1, in the first part of the proof, with $\Phi$ replacing $\theta$. Put $w=\psi-v=\arg\Phi-x-v$. Let again $w_n=w|_{J_n}$. Then $\widetilde{w}\in L^1_\Pi$, because $w_n$ are atoms with summable $L^1_\Pi$-norms.

As in the first part of the proof, we can use "atomic" estimates to show that if $\operatorname{dist}(J_k,J_n)>\max\{|J_k|,|J_n|\}$ and $x\in J_n$, then

$$|\widetilde{w}_k(x)|\lesssim\frac{|J_k|^3}{\operatorname{dist}^2(x,J_k)}.$$

By the monotonicity and the shortness of $J_k$, we conclude that

$$\sum_{\lambda_j\in J_n}\frac{|\widetilde{w}_k(\lambda_j)|}{\lambda_j^2}\lesssim\sum_{\lambda_j\in J_n}\frac{|\widetilde{w}_k(\lambda_j)|}{1+\operatorname{dist}^2(0,J_n)}\lesssim\frac{1}{1+\operatorname{dist}^2(0,J_n)}\int_{J_n}\frac{|J_k|^3}{\operatorname{dist}^2(x,J_k)}\,dx.$$

Hence, similarly to (6.23),

$$\begin{aligned}\sum_n\sum_{\substack{k\\ \operatorname{dist}(J_k,J_n)>\max\{|J_k|,|J_n|\}}}\sum_{\lambda_j\in J_n}\frac{|\widetilde{w}_k(\lambda_j)|}{\lambda_j^2}&\\ =2\sum_n\sum_{\substack{k<n\\ \operatorname{dist}(J_k,J_n)>\max\{|J_k|,|J_n|\}}}\sum_{\lambda_j\in J_n}\frac{|\widetilde{w}_k(\lambda_j)|}{\lambda_j^2}&\\ \lesssim\sum_n\sum_{\substack{k<n\\ \operatorname{dist}(J_k,J_n)>\max\{|J_k|,|J_n|\}}}\frac{1}{1+\operatorname{dist}^2(0,J_n)}\int_{J_n}\frac{|J_k|^3}{\operatorname{dist}^2(x,J_k)}\,dx<\infty.&\end{aligned}$$

In other words, on each $J_n$,

$$\sum_{\substack{k\\ \operatorname{dist}(J_k,J_n)>\max\{|J_k|,|J_n|\}}}|\widetilde{w}_k|\leqslant g_1,$$

where $g_1$ is a positive function satisfying

$$\sum_n\frac{g_1(\lambda_n)}{1+\lambda_n^2}<\infty.$$

Also, for any $x\in J_n$,

$$-\widetilde{w}_k(x)=\int_{J_k}\frac{w_k(t)\,dt}{t-x}=-\int_{J_k}\log|t-x|w'(t)\,dt.$$

If $k<n$ (the case $k>n$ is similar) then

$$\begin{aligned}-\int_{J_k}\log|t-x|w'(t)\,dt\geqslant -\int_{J_k}\log_+|t-x|(\arg\Phi)'(t)\,dt \\ +\int_{J_k}\log_+|t-x|(1+v')(t)\,dt-\mathrm{const}\gtrsim -|J_k|.\end{aligned}\tag{6.27}$$

Here we again used that

$$\int_{J_k}(\arg\Phi)'(t)\,dt=\int_{J_k}(1+v')(t)\,dt=|J_k|+O(1),$$

applied Lemma 7.9 (6) to the integral in the second line of (6.27) and used the estimate

$$-\int_{J_k}\log_+|t-x|(\arg\Phi)'(t)\,dt\geqslant -\log(x-b)\int_{J_k}(\arg\Phi)'(t)\,dt,\tag{6.28}$$

where $b$ is the left (right if $k>n$) endpoint of $J_k$, for the first integral.

Thus, for $x\in J_n$,

$$-\sum_{\substack{k\\ |k-n|>1\\ \mathrm{dist}(J_k,J_n)\leqslant\max\{|J_k|,|J_n|\}}}\widetilde{w}_k(x)\geqslant g_2(x),$$

where again

$$\sum_n\frac{|g_2(\lambda_n)|}{1+\lambda_n^2}<\infty.$$

Also, for $x\in J_n$ and $k\in\{n-1,n,n+1\}$, similarly to (6.27),

$$\begin{aligned}-\int_{J_k}\log_+|x-t|w'(t)\,dt=-\int_{J_k}\log_+|x-t|(\arg\Phi)'(t)\,dt \\ +\int_{J_k}\log_+|x-t|(1+v')(t)\,dt\gtrsim -|J_k|,\end{aligned}\tag{6.29}$$

by applying (6.28) to the first integral and Lemma 7.9 (5) to the second integral if $k=n$, or Lemma 7.9 (6) if $k=n\pm1$.

Hence for any $x\in\mathbb{R}$, if $x\in J_n$ for some $n$, then

$$-\widetilde{w}(x)\geqslant\int_{J_{n-1}\cup J_n\cup J_{n+1}}\log_-|x-t|w'(t)\,dt+g(x)=\int_{\mathbb{R}}\log_-|x-t|w'(t)\,dt+g(x)$$

for some function $g$ satisfying

$$\sum_n\frac{|g(\lambda_n)|}{1+\lambda_n^2}<\infty.$$

Therefore,

$$
\begin{aligned}
-\sum_n \frac{\widetilde{w}(\lambda_n)}{1+\lambda_n^2} &\geqslant \text{const}+\sum_n \frac{1}{1+\lambda_n^2}\int_{\lambda_n-1}^{\lambda_n+1} \log_-|\lambda_n-x|w'\,dx \\
&\geqslant \text{const}+\sum_n \frac{1}{1+\lambda_n^2}\int_{\lambda_{n-1}}^{\lambda_n} \log_-|\lambda_n-x|(\arg\Phi)'(x)\,dx \\
&\geqslant \text{const}+2\pi\sum_n \frac{\log_-|\lambda_n-\lambda_{n-1}|}{1+\lambda_n^2}.
\end{aligned}
$$

Let $h=(1+\Phi)K\nu$. Then $h$ is an outer function in $\mathbb{C}_+$ that belongs to $H^2$ and satisfies

$$h=e^{i\psi/2-\tilde{\psi}/2}.$$

Since

$$|h(\lambda_n)|=\frac{|\nu(\{\lambda_n\})|}{\mu(\{\lambda_n\})}=1,$$

we have that $\log|h(\lambda_n)|=2\tilde{\psi}(\lambda_n)=0$ for all $n$.

Recall that $\widetilde{w}(\lambda_n)=\tilde{\psi}(\lambda_n)+\tilde{v}(\lambda_n)=\tilde{v}(\lambda_n)$. It is left to show that

$$\sum_n \frac{-\tilde{v}(\lambda_n)}{1+\lambda_n^2}<\infty.$$

Recall that $v\in L^1_\Pi$, $\tilde{v}=\widetilde{w}-\tilde{\psi}=\widetilde{w}+2\log|h|\in L^1_\Pi$ and $v'$ is bounded on $\mathbb{R}$. Therefore the harmonic extension of $v$ into $\mathbb{C}_+$ has a bounded $x$-derivative in $\mathbb{C}_+$. Hence $\tilde{v}_y$ is bounded in $\mathbb{C}_+$ as well.

On each interval $J_n$ choose $\lambda_{k_n}$ so that

$$|\tilde{v}(\lambda_{k_n})|=\max_{\lambda_j\in J_n}|\tilde{v}(\lambda_j)|.$$

If the last sum is infinite, then so is

$$\sum_n |J_n|\frac{|\tilde{v}(\lambda_{k_n})|}{1+\operatorname{dist}^2(0,J_n)}.$$

Because of the boundedness of $\tilde{v}_y$, $|\tilde{v}(\lambda_{k_n}+i|J_n|)|\geqslant|\tilde{v}(\lambda_{k_n})|-C|J_n|$ and therefore

$$\sum_n |J_n|\frac{|\tilde{v}(\lambda_{k_n}+i|J_n|)|}{1+\operatorname{dist}^2(0,J_n)}=\infty.$$

Let $(\tilde{v})^M$ denote the maximal non-tangential function of $\tilde{v}$ in $\mathbb{C}_+$. The last equation implies that $(\tilde{v})^M\notin L^1_\Pi$. But this contradicts the fact that both $\tilde{v}$ and $v$ belong to $L^1_\Pi$. $\square$

Now let us return to the function $f=(1-\theta)K\nu$ defined before the claim. Recall that $f=e^{i\phi-\tilde{\phi}}$ in $\mathbb{C}_+$, where $2\phi=\arg\theta-x$. Again using Claim 6.1, we can find intervals $J_n$ and a function $v$ for $u=2\phi$. Let

$$w=\arg\theta-x-v=2\phi-v.$$

Recall that $\Gamma_n$ is the middle third of the interval $(\lambda_n,\lambda_{n+1})$. Notice that if $x\in\Gamma_n$, then

$$|f(x)|=|(1-\theta(x))K\nu(x)|\leqslant 2\left|\int_{\mathbb{R}}\frac{1}{t-x}\,d\nu(t)\right|\leqslant\frac{6\|\nu\|}{|\lambda_{n+1}-\lambda_n|}. \tag{6.30}$$

Since $\log|f|=-\tilde{\phi}$,

$$\begin{aligned}-\int_{\bigcup_n\Gamma_n}\phi'\tilde{\phi}\,\frac{dx}{x^2}&\lesssim\sum_n\frac{1}{1+\lambda_n^2}\int_{\Gamma_n}(\arg\theta)'\log_+|f|\,dx+\text{const}\\&\lesssim\sum_n\frac{1}{1+\lambda_n^2}\log_-|\lambda_{n+1}-\lambda_n|\int_{\Gamma_n}(\arg\theta)'\,dx+\text{const}\\&\lesssim\sum_n\frac{1}{1+\lambda_n^2}\log_-|\lambda_{n+1}-\lambda_n|+\text{const}<\infty,\end{aligned} \tag{6.31}$$

by (6.30) and Claim 6.3.

It follows that

$$\begin{aligned}-\int_{\bigcup_n\Gamma_n}w'\widetilde{w}\,\frac{dx}{x^2}=-4\int_{\bigcup_n\Gamma_n}\phi'\tilde{\phi}\,\frac{dx}{x^2}+2\int_{\bigcup_n\Gamma_n}\phi'\tilde{v}\,\frac{dx}{x^2}&\\+2\int_{\bigcup_n\Gamma_n}v'\tilde{\phi}\,\frac{dx}{x^2}-\int_{\bigcup_n\Gamma_n}v'\tilde{v}\,\frac{dx}{x^2}<\infty.&\end{aligned} \tag{6.32}$$

Indeed, arguing like at the end of the proof of the last claim, from the property that $(\tilde{v})^M\in L^1_\Pi$ we deduce that

$$\sum_n|J_n|\frac{\sup_{x\in J_n}|\tilde{v}(x)|}{1+\operatorname{dist}^2(0,J_n)}<\infty.$$

Therefore

$$\left|\int_{(\bigcup_k\Gamma_k)\cap J_n}\phi'\tilde{v}\,\frac{dx}{x^2}\right|\leqslant\int_{(\bigcup_k\Gamma_k)\cap J_n}|\phi'|\,dx\frac{\sup_{x\in J_n}|\tilde{v}(x)|}{1+\operatorname{dist}^2(0,J_n)}\asymp|J_n|\frac{\sup_{x\in J_n}|\tilde{v}(x)|}{1+\operatorname{dist}^2(0,J_n)},$$

and summing over all $n$ we get that

$$\left|\int_{\bigcup_k\Gamma_k}\phi'\tilde{v}\,\frac{dx}{x^2}\right|<\infty.$$

The third integral on the right-hand side of (6.32) is finite because $v'$ is bounded and $\tilde{\phi}=\log|f|$ is in $L^1_\Pi$. The last integral is finite because $v'$ is bounded and $\tilde{v}=\tilde{\phi}-\widetilde{w}\in L^1_\Pi$. The first integral is finite by (6.31).

As before, we put $w_n=w|_{J_n}$. Also let

$$L_n=\bigcup_{\substack{k\\ \operatorname{dist}(J_k,J_n)<\max\{|J_k|,|J_n|\}}} J_k,\quad q_n=w|_{L_n}\quad\text{and}\quad q_n^*=w-q_n.$$

Then

$$\int_{\bigcup_k\Gamma_k} w'\widetilde{w}\,\frac{dx}{x^2}=\sum_n\biggl(\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\tilde{q}_n^*\,\frac{dx}{x^2}+\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\tilde{q}_n\,\frac{dx}{x^2}\biggr).$$

The first integral can be, once again, estimated as in (6.7), i.e. using the property that each $w_j$ is an atom, and the sum of such integrals is shown to be finite. For the second integral we obtain

$$\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\tilde{q}_n\,\frac{dx}{x^2}=\sum_{J_l\subset L_n}\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\widetilde{w}_l\,\frac{dx}{x^2}.$$

Once again, in the "mid-range" case when

$$\min\{|J_n|,|J_l|\}\leqslant\operatorname{dist}(J_n,J_l)<\max\{|J_n|,|J_l|\},$$

we get

$$\biggl|\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\widetilde{w}_l\,\frac{dx}{x^2}\biggr|\lesssim\frac{|J_l|\,|J_n|}{1+\operatorname{dist}^2(J_l,0)},$$

see (6.9). In the case

$$\operatorname{dist}(J_n,J_l)<\min\{|J_n|,|J_l|\},$$

as in part I of the proof, we first notice that

$$\begin{aligned}-\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\widetilde{w}_l\,\frac{dx}{x^2}&=\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\biggl(\int_{J_l}\frac{w(t)\,dt}{t-x}\biggr)\frac{dx}{x^2}\\&=-\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\biggl(\int_{J_l}\log|t-x|w'(t)\,dt\biggr)\frac{dx}{x^2}.\end{aligned}$$

Similarly to the first part of the proof (see the paragraph after (6.11)) one can verify the conditions of Lemma 7.10 and apply the second inequality in (7.5), with $E=\bigl(\bigcup_k\Gamma_k\bigr)\cap J_n$ and $J_l=I$, to obtain

$$\begin{aligned}&-\int_{(\bigcup_k\Gamma_k)\cap J_n} w'\biggl(\int_{J_l}\log|t-x|w'(t)\,dt\biggr)\frac{dx}{x^2}\\&\qquad\gtrsim-\frac{1}{1+\operatorname{dist}^2(J_n,0)}\biggl(\iint_{\bigcup_k(\Gamma_k\cap J_n)\times J_l}\log|t-x|w'(x)w'(t)\,dx\,dt+C|J_n|^2\biggr).\end{aligned}$$

Altogether, we obtain

$$
\begin{aligned}
&-\int_{(\bigcup_k \Gamma_k)\cap J_n} w'\tilde{q}_n \frac{dx}{x^2} \qquad (6.33)\\
&\quad\gtrsim -\frac{1}{1+\operatorname{dist}^2(J_n,0)} \sum_{J_l\subset L_n} \biggl(\int_{(\bigcup_k \Gamma_k)\cap J_n} w'(x)\int_{J_l} \log|x-t|w'(t)\,dt\,dx - C|J_n|\,|J_l|\biggr).
\end{aligned}
$$

Furthermore, because of (6.27) (applied here with $\theta$ in place of $\Phi$),

$$
\begin{aligned}
&-\int_{(\bigcup_k \Gamma_k)\cap J_n} w'(x)\biggl(\sum_{J_l\subset L_n}\int_{J_l} \log|x-t|w'(t)\,dt\biggr)dx\\
&\qquad\gtrsim -\int_{(\bigcup_k \Gamma_k)\cap J_n} w'(x)\biggl(\int_{J_n} \log|x-t|w'(t)\,dt\biggr)dx - \sum_{J_l\subset L_n}|J_l|\,|J_n|.
\end{aligned}
$$

Let us remark right away that

$$
\begin{aligned}
&\sum_n \frac{1}{1+\operatorname{dist}^2(J_n,0)} \sum_{J_l\subset L_n}|J_l|\,|J_n|\\
&\qquad\lesssim \sum_n \sum_{\substack{l<n\\ \operatorname{dist}(J_l,J_n)<\max\{|J_l|,|J_n|\}}} \frac{|J_l|\,|J_n|}{1+\operatorname{dist}^2(J_n,0)} \lesssim \sum_n \frac{|J_n|}{1+\operatorname{dist}^2(J_n,0)} < \infty
\end{aligned}
$$

by the monotonicity and the shortness of $|J_n|$.

To continue the estimates let us split the last integral as

$$
\begin{aligned}
&-\int_{(\bigcup_k \Gamma_k)\cap J_n} w'(x)\biggl(\int_{J_n}\log|x-t|w'(t)\,dt\biggr)dx\\
&\quad= \int_{(\bigcup_k \Gamma_k)\cap J_n} (\arg\theta)'(x)\biggl(\int_{J_n}\log|x-t|(v'(t)+1)\,dt\biggr)dx\\
&\qquad -\int_{(\bigcup_k \Gamma_k)\cap J_n} (\arg\theta)'(x)\biggl(\int_{J_n}\log|x-t|(\arg\theta)'(t)\,dt\biggr)dx\\
&\qquad -\int_{(\bigcup_k \Gamma_k)\cap J_n} (v'(x)+1)\biggl(\int_{J_n}\log|x-t|(v'(t)+1)\,dt\biggr)dx\\
&\qquad +\int_{(\bigcup_k \Gamma_k)\cap J_n} (v'(x)+1)\biggl(\int_{J_n}\log|x-t|(\arg\theta)'(t)\,dt\biggr)dx\\
&\quad=: \mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}.
\end{aligned}
$$

To estimate III and IV denote by $D$ the constant satisfying

$$
\int_{(\bigcup_k \Gamma_k)\cap J_n} (v'(x)+1)\,dx = D|J_n|.
$$

Notice that because $1-2\varepsilon<v'+1<1+2\varepsilon$ and

$$\int_{J_n}(v'(x)+1)\,dx=\int_{J_n}(\arg\theta)'(x)\,dx=|J_n|,$$

for any $y\in J_n$,

$$\int_{J_n}\log|y-t|(v'(t)+1)\,dt=|J_n|\log|J_n|+O(|J_n|)$$

and

$$\begin{aligned}\mathrm{III}&=-\int_{(\bigcup_k\Gamma_k)\cap J_n}(v'(x)+1)\biggl(\int_{J_n}\log|x-t|(v'(t)+1)\,dt\biggr)\,dx\\&=-D|J_n|^2\log|J_n|+O(|J_n|^2).\end{aligned}$$

To estimate IV, observe that for any $t\in J_n$, if $\operatorname{dist}(t,(\lambda_k,\lambda_{k+1}))\geqslant 1$, then

$$\begin{aligned}&\int_{\Gamma_k}(v'(x)+1)\log_+|x-t|\,dx\\&\qquad\geqslant\int_{\Gamma_k}(v'(x)+1)\,dx\,\frac{1}{\lambda_{k+1}-\lambda_k}\int_{\lambda_k}^{\lambda_{k+1}}\log_+|x-t|\,dx-(\lambda_{k+1}-\lambda_k)\log 3\end{aligned}$$

(recall that $\Gamma_k$ is the middle third of $(\lambda_k,\lambda_{k+1})$ and that $\varepsilon$ is very small).

Consider a positive step function $\alpha(x)$ defined on each $(\lambda_k,\lambda_{k+1})$ as

$$\frac{1}{\lambda_{k+1}-\lambda_k}\int_{\Gamma_k}(v'(x)+1)\,dx.$$

Then $\bigl|\alpha-\frac{1}{3}\bigr|\leqslant\varepsilon$ on $J_n$. Hence one can apply Lemma 7.9 (5) to conclude that, for any $t\in J_n$,

$$\begin{aligned}\int_{(\bigcup_k\Gamma_k)\cap J_n}(v'(x)+1)\log_+|x-t|\,dx&\geqslant\int_{J_n}\alpha(x)\log_+|x-t|\,dx-\mathrm{const}|J_n|\\&\geqslant\biggl(\int_{J_n}\alpha(x)\,dx\biggr)\log|J_n|-\mathrm{const}|J_n|\\&=D|J_n|\log|J_n|-\mathrm{const}|J_n|.\end{aligned}$$

Also,

$$-\int_{(\bigcup_k\Gamma_k)\cap J_n}(v'(x)+1)\log_-|x-t|\,dx\geqslant-1-\varepsilon.$$

Therefore

$$\begin{aligned}\mathrm{IV}&=\int_{(\bigcup_k\Gamma_k)\cap J_n}(v'(x)+1)\biggl(\int_{J_n}\log|x-t|(\arg\theta)'(t)\,dt\biggr)\,dx\\&\geqslant\biggl(\int_{J_n}(\arg\theta)'(t)\,dt\biggr)(D|J_n|\log|J_n|-\mathrm{const}|J_n|-\mathrm{const})\\&\geqslant D|J_n|^2\log|J_n|-\mathrm{const}|J_n|^2.\end{aligned}$$

Combining the estimates we get

$$\mathrm{III}+\mathrm{IV}\gtrsim -|J_n|^2.$$

To estimate II notice that

$$\mathrm{II}=-\sum_{\Gamma_k\subset J_n}\int_{\Gamma_k}(\arg\theta)'(x)\,dx\sum_{\lambda_j,\lambda_{j+1}\in J_n}\int_{\lambda_j}^{\lambda_{j+1}}\log|x-t|(\arg\theta)'(t)\,dt.$$

If $t\in(\lambda_j,\lambda_{j+1})$ and $x\in\Gamma_k$ then

$$\log|x-t|\leqslant\begin{cases}\log|\lambda_j-\lambda_{k+1}|, & \text{if } j<k,\\ \log|\lambda_k-\lambda_{j+1}|, & \text{if } j>k,\\ \log|\lambda_{j+1}-\lambda_j|, & \text{if } j=k.\end{cases}$$

Put

$$\alpha_k=\int_{\Gamma_k}(\arg\theta)'(x)\,dx.$$

Then

$$\mathrm{II}\geqslant-\sum_{\Gamma_k\subset J_n}\alpha_k\sum_{\substack{\lambda_j\in J_n\\ j\neq k}}2\pi\log|\lambda_k-\lambda_j|+A_n,$$

where the constants $A_n$ satisfy

$$\sum_n\frac{|A_n|}{1+\operatorname{dist}^2(0,J_n)}<\infty.$$

Using (6.1), I can be rewritten as

$$\mathrm{I}=\sum_{\Gamma_k\subset J_n}\int_{\Gamma_k}(\arg\theta)'(x)\biggl(\int_{J_n}\log|x-t|(v'(t)+1)\,dt\biggr)\,dx=\biggl(\sum_{\Gamma_k\subset J_n}\alpha_k\biggr)|J_n|\log|J_n|+B_n,$$

where again

$$\sum_n\frac{|B_n|}{1+\operatorname{dist}^2(0,J_n)}<\infty.$$

By Lemma 7.8 (4),

$$\alpha_k=\int_{\Gamma_k}(\arg\theta)'(x)\,dx>c>0$$

for all $k$. Therefore, since there are $|J_n|/2\pi$ intervals $\Gamma_k$ in $J_n$,

$$\begin{aligned}\mathrm{I}+\mathrm{II}&=\biggl(\sum_{\Gamma_k\subset J_n}\alpha_k\biggr)|J_n|\log|J_n|-\sum_{\Gamma_k\subset J_n}\alpha_k\sum_{\substack{\lambda_j\in J_n\\ j\neq k}}2\pi\log|\lambda_k-\lambda_j|+A_n+B_n\\ &\gtrsim\frac{1}{2\pi}|J_n|^2\log|J_n|-\sum_{\substack{\lambda_j,\lambda_k\in J_n\\ j\neq k}}2\pi\log|\lambda_k-\lambda_j|+A_n+B_n.\end{aligned}$$

Now, going back to (6.33), we obtain

$$-\sum_n \int_{(\bigcup_k \Gamma_k)\cap J_n} w'\widetilde{w}_n \frac{dx}{x^2} \gtrsim \sum_n \frac{1}{1+\operatorname{dist}^2(J_n,0)} \Biggl( \frac{1}{4\pi^2}|J_n|^2 \log|J_n| - \sum_{\substack{\lambda_j,\lambda_k\in J_n\\ j\neq k}} \log|\lambda_k-\lambda_j| - |J_n|^2 - |A_n| - |B_n| \Biggr) + \text{const.}$$

The sum on the right-hand side is positive infinite, because otherwise $\Lambda$ would satisfy the energy condition (4.5) and $\mathcal{C}_X$ would be at least $1/2\pi$. This contradicts (6.32).

It remains to discuss the case when $\sup_n(\lambda_n-\lambda_{n-1})=\infty$. It can be reduced to the previous case by adding a sequence of small density to $\Lambda$.

Indeed, if $\Lambda$ is a sequence with arbitrarily large gaps, choose a large constant $C$ and consider the set of all gaps $R_k$ of $\Lambda$ of size larger than $C$:

$$R_k = (\lambda_{n_k}, \lambda_{n_k+1}), \quad \lambda_{n_k+1}-\lambda_{n_k} > C.$$

After that, one can add a separated set of points in every $R_k$ and consider a slightly larger sequence $\Lambda'=\{\lambda'_n\}_n\supset\Lambda$ that satisfies $\sup_n(\lambda'_n-\lambda'_{n-1})\leqslant C$ and

$$\inf_{\lambda'_n,\lambda'_{n-1}\in\Lambda'\setminus\Lambda}(\lambda'_n-\lambda'_{n-1}) \geqslant \tfrac{1}{2}C.$$

Since $\mathcal{C}_\Lambda<1/2\pi$, for large enough $C$ the sequence $\Lambda'$ will still satisfy $\mathcal{C}_{\Lambda'}<1/2\pi$.

The inner function $\theta$, given by Lemma 7.8, should then be chosen for the sequence $\Lambda'$ instead of $\Lambda$. Recall that $\nu$ is the measure supported on $\Lambda$, chosen at the beginning of the second part of the proof. It has a spectral gap of size 1 and its Cauchy integral $K\nu$ does not have any zeros. Consider the function

$$h=(1-\theta)K\nu\in K_\theta^{1,\infty}.$$

Since $\nu$ had a spectral gap of size 1, $h$ is divisible by $S$ (and $h/S$ is outer because $K\nu$ has no zeros). As $1-\theta$ has simple zeros at $\Lambda'$ and $K\nu$ has simple poles at $\Lambda$, $h$ has zeros at $\Upsilon=\Lambda'\setminus\Lambda$. Without loss of generality, $\Upsilon$ has bounded gaps. Since $\Upsilon$ is a separated sequence, there exists an inner function $I$, with $\operatorname{spec}_I=\Upsilon$, such that $(\arg I)'$ is bounded (by Lemma 7.8 below). If $C$ is large enough, $|(\arg I)'|\ll\varepsilon$.

Then the function

$$g=\frac{Ih}{1-I}$$

is divisible by $S$ and satisfies

$$\bar\theta g=\bar\theta\frac{Ih}{1-I}=\bar\theta h\frac{1}{1-\bar I}$$

on $\mathbb{R}$. Since the last function is antianalytic, $g \in K_\theta^+$. At the same time, $g$ no longer has zeros on $\mathbb{R}$. Denote $f = g/IS$. Then $f \in N^+[\bar{\theta}S]$ is an outer function whose argument on $\mathbb{R}$ is equal to $\frac{1}{2}(\arg\theta - x - \arg I)$. Now we can apply Claim 6.1 to $u = \arg\theta - x - \arg I$ to obtain functions $v = v_1 + v_2$ satisfying properties (1)–(5).

If one denotes by $\Gamma_n$ the middle third of $(\lambda'_n, \lambda'_{n+1})$, then similarly to (6.30),

$$|S^{-1}(x)h(x)| = |(1-\theta(x))K\nu(x)| \leqslant \frac{6\|\nu\|}{|\lambda'_{n+1} - \lambda'_n|}.$$

The argument of the function $h/S$ is $\arg\theta - x$. Note that Claim 6.3 still holds with $\Lambda'$ in place of $\Lambda$, because $\Upsilon$ is separated. Hence (6.31) still holds for $\phi = \arg\theta - x$.

After that, using that $|(\arg I)'| \ll \varepsilon$, one can "absorb" $\arg I$ into $v_1$ and replace $v_1$ by $y = v_1 + \arg I$. The remaining estimates, starting with (6.32), can be repeated with $v = y + v_2$ in place of $v = v_1 + v_2$, i.e. with $w = \arg\theta - x - y - v_2$.

## 7. Used technicalities

This section contains several lemmas and corollaries used in the previous sections.

If $\Lambda$ is a real sequence we define its (exterior) *Beurling–Malliavin density* as

$$d_{\mathrm{BM}}(\Lambda) = \sup\{d : \text{there is a long sequence } \{I_n\}_n \text{ so that } \#(\Lambda \cap I_n) \geqslant d|I_n| \text{ for all } n\}$$

if $\Lambda$ is discrete, and $d_{\mathrm{BM}}(\Lambda) = \infty$ otherwise.

An equivalent definition is given in [20, §4.6]:

$$d_{\mathrm{BM}}(\Lambda) = \sup\{a : N[\bar{S}^{2\pi a}\theta] = 0\},$$

where $\theta(z)$ denotes some/any meromorphic inner function with $\mathrm{spec}_\theta = \Lambda$. Note that the Beurling–Malliavin multiplier theorem implies that $N[\bar{S}^{2\pi a}\theta]$ in the above definition can be replaced by any $N^p[\bar{S}^{2\pi a}\theta]$, $0 < p \leqslant \infty$, the kernel in the Hardy space $H^p$, or by $N^+[\bar{S}^{2\pi a}\theta]$, the kernel in the Smirnov class, see [20, §4.2].

LEMMA 7.1. *Let $X \subset \mathbb{R}$ be a closed set and let $\Lambda$ be a discrete sequence. Then*

$$G_{X\cup\Lambda} \leqslant G_X + 2\pi d_{\mathrm{BM}}(\Lambda).$$

*Proof.* Let $d_{\mathrm{BM}}(\Lambda) = d_1$, $G_X = d_2$ and $G_{X\cup\Lambda} = d_3$. Let $\varepsilon > 0$ be a small number. By Theorem 3.2, $N[\bar{\theta}S^{d_3-\varepsilon}] \neq 0$ for some meromorphic inner $\theta$, $\mathrm{spec}_\theta \subset X \cup \Lambda$. Let

$$f \in N[\bar{\theta}S^{d_3-\varepsilon}].$$

Let $I$ be an inner function such that $\operatorname{spec}_I=\Lambda$.

By the above definition of the Beurling–Malliavin density, there exists a function

$$g\in N^\infty[\bar{S}^{2\pi d_1+\varepsilon}I].$$

Then the function $h=(1-I)g$ belongs to $N^\infty[\bar{S}^{2\pi d_1+\varepsilon}]$ and is equal to 0 on $\Lambda$. The function $fh$ belongs to $N[\bar{\theta}S^{d_3-2\pi d_1-2\varepsilon}]$ and is zero on $\Lambda$ (obviously, we assume that $d_3-2\pi d_1-2\varepsilon>0$). Finally, the function

$$l=S^{d_3-2\pi d_1-2\varepsilon}fh$$

belongs to $N[\bar{\theta}]=K_\theta$ and is still zero on $\Lambda$. By the Clark representation,

$$l=\frac{1}{2\pi i}(1-\theta)Kl\sigma,$$

where $\sigma$ is the Clark measure for $\theta$, with $\operatorname{supp}\sigma=\operatorname{spec}_\theta\subset\Lambda\cup X$. Since $l$ is divisible by $S^{d_3-2\pi d_1-2\varepsilon}$ in $\mathbb{C}_+$ and $1-\theta$ is an outer function in $\mathbb{C}_+$, $Kl\sigma$ is divisible by $S^{d_3-2\pi d_1-2\varepsilon}$ in $\mathbb{C}_+$. Equivalently, the measure $l\sigma$ has a spectral gap of size $d_3-2\pi d_1-2\varepsilon$. As $l$ is zero on $\Lambda$, the measure $l\sigma$ is supported on $X$. Hence

$$G_X\geqslant d_3-2\pi d_1-2\varepsilon=G_{X\cup\Lambda}-2\pi d_{\mathrm{BM}}(\Lambda)-2\varepsilon. \qquad \square$$

The following statement can be viewed as a version of the first Beurling–Malliavin theorem, see [20] and [21].

Lemma 7.2. *Let $\Lambda$ be a real sequence. Suppose that there exists a long sequence of intervals $I_n$ such that*

$$\#(\Lambda\cap I_n)\leqslant a|I_n| \tag{7.1}$$

*for all $n$, for some $a\geqslant 0$. Then $G_\Lambda\leqslant 2\pi a$.*

*Proof.* Suppose that $G_\Lambda=2\pi a+3\varepsilon$ for some $\varepsilon>0$. Then, by Theorem 3.2,

$$N[\bar{\theta}S^{2\pi a+2\varepsilon}]\neq 0$$

for some inner function $\theta$, with $\operatorname{spec}_\theta\subset\Lambda$. But (7.1) implies that the argument of the symbol increases greatly on $I_n$, which leads to a contradiction. More precisely, let

$$\gamma=\arg(\bar{\theta}S^{2\pi a+2\varepsilon})=(2\pi a+2\varepsilon)x-\arg\theta.$$

For each $I_n=(a_n,a_{n+1}]$ let

$$\delta_n=\inf_{I_n''}\gamma-\sup_{I_n'}\gamma,$$

where

$$I_n'=\left(a_n,a_n+\frac{\varepsilon|I_n|}{6(\pi a+\varepsilon)}\right) \quad\text{and}\quad I_n''=\left(a_{n+1}-\frac{\varepsilon|I_n|}{6(\pi a+\varepsilon)},a_{n+1}\right).$$

Then (7.1) implies that $\delta_n\geqslant\frac{1}{3}\varepsilon|I_n|$. Hence, by a version of the theorem in [20, §4.4], $N[\bar{\theta}S^{a+2\varepsilon}]$ has to be trivial. $\square$

Lemma 7.3. *Let* $I=[a,b]$ *be an interval on* $\mathbb{R}$ *and let* $\Lambda=\{\lambda_1,...,\lambda_N\}$ *be a set of points on* $I$, *with* $a\leqslant\lambda_1<...<\lambda_N\leqslant b$. *Let* $C>1$ *be a constant and suppose that for some subinterval* $J=[c,d]\subset I$,

$$\#(\Lambda\cap J)\leqslant\frac{|J|}{C}-1.$$

*Then one can spread the points of* $\Lambda$ *on* $J$ *without a large decrease in the energy* $E(\Lambda)$. *More precisely, if* $\Gamma=\{\gamma_1,...,\gamma_N\}$, $a\leqslant\gamma_1<...<\gamma_N\leqslant b$, *is another set of points on* $I$ *with the properties*

(1) $\gamma_k=\lambda_k$ *for all* $k$ *such that* $\lambda_k\notin J$;

(2) $c+C\leqslant\gamma_k\leqslant\gamma_{k+1}\leqslant d-C$ *and* $\gamma_{k+1}-\gamma_k\geqslant C$ *for all* $\gamma_k,\gamma_{k+1}\in J$;

*then*

$$E(\Gamma)\geqslant E(\Lambda)-D\frac{|J|N}{C}$$

*for some absolute constant* $D$, *where* $E$ *is defined by* (4.1).

*Proof.* Notice that

$$\sum_{\substack{\gamma_k\in I\\ \gamma_j\in J}}\log_-|\gamma_k-\gamma_j|=0.$$

If $\Gamma_1=\Gamma\setminus\Lambda$ and $\Gamma_2=\Gamma\cap\Lambda$ then $\#\Gamma_1<|J|/C$. Suppose that $\gamma_k=\lambda_k\in\Gamma_2$, $\gamma_k<c$. Then

$$\begin{aligned}\sum_{\gamma_j\in\Gamma_1}\log_+|\lambda_j-\lambda_k|-\sum_{\gamma_j\in\Gamma_1}\log_+|\gamma_j-\lambda_k| &\leqslant(\#\Gamma_1)\log_+|d-\lambda_k|-\sum_{\gamma_j\in\Gamma_1}\log_+|\gamma_j-\lambda_k|\\ &\leqslant(\#\Gamma_1)\log_+|d-c|-\sum_{\gamma_j\in\Gamma_1}\log_+|\gamma_j-c|\\ &\leqslant(\#\Gamma_1)\log_+|J|-\sum_{k=1}^{\#\Gamma_1}\log_+(kC)\\ &<\frac{|J|}{C},\end{aligned}$$

by Stirling's formula. The cases $\gamma_k>d$ and $\gamma_k\in\Gamma_1$ can be treated similarly. □

Corollary 7.4. *Let* $\Lambda$ *be a sequence of real points that satisfies the density and energy conditions* (4.4) *and* (4.5) *for some short partition* $I_n$ *and some* $a>0$. *Let* $C>1$. *Let* $J_k$ *be a sequence of disjoint intervals such that, for every* $k$, $J_k\subset I_n$ *for some* $n$ *and*

$$\#(\Lambda\cap J_k)\leqslant\frac{|J_k|}{C}-1$$

*for all* $k$. *Let* $\Gamma$ *be a sequence of points obtained from* $\Lambda$ *by spreading the points on each interval* $J_k$ *as in the last lemma. Then* $\Gamma$ *satisfies the density and energy conditions with the same partition* $I_n$ *and* $a$.

COROLLARY 7.5. *Let* $\Lambda=\{\lambda_n\}_n$ *be an increasing discrete sequence of real points such that* $\mathcal{C}_\Lambda\geqslant a>0$. *Then for every* $\varepsilon>0$ *there exists an increasing discrete sequence* $\Gamma=\{\gamma_n\}_n$ *such that*

(1) $\mathcal{C}_\Gamma\geqslant a$;

(2) $d_{\mathrm{BM}}(\Gamma\setminus\Lambda)<\varepsilon$;

(3) $\sup_n(\gamma_{n+1}-\gamma_n)<\infty$.

*Proof.* Choose $C>0$ so that $1/C\ll d$ and $1/C\ll\varepsilon$. Let $[\lambda_{n_k},\lambda_{n_k+1}]$ be a sequence of all "gaps" of $\Lambda$ satisfying $\lambda_{n_k+1}-\lambda_{n_k}>C$.

Since $\mathcal{C}_\Lambda\geqslant a$, there exists a partition $I_n$ such that $\Lambda$ satisfies (4.4) and (4.5) for $I_n$ and $a$. One can choose a sequence of disjoint intervals $J_k$ such that, for every $k$, $J_k\subset I_n$ for some $n$,

$$\bigcup_k[\lambda_{n_k},\lambda_{n_k+1}]\subset\bigcup_k J_k \quad\text{and}\quad \frac{|J_k|}{2C}\leqslant\#(\Lambda\cap J_k)\leqslant\frac{|J_k|}{C}-1 \text{ for all } k.$$

(The choice of the intervals $J_k$ can be made by a version of the "shading" algorithm, see for example [15, Volume 2, pp. 507–508].) Let $\Gamma$ be a sequence of points obtained from $\Lambda$ by spreading the points on each interval $J_k$ as in Lemma 7.3. Then (1) is satisfied by the previous corollary and the supremum in (3) is at most $2C$. Since the distances between the points of $\Gamma$ on $\bigcup_k J_k$ are at least $C$, we have

$$d_{\mathrm{BM}}(\Gamma\setminus\Lambda)\leqslant\frac{1}{C}<\varepsilon. \qquad\square$$

LEMMA 7.6. *Let* $\Lambda$ *be a sequence of real points and let* $\{I_n\}_n$ *be a short partition such that* $\Lambda$ *satisfies*

$$a|I_n|<\#(\Lambda\cap I_n)$$

*for all* $n$ *with some* $a>0$ *and the energy condition* (4.5) *on* $\{I_n\}_n$. *Then for any short partition* $\{J_n\}_n$, *there exists a subsequence* $\Gamma\subset\Lambda$ *that satisfies*

$$\#((\Lambda\setminus\Gamma)\cap J_n)=o(|J_n|)$$

*as* $n\to\pm\infty$, *and the energy condition* (4.5) *on* $\{J_n\}_n$.

*Proof.* To simplify the estimates, we will assume that the endpoints of $I_n$ belong to $\Lambda$, i.e. that $I_n=(\lambda_{k_n},\lambda_{k_{n+1}}]$ for each $n$, and that the energy condition (4.5) is satisfied on $I_n$ with $E_n$ defined by (6.1), see the explanation there.

(To include the endpoints in $E_n$, one may need to compensate by deleting a point on each $I_n$, as explained in the beginning of the proof of Theorem 4.7. Here one may need to pass from $\Lambda$ to a subsequence $\Gamma$. As $|I_n|\to\infty$, $\Gamma$ will satisfy $\#((\Lambda\setminus\Gamma)\cap J_n)=o(|J_n|)$.)

We will also assume that $\#(\Lambda\cap I_n)=|I_n|$ for all $n$. In this case one can choose $\Gamma=\Lambda$. Fix $n$ and suppose that the intervals $I_l, ..., I_{l+N}$ cover $J_n$. To estimate the energy expression for $J_n$ let us first consider the case when $\bigcup_{j=l}^{l+N} I_j=J_n$. Denote by $u$ a continuous, piecewise linear function on $J_n$, which is zero at the left endpoint of $J_n$ and grows linearly by 1 between each pair of points $\lambda_n$ and $\lambda_{n+1}$. Let

$$p(x)=\begin{cases} u(x)-x+\lambda_{k_l}, & \text{on } J_n=(\lambda_{k_l}, \lambda_{k_{l+N+1}}],\\ 0, & \text{on } \mathbb{R}\setminus J_n.\end{cases}$$

Then $p(\lambda_{k_n})=0$ for all $l\leqslant n\leqslant l+N+1$. Denote by $p_n$ the restriction $p|_{I_n}$.

On each $(\lambda_j, \lambda_{j+1})$ the function $u'$ satisfies the same estimates as $|\theta'|$ from the statement of Lemma 7.8. Therefore for the function $p$ one can apply the same argument as in the first part of the proof of Theorem 4.7, where $p$ was defined as $\arg\theta-x-v_1$ (we will simply assume that $v_1\equiv 0$).

First, one can show that

$$\begin{aligned}-\iint_{J_n\times J_n} &\log|t-x|p'(t)p'(x)\,dt\,dx\\ &=|J_n|^2\log|J_n|-\sum_{\substack{\lambda_{k_l}\leqslant\lambda_j,\lambda_{j'}\leqslant\lambda_{k_{l+N+1}}\\ \lambda_j\neq\lambda_{j'}}}\log|\lambda_j-\lambda_{j'}|+\text{const}|J_n|^2.\end{aligned}$$

To estimate the last integral, rewrite it as

$$-\iint_{J_n\times J_n}\log|t-x|p'(t)p'(x)\,dt\,dx=\sum_{I_j\subset J_n}\sum_{I_{j'}\subset J_n}\biggl(-\iint_{I_j\times I_{j'}}\log|t-x|p'(t)p'(x)\,dt\,dx\biggr).$$

For the last integral, when $j=j'$, by (6.16) and (6.17) we have

$$-\iint_{I_j\times I_j}\log|t-x|p'(t)p'(x)\,dt\,dx\lesssim|I_j|^2\log|I_j|-E_j+\text{const}|I_j|^2.$$

As usual, we assume that $|I_n|\gg 1$. If $I_j$ does not intersect $2I_{j'}$, then

$$\operatorname{dist}(I_j, I_{j'})\geqslant|I_{j'}|>1$$

and the integral over $I_j\times I_{j'}$ can be estimated by first noticing that the $\log_-$ part is zero, because $\log_-|x-t|=0$ when $|x-t|>1$, as in (6.18). For the $\log_+$ part we have (6.19). Altogether we obtain

$$-\iint_{I_j\times I_{j'}}\log|t-x|p'(t)p'(x)\,dt\,dx\lesssim|I_j|\,|I_{j'}|.$$

For the case when $I_j$ intersects $2I_{j'}$ but is not contained in $2I_{j'}$, or when $I_j$ is adjacent to $I_{j'}$ (note that there are at most four such $I_j$ for each $I_{j'}$), we can estimate the integral as in (6.21) to conclude that

$$-\iint_{I_j\times I_{j'}} \log|t-x|p'(t)p'(x)\,dt\,dx \\ \lesssim (|I_j|^2\log|I_j|-E_j)+(|I_{j'}|^2\log|I_{j'}|-E_{j'})+|I_j|^2+|I_{j'}|^2.$$

Finally, in the case when $I_j\subset 2I_{j'}$, $j'>j+1$ (the case $j'<j-1$ is similar), again we can use that $\operatorname{dist}(x,I_{j'})\geqslant|I_{j+1}|>1$ to skip the estimates of $\log_-$. The $\log_+$ part can be estimated by the integral over $I_{j'}\times I_{j'}$. Notice that

$$\begin{aligned}
-\iint_{I_{j'}\times I_{j'}} &\log_+|s-t|p'(t)p'(s)\,dt\,ds \\
&=2\int_{I_{j'}} u'(s)\int_{I_{j'}}\log_+|s-t|\,dt\,ds-\iint_{I_{j'}\times I_{j'}}\log_+|s-t|\,dt\,ds \\
&\quad-\int_{I_{j'}} u'(s)\int_{I_{j'}}\log_+|s-t|u'(t)\,dt\,ds \\
&\geqslant |I_{j'}|^2\log|I_{j'}|-|I_{j'}|\int_{I_{j'}}\log_+|x-t|u'(t)\,dt+\mathrm{const}|I_{j'}|^2,
\end{aligned}$$

because $\int_{I_{j'}}u'(t)\,dt=|I_{j'}|$. Also, for any $x\in I_j$ (recall that $j'>j+1$),

$$\int_{I_{j'}}\log_+|x-t|\,dt-\int_{I_{j'}}\log_+|x-t|u'(t)\,dt \\ \geqslant (|I_{j'}|\log(\lambda_{k_{j'+1}}-x)-|I_{j'}|)-\log(\lambda_{k_{j'+1}}-x)\int_{I_{j'}}u'(t)\,dt\geqslant -|I_{j'}|.$$

Therefore

$$\begin{aligned}
-\iint_{I_j\times I_{j'}} &\log_+|t-x|p'(t)p'(x)\,dt\,dx \\
&\leqslant \int_{I_j}|p'(x)|\left(\int_{I_{j'}}\log_+|x-t|\,dt-\int_{I_{j'}}\log_+|x-t|u'(t)\,dt+\mathrm{const}|I_{j'}|\right)dx \\
&\leqslant 2|I_j|\left(|I_{j'}|\log|I_{j'}|-\int_{I_{j'}}\log_+|x-t|u'(t)\,dt\right)+\mathrm{const}|I_j|\,|I_{j'}| \\
&\leqslant -2\frac{|I_j|}{|I_{j'}|}\iint_{I_{j'}\times I_{j'}}\log_+|s-t|p'(t)p'(s)\,dt\,ds+\mathrm{const}|I_j|\,|I_{j'}| \\
&=2\frac{|I_j|}{|I_{j'}|}\|p_{j'}\|_{\mathcal{D}}+\mathrm{const}|I_j|\,|I_{j'}| \\
&\lesssim \frac{|I_j|}{|I_{j'}|}(|I_{j'}|^2\log|I_{j'}|-E_{j'})+|I_j|\,|I_{j'}|.
\end{aligned}$$

Combining the estimates and using the shortness of $\{J_n\}_n$, we obtain that $\Lambda$ satisfies the energy condition on $\{J_n\}_n$.

In the case when the intervals $I_l, ..., I_{l+N}$ cover $J_n$ but $\bigcup_{j=l}^{l+N} I_j \neq J_n$, i.e. when $I_l, I_{l+N} \cap J_n \neq \varnothing$ but at least one of $I_l$ and $I_{l+N}$ is not a subset of $J_n$, set $I_l^* = I_l \cap J_n$ and $I_{l+N}^* = I_{l+N} \cap J_n$. Notice that, by Remark 4.3 and the fact that $\log |I_l^*| < \log |I_l|$,

$$|I_l^*|^2 \log |I_l^*| - E_l^* \leqslant |I_l|^2 \log |I_l| - E_l.$$

Similarly,

$$|I_{l+N}^*|^2 \log |I_{l+N}^*| - E_{l+N}^* \leqslant |I_{l+N}|^2 \log |I_{l+N}| - E_{l+N}.$$

Now we can use the previous case with $I_l^*$ and $I_{l+N}^*$ in place of $I_l$ and $I_{l+N}$, respectively. $\square$

Corollary 7.7. *Let $\Lambda$ be a sequence of real points and let $\{I_n\}_n$ be a short partition such that $\Lambda$ satisfies the density condition* (4.4), *with some $a>0$, and the energy condition* (4.5). *Then for any $\varepsilon>0$ there exists a subsequence $\Gamma \subset \Lambda$ and a short monotone partition $J_n$ such that $\Gamma$ satisfies* (4.4), *with $a-\varepsilon$ in place of $a$, and* (4.5) *on $J_n$.*

*Proof.* One can choose a short monotone partition $\{J_n\}_n$ satisfying

$$\left(a - \tfrac{1}{2}\varepsilon\right)|J_n| \leqslant \#(\Lambda \cap J_n)$$

for all $n$. Such a partition can be constructed in the same way as in the second part of the proof of Theorem 4.7, see the second paragraph before Claim 6.3. Then $\Gamma$ can be found by Lemma 7.6. $\square$

Lemma 7.8. *Let $A = \{a_n\}_{n\in\mathbb{Z}}$ be a real sequence satisfying*

$$a_n < a_{n+1}, \quad a_{n+1} - a_n < C < \infty$$

*and $a_n \to \pm\infty$ as $n \to \pm\infty$. Set $I_n = (a_n, a_{n+1})$ and let $J_n$ be the middle third of $I_n$. Then there exists an inner function $\theta$ satisfying*

(1) $\operatorname{spec}_\theta = A$;

(2) $|\theta'| \lesssim |I_n|^{-2}$ *on $J_n$ for all $n$*;

(3) $|\theta'| \lesssim \min\{|I_{n-1}|, |I_n|, |I_{n+1}|\}^{-1}$ *on the rest of $I_n$ for all $n$*;

(4) $\int_{J_n} (\arg\theta)'(x)\, dx \geqslant c$ *for some $c>0$ and all $n$.*

*Proof.* Define a second sequence $B = \{b_n\}_{n\in\mathbb{Z}}$ as the sequence of midpoints of complementary intervals of $A$ in $\mathbb{R}$, i.e. $b_n = \frac{1}{2}(a_n + a_{n+1})$.

Define the inner function $\theta$ to satisfy

$$\frac{1-\theta}{1+\theta} = \mathrm{const}\, e^{Ku}, \tag{7.2}$$

where $u=1_E-\frac{1}{2}$,

$$E=\bigcup_k (a_k, b_k),$$

and $Ku$ is the *improper* integral

$$Ku(z)=\int \frac{u(t)\,dt}{t-z}, \quad z\in\mathbb{C}_+.$$

The integral converges, since $u$ is a convergent sum of atoms $u|_{[a_n,a_{n+1}]}$.

(Formulas similar to (7.2) are often used in perturbation theory. In those settings, $u$ is the Kreĭn–Lifshits shift function and $\theta$ is the characteristic function of the perturbed operator, see for example [27] and [31].)

Let $\mu_1$ and $\mu_{-1}$ be the Clark measures for $\theta$ defined by the Herglotz representations

$$\frac{1+\theta}{1-\theta} = \frac{1}{\pi i}\int_{\mathbb{R}} \left(\frac{1}{t-z} - \frac{t}{1+t^2}\right) d\mu_1(t) + \mathrm{const},$$
$$\frac{1-\theta}{1+\theta} = \frac{1}{\pi i}\int_{\mathbb{R}} \left(\frac{1}{t-z} - \frac{t}{1+t^2}\right) d\mu_{-1}(t) + \mathrm{const}.$$

The measures $\mu_1$ and $\mu_{-1}$ have the form

$$\mu_1=\sum_n \alpha_n\delta_{a_n} \quad \text{and} \quad \mu_{-1}=\sum_n \beta_n\delta_{b_n}$$

for some positive numbers $\alpha_n$ and $\beta_n$. (It is easy to see that $\mu_{\pm 1}(\{\infty\})=0$ although we do not actually need this fact.)

Put $\delta_n=a_{n+1}-a_n$. We claim that

$$\delta_n^2 \lesssim \beta_n \lesssim \delta_n. \tag{7.3}$$

Assuming that this estimate holds, we could finish as follows. Since

$$|\theta'| \asymp |1-\theta|^2\,|(\mathcal{S}\mu_1)'| \quad \text{and} \quad |\theta'| \asymp |1+\theta|^2\,|(\mathcal{S}\mu_{-1})'|,$$

we have

$$|\theta'(x)| \asymp \min\left\{\sum_n \frac{\alpha_n}{(x-a_n)^2}, \sum_n \frac{\beta_n}{(x-b_n)^2}\right\}, \quad x\in\mathbb{R}.$$

Now if $x$ belongs to the middle third of one of the intervals $(a_m, a_{m+1})$, then $|\theta'(x)|$ can be estimated as

$$|\theta'(x)| \lesssim \sum_n \frac{\alpha_n}{(x-a_n)^2} \asymp \sum_n \frac{\alpha_n}{(b_m-a_n)^2} = \frac{1}{\beta_m}$$

and the estimate in part (2) follows from the left half of (7.3).

On the rest of the interval $|\theta'(x)|$ can be estimated by

$$\sum_n \frac{\beta_n}{(x-b_n)^2}$$

which together with the right half of (7.3) gives the desired estimate.

To establish part (4), notice that if $|1-\theta|>|1+\theta|$ on $J_n$, then

$$|\theta'(x)| \asymp \sum_n \frac{\alpha_n}{(x-a_n)^2} \asymp \sum_n \frac{\alpha_n}{(b_m-a_n)^2} = \frac{1}{\beta_m},$$

which implies that $(\arg\theta)' \gtrsim 1/\delta_n$, by the right half of (7.3). This implies the inequality for the integral. If, however, $|1-\theta| \leqslant |1+\theta|$ at some point $c_n \in J_n$, then the integral taken between $c_n$ and $b_n$ is at least $\frac{1}{4}\pi$. Since $(\arg\theta)'>0$, we again obtain the desired estimate.

It remains to prove (7.3). As follows from (7.2),

$$\beta_n = \operatorname{const} \operatorname{Res}_{b_n} e^{-Ku}.$$

Let

$$g_n(z) = \exp\left(-\int_{a_n}^{a_{n+1}} \frac{u(t)}{t-z}\,dt\right) = \frac{\sqrt{(a_n-z)(a_{n+1}-z)}}{b_n-z}$$

and

$$A_n = \exp\left(-\int_{\mathbb{R}\setminus(a_n,a_{n+1})} \frac{u(t)}{t-b_n}\,dt\right),$$

so

$$\operatorname{Res}_{b_n} e^{-Ku} = A_n \operatorname{Res}_{b_n} g_n \quad \text{and} \quad |\operatorname{Res}_{b_n} g_n| = \tfrac{1}{2}\delta_n.$$

To prove the right half of (7.3) notice that $A_n \lesssim 1$. Indeed, to the right from $a_{n+1}$, on each $(a_j, a_{j+1})$ the function $u$ is positive on the half of the interval which is closer to $b_n$ and negative on the half which is further from it. Thus

$$-\int_{a_{n+1}}^{\infty} \frac{u(t)}{t-b_n}\,dt < 0.$$

Similarly

$$-\int_{-\infty}^{a_n} \frac{u(t)}{t-b_n}\,dt < 0.$$

To prove the left half of (7.3) one needs to show that $\delta_n \lesssim A_n$. Notice that, since $\delta_n < C$,

$$-\sum_{\operatorname{dist}(b_n,(a_j,a_{j+1}))\geqslant 1} \int_{a_j}^{a_{j+1}} \frac{u(t)}{t-b_n}\,dt > \mathrm{const} > -\infty.$$

As for the remaining part,

$$-\sum_{0<\operatorname{dist}(b_n,(a_j,a_{j+1}))\leqslant 1} \int_{a_j}^{a_{j+1}} \frac{u(t)}{t-b_n}\,dt > -\int_{\delta_n/2}^{1+C} \frac{dx}{x} = \log\delta_n + \mathrm{const}. \qquad \square$$

Our next lemma can be easily verified. We state it without a proof.

LEMMA 7.9. *Let $a_1<a_2$ and $b_1<b_2$ be points on the real line. Let $\alpha$ and $\beta$ be non-negative functions on the intervals $(a_1,a_2)$ and $(b_1,b_2)$ correspondingly satisfying*

$$\int_{a_1}^{a_2} \alpha(t)\,dt = \int_{b_1}^{b_2} \beta(t)\,dt = 1, \quad \alpha < A \quad \text{and} \quad \beta < B,$$

*where $A, B>1$ are constants. Then*

(1)

$$\log_-(a_2-a_1) \leqslant \int_{a_1}^{a_2}\int_{a_1}^{a_2} \log_-(x-y)\alpha(x)\alpha(y)\,dx\,dy \leqslant \log_- \frac{1}{A}+1.$$

(2) *If $a_2<b_1$ then*

$$\log_-(b_2-a_1) \leqslant \int_{a_1}^{a_2}\int_{b_1}^{b_2} \log_-(x-y)\alpha(x)\beta(y)\,dx\,dy \leqslant \log_-(b_1-a_2).$$

(3) *If $a_2=b_1$ then*

$$\int_{a_1}^{a_2}\int_{b_1}^{b_2} \log_-(x-y)\alpha(x)\beta(y)\,dx\,dy \leqslant \min\left\{\log_- \frac{1}{A}, \log_- \frac{1}{B}\right\}+1.$$

(4) *If $a_2\leqslant b_1$ then*

$$\log_+(b_1-a_2) \leqslant \int_{a_1}^{a_2}\int_{b_1}^{b_2} \log_+(x-y)\alpha(x)\beta(y)\,dx\,dy \leqslant \log_+(b_2-a_1).$$

(5) *If $\frac{1}{2}A\leqslant\alpha(x)\leqslant A$ on $(a_1,a_2)$ then for any $y\in(a_1,a_2)$,*

$$\log_+|a_2-a_1|-C \leqslant \int_{a_1}^{a_2} \log_+(x-y)\alpha(x)\,dx \leqslant \log_+|a_2-a_1|+C$$

*for some absolute constant $C$.*

(6) *If $\frac{1}{2}A\leqslant\alpha(x)\leqslant A$ on $(a_1,a_2)$ then for any $y>a_2$,*

$$\log_+|y-a_1|-C \leqslant \int_{a_1}^{a_2} \log_+(x-y)\alpha(x)\,dx \leqslant \log_+|y-a_1|$$

*for some absolute constant $C$.*

Lemma 7.10. *Let* $E\subset\mathbb{R}$ *be a set and let* $I\subset\mathbb{R}$ *be an interval such that*

$$d=\operatorname{dist}(E\cup I,0)>0 \quad \textit{and} \quad E\cup I\subset[d,2d].$$

*Suppose that the function* $h$ *on* $E\cup I$ *satisfies* $h=f-g$, *where*

$$f>0 \quad \textit{and} \quad |g-1|<\varepsilon \quad \textit{on } E\cup I$$

*for some constant* $0<\varepsilon<\frac{1}{3}$ *and*

$$\int_I f(x)\,dx=\int_I g(x)\,dx. \tag{7.4}$$

*Write*

$$\int_E |h(x)|\,dx=D_1 \quad \textit{and} \quad \int_E\left|\int_I \log|t-x|h(t)\,dt\right|dx=D_2.$$

*Then*

$$\begin{aligned} \frac{1}{d^2}&\left(\iint_{E\times I}\log|t-x|h(x)h(t)\,dx\,dt-DD_1|I|-2D_1-4D_2\right)\\ &\leqslant \int_E h(x)\left(\int_I \log|t-x|h(t)\,dt\right)\frac{dx}{x^2}\\ &\leqslant \frac{4}{d^2}\left(\iint_{E\times I}\log|t-x|h(x)h(t)\,dx\,dt+DD_1|I|+2D_1+4D_2\right) \end{aligned} \tag{7.5}$$

*for some absolute constant* $D$.

*Proof.* Since $d<x<2d$ for $x\in E\cup I$, this estimate would be obvious if the product of the functions under the integral were negative. To prove (7.5) in the general case, notice that

$$\int_I \log|t-x|h(t)\,dt\leqslant D|I|+2 \tag{7.6}$$

for any $x\in E$. Indeed,

$$\int_I \log|t-x|h(t)\,dt=\int_I \log_+|t-x|h(t)\,dt-\int_I \log_-|t-x|h(t)\,dt,$$

where

$$-\int_I \log_-|t-x|h(t)\,dt\leqslant 2$$

because $h\geqslant -2$. Also,

$$\int_I \log_+|t-x|h(t)\,dt=\int_I \log_+|t-x|f(t)\,dt-\int_I \log_+|t-x|g(t)\,dt.$$

Let $I=(a,b)$. If $x>b$, then by Lemma 7.9 (6) we obtain

$$-\int_I \log_+|t-x|g(t)\,dt \leqslant -(\log_+|x-a|-C)\int_I g(x)\,dx$$

and, since $f>0$ and $\log_+|t-x|\leqslant\log_+|x-a|$,

$$\int_I \log_+|t-x|f(t)\,dt \leqslant \log_+|x-a|\int_I f(x)\,dx.$$

Together the last three relations give

$$\int_I \log_+|t-x|h(t)\,dt \leqslant \log_+|x-a|\int_I f(x)\,dx-(\log_+|x-a|-C)\int_I g(x)\,dx \leqslant 2C|I|,$$

because $g<2$ and $\int_I f(x)\,dx=\int_I g(x)\,dx$. This establishes (7.6). Similar estimates can be applied for $x<a$. For $x\in I$ the same relation can be obtained using part (5) of Lemma 7.9, instead of part (6).

To finish the proof set

$$E^+=\{x\in E:h(x)>0\} \quad \text{and} \quad E^-=\{x\in E:h(x)\leqslant 0\}.$$

Notice that

$$\begin{aligned}
&\int_E h(X)\biggl(\int_I \log|t-x|h(t)\,dt\biggr)\frac{dx}{x^2}\\
&\qquad=\int_{E^+} h(x)\biggl(\int_I \log|t-x|h(t)\,dt\biggr)\frac{dx}{x^2}+\int_{E^-} h(x)\biggl(\int_I \log|t-x|h(t)\,dt\biggr)\frac{dx}{x^2}\\
&\qquad=\int_{E^+} h(x)\max\biggl\{\int_I \log|t-x|h(t)\,dt,0\biggr\}\frac{dx}{x^2}\\
&\qquad\quad\;+\int_{E^+} h(x)\min\biggl\{\int_I \log|t-x|h(t)\,dt,0\biggr\}\frac{dx}{x^2}\\
&\qquad\quad\;+\int_{E^-} h(x)\biggl(\int_I \log|t-x|h(t)\,dt\biggr)\frac{dx}{x^2}\\
&\qquad=:\mathrm{I}+\mathrm{II}+\mathrm{III}.
\end{aligned}$$

If one replaces $dx/x^2$ by $dx/d^2$ in II, the integral will not increase because the function under the integral is negative and $x^2\geqslant d^2$. It also will not decrease much because $x^2\leqslant 4d^2$. Under the same operation, the positive integral $I$ will become at most

$$(D|I|+2)\int_E |h(x)|\,\frac{dx}{d^2}\leqslant\frac{D_1(D|I|+2)}{d^2},$$

in view of (7.6). Hence it will not change by more than the last quantity. Finally, III satisfies

$$|\mathrm{III}| \leqslant \frac{1}{d^2}\int_{E^-}\left|h(x)\int_I \log|t-x|w'(t)\,dt\right|dx,$$

and after replacing $dx/x^2$ by $dx/d^2$ we will still have

$$\left|\int_{E^-} h(x)\left(\int_I \log|t-x|h(t)\,dt\right)\frac{dx}{d^2}\right| \leqslant \frac{1}{d^2}\int_{E^-}\left|h(x)\int_I \log|t-x|w'(t)\,dt\right|dx.$$

Therefore III will change at most by

$$\frac{2}{d^2}\int_{E^-}\left|h(x)\int_I \log|t-x|w'(t)\,dt\right|dx \leqslant \frac{4D_2}{d^2},$$

because $0>h>-2$ on $E^-$. □

## 8. de Branges' theorem in the Toeplitz form

One of the classical results on the gap problem is de Branges' theorem [7, Theorem 66, p. 271]. There the answer is given not in terms of the set $X$ but in terms of the existence of a certain entire function. In this section we discuss two versions of that theorem, see Lemma 8.3 and Corollary 8.4.

Before we formulate de Branges' theorem we need the following definitions.

Recall that $\mathcal{N}(\mathbb{C}_+)$ stands for the Nevanlinna class in the upper half-plane consisting of analytic functions $f(z)$ that can be represented as a ratio $g(z)/h(z)$ of two bounded analytic functions. The mean type of a function $f(z)$ in $\mathcal{N}(\mathbb{C}_+)$ is defined as

$$\limsup_{y\to\infty}\frac{\log|F(iy)|}{y}.$$

THEOREM 8.1. ([7, Theorem 66]) *Let $a>0$ be a given number and let $X$ be a closed subset of the real line. A necessary and sufficient condition that $G_X \geqslant 2a$ is that there exists an entire function $E(z)$, which is real for real $z$ and has only real simple zeros, all in $X$, such that $E(z)$ belongs to $\mathcal{N}(\mathbb{C}_+)$ and has mean type $a$ in the upper half-plane, and such that*

$$\sum_{t:E(t)=0}\frac{1}{|E'(t)|}<\infty. \tag{8.1}$$

Despite the fact that the existence of such an entire function $E$ is not easy to verify, this theorem has been successfully applied in the areas adjacent to the gap problem, see for example [32] for a discussion of such applications and further references.

Before stating and proving an extension of Theorem 8.1, we need the following definitions and a lemma proved by Aleksandrov in [1]. For the reader's convenience we supply the proof.

We say that a finite measure $\mu$ on $\mathbb{R}$ annihilates $K_\theta$ if $\int_{\mathbb{R}} f\,d\mu=0$ for a dense set of $f\in K_\theta$. Note that the integral always exists for a dense set of functions since, for instance, the space $C_A(\mathbb{C}_+)$ of bounded analytic functions in $\mathbb{C}_+$ continuous up to the boundary, is dense in every $K_\theta$.

We say that the Cauchy integral $K\mu$ is divisible by an inner function $\theta$ if $K\mu/\theta=K\eta$ in $\mathbb{C}_+$ for some finite complex measure $\eta$ on $\mathbb{R}$. Equivalently, $K\mu$ is divisible by $\theta$ if $K\mu/\theta\in H^p(\mathbb{C}_+)$ for some $p>0$, see [33].

LEMMA 8.2. ([1]) *Let $\mu$ be a finite complex measure on $\mathbb{R}$ and let $\theta$ be an inner function in $\mathbb{C}_+$. Then the following statements are equivalent*:

(i) *$\mu$ annihilates $K_\theta$*;

(ii) *The Cauchy integral $K\bar{\mu}$ of the conjugate measure $\bar{\mu}$ is divisible by $\theta$.*

*Proof.* (i) ⇒ (ii) We will assume that the reproducing kernels of $K_\theta$ belong to the dense set annihilated by $\mu$ (otherwise one needs to use a standard limiting procedure). If $\lambda\in\mathbb{C}_+$ then

$$\overline{\int \frac{1-\bar{\theta}(\lambda)\theta(z)}{\bar{\lambda}-z}\,d\mu(z)} = \theta(\lambda)K\bar{\theta}\bar{\mu}(\lambda)-K\bar{\mu}(\lambda), \tag{8.2}$$

which implies the statement because the initial integral is zero.

(ii) ⇒ (i) If $\eta$ is the measure such that $K\eta=K\bar{\mu}/\theta$ in $\mathbb{C}_+$, then $\eta$ can be chosen as $\bar{\theta}\bar{\mu}$, see for instance [26, Theorem 3.4]. We may assume that the boundary values of $\theta$ exist $\mu$-a.e. (otherwise $\theta$ can be replaced by a divisor). Then the right-hand side of (8.2) is zero because

$$K\bar{\mu}(\lambda)=\theta(\lambda)K\eta(\lambda)=\theta(\lambda)K\bar{\theta}\bar{\mu}(\lambda).$$

Since reproducing kernels are dense in $K_\theta$, we obtain the statement. □

Note that the condition (8.1) implies that $1/E$ is a Cauchy integral of a finite measure $\mu$ concentrated on the zero set of $E$. The pointmass of $\mu$ at a zero $t$ of $E$ is equal to $1/|E'(t)|$.

Thus the existence of $E$ as in the statement of Theorem 8.1 is equivalent to the existence of a finite discrete real measure $\mu$ supported on $X$ such that $K\mu$ does not have any zeros in $\mathbb{C}$ and is divisible by $S^a$ in the upper half-plane. The theorem says that if $X$ supports any measure whose Cauchy integral is divisible by $S^a$, then it also supports such a $\mu$ with all the above properties. Our next lemma shows that a similar statement can be formulated for any inner $\theta$ in place of $S^a$. As in de Branges' proof, we use Kreĭn–Milman's theorem on the existence of extreme points in a weak-star closed convex set.

LEMMA 8.3. *Let $\theta$ be an inner function in $\mathbb{C}_+$. Let $\mu$ be a finite complex measure whose Cauchy integral $K\mu$ is divisible by $\theta$ (or, equivalently, $\bar{\mu}$ annihilates $K_\theta$). Then there exists a finite singular complex measure $\nu$ such that*

(1) $\operatorname{supp}\nu\subset\operatorname{supp}\mu$;

(2) *$K\nu$ is divisible by $\theta$ ($\bar{\nu}$ annihilates $K_\theta$);*

(3) *$K\nu/\theta$ is outer in $\mathbb{C}_+$ and $K\nu$ is outer in $\mathbb{C}_-$; $K\nu$ has no zeros outside* $\operatorname{supp}\nu$, *except the zeros of $\theta$ in $\mathbb{C}_+$;*

(4) *if $\theta$ is a meromorphic inner function, then $\nu$ is concentrated on a discrete set.*

*Proof.* First, let us symmetrize $\mu$. Since together with any $f\in K_\theta$ one has $\theta\bar{f}\in K_\theta$, the measure $\bar{\theta}\mu$, just like $\bar{\mu}$, annihilates $K_\theta$ and $K\theta\bar{\mu}$ is divisible by $\theta$. Consider $\eta=\mu+\theta\bar{\mu}$. Without loss of generality, we may assume that $\|\eta\|\leqslant 1$.

Let $\Sigma=\operatorname{supp}\mu$. Let $A^\theta_\Sigma$ be the set of all finite complex measures $\sigma$ such that $\|\sigma\|\leqslant 1$, $\operatorname{supp}\sigma\subset\Sigma$, the Cauchy integral of $\sigma$ is divisible by $\theta$ and

$$\theta\bar{\sigma}=\sigma. \tag{8.3}$$

Since $\eta\in A^\theta_\Sigma$, this set is not empty. It is also weak-star closed and convex. By the Kreĭn–Milman theorem, it contains a non-zero extremal point $\nu$. We claim that this is the desired measure.

First, let us show that the set of real $L^\infty(|\nu|)$-functions $h$ such that $Kh\nu$ is divisible by $\theta$ is 1-dimensional, and therefore $h=c\in\mathbb{R}$. (This is equivalent to the statement that the closure of $K_\theta$ in $L^1(|\nu|)$ has deficiency 1, i.e. the space of its annihilators is 1-dimensional).

Let $h$ be a bounded real function such that $Kh\nu$ is divisible by $\theta$. Without loss of generality, we may assume that $h\geqslant 0$, since one can add constants, and $\|h\nu\|=1$. Choose $0<\alpha<1$ so that $|\alpha h|<1$. Consider probability measures

$$\nu_1=h\nu \quad \text{and} \quad \nu_2=(1-\alpha)^{-1}(\nu-\alpha\nu_1).$$

Then both of them belong to $A^\theta_\Sigma$ and

$$\nu=\alpha\nu_1+(1-\alpha)\nu_2,$$

which contradicts the extremality of $\nu$.

Now let us show that $\nu$ is a singular measure. Let $g$ be a continuous compactly supported real function such that $\int_{\mathbb{R}} g\,d\nu=0$. By the previous part, there exists a sequence $f_n\in K_\theta$, with $f_n\to g$ in $L^1(|\nu|)$ (otherwise the defect is larger than 1). Since $\bar{\nu}$ annihilates $K_\theta$ and $(f_n(z)-f_n(w))/(z-w)\in K_\theta$ for every fixed $w\in\mathbb{C}\setminus\mathbb{R}$,

$$0=\int_{\mathbb{R}}\frac{f_n(z)-f_n(w)}{z-w}\,d\bar{\nu}(z)=Kf_n\bar{\nu}(w)-f_n(w)K\bar{\nu}(w)$$

and therefore

$$f_n(w)=\frac{Kf_n\bar{\nu}}{K\bar{\nu}}(w).$$

Taking the limit,

$$f=\lim_{n\to\infty}f_n=\lim_{n\to\infty}\frac{Kf_n\bar{\nu}}{K\bar{\nu}}=\frac{Kg\bar{\nu}}{K\bar{\nu}},$$

where convergence is uniform on compact subsets of $\mathbb{C}\setminus\mathbb{R}$. Since all $f_n$ have pseudocontinuations, one can show that the limit function $f$ must have one as well, i.e. the radial limits of $f$ taken from the upper and lower half-planes match a.e. on $\mathbb{R}$. Indeed, denote by $f_n^\pm$ and $f^\pm$ the radial limits of $f_n$ and $f$ on $\mathbb{R}$ taken from $\mathbb{C}_\pm$ correspondingly. Then, as $f_n$ converge in $L^1(|\nu|)$, $f_n^\pm\to f^\pm$ in measure (with respect to the Lebesgue measure) on $\mathbb{R}$. Since $f_n^+-f_n^-=0$ a.e. on $\mathbb{R}$ for all $n$, $f^+-f^-=0$ a.e. on $\mathbb{R}$.

As the numerator in the representation

$$f=\frac{Kg\bar{\nu}}{K\bar{\nu}}$$

is analytic outside the compact support of $g$, the measure in the denominator must be singular outside that support: Cauchy integrals of non-singular measures have jumps at the real line on the support of the absolutely continuous part, which would contradict the existence of pseudocontinuation. Choosing two different functions $g$ with disjoint supports, we conclude that $\nu$ is singular.

Moreover, $f$ must be analytically continuable through the real line outside $\operatorname{clos}\operatorname{spec}_\theta$, like all of $f_n$. In particular, if $\theta$ is meromorphic, the zero set of $f$ has to be discrete. Since $\nu$ is singular, $K\nu$ tends to $\infty$ at $\nu$-a.e. point and $f=0$ at $\nu$-a.e. point outside the support of $g$. Again, by choosing two different $g$ with disjoint supports, we see that if $\theta$ is meromorphic, then $\nu$ is concentrated on a discrete set.

It remains to verify (3). Let $J$ be the inner function corresponding to $|\nu|$ ($|\nu|$ is the Clark measure for $J$). Set

$$G=\frac{1}{2\pi i}(1-J)K\nu\in K_J.$$

As was mentioned in §3, $G$ has non-tangential boundary values $|\nu|$-a.e. and

$$\nu=G|\nu|,$$

by a result from [26]. Since $K\nu$ is divisible by $\theta$, $G$ is divisible by $\theta$. Let us first show that $G/\theta$ does not have an inner component in the upper half-plane. Suppose that $G=\theta UH$ for some inner $U$. Since the measure $\nu$ satisfies (8.3), $\overline{G}=G/\theta$ $|\nu|$-a.e.

If $F\in K_J$ is the function such that $\bar{J}G=\overline{F}$, then $F=\overline{G}=G/\theta$ $|\nu|$-a.e., as $J=1$ $|\nu|$-a.e. Since functions in $K_J$ are uniquely determined by their traces on the Clark measure $|\nu|$, $F=G/\theta=UH$. Notice that the function $h=\theta(1+U)^2H$ also belongs to $K_J$:

$$\bar{J}h=\bar{J}\theta(1+U)^2H=(\bar{J}G)\overline{U}(1+U)^2=\overline{F}\overline{U}(1+U)^2=\overline{(1+U)^2H}=\overline{h/\theta}\in\overline{H^2}(\mathbb{C}_+),$$

because $\overline{U}(1+U)^2$ is real a.e. on $\mathbb{R}$. Denote by $\gamma$ the measure from the Clark representation of $h$, i.e.

$$\gamma = h|\nu|, \quad h = \frac{1}{2\pi i}(1-J)K\gamma.$$

Then

$$\gamma = h|\nu| = \overline{U}(1+U)^2 G|\nu| = \overline{U}(1+U)^2\nu.$$

The Cauchy integral of $\gamma$ is divisible by $\theta$, because $h$ is divisible by $\theta$. Since $\overline{U}(1+U)^2$ is real, a constant multiple of $\gamma$ belongs to $A^\theta_\Sigma$. As $U$ is non-constant and $|\nu|$ is the Clark measure for $J$, $\gamma$ is not a constant multiple of $\nu$. We obtain a contradiction with the property that the space of annihilators is 1-dimensional.

Thus $G/\theta \in K_J$ is outer in $\mathbb{C}_+$. Since

$$J\overline{G} = F = G/\theta,$$

the pseudocontinuation of $G$ does not have an inner factor in $\mathbb{C}_-$ either. Hence $K\nu/\theta$ is outer in $\mathbb{C}_+$ and $K\nu$ is outer in $\mathbb{C}_-$.

If $G$ has a zero at $x=a\in\mathbb{R}$ outside $\operatorname{spec}_J$, then

$$\frac{G}{x-a} \in K_J$$

and the measure

$$\gamma = \frac{G}{x-a}|\nu|$$

leads to a similar contradiction with the property that the space of annihilators is 1-dimensional, as $1/(x-a)$ is bounded and real on the support of $\nu$. Since

$$G = \frac{1}{2\pi i}(1-J)K\nu,$$

$K\nu$ does not have any extra zeros. □

Our last statement is a Toeplitz version of Theorem 8.1. Recall that a function $f\in N[\phi]$ is said to be *purely outer* if $f$ is outer in the upper half-plane and $\phi f=\bar{g}$ is outer in the lower half-plane.

COROLLARY 8.4. *Let $I$ and $\theta$ be inner functions in $\mathbb{C}_+$. Suppose that the kernel $N[\bar{I}\theta]$ is non-trivial.*

*Then there exists an inner function $J$ in $\mathbb{C}_+$ such that $\operatorname{spec}_J \subset \operatorname{spec}_I$ and the kernel $N[\bar{J}\theta]$ contains a purely outer function $f$ that does not have any zeros on $\mathbb{R}\setminus\operatorname{spec}_J$.*

*If $\sigma_1$ is the Clark measure of $J$, then $f$ is also non-zero $\sigma_1$-a.e. on $\operatorname{spec}_J$.*

*If $\theta$ is a meromorphic function, then $J$ can be chosen as a meromorphic function.*

*Proof.* We will consider the case $\operatorname{spec}_I \subset \mathbb{R}$. The general case can be treated similarly.

If $f \in N[\bar{I}\theta]$ then the function $g=\theta f$ belongs to $K_I$. Consider its Clark representation

$$g=\frac{1}{2\pi i}(1-I)Kg\sigma,$$

where $\sigma$ is the Clark measure corresponding to $I$. Since $1-I$ is outer, the Cauchy integral $Kg\sigma$ is divisible by $\theta$. By Lemma 8.3, there exists a finite singular measure $\nu$, with

$$\operatorname{supp}\nu \subset \operatorname{supp} g\sigma \subset \operatorname{spec}_I,$$

satisfying properties (1)–(4). Let $J$ be the inner function whose Clark measure is $|\nu|$. Then the function $h=(1-J)K\nu$ belongs to $K_J$ and is divisible by $\theta$. Therefore $l=h/\theta$ belongs to $N[\bar{J}\theta]$. Note that $l$ is purely outer by Lemma 8.3 (3). Since $2\pi i\nu=h|\nu|$, $l$ has no zeros a.e. with respect to $|\nu|$, the Clark measure of $J$.

If $\theta$ is a meromorphic function then, by Lemma 8.3 (4), $\nu$ can be chosen to be discrete. Then $J$ is meromorphic. □

## References


[1] ALEKSANDROV, A. B., Isometric embeddings of co-invariant subspaces of the shift operator. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*), 232 (1996), 5–15, 213 (Russian); English translation in *J. Math. Sci.* (*N. Y.*), 92 (1998), 3543–3549.

[2] BENEDICKS, M., The support of functions and distributions with a spectral gap. *Math. Scand.*, 55 (1984), 285–309.

[3] BEURLING, A., *On Quasianalyticity and General Distributions.* Summer Institute, Stanford University, Stanford, CA, 1961.

[4] BEURLING, A. & MALLIAVIN, P., On Fourier transforms of measures with compact support. *Acta Math.*, 107 (1962), 291–309.

[5] — On the closure of characters and the zeros of entire functions. *Acta Math.*, 118 (1967), 79–93.

[6] BORICHEV, A. & SODIN, M., Weighted exponential approximation and non-classical orthogonal spectral measures. *Adv. Math.*, 226 (2011), 2503–2545.

[7] DE BRANGES, L., *Hilbert Spaces of Entire Functions.* Prentice-Hall, Englewood Cliffs, NJ, 1968.

[8] COIFMAN, R. R. & WEISS, G., Extensions of Hardy spaces and their use in analysis. *Bull. Amer. Math. Soc.*, 83 (1977), 569–645.

[9] CORNU, F. & JANCOVICI, B., On the two-dimensional Coulomb gas. *J. Stat. Phys.*, 49 (1987), 33–56.

[10] DYM, H., On the span of trigonometric sums in weighted $l^2$ spaces, in *Linear and Complex Analysis Problem Book* 3, Part II, Lecture Notes in Math., 1574, pp. 87–88. Springer, Berlin–Heidelberg, 1994.

[11] DYM, H. & MCKEAN, H. P., *Gaussian Processes*, *Function Theory*, *and the Inverse Spectral Problem.* Probability and Mathematical Statistics, 31. Academic Press, New York, 1976.

[12] Gel′fand, I. M. & Levitan, B. M., On the determination of a differential equation from its spectral function. *Izv. Akad. Nauk SSSR Ser. Mat.*, 15 (1951), 309–360 (Russian); English translation in *Amer. Math. Soc. Transl.*, 1 (1955), 253–304.

[13] Havin, V. & Jöricke, B., *The Uncertainty Principle in Harmonic Analysis*. Ergebnisse der Mathematik und ihrer Grenzgebiete, 28. Springer, Berlin–Heidelberg, 1994.

[14] Kerov, S. V., Equilibrium and orthogonal polynomials. *Algebra i Analiz*, 12:6 (2000), 224–237 (Russian); English translation in *St. Petersburg Math. J.*, 12 (2001), 1049–1059.

[15] Koosis, P., *The Logarithmic Integral*. I. Cambridge Studies in Advanced Mathematics, 12. Cambridge University Press, Cambridge, 1988.

[16] Kreĭn, M. G., On a problem of extrapolation of A. N. Kolmogoroff. *Dokl. Acad. Nauk SSSR*, 46 (1945), 306–309 (Russian).

[17] — On the transfer function of a one-dimensional boundary problem of the second order. *Dokl. Akad. Nauk SSSR*, 88 (1953), 405–408 (Russian).

[18] — On a basic approximation problem of the theory of extrapolation and filtration of stationary random processes. *Dokl. Akad. Nauk SSSR*, 94 (1954), 13–16 (Russian).

[19] Levinson, N., *Gap and Density Theorems*. American Mathematical Society Colloquium Publications, 26. Amer. Math. Soc., New York, 1940.

[20] Makarov, N. & Poltoratski, A., Meromorphic inner functions, Toeplitz kernels and the uncertainty principle, in *Perspectives in Analysis*, Math. Phys. Stud., 27, pp. 185–252. Springer, Berlin–Heidelberg, 2005.

[21] — Beurling–Malliavin theory for Toeplitz kernels. *Invent. Math.*, 180 (2010), 443–480.

[22] Mitkovski, M. & Poltoratski, A., Pólya sequences, Toeplitz kernels and gap theorems. *Adv. Math.*, 224 (2010), 1057–1070.

[23] Nienhuis, B., Coulomb gas formulation of two-dimensional phase transitions, in *Phase Transitions and Critical Phenomena*, Vol. 11, pp. 1–53. Academic Press, London, 1987.

[24] Nikol′skii, N. K., Bases of exponentials and values of reproducing kernels. *Dokl. Akad. Nauk SSSR*, 252 (1980), 1316–1320 (Russian); English translation in *Soviet Math. Dokl.*, 21 (1980), 937–941.

[25] — *Treatise on the Shift Operator*. Grundlehren der Mathematischen Wissenschaften, 273. Springer, Berlin–Heidelberg, 1986.

[26] Poltoratski, A., On the boundary behavior of pseudocontinuable functions. *Algebra i Analiz*, 5:2 (1993), 189–210 (Russian); English translation in *St. Petersburg Math. J.*, 5 (1994), 389–406.

[27] — Kreĭn's spectral shift and perturbations of spectra of rank one. *Algebra i Analiz*, 10:5 (1998), 143–183 (Russian); English translation in *St. Petersburg Math. J.*, 10 (1999), 833–859.

[28] — A problem on completeness of exponentials. Preprint, 2010. `arXiv:1006.1840 [math.CA]`.

[29] Poltoratski, A. & Sarason, D., Aleksandrov–Clark measures, in *Recent Advances in Operator-Related Function Theory*, Contemp. Math., 393, pp. 1–14. Amer. Math. Soc., Providence, RI, 2006.

[30] Šamaj, L., The statistical mechanics of the classical two-dimensional Coulomb gas is exactly solved. *J. Phys. A*, 36 (2003), 5913–5920.

[31] Simon, B., Spectral analysis of rank one perturbations and applications, in *Mathematical Quantum Theory*. II. Schrödinger Operators (Vancouver, BC, 1993), CRM Proc. Lecture Notes, 8, pp. 109–149. Amer. Math. Soc., Providence, RI, 1995.

[32] Sodin, M. & Yuditskii, P., Another approach to de Branges' theorem on weighted polynomial approximation, in *Proceedings of the Ashkelon Workshop on Complex Function Theory* (1996), Israel Math. Conf. Proc., 11, pp. 221–227. Bar-Ilan University, Ramat Gan, 1997.

[33] Vinogradov, S. A., Properties of multipliers of integrals of Cauchy–Stieltjes type, and some problems of factorization of analytic functions, in *Mathematical Programming and Related Questions* (Proc. Seventh Winter School, Drogobych, 1974), pp. 5–39 (Russian). Central Èkonom.-Mat. Inst. Akad. Nauk SSSR, Moscow, 1976.

Alexei Poltoratski
Texas A&M University
Department of Mathematics
College Station, TX 77843-3368
U.S.A.
alexeip@math.tamu.edu

*Received August 31, 2009*
*Received in revised form January 7, 2011*