CHAPTER 2

Gap Theorems

In this chapter we discuss the problem of finding the maximal possible size of the gap in the Fourier spectrum of a measure. The Fourier spectrum of a finite measure is the closed support of its Fourier transform on $\mathbb{R}$. Conditions imposed on the measure may take several different forms, as was discussed in the first chapter. Classical theorems by Krein-McKean, Levinson and de Branges show that a measure that decays fast at infinity and satisfies some regularity conditions may not have *any* spectral gaps. Our first goal in this chapter is to review these classical results and give them simple proofs following one universal approach, see section 1.

The rest of the chapter is devoted to the Gap Problem stemming from Beurling's theorem discussed in subsection 2.2 in chapter 1. Following the general scope of UP, the problem asks to show that if the support of the measure is porous then its Fourier spectrum cannot have large gaps.

Beurling's theorem, as well as other classical gap theorems mentioned above, give sufficient conditions for the absence of spectral gaps. Our methods allow us to take on the next logical stage of the problem and find the maximal possible size of the spectral gap for a measure with a given support. More precisely, for a closed set X on the real line we give a formula for its gap characteristic, the supremum of the size of the spectral gap taken over all non-zero finite complex measures supported on X, see section 4.

The Gap Problem has many connections and applications in adjacent fields of UP. In particular, the formula for the gap characteristic of a set proves to be useful in several classical problems discussed in other chapters of these notes: a problem on completeness of complex exponentials in L^2-spaces (the Type Problem, chapter 6), a problem on sampling sets for entire functions of zero exponential type (the Polya-Levinson problem, chapter 3) and a problem on oscillations of high-pass signals (the problem by Grinevich, chapter 4).

We utilize close connections between most problems from this area of harmonic analysis and the problem of injectivity of Toeplitz operators. Such connections are further discussed in chapters 7 and 8. In the case of the Gap Problem, this connection is expressed by theorem 8 below. The Toeplitz approach for similar problems was first suggested by Nikolski [**116**], see also [**68**]. Our main proof utilizes several important ideas of the Beurling–Malliavin theory, 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. The relations between the Gap Problem and the Beurling–Malliavin theory on 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

$$\phi = \exp(-iax)\Theta,$$

for a suitable meromorphic inner function Θ, while the Gap Problem reduces to the triviality of the kernel of the Toeplitz operator with the symbol

$$\bar{\phi} = \exp(iax)\bar{\Theta},$$

see chapter 8 and theorem 8 below.

This chapter is based on [**127**].

1. Classical Gap Theorems

The goal of this section is to discuss gap theorems by Krein, Levinson-McKean, Beurling and de Branges. We formulate theorem 6, that can be viewed as a hybrid of Beurling's and Levinson's theorems, and give it a short elementary proof. We then show how to deduce the classical results from theorem 6. In some cases, instead of deducing classical theorems from each other we prefer to give each a direct closed proof through theorem 6.

Recall that a sequence of disjoint intervals $\{I_n\}$ on the real line is called long (in the sense of Beurling and Malliavin) if

$$\sum_n \frac{|I_n|^2}{1+\text{dist}^2(0, I_n)} = \infty \tag{2.1}$$

where $|I_n|$ stands for the length of I_n. If the sum is finite we call $\{I_n\}$ short. If I is an interval on $\mathbb{R}$ and $C > 0$ we denote by CI the interval with the same center as I of length $C|I|$. Sometimes we call $\cup I_n$ long if $\{I_n\}$ is long, see the statement of our next theorem.

This definition appears in many results including several statements discussed in these notes. For readers unfamiliar with this notion of 'longness,' let us point out some of its simple properties. If the intervals I_n are uniformly bounded in length then the sequence is short because the sum in the definition is $\lesssim \sum \frac{1}{n^2}$. At the same time, if I_n is a subsequence of diadic intervals, $< 2^{k_n}, 2^{k_n+1} >$, then the sum is infinite, no matter how rare the subsequence is. More subtle examples lie in between these two extreme cases. The condition of shortness of $\{I_n\}$ is equivalent to Poisson-summability of the sum of 'tent-functions'

$$\sum T_n, \quad T_n(x) = \text{dist}(x, \mathbb{R} \setminus I_n).$$

If the sequence is long then not only is the sum of tent functions Poisson-unsummable, but it is not a harmonic conjugate of a Poisson-summable function (this is a version of one of the BM theorems, e.g. lemma 26 in chapter 6). Via this property the longness condition enters into proofs, including its use in the Beurling–Malliavin theory.

As was mentioned in chapter 1, Beurling's version of a Gap Theorem says that if the support of μ is porous, then the support of $\hat{\mu}$ cannot have any gaps. Note that a support of a measure on $\mathbb{R}$ is a closed set and its complement is a union of disjoint open intervals (gaps).

THEOREM 1. *[Beurling's Gap Theorem] Let $\mu \not\equiv 0$ be a finite complex measure on $\mathbb{R}$. If the complement of the support of μ is long then the support of $\hat{\mu}$ does not have any gaps, i.e. there is no interval on $\mathbb{R}$ where $\hat{\mu}$ is identically 0.*

Levinson's version of the theorem says that if instead of having porous support μ decays fast at infinity, one can arrive at the same conclusion:

THEOREM 2. *[Levinson's Gap Theorem] Let μ be a finite measure on $\mathbb{R}$. Denote*

$$M(x) = |\mu|((x,\infty)).$$

Suppose that $\log M$ is not Poisson-summable on $\mathbb{R}_+$. If $\hat{\mu}$ vanishes on an interval then $\mu \equiv 0$.

We will deduce the last two theorems from the following more technical 'hybrid' statement:

THEOREM 6. *Let μ be a finite measure on $\mathbb{R}$ whose Fourier transform vanishes on an interval. Suppose that there exists a sequence of disjoint intervals $\{I_n\}$ such that*

$$\sum \frac{|I_n|}{1+\operatorname{dist}^2(I_n,0)} \min\left(|I_n|, \log\frac{1}{|\mu|(I_n)}\right) = \infty. \tag{2.2}$$

Then $\mu \equiv 0$.

In comparison with Beurling's theorem, our statement says that the measure does not have to vanish on a long sequence of intervals, it just has to be small on it. In comparison with Levinson's version, we say that the measure does not have to decay fast on the whole line, just on a large enough subset of the line. These remarks are made precise in the proofs of theorems 1 and 2 below.

The proof of theorem 6 borrows an idea from the proof of Beurling's Gap Theorem by Benedicks in [**12**].

PROOF OF THEOREM 6. Without loss of generality $|I_n| > 1$ for all n, because the sum in (2.2) taken over all intervals of length less than 1 is finite. Suppose that $\hat{\mu}$ vanishes on $[-a,a]$. Then its Cauchy integral $K\mu$ (defined as in (1.8)) is divisible by $e^{2\pi iaz}$ in $\mathbb{C}_+$, in the sense that

$$K\mu = e^{2\pi iaz}K\nu,$$

where ν is a finite measure, $\nu = e^{-2\pi iaz}\mu$, see for instance lemma 11 in chapter 3 in addition to the results from [**122**].

Denote by J_n the interval on $\mathbb{R}+i$:

$$J_n = \left\{ z \mid \Im z = 1,\ \Re z \in \frac{1}{2}I_n \right\}.$$

Denote by μ_n the restriction of μ on I_n and put $\eta_n = \mu - \mu_n$. Notice that $K\eta_n(z)$ is holomorphic in $(\mathbb{C}\setminus\mathbb{R})\cup I_n$. Hence $-\log|K\eta_n(z)|$ is superharmonic in $\{|z-\xi| \leqslant |I_n|/4\}$ for any $\xi \in J_n$. Since

$$-\log|K\mu(z)| = -\log|K\nu(z)| - \log|e^{2\pi iaz}| \gtrsim a|I_n|$$

in the half-plane

$$\{\Im z > |I_n|/8\},$$

we obtain

$$-\log|K\eta_n(\xi)| \geqslant -\frac{1}{2\pi}\int_0^{2\pi}\log\left|K\eta_n\left(\xi+\frac{|I_n|}{4}e^{i\phi}\right)\right|d\phi$$
$$= -\frac{1}{2\pi}\int_0^{2\pi}\log\left|K\mu\left(\xi+\frac{|I_n|}{4}e^{i\phi}\right)-K\mu_n\left(\xi+\frac{|I_n|}{4}e^{i\phi}\right)\right|d\phi$$
$$\gtrsim \min\left(a|I_n|, -\log\frac{|\mu|(I_n)}{|I_n|}\right)$$

for any $\xi \in J_n$. On the other hand,

$$|K\mu(\xi)| = |K\eta_n(\xi) + K\mu_n(\xi)| \leqslant |K\eta_n(\xi)| + |\mu|(I_n)|$$

and

$$-\log|K\mu(\xi)| \gtrsim \min(|I_n|, -\log\frac{|\mu|(I_n)}{|I_n|}, -\log|\mu_n|(I_n)) \gtrsim \min(|I_n|, -\log|\mu|(I_n))$$

(recall that $|I_n| > 1|$).

Now (2.2) implies that $\log|K\mu|$ is not Poisson-summable on the line

$$\{\Im z = 1\}.$$

Therefore it is identically zero in $\mathbb{C}_+$. Similarly, it is zero in $\mathbb{C}_-$. If $K\mu$ is zero in both half-planes, $\mu \equiv 0$. □

Now Beurling's theorem follows immediately:

PROOF OF THEOREM 1. Assume that the complement of $\operatorname{supp}\mu$ is long. Then the complement can be taken as $\{I_n\}$ in (2.2). □

Next, we deduce Levinson's theorem:

PROOF OF THEOREM 2. Suppose that $\log M$ is not Poisson-summable on $\mathbb{R}_+$. Without loss of generality, $M(0) = 1$. Let $0 = a_0 < a_1 < a_2 < \ldots$ be the points such that $M(a_n) = 2^{-n}$ and denote by $I_n = (a_n, a_{n+1}]$ the corresponding partition of $\mathbb{R}_+$. If

$$\sum \frac{n|I_n|}{1+\operatorname{dist}^2(I_n, 0)} < \infty$$

then $\log M$ is Poisson-summable and we have a contradiction.

If the last sum is infinite, but the sum in 2.2 is finite, i.e. the partition I_n is short, then any long super-partition of I_n will satisfy (2.2). If the last sum is infinite and I_n is long, then (2.2) is satisfied. □

As we discussed in section 2.4 of chapter 1, Levinson's result above was later improved by Beurling [**14**] who showed that an interval can be replaced with a set of positive Lebesgue measure. It seems logical to consider such an improvement after every gap result. Note that Beurling's theorem itself does not admit such a strengthening, see [**85**]. A further step would be to improve the condition of vanishing on a set of positive measure to summability of the logarithm of the density, which is possible in some situations, see for instance [**146**]. In this text we will not follow this path focusing just on the Gap Problem.

A stronger gap theorem, that implies the classical statements discussed above, belongs to de Branges (theorem 63 [**26**]):

Theorem 7 (de Branges' Gap Theorem). *Let $K(x)$ be a continuous function on $\mathbb{R}$ such that $K(x) \geqslant 1$, $\log K$ is uniformly continuous and Poisson-unsummable. Then there is no nonzero finite measure μ on $\mathbb{R}$ such that*

$$\int_{-\infty}^{\infty} K d|\mu| < \infty \tag{2.3}$$

and $\hat{\mu}$ vanishes on an interval.

Proof. Without loss of generality $K \geqslant 2$ and K is Poisson-unsummable on $\mathbb{R}_+$. Choose points $a_0, a_1, ...$ on $\mathbb{R}_+$ in the following way. Put $a_0 = 0$. After a_{n-1} is chosen, choose a_n to be the smallest point greater than a_{n-1} such that

$$\log K(a_n) \notin \left(\frac{\log K(a_{n-1})}{2}, 2\log K(a_{n-1})\right).$$

Note that such a_n always exists because K is unbounded on any ray $[x, \infty)$. Denote by L the step function, minorating $\log K$ defined as

$$L(x) = L_n = \min_{I_n} \log K$$

on each $I_n = (a_{n-1}, a_n]$. Notice that by the choice of $\{I_n\}$, $\log L \asymp \log K$. In particular, $\log L$ is Poisson-unsummable. By (2.3), $\mu(I_n) \lesssim 1/L_n$. Also, because of uniform continuity of $\log K$, $\log L_n \lesssim |I_n|$. Hence the sum in (2.2) is minorated by

$$\sum \frac{|I_n| \log L_n}{1 + \text{dist}^2(I_n, 0)} \gtrsim \int \log L(x) \frac{dx}{1 + x^2} = \infty.$$

□

For a finite positive measure μ on $\mathbb{R}$ define

$$\mathbf{G}^p_\mu = \sup\{a\ > 0 | \ \exists\ f \in L^p(\mu), \forall\ \lambda \in [0, a], \int f(x) e^{2\pi i \lambda x} d\mu(x) = 0\}, \tag{2.4}$$

or 0 if there are no such a. By duality, for $1 < p \leqslant \infty$, $\mathbf{G}^p_\mu$ can be defined as the infimum of a such that the system of exponential functions

$$\mathcal{E}_a = \{e^{2\pi i \lambda t}\}_{\lambda \in [0,a]}$$

is complete in $L^q(\mu)$, $\frac{1}{p} + \frac{1}{q} = 1$. For $p = 2$, the problem of finding $\mathbf{G}^2_\mu$ constitutes the well-known Type Problem discussed in chapter 6. Cases $p \neq 2$ were considered in several papers, see for instance articles by Koosis [**87**] or Levin [**96**] for the case $p = \infty$. As will be discussed in the next section, the case $p = 1$ is an equivalent reformulation of the Gap Problem.

Theorem 6 has the following partial inverse.

Proposition 1. *Let $\mu = w(x)dx$ be an absolutely continuous finite measure with $w > 0$ and $\log |w|$ absolutely continuous. Suppose that the sequence of intervals I_n satisfying (2.2) does not exist. Then $\mathbf{G}^\infty_\mu = \infty$.*

Proof. Similarly to the last proof, it is not difficult to show that $\log |w|$ is Poisson-summable. After that for any $C > 0$ consider the measure $u\mu$ with

$$u = e^{iCx} F/w,$$

where F is the outer function in the upper half-plane satisfying $|F| = w$. □

We will return to the discussion of the classical gap theorems of this section in chapter 4 to formulate stronger versions of these results.

2. Spectral gap as a property of the support

The classical Gap Theorems of the last section give sufficient conditions for the absence of gaps in the Fourier spectrum of a measure. In this section we continue in the direction of Beurling's result and study further relations between gaps in the supports of μ and $\hat{\mu}$, thus following the 'support/support' version of UP as discussed in section 2.2 of chapter 1.

As usual, the ultimate challenge in the problems of UP is to obtain a quantative estimate for the uncertainty. In our settings, we would like to find a formula for the maximal possible size of the gap in the Fourier spectrum of a non-zero measure in terms of the metric properties of its support.

Recall that M denotes the set of all finite Borel complex measures on the real line. If X is a closed subset of the real line we denote by $\mathbf{G}_X$ its gap characteristic defined as

$$\mathbf{G}_X = \sup\{a \mid \exists\ \mu \in M,\ \mu \not\equiv 0,\ \operatorname{supp}\mu \subset X, \text{ such that } \hat{\mu} = 0 \text{ on } [0,a]\}. \tag{2.5}$$

Our goal for the remainder of this chapter will be to find a formula for the gap characteristic of X. First, let us discuss our choice of the set-up in more details.

If $\mu \in M$ it seems natural, for instance, to look at the quantity $\mathbf{G}_\mu = \mathbf{G}^1_\mu$ as defined in the last section and try to find it instead of $\mathbf{G}_X$. However, one soon realizes that the latter problem reduces to the former:

PROPOSITION 2.

$$\mathbf{G}_\mu = \mathbf{G}_{\operatorname{supp}\mu}. \tag{2.6}$$

PROOF. Obviously, $\mathbf{G}_{\operatorname{supp}\mu} \geq \mathbf{G}_\mu$. To prove the opposite inequality, notice that by de Branges theorem 66 (see theorem 10 discussed in section 8), there exists a finite discrete measure

$$\nu = \sum \alpha_n \delta_{x_n}, \{x_n\} \subset \operatorname{supp}\mu,$$

such that $\hat{\mu}$ has a gap of the size greater than $\mathbf{G}_{\operatorname{supp}\mu} - \varepsilon$ (see the remark before lemma 9). Around each x_n choose a small neighborhood $V_n = (a_n, b_n)$ so that for any sequence of points

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

there exists a non-trivial measure $\eta_Y = \sum \beta_n \delta_{y_n}$ such that $\hat{\eta}$ has a gap of the size greater than $\mathbf{G}_{\operatorname{supp}\eta} - \varepsilon$. The existence of such a collection of neighborhoods follows from the results of [**12**] (for some sequences) and [**21**] as well as from theorem 9 below.

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

- for each τ,
  $$\eta_\tau = \sum \beta^\tau_n \delta_{y^\tau_n}$$
  such that $y^\tau_n \in V_n$ and $\hat{\eta}_\tau$ has a gap of the size greater than $\mathbf{G}_{\operatorname{supp}\eta} - \varepsilon$ centered at 0;

- the measure
$$\gamma = \int_0^1 \eta_\tau d\tau$$
is non-trivial and absolutely continuous with respect to μ.

It remains to notice that then the support of $\hat{\gamma}$ has a gap of the size at least $\mathbf{G}_{\operatorname{supp}\eta} - \varepsilon$. □

3. Toeplitz kernels and uniform approximation

Recall that for a function $\phi \in L^\infty(\mathbb{R})$ we denote by $N[\phi]$ the kernel of the Toeplitz operator T_ϕ in $H^2(\mathbb{C}_+)$. The notation S^a is used for the singular inner function $S^a(z) = e^{iaz}$ in $\mathbb{C}_+$. For a meromorphic inner function Θ in $\mathbb{C}_+$, $\operatorname{spec}_\Theta$ denotes the set $\{\Theta = 1\} \subset \hat{\mathbb{R}}$.

One of the standard classes of Toeplitz operators consists of those with symbols of the form $\bar{I}J$, where I and J are inner functions. A subclass with $J = S^a$ plays a key role in the Gap Problem.

Definition 1. *If $X \subset \mathbb{R}$ is a closed set, denote*
$$\mathcal{T}_X = \sup\{\ a \mid N[\bar{\Theta}S^{2\pi a}] \neq 0 \textit{ for some meromorphic inner } \Theta,\ \operatorname{spec}_\Theta \subset X\}.$$

The following theorem will be proved in section 5 chapter 3. It shows the connection between the Gap Problem and the problem of triviality of Toeplitz kernels which will be used throughout these notes.

Theorem 8.
$$\mathbf{G}_X = \mathcal{T}_X.$$

The Gap Problem is closely related to problems of weighted uniform approximation of continuous functions by trigonometric polynomials. This topic will be discussed in chapters 5 and 6.

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 Payley-Wiener space
$$\mathcal{PW}_a = \{\hat{f} \mid f \in L^2([-a,a])\}.$$

Let us define
$$\mathcal{A}_X = \inf\{\ a > 0 \mid \mathcal{PW}_a \text{ is dense in } C_0(X)\}$$
or ∞ if the set is empty. The following statement is a product of the standard duality argument.

Proposition 3.
$$\mathbf{G}_X = 2\mathcal{A}_X.$$

Together with theorem 9 this statement gives a formula for $\mathcal{A}_X$.

4. A formula for the gap characteristic of a set

Until recently we could calculate the gap characteristic $\mathbf{G}_X$ only for a small collection of sets X. For instance, to calculate $\mathbf{G}_\mathbb{Z}$, notice that the function $z(z-1)\csc \pi z$ is a Cauchy integral $K\mu$ of a finite measure μ supported on $\mathbb{Z}\setminus\{0,1\}$. The Cauchy integral $K\mu(iy)$ decays as $e^{-|y|}$ along the imaginary axis. It is well known that such a decay indicates a spectral gap, see lemma 11 in section 6, chapter 3. Hence, $\mathbf{G}_\mathbb{Z}\geq 1$. The opposite inequality we leave to the reader as an exercise. The conclusion is that $\mathbf{G}_\mathbb{Z}=1$.

Furthermore, the gap characteristic easily adjusts to the affine transformations of the set, which allows us to calculate the gap characteristic of any infinite arithmetic progression $P=\{an+C\}_{n\in\mathbb{Z}},\ a>0, C\in\mathbb{R}$:

$$\mathbf{G}_P = 1/a.$$

Our next trivial observation is that the gap characteristic is monotone, i.e. $\mathbf{G}_X\geq \mathbf{G}_Y$ if $Y\subset X$. Hence for an arbitrary closed X,

$$\mathbf{G}_X\geq \sup\{a|\ X \text{ contains } \{an+C\}_{n\in\mathbb{Z}}\}.$$

As it turns out, our simple observations are the first step towards a formula for $\mathbf{G}_X$. The last inequality can be made an equation if instead of arithmetic progressions we consider a wider class of sequences, whose gap characteristic we are able to evaluate. The proper generalization of an arithmetic progression in this context is a sequence that we call d-uniform. Before we give a definition of such sequences we need some preparation.

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

$$E(\Lambda)=\sum_{\lambda_k,\lambda_j\in\Lambda,\ \lambda_k\neq\lambda_j}\log|\lambda_k-\lambda_l|. \tag{2.7}$$

Physical interpretation:

According to the 2-dimensional Coulomb's law, $E(\Lambda)$ is the potential energy of a system of 'flat' electrons placed at the points of Λ. The 2D Coulomb-gas formalism corresponds to the planar potential theory with logarithmic potential and assumes the potential energy at infinity to be equal to $-\infty$, e.g. [**37, 113, 137**].

Physically, the 2D Coulomb's law can be derived from the standard 3D 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 3D law and a renormalization procedure.

Thus, if Λ and Γ are two finite sequences with equal number of points, the difference $E(\Lambda)-E(\Gamma)$ presents work needed to move the the electrons from Γ to Λ.

The following example will help us interpret the main definition.

Key example:

Let $I \subset \mathbb{R}$ be an interval, $C > 0$ and let Λ be a set of k points uniformly spread across I:

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

Then

$$E(\Lambda) = \sum_{1 \le m \le k} \log\left[C^{k-1}(m-1)!(k-m)!\right] = k^2 \log|I| + O(|I|^2) \tag{2.8}$$

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

Note that the energy of k points on I will never exceed the main term $k^2 \log|I|$ in the last equation as the energy is the sum of less than k^2 terms each no greater than $\log|I|$. Thus, even though the uniform distribution of points on the interval does not maximize the energy $E(\Lambda)$, it 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 true maximum for the energy of k electrons on I is achieved when they are placed at the endpoints of I and the zeros of the Jacobi $(1,1)$-polynomial of degree $k-2$, see for example [**74**].

Let

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

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 $\{I_n\}$ is short, i.e. the sum in (2.1) is finite.

As usual, a sequence of points $\Lambda = \{\lambda_n\} \subset \mathbb{C}$ is called discrete if it has no finite accumulation points.

Main definition:

Let $\Lambda = \{\lambda_n\}$ be a discrete sequence of distinct real points and let d be a positive number. We say that Λ is a d-uniform sequence if there exists a short partition $\{I_n\}$ such that

$$\Delta_n = \#(\Lambda \cap I_n) = d|I_n| + o(|I_n|) \text{ as } n \to \pm\infty \text{ (density condition)} \tag{2.9}$$

and

$$\sum_n \frac{\Delta_n^2 \log|I_n| - E_n}{1 + \operatorname{dist}^2(0, I_n)} < \infty \qquad \text{(energy condition)} \tag{2.10}$$

where

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

This definition will play a key role not only in our gap theorem below but in the main results of chapters 4 and 3. Roughly, it can be summarized as follows: a sequence is d-uniform if it is close to the arithmetic progression $\frac{1}{d}\mathbb{Z}$ in the precise sense given by the two conditions.

The density condition says that on each interval of the short partition I_n our sequence has about the same number of points, $d|I_n|$, as the said arithmetic progression. The example before the definition shows that the numerator in the energy condition is equal (up to lower order terms) to the difference between the energy of the arithmetic progression and the energy of Λ on I_n. Thus the energy condition is a requirement that the placement of the points of Λ is close to uniform, in the sense that the total work needed to move the points of Λ into the points of an arithmetic progression on each interval is finite (with respect to the Poisson weight).

Notice that the series in the energy condition is positive since every term in the sum defining E_n is at most $\log |I_n|$ and there are less than Δ_n^2 terms. Convergence of positive series is usually easier to analyze.

One of the features of the Gap Problem is that the density of a sequence by itself turns out to be insufficient to find the correct formula. Out of all the sequences with correct densities we consider only those satisfying the energy condition that prohibits clustering of points.

The following remarks will be used in the proofs.

REMARK 1. *The simple inequality* $\log |I_n| - \log |\lambda_k - \lambda_j| > 0$, *which holds for any* $\lambda_k, \lambda_j \in I_n$, *also implies that if a sequence* Λ *satisfies the energy condition* (2.10), *then any subsequence of* Λ *also satisfies* (2.10) *on* $\{I_n\}$. *A deletion of points from* Λ *will eliminate some of such positive differences from the numerator in* (2.10) *which can only make the sum smaller.*

REMARK 2. *We say that a partition* $\{I_n\}$ *is monotone if* $|I_n| \le |I_{n+1}|$ *for* $n \ge 0$ *and* $|I_{n+1}| \le |I_n|$ *for* $n < 0$. *Corollary* 3 *in section* 7 *shows that in the above definition the words "short partition" can be replaced with "short monotone partition." Since monotone partitions are easier to work with, this modified definition will be used in the proof of theorem* 9.

REMARK 3. *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* Δ_n *and* E_n *in* (2.10). *One could, for instance, use*

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

and

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

In [**127**] the density condition (2.9) was given in a slightly weaker form:

$$\Delta_n = \#(\Lambda \cap I_n) \ge d|I_n| \tag{2.11}$$

for all n. In view of remark 1, one can easily see that either one of the definitions can be used in the statement of theorem 9 below.

Those readers familiar with various definitions of densities used in the area of UP may notice similarities with our density condition. In fact, existence of a short partition $\{I_n\}$ on which Λ satisfies (2.11) is equivalent to the property that the lower (interior) density of the sequence in the sense of Beurling and Malliavin, is at least d. Such a density can be defined in several different ways.

- If Λ 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 satisfying (2.11).

- Denote by $d_2(\Lambda)$ the supremum of all a such that there exists a short (not necessarily monotone) partition I_n satisfying (2.11).

- Define $d_3(\Lambda)$ to be the supremum of all a such that there exists a subsequence of Λ whose counting function $n(x)$ satisfies
$$\int \frac{|n - ax|}{1+x^2}dx < \infty.$$
This definition was used in [**26**].

- 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 will be used in chapter 3.

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 [**17**], where it was called *interior* density. A closely related notion of *exterior* density appears in the Beurling–Malliavin theorem on completeness of exponential functions in L^2 on an interval discussed in chapter 5. We give a definition of exterior density in section 7. For further discussion of BM densities see section 3 in chapter 3.

Now we are ready to return to the Gap Problem and give the following formula for the gap characteristic $\mathbf{G}_X$ of a closed subset of $\mathbb{R}$.

THEOREM 9.
$$\mathbf{G}_X = \sup\{d \mid X \textit{ contains a d-uniform sequence}\},$$
if the set on the right is non-empty and $\mathbf{G}_X = 0$ *otherwise.*

It will be proved in section 6.

5. Examples and applications

In this section we discuss examples related to theorem 9, including some of its relations with existing results.

EXAMPLE 1. *As discussed above, if the points of the sequence are spread uniformly over the interval, then* $E_n = \sum_{\lambda_i,\lambda_j \in I_n} \log|\lambda_i - \lambda_j|$ *is roughly (up to* $O(|I_n|^2)$*, which is small for short sequences of* I_n*) equal to* $\Delta_n^2 \log |I_n|$*. This happens for instance when the sequence* Λ *is separated, i.e. satisfies* $|\lambda_n - \lambda_{n+1}| > \delta > 0$ *for all* n*. Thus for separated sequences* Λ *the energy condition disappears and*
$$\mathbf{G}_\Lambda = d_i(\Lambda)$$
where $d_i, i = 1, 2, 3, 4$ *is any of the equivalent densities defined in the last section, i.e. the interior density of* Λ*. This is one of the results of* [**129**] *discussed in chapter* 3 *of these notes.*

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

EXAMPLE 2. *Let Λ be a real sequence such that the density condition (2.9) holds for some $d > 0$ and some partition $\{I_n\}$ that satisfies a stronger log-shortness condition:*

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

Then we will automatically have that

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

Accordingly, condition (2.10) will be significantly simplified and one will only need to check that

$$\sum_{\lambda_i, \lambda_j \in \Lambda,\ \lambda_i \neq \lambda_j} \frac{\log_- |\lambda_i - \lambda_j|}{1 + \lambda_j^2} < \infty$$

to conclude that $\mathbf{G}_\Lambda = d$. Note that since $\log_-$ is supported on $[0, 1]$, only points that are close to each other contribute to the last sum.

Consider, for instance, 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 $\Lambda = \{\lambda_n\}$ be such that

$$\alpha |n|^{\alpha - 1} \leq \#(\Lambda \cap I_n) \leq \alpha |n|^{\alpha - 1} + 1 \tag{2.12}$$

for all n and

$$\lambda_{k+1} - \lambda_k \gtrsim e^{-|k|/\log^2 |k|} \tag{2.13}$$

for all $k, |k| > 1$. Then by the previous discussion $\mathbf{G}_\Lambda = 1$.

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

On the other hand, if (2.12) holds, but instead of (2.13) we have that on each I_n,

$$\lambda_{k+1} - \lambda_k \lesssim e^{-|k|/\log |k|}$$

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

Theorem 9 easily implies Beurling's Gap Theorem 1. Notice that if the support of μ has long complement, then for any short partition of $\mathbb{R}$ infinitely many intervals of the partition will be contained in the gaps of $X = \text{supp}\,\mu$. Hence X does not contain a sequence Λ that satisfies the density condition (2.9) on a short partition with $d > 0$. Therefore $\mathbf{G}_X = 0$.

Another application of theorem 9 in the 'positive' direction produces examples given by the result of Benedicks in [**12**]. Benedicks' theorem provided some of the very few examples of sets with positive gap characteristic that existed in the literature before [**127**]. This result is discussed in chapter 6.

6. Appendix: Proof of the gap formula

The rest of this chapter is occupied by the proof of theorem 9 and can be skipped by those readers not willing to go into hard technical details. The first step of the proof follows the approach of de Branges which relies on the Krein-Milman theorem on the existence of extremal points of a star-weakly compact convex set. That step is not very technical but quite important for several results in these notes. For that purpose we formulate a suitable version of de Branges' theorem 66 from [**26**], see theorem 10 in section 7, which may be of independent interest. Further versions of that result are included in other chapters of these notes. Switching our attention to extreme points of the set of measures with a fixed spectral gap allows us to discretize the problem and ultimately reduce it to measures concentrated on d-uniform sequences. That reduction is made via the Toeplitz approach with the use of the real Dirichlet space in the upper half-plane and the Beurling–Malliavin multiplier theorem.

Before starting the proof, let us introduce the following notations. 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 $\mathbb{R} \setminus I$.

Recall that by Π 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,

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

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

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

where $\widetilde{h}'$ denotes the derivative of a harmonic conjugate function $\widetilde{h}$.

The proof relies on several technical lemmas supplied in section 7.

Proof of theorem 9, part I:

First suppose that X contains a d-uniform sequence $\Lambda = \{\lambda_n\}$ with $d > \frac{1}{2\pi}$. We will show that $\mathbf{G}_X \geq \frac{1}{2\pi}$. (Obviously, it is enough to prove this for any positive constant in place of $\frac{1}{2\pi}$; our choice is due to purely technical reasons.)

Choose $\varepsilon > 0$. Let

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

be the short monotone partition from the definition of d-uniform sequences corresponding to Λ, see remark 2. We will assume that

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

for all n, which can be achieved by deletion of some of the points from Λ, see remark 1, and uniting finitely many I_n into one, if necessary. We will also assume that $|I_n| >> 1/\varepsilon >> 1$ for all n.

By lemma 1 and corollary 2 we can suppose 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 Λ, i.e. that $a_n = \lambda_{k_n}$ and $a_{n+1} = \lambda_{k_{n+1}}$ 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

$$E_n = \sum_{\lambda_{k_n} \le \lambda_k, \lambda_l \le \lambda_{k_{n+1}},\ \lambda_k \ne \lambda_l} \log |\lambda_k - \lambda_l| \tag{2.14}$$

and assuming that (2.10) is satisfied with these E_n. Such an assumption can be made because if the sum in (2.10) becomes infinite with E_n defined by (2.14), one can, for instance, delete the first point λ_{k_n+1} from Λ on all I_n for large n. After the addition of λ_{k_n} and deletion of λ_{k_n+1} in the sum defining E_n, each term in (2.10) 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 satisfy the density condition.

Our goal is to show that $\mathbf{G}_\Lambda \ge 1$ by producing a measure on Λ with spectral gap of the size arbitrarily close to 1. Due to connections discussed in section 3, existence of such a measure will follow from non-triviality of a certain Toeplitz kernel.

Since the lengths of $(\lambda_n, \lambda_{n+1})$ are bounded from above, we can apply lemma 5. Denote by Θ 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 so that $u - v$ becomes an atom on each J_n:

Claim 1. *There exists a subsequence $\{b_n\}$ of the sequence $\{a_n\}$ and smooth functions v_1, v_2 such that:*

1) $|v_1'| < \varepsilon/2$ and $u - v_1 = 0$ at all a_n;
2) $J_n = (b_n, b_{n+1})$ is a short monotone partition;
3) $|v_2'| < \varepsilon/2$ 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 of claim. First, choose a smooth function v_1 satisfying 1. Such a function exists because

$$|2\pi\Delta_n - |I_n|| \le 2\pi << \frac{\varepsilon}{2}|I_n|.$$

Notice that because the sequence I_n is short and

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

condition 1 implies

$$u - v_1 \in L^1_\Pi. \tag{2.15}$$

Choose $b_0 = a_0 = 0$. Choose $b_1 = a_{n_1} > b_0$ to be the smallest element of $\{a_k\}$ 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 (2.15) such an a_{n_1} will always exist. After that proceed choosing $b_2, b_3, ...$ in the following way: If b_i is chosen, choose $b_{i+1} = a_{n_{i+1}}$ to be the smallest element of $\{a_k\}$ satisfying $a_{n_{i+1}} > b_i$,

$$\left| \int_{b_i}^{a_{n_{i+1}}} (u - v_1) dx \right| < \frac{\varepsilon}{8} (a_{n_{i+1}} - b_i)^2 \tag{2.16}$$

and

$$a_{n_{i+1}} - b_i \geq b_i - b_{i-1}.$$

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

$$|J_k| < \left(\left[\frac{10}{\varepsilon} \right] + 1 \right) |I_{n_k}| \tag{2.17}$$

where $[\cdot]$ stands for the integer part of a real number.

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

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

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

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

and

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

It follows that

$$|J_1| = (a_{n_1 - 1} - b_0) + |I_{n_1}| \leq 9 \varepsilon^{-1} |I_{n_1 - 1}| + |I_{n_1 - 1}| \leq \frac{10}{\varepsilon} |I_{n_1 - 1}| \tag{2.18}$$

(if ε is small enough) . For the inductional step, assume that (2.17) holds for $k = l - 1$. For $J_l = (b_{l-1}, b_l), b_l = a_{n_l}$ there are two possibilities:

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

or

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

In the first case we prove (2.18) in the same way as in the basic step. In the second case we notice that, by monotonicity of I_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 $|J_{l-1}| / |I_{n_{l-1}}| \leq [10/\varepsilon] + 1$. Accordingly, the number of intervals in $(b_{l-1}, a_{n_l - 1})$ is at most $[10/\varepsilon]$. Therefore the number of intervals in $J_l = (b_{l-1}, b_l)$ is at most $[10/\varepsilon] + 1$. Now, since I_{n_l} is the largest interval in J_l, we again get (2.17), which implies shortness of J_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

$$T_k(x) = \frac{\varepsilon}{4}\text{dist}(x, \mathbb{R} \setminus J_k).$$

Notice that because of (2.16), for each k there exists a constant $C_k, |C_k| \leq 1$ such that

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

Now define v_2 as a smoothed-out sum $\sum C_k T_k$ that satisfies $|v_2'| < \varepsilon/2$ 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_Π-norms:

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

(for more on atomic decompositions see [**36**]). $\triangle$

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 5.

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 belong to L^1_Π.

Without loss of generality $\arg(\Theta \bar{S} I_+ \bar{I}_-) = 0$ at 0.

Claim 2. *The function*

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

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

Proof of claim. 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, $w(0) = 0$ with large multiplicity). The function

$$[v - (\arg I_- - \arg I_+)]\, /x$$

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 $q + i\widetilde{q}$ in $\mathbb{C}_+$ is finite, i.e. that

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

where $D(r)$ is the semidisc $\{|z| < r\} \cap \mathbb{C}_+$.

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

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

First, notice that the harmonic conjugate of $\frac{w}{x} = q$ is $\frac{\widetilde{w}}{x} = \widetilde{q}$ (we can assume that $\widetilde{w}(0) = 0$) and $\left(\frac{w}{x}\right)' = \frac{w'}{x} - \frac{w}{x^2}$. Recall that by our construction (see claim 1) w' is bounded from below and w is zero at the endpoints of every J_n. Hence $|w| \lesssim |J_n|$ on every J_n. Since the partition $\{J_n\}$ is short, it follows that $|w(x)| = o(|x|)$ and that $\frac{w}{x^2}$ is a bounded function. Therefore,

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

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

$$\int_{\mathbb{R}} w' \tilde{w} \frac{dx}{x^2} = \sum_n \sum_k \int_{J_n} w' \tilde{w}_k \frac{dx}{x^2}. \tag{2.19}$$

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

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

In this case

$$\begin{aligned} \left| \int_{J_n} w' \tilde{w}_k \frac{dx}{x^2} \right| &\lesssim \int_{J_n} |w'| \frac{|J_k|^3}{\operatorname{dist}^2(J_k, x)} \frac{dx}{x^2} \\ &\lesssim \frac{|J_k|^3}{1 + \operatorname{dist}^2(J_n, 0)} \int_{J_n} \frac{dx}{\operatorname{dist}^2(J_k, x)}. \end{aligned} \tag{2.20}$$

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 [**36**]. In the last inequality we used the property

$$\int_{J_n} |w'(x)| dx \lesssim |J_n|. \tag{2.21}$$

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

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

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

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

$$\le \frac{|J_k|}{1+\operatorname{dist}^2(J_k,0)} \frac{|J_k|^2}{\operatorname{dist}^2(J_k,J_n)} \int_{J_n} |w'| \lesssim \frac{|J_k||J_n|}{1+\operatorname{dist}^2(J_k,0)}. \tag{2.22}$$

Finally, the last case is

$$\operatorname{dist}(J_n,J_k) < \min(|J_n|,|J_k|). \tag{2.23}$$

Again we assume that $n \ge k > 0$. 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

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

By the first inequality of (2.52) in lemma 7, applied to $w' = h, (\arg\Theta)' = f$ and $(x+v)' = g$, for $1 << k \le n$ we have

$$-\int_{J_n} w' \left[\int_{J_k} \log|t-x|w'(t)dt\right] \frac{dx}{x^2} \tag{2.24}$$
$$\lesssim -\frac{1}{1+\operatorname{dist}^2(J_n,0)} \left[\iint_{J_n\times J_k} \log|t-x|w'(x)w'(t)dxdt + C|J_n|^2\right]$$

and we can work with the latter integral instead of the former. To verify the conditions of lemma 7, note that if $E = J_n$ and $I = J_k$, then for large enough k,n we will have $E \cup I \subset [d,2d]$ as a consequence of the shortness condition and (2.23). The relation (2.51) will be satisfied because $w = 0$ at the endpoints of J_k, see condition 3 of Claim 1. The constant D_1 satisfies

$$D_1 \le \int_{J_n} (\arg\Theta)' + \int_{J_n} (x+v)' \lesssim |J_n|.$$

Finally,

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

because w_k is an atom.

To estimate the integral in the right-hand side of (2.24), denote

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

where the functions v_1, v_2 are from claim 1. Also denote $p_n = p|_{J_n}$ 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 that $|v_1'|, |v_2'| < \varepsilon$. Then

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

Notice that

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

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

$$\left|\int_{J_k} p' \widetilde{v_2^n} dx\right| = |< p_k, v_2^n >_{\mathcal{D}} | = \left|\int_{J_n} \widetilde{p_k} v_2' dx\right| \lesssim |J_n|^2$$

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

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

Hence

(2.25)

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

For the last integral we have

(2.26)

$$-\iint_{J_n\times J_k} \log|t-x|p'(x)p'(t)dxdt = -\sum_{I_i\subset J_k}\sum_{I_j\subset J_n}\iint_{I_i\times I_j} \log|t-x|p'(x)p'(t)dxdt.$$

To estimate

$$(2.27)\qquad -\iint_{I_i\times I_j} \log|t-x|p'(x)p'(t)$$
$$= \iint_{I_i\times I_j} \log_-|t-x|p'(x)p'(t) - \iint_{I_i\times I_j} \log_+|t-x|p'(x)p'(t)$$

we consider 3 cases. First, to estimate the integral in the case when $i = j$, 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

$$(2.28)\qquad \int_{I_l} (\arg\Theta)' = \int_{I_l} (v_1'+1) = 2\pi\Delta_l + O(1) = |I_l| + O(1).$$

Hence

$$\iint_{I_j\times I_j} \log_-|t-x|p'(x)p'(t)$$
$$= \iint_{I_j\times I_j} \log_-|t-x|(\arg\Theta)'(x)(\arg\Theta)'(t)$$
$$-2\iint_{I_j\times I_j} \log_-|t-x|(1+v_1'(x))(\arg\Theta)'(t)$$
$$+\iint_{I_j\times I_j} \log_-|t-x|(1+v_1'(x))(1+v_1'(t))$$
$$+\iint_{I_j\times I_j} \log_-|t-x|(\arg\Theta)'(x)(\arg\Theta)'(t) + O(|I_j|).$$

For the last integral we have

$$\iint_{I_j\times I_j} \log_- |t-x|(\arg\Theta)'(x)(\arg\Theta)'(t)$$
$$= \sum_{(\lambda_l,\lambda_{l+1})\subset I_j}\ \sum_{(\lambda_m,\lambda_{m+1})\subset I_j} \int_{\lambda_l}^{\lambda_{l+1}}\int_{\lambda_m}^{\lambda_{m+1}} (\arg\Theta)'(x)(\arg\Theta)'(t).$$

Using the properties that

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

and that

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

for all s by lemma 5, we can apply lemma 6, parts 1-3. Assuming that $\lambda_l \le \lambda_m$ we conclude that

$$\int_{\lambda_l}^{\lambda_{l+1}}\int_{\lambda_m}^{\lambda_{m+1}} \log_- |t-x|(\arg\Theta)'(x)(\arg\Theta)'(t) \lesssim$$

$$\begin{cases} \log_-(\lambda_m-\lambda_{l+1}), \\ \qquad \text{if} \quad \lambda_m > \lambda_{l+1}; \\ \max(\log_-(\lambda_l-\lambda_{l-1}), \log_-(\lambda_{l+1}-\lambda_l), \log_-(\lambda_{l+2}-\lambda_{l+1}), \log_-(\lambda_{l+3}-\lambda_{l+2})) + 1, \\ \qquad \text{if} \quad \lambda_m = \lambda_{l+1}; \\ \max(\log_-(\lambda_l-\lambda_{l-1}), \log_-(\lambda_{l+1}-\lambda_l), \log_-(\lambda_{l+2}-\lambda_{l+1})) + 1, \\ \qquad \text{if} \quad \lambda_m = \lambda_l; \end{cases}$$

which implies

$$\iint_{I_j\times I_j} \log_- |t-x|p'(x)p'(t)dxdt \lesssim$$

$$\sum_{a_j\le\lambda_k,\lambda_l\le a_{j+1},\ \lambda_k\ne\lambda_l} \log_- |\lambda_k-\lambda_l| + |I_j|. \tag{2.29}$$

To estimate the integral of $\log_+$, first notice that by lemma 6, part 5, and (2.28),

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

for any $t \in I_j$.

Together with part 4 of lemma 6 and (2.28) we get:

(2.30)
$$\begin{aligned}
&\iint_{I_j\times I_j} \log_+|t-x|p'(x)p'(t) = \iint_{I_j\times I_j} \log_+|t-x|(\arg\Theta)'(x)(\arg\Theta)'(t)\\
&\quad -2\iint_{I_j\times I_j} \log_+|t-x|(v_1'(x)+1)(\arg\Theta)'(t)\\
&\quad +\iint_{I_j\times I_j} \log_+|t-x|(v_1'(x)+1)(v_1'(t)+1)\\
&= \sum_{a_j\le\lambda_l,\lambda_m<a_{j+1}} \int_{\lambda_l}^{\lambda_{l+1}}\int_{\lambda_m}^{\lambda_{m+1}} \log_+|t-x|(\arg\Theta)'(x)(\arg\Theta)'(t) - |I_j|^2\log|I_j|\\
&\quad + O(|I_j|^2)\\
&\ge 4\pi^2 \sum_{a_j\le\lambda_l,\lambda_m<a_{j+1}} \log_+|\lambda_l-\lambda_m| - |I_j|^2\log|I_j| + O(|I_j|^2).
\end{aligned}$$

Next, let us consider the case when $i \neq j$ and the intervals I_i, I_j are not adjacent. This estimate is similar to (2.22), but we will do it using a different technique. Assume for example that $j > i+1$. For $\log_-$, recalling that $|I_k| > 1$ for all k, we get

$$-\iint_{I_i\times I_j} \log_-|t-x|p'(x)p'(t)dxdt = 0. \tag{2.31}$$

For $\log_+$ we have

(2.32)
$$\begin{aligned}
&-\iint_{I_i\times I_j} \log_+|t-x|p'(x)p'(t) = -\int_{a_i}^{a_{i+1}}\int_{a_j}^{a_{j+1}} \log_+|t-x|p'(x)p'(t)\\
&= -\int_{a_i}^{a_{i+1}}\int_{a_j}^{a_{j+1}} \log_+|t-x|(\arg\Theta - x - v_1)'(x)(\arg\Theta - x - v_1)'(t)\\
&\le -\int_{I_j}\left(\log|a_{i+1}-t|\int_{I_i}(\arg\Theta)'(x)dx - \log|a_i-t|\int_{I_i}(v_1'+1)(x)dx\right)(\arg\Theta)'(t)dt\\
&\quad +\int_{I_j}\left(\log|a_i-t|\int_{I_i}(\arg\Theta)'(x)dx - \log|a_{i+1}-t|\int_{I_i}(v_1'+1)(x)dx\right)(v_1'+1)(t)dt\\
&\le 2|I_i||I_j|
\end{aligned}$$

Here we used the properties that $\operatorname{dist}(I_i, I_j) \ge |I_i|$ by monotonicity and (2.28).

In the case when I_i and I_j are adjacent, i.e. $j = i+1$, 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 (2.29) and (2.30) yield

$$||p_n||^2_{\mathcal{D}} \lesssim \frac{1}{4\pi^2}|I_n|^2\log|I_n| - E_n + |I_n|^2. \tag{2.33}$$

Hence

$$\iint_{I_i\times I_{i+1}}\log|t-x|p'(x)p'(t) =< p_i, p_{i+1} >_{\mathcal{D}}\le ||p_i||^2+||p_{i+1}||^2$$
$$\lesssim \left(\frac{1}{4\pi^2}|I_i|^2\log|I_i|-E_i\right)+\left(\frac{1}{4\pi^2}|I_{i+1}|^2\log|I_{i+1}|-E_{i+1}\right)$$
$$+|I_i|^2+|I_{i+1}|^2. \tag{2.34}$$

Now we can return to estimating

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

in the case when $|k-n|\le 1$. Using the estimates (2.24), (2.25) and (2.26) we obtain

$$-\int_{J_n} w'\widetilde{w_k}\frac{dx}{x^2}=\frac{-\sum_{I_i\subset J_k}\sum_{I_j\subset J_n}\iint_{I_i\times I_j}\log|t-x|p'(x)p'(t)dxdt+O(|J_n|^2)}{1+\text{dist}^2(0,J_n)}.$$

The estimates (2.29 – 2.34) yield

$$-\sum_{I_i\subset J_n}\sum_{I_j\subset J_k}\iint_{I_i\times I_j}\log|t-x|p'(x)p'(t)dxdt\lesssim$$
$$\sum_{I_i\subset J_k\cup J_n}\left(\frac{1}{4\pi^2}|I_i|^2\log|I_i|-E_i+|I_i|^2\right)+\sum_{I_i,I_j\subset J_k\cup J_n}|I_i||I_j|\le$$
$$\sum_{I_i\subset J_k\cup J_n}\left(\frac{1}{4\pi^2}|I_i|^2\log|I_i|-E_i\right)+|J_n|^2+|J_k|^2.$$

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

$$-\int_{J_n} w'\widetilde{w_k}\frac{dx}{x^2}\lesssim$$

$$\frac{1}{1+\text{dist}^2(0,J_n)}\left[\sum_{I_i\subset J_k\cup J_n}\left(\frac{1}{4\pi^2}|I_i|^2\log|I_i|-E_i\right)+|J_n|^2+|J_k|^2\right]. \tag{2.35}$$

To continue the estimate of the right-hand side of (2.19) notice that

$$\sum_n\sum_k-\int_{J_n} w'\widetilde{w_k}\frac{dx}{x^2}$$
$$=\sum_n\left[\sum_{k\ :\ \max(|J_k|,|J_n|)\le\text{dist}(J_k,J_n)}\right]$$
$$+\sum_n\left[\sum_{k\ :\ \min(|J_k|,|J_n|)\le\text{dist}(J_k,J_n)<\max(|J_k|,|J_n|)}\right]$$
$$+\sum_n\left[\sum_{k\ :\ \text{dist}(J_k,J_n)<\min(|J_k|,|J_n|)}\right]$$
$$=I+II+III.$$

For the first sum, by (2.20), we get

$$I \lesssim \sum_k \left[\sum_{n\,:\,n>k,\ \mathrm{dist}(J_k,J_n)>|J_k|} \frac{|J_k|^3}{1+\mathrm{dist}^2(J_n,0)} \int_{J_n} \frac{dx}{\mathrm{dist}^2(J_k,x)} \right] \tag{2.36}$$

$$\leq \sum_k \frac{|J_k|^3}{1+\mathrm{dist}^2(J_k,0)} \frac{1}{|J_k|} = \sum_k \frac{|J_k|^2}{1+\mathrm{dist}^2(J_k,0)} < \infty.$$

For the second sum, by (2.22),

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

Recall that by our assumption $|J_n| \leq \mathrm{dist}(J_n, 0)$ for all $n \neq 0, -1$. We can 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, n.

$$\sum_{n\geq 0} \left[\sum_{k\geq 0\,:\,k\neq n,\ \mathrm{dist}(J_k,J_n)<\max(|J_k|,|J_n|)} \frac{|J_k||J_n|}{1+\mathrm{dist}^2(J_n,0)} \right]$$

$$= 2\sum_{n\geq 0} \left[\sum_{k\geq 0\,:\,k<n,\ \mathrm{dist}(J_k,J_n)<|J_n|} \frac{|J_k||J_n|}{1+\mathrm{dist}^2(J_n,0)} \right]$$

$$\leq 4 \sum_{n\geq 0} \frac{|J_n|^2}{1+\mathrm{dist}^2(J_n,0)} < \infty.$$

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

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

Finally, for the third sum by (2.35),

$$III \lesssim \sum_n \frac{\sum_{k\,:\,|k-n|\leq 1} \left[\sum_{I_i\subset J_k\cup J_n} \left(\frac{1}{4\pi^2}|I_i|^2\log|I_i| - E_i\right) + |J_n|^2 + |J_k|^2\right]}{1+\mathrm{dist}^2(0,J_n)}$$

$$\lesssim \sum_n \frac{1}{1+\mathrm{dist}^2(0,J_n)} \left[\sum_{I_i\subset J_n} \left(\frac{1}{4\pi^2}|I_i|^2\log|I_i| - E_i\right) + |J_n|^2 \right] < \infty$$

because Λ satisfies the energy condition on I_n. Altogether these estimates give us

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

Now let us estimate the integrals over the circular part of $\partial D(r)$. We need to show that the integrals

$$-\int_{\partial D(r)\setminus\mathbb{R}} \widetilde{q}\, dq = rI'(r), \qquad I(r) = \frac{1}{2}\int_0^\pi \widetilde{q}^2\left(re^{i\phi}\right) d\phi,$$

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

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

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

As we will see shortly, $I(r) \not\to \infty$ for *any* $h \in L^1(1+|x|^{-1})$ in place of $\widetilde{q}$ (recall that $\widetilde{q} = \widetilde{w}/x$ 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 $h + i\widetilde{h}$ 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

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

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

$$l^M \in L^{1,\infty}(\mathbb{T}). \tag{2.37}$$

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

$$\frac{1}{\epsilon}\int_{C_\epsilon} |h + i\widetilde{h}|^2 |dz| \not\to \infty, \qquad C_\epsilon = \{|1-z| = \epsilon\} \cap \mathbb{D}.$$

We have

$$\frac{1}{\epsilon}\int_{C_\epsilon} |h + i\widetilde{h}|^2 \leq \epsilon \int_{C_\epsilon} |l|^2 \lesssim \left[\epsilon l^M(\zeta)\right]^2 + \left[\epsilon l^M(\bar{\zeta})\right]^2,$$

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

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

on an interval of length ϵ, which would contradict (2.37). △

Let

$$\phi = \arg(\Theta \bar{S} I_+ \bar{I}_-)/2.$$

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

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

see for instance lemma 12 in [**26**]. Similar statements are also discussed in chapter 8.

In other words, if Φ and M are outer functions,

$$\Phi = \exp(i\phi - \widetilde{\phi}),\ M = \exp(im - \widetilde{m}),$$

then ΦM is bounded in $\mathbb{C}_+$.

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

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

Denote

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

Then the difference

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

satisfies $|d| < \pi$.

Put

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

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

Consider an outer function $\Psi = \exp(il - \widetilde{l})$. Then

$$\bar{S}^{2\varepsilon}\Psi = \bar{J}\bar{\Psi}$$

or equivalently

$$\bar{S}^{2\varepsilon}J\Psi = \bar{\Psi}$$

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

Moreover, the ratio Ψ/M is equal to $\exp\left(i\frac{d}{2} - \frac{\tilde{d}}{2}\right)$. Since $|d| < \pi$, Ψ/M belongs to any L^p_Π, $p < 1$. 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

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

where $[\cdot]$ again denotes the entire part of a real number. Then

$$|d - \alpha| < \frac{2\pi}{5}. \tag{2.38}$$

Since $d_1 = d_2 = \pi n$ at the points $\{c_n\} = \{J = \pm 1\}$, the function α only takes values $k\frac{\pi}{5}$, $k = -4, -3, ..., 4$. Therefore α can be represented as

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

where β_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 $[c_n, c_{n+1}]$. For each $n = 1, 2, ..., 8$ one can choose an inner function Q_n so that

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

Notice that then

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

Because of (2.38) we have

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

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 $\{J=1\}$ has Beurling–Malliavin exterior density at most 2ε, see section 4.6 of [**103**]. This also follows from theorem 52 in chapter 8 with $\kappa=0$. See section 7 for the definition of the density or section 3 in chapter 3 for a more detailed discussion.

By our construction the Beurling–Malliavin exterior density of each of the sets $\{Q_n=1\}$ is the same as that of $\{J=1\}$, i.e. at most 2ε. Consequently, the kernel

$$N^\infty[\bar{S}^{17\varepsilon}\prod_n Q_n]$$

contains a non-zero function τ, see sections 4.2 and 4.6 of [**103**] or theorem 52 in chapter 8.

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

Therefore

$$\begin{aligned}&\bar{\Theta}S^{1-20\varepsilon}\kappa\tau\prod_{n=1}^{4}(1+Q_n)\prod_{n=5}^{8}(1-Q_n)\Psi\Phi\\&\qquad=\left(\bar{S}^\varepsilon I_+\kappa\right)\left(\bar{S}^{17\varepsilon}\prod_{n=1}^{4}(1+Q_n)\prod_{n=5}^{8}(1-Q_n)\tau\right)\left(\bar{S}^{2\varepsilon}\Psi\right)\left(\bar{\Theta}S^1\bar{I}_+\Phi\right)\in\bar{H}^2.\end{aligned}$$

Accordingly, the space K_Θ contains the function

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

Now we could simply refer to theorem 8 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 σ_1 is the Clark measure corresponding to Θ concentrated on $\{\Theta=1\}=\Lambda$. Since $1+\Theta$ is outer and f is divisible by $S^{1-20\varepsilon}$ in the upper half-plane, $Kf\sigma_1$ decreases faster than $\exp[-(1-21\varepsilon)y]$ along the positive y-axis. Hence $\mu=f\sigma_1$ is the measure concentrated on Λ with the spectral gap at least $(1-21\varepsilon)$. ■

Proof of theorem 9, part II:

Now suppose that $\mathbf{G}_X>\frac{1}{2\pi}$ but X does not contain a $\frac{1}{2\pi}$-uniform sequence.

By lemma 9 there exists a discrete increasing sequence $\Lambda=\{\lambda_n\}_{n\in\mathbb{Z}}\subset X$ and a measure ν, $\operatorname{supp}\nu=\Lambda$ such that ν has a spectral gap of the size greater than $1/2\pi$ 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 5 and consider the inner function Θ corresponding to Λ. 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 section 8. Note that $f = \exp(i\phi - \widetilde{\phi})$ in $\mathbb{C}_+$, where $2\phi = \arg\Theta - x$.

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

$$\int_{\cup\Gamma_n} \phi'\widetilde{\phi}\frac{dx}{x^2} \tag{2.39}$$

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

First let us choose a short monotone partition $\{I_n\}$ of $\mathbb{R}$ such that Λ has a subsequence which satisfies the density condition (2.9) with $d = 1/2\pi$ on that partition:

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

$$\#(\Lambda \cap (a_0, a_1]) \geq \frac{1}{2\pi}(a_1 - a_0).$$

Note that such a point always exists because Λ supports a measure with a spectral gap greater than $\frac{1}{2\pi}$: otherwise we would be able to choose a long sequence of intervals satisfying (2.48) in lemma 2 with $a = \frac{1}{2\pi}$ and arrive at a contradiction. After $a_i,\ i \geq 1$ is chosen, choose $a_{i+1} > a_i$ as the smallest point such that

$$\#(\Lambda \cap (a_i, a_{i+1}]) \geq \frac{1}{2\pi}(a_{i+1} - a_i) \quad \text{and} \quad (a_{i+1} - a_i) \geq (a_i - a_{i-1}).$$

Choose $a_i, i < 0$ in a similar way. Put $I_n = (a_n, a_{n+1}]$. Again by lemma 2, $\{I_n\}$ has to be short.

In what follows we will assume, without loss of generality, that $\frac{1}{2\pi}|I_n| = \#(\Lambda \cap I_n)$.

Note that since X does not contain a $\frac{1}{2\pi}$-uniform sequence, the sum in the energy condition (2.10), for Λ and $\{I_n\}$, has to be infinite. At the same time, a part of that sum has to be finite:

Claim 3.

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

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

Define the intervals J_n and the function v like in Claim 1 of part I of the proof, with Φ replacing Θ. 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_Π-norms.

Like in part I, we can use 'atomic' estimates to show that if $\text{dist}(J_k, J_n) > \max(|J_k|, |J_n|)$ and $x \in J_n$, then

$$|\widetilde{w_k}(x)| \lesssim \frac{|J_k|^3}{\text{dist}^2(x, J_k)}.$$

By monotonicity and shortness of J_k we conclude that

$$\sum_{\lambda_i \in J_n} \frac{|\widetilde{w_k}(\lambda_i)|}{\lambda_i^2} \lesssim \sum_{\lambda_i \in J_n} \frac{|\widetilde{w_k}(\lambda_i)|}{1 + \text{dist}^2(0, J_n)} \lesssim \frac{1}{1 + \text{dist}^2(0, J_n)} \int_{J_n} \frac{|J_k|^3}{\text{dist}^2(x, J_k)} dx.$$

Hence, similarly to (2.36),

$$\sum_n \left[\sum_{k\ :\ \mathrm{dist}(J_k,J_n)>\max(|J_k|,|J_n|)} \left[\sum_{\lambda_i\in J_n} \frac{|\widetilde{w_k}(\lambda_i)|}{\lambda_i^2} \right]\right]$$

$$= 2\sum_n \left[\sum_{k\ :\ k<n,\ \mathrm{dist}(J_k,J_n)>\max(|J_k|,|J_n|)} \left[\sum_{\lambda_i\in J_n} \frac{|\widetilde{w_k}(\lambda_i)|}{\lambda_i^2} \right]\right]$$

$$\lesssim \sum_n \left[\sum_{k\ :\ k<n,\ \mathrm{dist}(J_k,J_n)>\max(|J_k|,|J_n|)} \frac{1}{1+\mathrm{dist}^2(0,J_n)} \int_{J_n} \frac{|J_k|^3}{\mathrm{dist}^2(x,J_k)} dx \right] < \infty.$$

In other words, on each J_n

$$\sum_{k\ :\ \mathrm{dist}(J_k,J_n)>\max(|J_k|,|J_n|)} |\widetilde{w_k}| \le g_1$$

where g_1 is a positive function satisfying

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

Also for any $x \in J_n$

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

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

$$\begin{aligned} -\int_{J_k} \log|t-x|w'(t)dt &\ge -\int_{J_k} \log_+|t-x|(\arg\Phi)'(t)dt \\ &\quad + \int_{J_k} \log_+|t-x|(1+v')(t)dt - \text{const} \gtrsim -|J_k|. \end{aligned} \tag{2.40}$$

Here we again used that $\int_{J_k}(\arg\Phi)' = \int_{J_k}(1+v') = |J_k| + O(1)$, applied lemma 6, part 6, to the second integral in the second line and used the estimate

$$-\int_{J_k} \log_+|t-x|(\arg\Phi)'(t)dt \ge -\log(x-b)\int_{J_k}(\arg\Phi)', \tag{2.41}$$

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

Thus for $x \in J_n$

$$-\sum_{k\ :\ |k-n|>1,\ \mathrm{dist}(J_k,J_n)\le\max(|J_k|,|J_n|)} \widetilde{w_k}(x) \ge g_2(x),$$

where again

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

Also, for $x \in J_n$ and $k = n$ or $k = n \pm 1$, similar to (2.40),

$$\begin{aligned} &-\int_{J_k} \log_+|x-t|w'(t)dt \\ &\quad = -\int_{J_k} \log_+|x-t|(\arg\Phi)'(t)dt + \int_{J_k} \log_+|x-t|(1+v')(t)dt \gtrsim -|J_k|. \end{aligned} \tag{2.42}$$

by applying (2.41) to the first integral and lemma 6, part 5, to the second integral if $n = k$, or part 6 if $k = n \pm 1$.

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

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

for some function g satisfying

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

Therefore,

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

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

$$h = \exp\left(i\frac{\psi}{2} - \frac{\widetilde{\psi}}{2}\right).$$

Since $|h(\lambda_n)| = |\nu(\{\lambda_n\})|/\mu(\{\lambda_n\}) = 1$, we have that $\log|h(\lambda_n)| = 2\widetilde{\psi}(\lambda_n) = 0$ for all n.

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

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

Recall that $v \in L^1_\Pi, \widetilde{v} = \widetilde{w} - \widetilde{\psi} = \widetilde{w} - \log|h|/2 \in L^1_\Pi$ 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 $\widetilde{v}_y$ is bounded in $\mathbb{C}_+$ as well.

On each interval J_n choose λ_{k_n} so that

$$|\widetilde{v}(\lambda_{k_n})| = \max_{\lambda_i \in J_n} |\widetilde{v}(\lambda_i)|.$$

If the last sum is infinite then so is

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

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

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

Denote by $(\widetilde{v})^M$ the maximal non-tangential function of $\widetilde{v}$ in $\mathbb{C}_+$. The last equation implies that $(\widetilde{v})^M \notin L^1_\Pi$. But that contradicts the property that both $\widetilde{v}$ and v belong to L^1_Π. △

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

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

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

$$|f(x)| = |(1-\Theta(x))K\nu(x)| \le 2\left|\int \frac{1}{t-x}d\nu(t)\right| \le 6||\nu||\ |\lambda_{n+1}-\lambda_n|^{-1}. \tag{2.43}$$

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

$$\begin{aligned}
-\int_{\cup\Gamma_n}\phi'\widetilde{\phi}\frac{dx}{x^2} &\lesssim \sum_n \frac{1}{1+\lambda_n^2}\int_{\Gamma_n}(\arg\Theta)'\log_+|f|dx + \text{const}\\
&\lesssim \sum_n \frac{1}{1+\lambda_n^2}\log_-|\lambda_{n+1}-\lambda_n|\int_{\Gamma_n}(\arg\Theta)'dx + \text{const}\\
&\lesssim \sum_n \frac{1}{1+\lambda_n^2}\log_-|\lambda_{n+1}-\lambda_n| + \text{const} < \infty \qquad (2.44)
\end{aligned}$$

by (2.43) and claim 3.

It follows that

$$\begin{aligned}
-\int_{\cup\Gamma_n} w'\widetilde{w}\frac{dx}{x^2} = -4\int_{\cup\Gamma_n}\phi'\widetilde{\phi}\frac{dx}{x^2} + 2\int_{\cup\Gamma_n}\phi'\widetilde{v}\frac{dx}{x^2}&\\
+ 2\int_{\cup\Gamma_n} v'\widetilde{\phi}\frac{dx}{x^2} - \int_{\cup\Gamma_n} v'\widetilde{v}\frac{dx}{x^2} < \infty.&
\end{aligned} \tag{2.45}$$

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

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

Therefore

$$\left|\int_{\cup\Gamma_k\cap J_n}\phi'\widetilde{v}\frac{dx}{x^2}\right| \le \int_{\cup\Gamma_k\cap J_n}|\phi'|dx\frac{\sup_{x\in J_n}|\widetilde{v}(x)|}{1+\text{dist}^2(0,J_n)} \asymp |J_n|\frac{\sup_{x\in J_n}|\widetilde{v}(x)|}{1+\text{dist}^2(0,J_n)}$$

and summing over all n we get

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

The third integral on the right-hand side of (2.45) is finite because v' is bounded and $\widetilde{\phi} = \log|f|$ is in L^1_Π. The last integral is finite because v' is bounded and $\widetilde{v} = \widetilde{\phi} - \widetilde{w} \in L^1_\Pi$. The first integral is finite by (2.44).

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

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

Then

$$\int_{\cup\Gamma_k} w'\widetilde{w}\frac{dx}{x^2} = \sum_n\left[\int_{\cup\Gamma_k\cap J_n} w'\widetilde{q_n^*}\frac{dx}{x^2} + \int_{\cup\Gamma_k\cap J_n} w'\widetilde{q_n}\frac{dx}{x^2}\right].$$

The first integral can be, once again, estimated like in (2.20), i.e. using the property that each w_i is an atom, and the sum of such integrals shown to be finite. For the second integral we obtain

$$\int_{\cup\Gamma_k\cap J_n} w'\widetilde{q_n}\frac{dx}{x^2} = \sum_{J_l\subset L_n}\int_{\cup\Gamma_k\cap J_n} w'\widetilde{w_l}\frac{dx}{x^2}.$$

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

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

we get

$$\left|\int_{\cup\Gamma_k\cap J_n} w'\widetilde{w_l}\frac{dx}{x^2}\right| \lesssim \frac{|J_l||J_n|}{1+\operatorname{dist}^2(J_l,0)},$$

see (2.22). In the case

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

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

$$-\int_{\cup\Gamma_k\cap J_n} w'\widetilde{w_l}\frac{dx}{x^2} = \int_{\Gamma_k\cap J_n} w'\left[\int_{J_l}\frac{w(t)dt}{t-x}\right]\frac{dx}{x^2}$$
$$= -\int_{\cup\Gamma_k\cap J_n} w'\left[\int_{J_l}\log|t-x|w'(t)dt\right]\frac{dx}{x^2}.$$

Similarly to the first part of the proof (see the paragraph after (2.24)) one can verify the conditions of lemma 7 and apply the second inequality in (2.52), with $E=\cup\Gamma_k\cap J_n$ and $J_l=I$, to obtain

$$-\int_{\cup\Gamma_k\cap J_n} w'\left[\int_{J_l}\log|t-x|w'(t)dt\right]\frac{dx}{x^2}$$
$$\gtrsim -\frac{1}{1+\operatorname{dist}^2(J_n,0)}\left[\iint_{(\Gamma_k\cap J_n)\times J_l}\log|t-x|w'(x)w'(t)dxdt + C|J_n|^2\right].$$

Altogether, we obtain

$$-\int_{\cup\Gamma_k\cap J_n} w'\widetilde{q_n}\frac{dx}{x^2} \tag{2.46}$$
$$\gtrsim -\frac{1}{1+\operatorname{dist}^2(J_n,0)}\sum_{J_l\subset L_n}\left[\int_{\cup\Gamma_k\cap J_n} w'(x)\int_{J_l}\log|x-t|w'(t)dtdx - C|J_n||J_l|\right].$$

Furthermore, because of (2.40) (applied here with Θ in place of Φ),

$$-\int_{\cup\Gamma_k\cap J_n} w'(x)\left[\sum_{J_l\subset L_n}\int_{J_l}\log|x-t|w'(t)dt\right]dx$$
$$\gtrsim -\int_{\cup\Gamma_k\cap J_n} w'(x)\left[\int_{J_n}\log|x-t|w'(t)dt\right]dx - \sum_{J_l\subset L_n}|J_l||J_n|.$$

Let us remark right away that

$$\sum_n \frac{1}{1+\operatorname{dist}^2(J_n,0)} \sum_{J_l\subset L_n} |J_l||J_n|$$
$$\lesssim \sum_n \left[\sum_{l<n,\ \operatorname{dist}(J_l,J_n)<\max(|J_l|,|J_n|)} \frac{|J_l||J_n|}{1+\operatorname{dist}^2(J_n,0)}\right] \lesssim \sum_n \frac{|J_n|^2}{1+\operatorname{dist}^2(J_n,0)} < \infty$$

by the monotonicity and shortness of $\{J_n\}$.

To continue the estimates let us split the last integral:

$$\begin{aligned}
-\int_{\cup\Gamma_k\cap J_n} & w'(x)\left[\int_{J_n}\log|x-t|w'(t)dt\right]dx\\
&= \int_{\cup\Gamma_k\cap J_n}(\arg\Theta)'(x)\left[\int_{J_n}\log|x-t|(v'(t)+1)dt\right]dx\\
&\quad - \int_{\cup\Gamma_k\cap J_n}(\arg\Theta)'(x)\left[\int_{J_n}\log|x-t|(\arg\Theta)'(t)dt\right]dx\\
&\quad - \int_{\cup\Gamma_k\cap J_n}(v'(x)+1)\left[\int_{J_n}\log|x-t|(v'(t)+1)dt\right]dx\\
&\quad + \int_{\cup\Gamma_k\cap J_n}(v'(x)+1)\left[\int_{J_n}\log|x-t|(\arg\Theta)'(t)dt\right]dx\\
&= I+II+III+IV.
\end{aligned}$$

To estimate III and IV denote by D the constant satisfying

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

Notice that because $1-2\varepsilon < v'+1 < 1+2\varepsilon$ and $\int_{J_n} v'+1 = \int_{J_n}(\arg\Theta)' = |J_n|$, for any $y\in J_n$,

$$\int_{J_n}\log|y-t|(v'(t)+1)dt = |J_n|\log|J_n| + O(|J_n|) \tag{2.47}$$

and

$$\begin{aligned}
III &= -\int_{\cup\Gamma_k\cap J_n}(v'(x)+1)\left[\int_{J_n}\log|x-t|(v'(t)+1)dt\right]dx\\
&= -D|J_n|^2\log|J_n| + O(|J_n|^2).
\end{aligned}$$

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

$$\int_{\Gamma_k}(v'(x)+1)\log_+|x-t|dx$$
$$\geq \int_{\Gamma_k}(v'(x)+1)dx\ \frac{\int_{\lambda_k}^{\lambda_{k+1}}\log_+|x-t|dx}{\lambda_{k+1}-\lambda_k} - (\lambda_{k+1}-\lambda_k)\log 3$$

(recall that Γ_k is the middle one-third of $(\lambda_k,\lambda_{k+1})$ and that ε is very small).

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

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

Then $|\alpha - \frac{1}{3}| \leq \varepsilon$ on J_n. Hence one can apply lemma 6 part 5 to conclude that, for any $t \in J_n$,

$$\int_{\cup\Gamma_k \cap J_n} (v'(x)+1)\log_+|x-t|dx$$
$$\geq \int_{J_n} \alpha(x)\log_+|x-t|dx - \text{const}|J_n| \geq \left(\int_{J_n}\alpha(x)\right)\log|J_n| - \text{const}|J_n|$$
$$= D|J_n|\log|J_n| - \text{const}|J_n|.$$

Also,

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

Therefore

$$IV = \int_{\cup\Gamma_k \cap J_n} (v'(x)+1)\left[\int_{J_n}\log|x-t|(\arg\Theta)'(t)dt\right]dx$$

$$\geq \left(\int_{J_n}(\arg\Theta)'\right)(D|J_n|\log|J_n| - \text{const}|J_n| - \text{const}) \geq D|J_n|^2\log|J_n| - \text{const}|J_n|^2.$$

Combining the estimates we get

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

To estimate II notice that

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

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

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

Put $\alpha_k = \int_{\Gamma_k}(\arg\Theta)'$. Then

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

where the constants A_n satisfy

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

Using (2.47), I can be rewritten as

$$I = \sum_{\Gamma_k \subset J_n}\int_{\Gamma_k}(\arg\Theta)'(x)\left[\int_{J_n}\log|x-t|(v'(t)+1)dt\right]dx$$
$$= \left(\sum_{\Gamma_k\subset J_n}\alpha_k\right)|J_n|\log|J_n| + B_n$$

where again

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

By part 4 of lemma 5,

$$\alpha_k = \int_{\Gamma_k} (\arg \Theta)' > c > 0$$

for all k. Therefore, since there are $\frac{1}{2\pi}|J_n|$ intervals Γ_k in J_n,

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

Now, going back to (2.46), we obtain

$$- \sum_n \int_{\cup \Gamma_k \cap J_n} w' \widetilde{w_n} \frac{dx}{x^2}$$

$$\gtrsim \sum_n \frac{\frac{1}{4\pi^2}|J_n|^2 \log |J_n| - \sum_{\lambda_j, \lambda_k \in J_n, j \neq k} \log |\lambda_k - \lambda_j| - |J_n|^2 - |A_n| - |B_n|}{1 + \operatorname{dist}^2(J_n, 0)}$$

$$+\text{const}.$$

It remains to notice that the sum on right-hand side is positive infinite because otherwise Λ would satisfy the energy condition (2.10) and the density condition with $d = \frac{1}{2\pi}$. This contradicts (2.45).

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 Λ.

Indeed, if Λ is a sequence with arbitrarily large gaps, choose a large constant C and consider the set of all gaps R_k of Λ of the size larger than C.

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

After that one can add a separated set of points in every R_k and consider a slightly larger sequence $\Lambda' = \{\lambda'_n\} \supset \Lambda$ that satisfies $\sup_n(\lambda'_n - \lambda'_{n-1}) \le C$, $\operatorname{dist}(\Lambda, \Lambda' \setminus \Lambda) > C/2$, and

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

The inner function Θ, given by lemma 5, should then be chosen for the sequence Λ' instead of Λ. Recall that ν is the measure supported on Λ, chosen at the beginning of the second part of the proof. It has a spectral gap of size $1/2\pi$ 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 ν had a spectral gap of size $1/2\pi$, h is divisible by S (and h/S is outer because $K\nu$ has no zeros). Since $1 - \Theta$ has simple zeros at Λ' and $K\nu$ has simple poles at Λ, h has zeros at $\Upsilon = \Lambda' \setminus \Lambda$. Without loss of generality, Υ has bounded gaps. Since Υ is a separated sequence, there exists an inner function I, $\operatorname{spec}_I = \Upsilon$ such that $(\arg I)'$ is bounded (by lemma 5). If C is large enough, $|(\arg I)'| << \varepsilon$.

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 $(\arg\Theta - x - \arg I)/2$. Now we can apply claim 1 to $u = \arg\Theta - x - \arg I$ to obtain functions $v = v_1 + v_2$ satisfying the properties 1-5.

If one denotes by Γ_n the middle one-third of $(\lambda'_n, \lambda'_{n+1})$, then similarly to (2.43),

$$|S^{-1}(x)h(x)| = |(1-\Theta(x))K\nu(x)| \leq 6||\nu||\ |\lambda'_{n+1} - \lambda'_n|^{-1}$$

for any $x \in \Gamma_n$. The argument of the function h/S is $\arg\Theta - x$. Note that claim 3 still holds with Λ' in place of Λ, because Υ is separated and separated from Λ. Hence (2.44) still holds for $\phi = \arg\Theta - x$.

After that, using the property that $|(\arg I)'| << \varepsilon$, one can 'absorb' $\arg I$ into v_1 and replace v_1 with $y = v_1 + \arg I$. The remaining estimates, starting with (2.45), 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. Appendix: Technical lemmas

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

If Λ is a real sequence we define its exterior Beurling–Malliavin density as

$$D^*(\Lambda) = \sup\{d \mid \exists \text{ long } \{I_n\} \text{ such that } \#(\Lambda \cap I_n) \geq d|I_n|\ \forall n\}$$

if Λ is discrete and as ∞ otherwise.

An equivalent definition in terms of Toeplitz kernels is given in [**103**], section 4.6:

$$D^*(\Lambda) = \sup\{a : N[\bar{S}^{2\pi a}\Theta] = 0\},$$

where $\Theta(z)$ denotes some/any meromorphic inner function with $\text{spec}_\Theta = \Lambda$.

BM densities will appear in several other places in these notes. For a more detailed discussion see section 3 of chapter 3. Equivalence of the two definitions can be deduced from the material of chapter 8. It follows from the Beurling–Malliavin theorem discussed in chapter 5.

Note that the Beurling–Malliavin multiplier theorem implies that $N[\bar{S}^{2\pi a}\Theta]$ in the above definition can be replaced with any $N^p[\bar{S}^{2\pi a}\Theta]$, $0 < p \leq \infty$, the kernel in the Hardy space H^p, or with $N^+[\bar{S}^{2\pi a}\Theta]$, the kernel in the Smirnov class, see section 4.2 in [**103**].

LEMMA 1. *Let $X \subset \mathbb{R}$ be a closed set and let Λ be a discrete sequence. Then*

$$\mathbf{G}_{X\cup\Lambda} \leq \mathbf{G}_X + D^*(\Lambda).$$

PROOF. Let $D^*(\Lambda) = d_1$, $\mathbf{G}_X = d_2$ and $\mathbf{G}_{X\cup\Lambda} = d_3$. Let $\varepsilon > 0$ be a small number. By theorem 8,

$$N[\bar{\Theta}S^{2\pi d_3 - \varepsilon}] \neq 0$$

for some meromorphic inner Θ, $\text{spec}_\Theta \subset X \cup \Lambda$. Let

$$f \in N[\bar{\Theta}S^{2\pi d_3 - \varepsilon}].$$

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

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 Λ. The function fh belongs to

$$N[\bar{\Theta} S^{2\pi d_3 - 2\pi d_1 - 2\varepsilon}]$$

and is zero on Λ (obviously, we assume that $2\pi d_3 - 2\pi d_1 - 2\varepsilon > 0$). Finally, the function

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

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

$$l = \frac{1}{2\pi i}(1-\Theta) K l\sigma$$

where σ is the Clark measure for Θ, $\mathrm{supp}\,\sigma = \mathrm{spec}_\Theta \subset \Lambda \cup X$. Since l is divisible by

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

in $\mathbb{C}_+$ and $(1-\Theta)$ is an outer function in $\mathbb{C}_+$, $Kl\sigma$ is divisible by

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

in $\mathbb{C}_+$. Equivalently, the measure $l\sigma$ has a spectral gap of the size $d_3 - d_1 - \frac{\varepsilon}{\pi}$. Since l is zero on Λ, the measure $l\sigma$ is supported on X. Hence

$$\mathbf{G}_X \geq d_3 - d_1 - \frac{\varepsilon}{\pi} = \mathbf{G}_{X\cup\Lambda} - D^*(\Lambda) - \frac{\varepsilon}{\pi}.$$

□

The following statement can be viewed as a version of the first Beurling–Malliavin theorem, see [**103, 104**] or chapter 8.

Lemma 2. *Let Λ be a real sequence. Suppose that there exists a long sequence of intervals I_n such that*

$$\#(\Lambda \cap I_n) \leq a|I_n| \tag{2.48}$$

for all n, for some $a \geq 0$. Then $\mathbf{G}_\Lambda \leq a$.

Proof. Suppose that $\mathbf{G}_\Lambda = a + 3\varepsilon$ for some $\varepsilon > 0$. Then by theorem 8,

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

for some inner function Θ, $\mathrm{spec}_\Theta \subset \Lambda$. But (2.48) implies that the argument of the symbol increases greatly on I_n, which leads to a contradiction. More precisely, denote

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

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

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

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 (2.48) implies that $\delta_n \geq \frac{\varepsilon}{3}|I_n|$. Hence by a version of the theorem in [**103**], section 4.4, $N[\bar{\Theta}S^{2\pi a+2\varepsilon}]$ has to be trivial. □

LEMMA 3. *Let $I = [a, b]$ be an interval on $\mathbb{R}$ and let*

$$\Lambda = \{\lambda_1, ..., \lambda_N\}, \; a \leq \lambda_1 < ... < \lambda_N \leq b,$$

be a set of points on I. Let $C > 1$ be a constant and suppose that for some subinterval $J = [c, d] \subset I$,

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

Then one can spread the points of Λ on J without a large decrease in the energy $E(\Lambda)$ defined by (2.7). More precisely, if

$$\Gamma = \{\gamma_1, ..., \gamma_N\}, \; a \leq \gamma_1 < ... < \gamma_N \leq b,$$

is another set of points on I with the properties that

1) $\gamma_k = \lambda_k$ for all k such that $\lambda_k \notin J$ and
2) $c + C \leq \gamma_k \leq \gamma_{k+1} \leq d - C$ and $\gamma_{k+1} - \gamma_k \geq C$ for all $\gamma_k, \gamma_{k+1} \in J$,

then

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

for some absolute constant D.

PROOF. Notice that

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

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

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

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

COROLLARY 1. *Let Λ be a sequence of real points that satisfies the density* (2.9) *and energy* (2.10) *conditions for some short partition I_n and $a > 0$. 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*

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

for all k. Let Γ be a sequence of points obtained from Λ by spreading the points on each interval J_k like in the last lemma. Then Γ satisfies the density and energy conditions with the same partition I_n and a.

COROLLARY 2. *Let $\Lambda = \{\lambda_n\}$ be an increasing discrete sequence of real points that contains a d-uniform sequence for some $d > 0$. Then for any $\varepsilon > 0$ there exists an increasing discrete sequence $\Gamma = \{\gamma_n\}$ such that*

1) Γ contains a d-uniform sequence,
2) $D^*(\Gamma \setminus \Lambda) < \varepsilon$ and
3) $\sup_n(\gamma_{n+1} - \gamma_n) < \infty$.

PROOF. Choose $C > 0$ so that $1/C << d$ and $1/C << \varepsilon$. Let $[\lambda_{n_k}, \lambda_{n_k+1}]$ be a sequence of all 'gaps' of Λ satisfying $\lambda_{n_k+1} - \lambda_{n_k} > C$.

Without loss of generality, Λ itself is d-uniform. Then there exists a partition I_n such that Λ satisfies (2.9) and (2.10) for I_n and d. One can choose a sequence of disjoint intervals J_k such that for every k, $J_k \subset I_n$ for some n,

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

(the choice of J_k can be made by a version of the 'shading' algorithm, e.g. [**85**], volume 2, pp 507–508). Let Γ be a sequence of points obtained from Λ by spreading the points on each interval J_k like in lemma 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 Γ on $\cup J_k$ are at least C,

$$D^*(\Gamma \setminus \Lambda) \leq C^{-1} < \varepsilon.$$

□

LEMMA 4. *Let Λ be a discrete sequence of real points and let $\{I_n\}$ be a short partition such that Λ satisfies the density condition* (2.9) *with some $a > 0$ and the energy condition* (2.10) *on $\{I_n\}$. Then for any short partition $\{J_n\}$, 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 (2.10) *on $\{J_n\}$.*

Proof. To simplify the estimates we will assume that the endpoints of I_n belong to Λ, i.e. that $I_n = (\lambda_{k_n}, \lambda_{k_{n+1}}]$ for each n, and that the energy condition (2.10) is satisfied on I_n with E_n defined by (2.14), 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 9. This is where one may need to pass from Λ to a subsequence Γ. Since $|I_n| \to \infty$, Γ 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 and $\cup_l^{l+N} I_j = J_n$. Denote by u a continuous, piecewise linear function on J_n, that is zero at the left endpoint of J_n and grows linearly by 1 between each pair of points λ_n and λ_{n+1}. Denote

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

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

On each $(\lambda_i, \lambda_{i+1})$ the function u' satisfies the same estimates as $|\Theta'|$ from the statement of lemma 5. Therefore for the function p one can apply the same argument as in the first part of the proof of theorem 9, where p was defined as $\arg \Theta - x - v_1$ (we will simply assume that $v_1 \equiv 0$).

First, one can show that

$$- \iint_{J_n \times J_n} \log |t - x| p'(t) p'(x) dt dx$$
$$= |J_n|^2 \log |J_n| - \sum_{\lambda_{k_l} \le \lambda_i, \lambda_j \le \lambda_{k_{l+N+1}},\ \lambda_i \ne \lambda_j} \log |\lambda_i - \lambda_j| + \text{const}|J_n|^2.$$

To estimate the last integral rewrite it as

$$- \iint_{J_n \times J_n} \log |t - x| p'(t) p'(x) dt dx = \sum_{I_i \subset J_n} \sum_{I_j \subset J_n} - \iint_{I_i \times I_j} \log |t - x| p'(t) p'(x) dt dx.$$

For the last integral, when $i = j$ by (2.29) and (2.30) we have

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

As usual, we assume that $|I_n| >> 1$. If I_i does not intersect $2I_j$ then

$$\text{dist}(I_i, I_j) \ge |I_j| > 1$$

and the integral over $I_i \times I_j$ can be estimated by first noticing that the $\log_-$ part is zero, because $\log_- |x - t| = 0$ when $|x - t| > 1$, like in (2.31). For the $\log_+$ part we have (2.32). Altogether we obtain

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

For the case when I_i intersects $2I_j$ but is not contained in $2I_j$, or when I_i is adjacent to I_j, (note that there are at most four of such I_i for each I_j) we can estimate the integral like in (2.34) to conclude that

$$-\iint_{I_i\times I_j}\log|t-x|p'(t)p'(x)dtdx$$
$$\lesssim\left(|I_i|^2\log|I_i|-E_i\right)+\left(|I_j|^2\log|I_j|-E_j\right)+|I_i|^2+|I_j|^2.$$

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

$$\begin{aligned}
-&\iint_{I_j\times I_j}\log_+|s-t|p'(t)p'(s)dtds\\
&=\left(2\int_{I_j}u'(s)\int_{I_j}\log_+|s-t|dtds-\iint_{I_j\times I_j}\log_+|s-t|dtds\right)\\
&\quad-\int_{I_j}u'(s)\int_{I_j}\log_+|s-t|u'(t)dtds\\
&\geq|I_j|^2\log|I_j|-|I_j|\int_{I_j}\log_+|x-t|u'(t)dt+\text{const}|I_j|^2
\end{aligned}$$

because $\int_{I_j}u'=|I_j|$. Also, for any $x\in I_i$ (recall that $j>i+1$),

$$\int_{I_j}\log_+|x-t|dt-\int_{I_j}\log_+|x-t|u'(t)dt$$
$$\geq\left(|I_j|\log(\lambda_{k_{j+1}}-x)-|I_j|\right)-\log(\lambda_{k_{j+1}}-x)\int_{I_j}u'\geq-|I_j|.$$

Therefore

$$\begin{aligned}
-&\iint_{I_i\times I_j}\log_+|t-x|p'(t)p'(x)dtdx\\
&\leq\int_{I_i}|p'(x)|\left(\int_{I_j}\log_+|x-t|dt-\int_{I_j}\log_+|x-t|u'(t)dt+\text{const}|I_j|\right)dx\\
&\leq 2|I_i|\left(|I_j|\log|I_j|-\int_{I_j}\log_+|x-t|u'(t)dt\right)+\text{const}|I_i||I_j|\\
&\leq-2\frac{|I_i|}{|I_j|}\iint_{I_j\times I_j}\log_+|s-t|p'(t)p'(s)dtds+\text{const}|I_i||I_j|\\
&=2\frac{|I_i|}{|I_j|}||p_j||_{\mathcal{D}}+\text{const}|I_i||I_j|\\
&\lesssim\frac{|I_i|}{|I_j|}\left(|I_j|^2\log|I_j|-E_j\right)+|I_i||I_j|.
\end{aligned}$$

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

In the case when the intervals $I_l,...,I_{l+N}$ cover J_n but $\cup_l^{l+N}I_j\neq J_n$, i.e. when $I_l,I_{l+N}\cap J_n\neq\emptyset$ but at least one of I_l,I_{l+N} is not a subset of J_n, denote $I_l^*=I_l\cap J_n$ and $I_{l+N}^*=I_{l+N}\cap J_n$. Notice that by remark 1 and the fact that $\log|I_l^*|<\log|I_l|$,

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

Similarly,

$$|I^*_{l+N}|^2 \log |I^*_{l+N}| - E^*_{l+N} \le |I_{l+N}|^2 \log |I_{l+N}| - E_{l+N}.$$

Now we can use the previous case with I^*_l, I^*_{l+N} in place of I_l, I_{l+N}. □

COROLLARY 3. *Let Λ be a sequence of real points and let $\{I_n\}$ be a short partition such that Λ satisfies the density condition* (2.9) *with some $a > 0$ and the energy condition* (2.10)*. Then for any $\varepsilon > 0$ there exists a subsequence $\Gamma \subset \Lambda$ and a short monotone partition J_n such that Γ satisfies* (2.9)*, with $a - \varepsilon$ in place of a, and* (2.10) *on J_n.*

PROOF. One can choose a short monotone partition $\{J_n\}$ satisfying

$$\left(a - \frac{\varepsilon}{2}\right) |J_n| \le \#(\Lambda \cap J_n)$$

for all n. Such a partition can be constructed in the same way as in the second part of the proof of theorem 9, see the second paragraph before claim 3. Then Γ can be found by lemma 4 and remark 1. □

LEMMA 5. *Let $A = \{a_n\}_{n\in\mathbb{Z}}$ be a real sequence satisfying*

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

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

1) $\mathrm{spec}_\Theta = A$*;*
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) $\int_{J_n} (\arg\Theta)'(x)dx \ge c$ *for some $c > 0$ and all n.*

PROOF. Define a second sequence B as the sequence of midpoints of complementary intervals of A in $\mathbb{R}$: $b_n = (a_n + a_{n+1})/2$.

Define the inner function Θ to satisfy

$$\frac{1-\Theta}{1+\Theta} = \text{const } e^{Ku}, \tag{2.49}$$

where $u = 1_E - 1/2$,

$$E = \bigcup_{k=-\infty}^{\infty} (a_k, b_k),$$

and Ku is the *improper* integral

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

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

(Formulas similar to (2.49) are often used in perturbation theory. In those settings, u is the Krein-Lifshits shift function and Θ is the characteristic function of the perturbed operator, e.g. [**125, 136**].)

Let μ_1, μ_{-1} be the Clark measures for Θ defined by the Herglotz representation

$$\frac{1+\Theta}{1-\Theta} = \frac{1}{\pi i}\int_{\mathbb{R}}\left[\frac{1}{t-z} - \frac{t}{1+t^2}\right] d\mu_1(t) + \text{const},$$

$$\frac{1-\Theta}{1+\Theta} = \frac{1}{\pi i}\int_{\mathbb{R}}\left[\frac{1}{t-z} - \frac{t}{1+t^2}\right] d\mu_{-1}(t) + \text{const}.$$

The measures μ_1, μ_{-1} have the following form:

$$\mu_1 = \sum \alpha_n \delta_{a_n}, \qquad \mu_{-1} = \sum \beta_n \delta_{b_n}$$

for some positive numbers α_n, β_n. (It is easy to see that $\mu_{\pm 1}\{\infty\} = 0$ although we don't actually need this fact.)

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

$$\delta_n^2 \lesssim \beta_n \lesssim \delta_n. \tag{2.50}$$

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

$$|\Theta'| \asymp |1-\Theta|^2 \; |(\mathcal{S}\mu_1)'|, \qquad |\Theta'| \asymp |1+\Theta|^2 \; |(\mathcal{S}\mu_{-1})'|,$$

we have

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

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

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

and the estimate in part 2 follows from the left half of (2.50).

On the rest of the interval $|\Theta'(x)|$ can be estimated by $\sum \frac{\beta_n}{(x-b_n)^2}$ which together with the right half of (2.50) gives the desired estimate.

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

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

which implies $(\arg\Theta)' \gtrsim \delta_n^{-1}$ by the right half of (2.50). That implies the inequality for the integral. If, however, $|1-\Theta| \le |1+\Theta|$ at some point $c_n \in J_n$ then the integral taken between c_n and b_n is at least $\pi/4$. Since $(\arg\Theta)' > 0$, we again obtain the desired estimate.

It remains to prove (2.50). As follows from (2.49),

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

Denote

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

and

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

so

$$\text{Res}_{b_n} e^{-Ku} = A_n \,\text{Res}_{b_n} g_n, \qquad |\text{Res}_{b_n} g_n| = \frac{1}{2}\delta_n.$$

To prove the right half of (2.50) 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 that is closer to b_n and negative on the half that is further from it. Thus

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

Similarly

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

To prove the left half of (2.50) one needs to show that $\delta_n \lesssim A_n$. Notice that, since $\delta_n < C$,

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

As for the remaining part,

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

□

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

LEMMA 6. *Let $a_1 < a_2$ and $b_1 < b_2$ be points on the real line. Let α and β be non-negative functions on the intervals (a_1, a_2) and (b_1, b_2) correspondingly satisfying*

$$\int_{a_1}^{a_2} \alpha = \int_{b_1}^{b_2} \beta = 1,\ \alpha < A \quad \textit{and} \quad \beta < B$$

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

1) $\log_-(a_2-a_1) \leq \int_{a_1}^{a_2}\int_{a_1}^{a_2} \log_-(x-y)\alpha(x)\alpha(y)dxdy \leq \log_- \frac{1}{A} + 1.$

2) If $a_2 < b_1$ then

$$\log_-(b_2-a_1) \leq \int_{a_1}^{a_2}\int_{b_1}^{b_2} \log_-(x-y)\alpha(x)\beta(y)dxdy \leq \log_-(b_1-a_2).$$

3) If $a_2 = b_1$ then

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

4) If $a_2 \leq b_1$ then

$$\log_+(b_1-a_2) \leq \int_{a_1}^{a_2}\int_{b_1}^{b_2} \log_+(x-y)\alpha(x)\beta(y)dxdy \leq \log_+(b_2-a_1).$$

5) If $A/2 \leq \alpha(x) \leq A$ on (a_1, a_2) then for any $y \in (a_1, a_2)$

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

for some absolute constant C.

6) If $A/2 \le \alpha(x) \le A$ on (a_1, a_2) then for any $y > a_2$

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

for some absolute constant C.

LEMMA 7. *Let $E \subset \mathbb{R}$ be a set and let $I \subset \mathbb{R}$ be an interval such that*

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

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

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

for some constant $0 < \varepsilon < 1/3$ and

$$\int_I f(x)dx = \int_I g(x)dx. \tag{2.51}$$

Denote

$$\int_E |h(x)|dx = D_1, \qquad \int_E \left| \int_I \log|t - x| h(t)dt \right| dx = D_2.$$

Then

$$\begin{aligned} &\frac{1}{d^2}\left[\iint_{E\times I} \log|t - x|h(x)h(t)dxdt - DD_1|I| - 2D_1 - 4D_2\right] \\ &\quad \le \int_E h(x)\left[\int_I \log|t - x|h(t)dt\right]\frac{dx}{x^2} \\ &\quad \le 4\frac{1}{d^2}\left[\iint_{E\times I} \log|t - x|h(x)h(t)dxdt + DD_1|I| + 2D_1 + 4D_2\right] \end{aligned} \tag{2.52}$$

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 functions under the integral were negative. To prove (2.52) in the general case, notice that

$$\int_I \log|t - x|h(t)dt \le D|I| + 2 \tag{2.53}$$

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

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

because $h \ge -2$. 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 $x > b$, by lemma 6 part 6 we obtain

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

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

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

Together the last three relations give

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

because $g < 2$ and $\int_I f = \int_I g$. This establishes (2.53). Similar estimates can be applied for $x < a$. For $x \in I$ the same relation can be obtained using lemma 6 part 5 instead of part 6.

To finish the proof denote

$$E^+ = \{h > 0\} \cap E, \quad E^- = \{h \le 0\} \cap E.$$

Notice that

$$\begin{aligned} &\int_E h \left[\int_I \log |t-x| h(t) dt \right] \frac{dx}{x^2} \\ &= \int_{E^+} h \left[\int_I \log |t-x| h(t) dt \right] \frac{dx}{x^2} + \int_{E^-} h \left[\int_I \log |t-x| h(t) dt \right] \frac{dx}{x^2} \\ &= \int_{E^+} h(x) \max \left(\left[\int_I \log |t-x| h(t) dt \right], 0 \right) \frac{dx}{x^2} \\ &+ \int_{E^+} h(x) \min \left(\left[\int_I \log |t-x| h(t) dt \right], 0 \right) \frac{dx}{x^2} \\ &+ \int_{E^-} h \left[\int_I \log |t-x| h(t) dt \right] \frac{dx}{x^2} = I + II + III. \end{aligned}$$

If one replaces $\frac{dx}{x^2}$ with $\frac{dx}{d^2}$ in II, the integral will not increase because the function under the integral is negative and $x^2 \ge d^2$. It also will not decrease much because $x^2 \le 4d^2$. Under the same operation the positive integral I will become at most

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

in view of (2.53). Hence it will not change by more than the last quantity. Finally, III satisfies

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

and after replacing $\frac{dx}{x^2}$ with $\frac{dx}{d^2}$ we will still have

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

Therefore III will change at most by

$$2 \frac{1}{d^2} \int_{E^-} \left| h \int_I \log |t-x| w'(t) dt \right| dx \le \frac{4D_2}{d^2}$$

because $0 > h > -2$ on E^-. □

8. Appendix: De Branges' theorem 66 in Toeplitz form

One of the classical results on the Gap Problem is de Branges' theorem from [**26**], page 271, often referred to as 'theorem 66' by the experts. There the answer is given not in terms of the set X but in terms of existence of a certain entire function. In this section we discuss two versions of that theorem, see lemma 9 and corollary 4. De Branges' idea to apply the Krein-Milman theorem and look for an extreme point in the proper set of measures was used in several of his papers [**27, 28, 29**]. The extreme point approach proves to be an important tool in this area of UP. Further variations of theorem 66, along with a detailed discussion of applications, can be found in [**21, 138, 140**].

To formulate theorem 66 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

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

THEOREM 10 (Theorem 66, [**26**]). *Let $a > 0$ be a given number and let X be a closed subset of the real line. A necessary and sufficient condition that $\mathbf{G}_X \geq 2a$ 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*

$$\sum_{E(t)=0} \frac{1}{|E'(t)|} < \infty. \tag{2.54}$$

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 [**140**] for some of such applications and further references.

Before stating and proving an extension of theorem 10 we need the following definitions and a lemma proved by Aleksandrov in [**1**].

We say that a finite measure μ on $\mathbb{R}$ annihilates K_Θ if $\int f d\mu = 0$ 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_Θ.

We say that the Cauchy integral $K\mu$ is divisible by an inner function Θ if $K\mu/\Theta = K\eta$ in $\mathbb{C}_+$ for some finite complex measure η on $\mathbb{R}$. Equivalently, $K\mu$ is divisible by Θ if $K\mu/\Theta \in H^p(\mathbb{C}_+)$ for some $p > 0$, see [**145**].

LEMMA 8. [**1**] *Let μ be a finite complex measure on $\mathbb{R}$ and let Θ be an inner function in $\mathbb{C}_+$. Then the following statements are equivalent:*

(i) μ annihilates K_Θ;

(ii) The Cauchy integral of the conjugate measure $\bar\mu$, $K\bar\mu$, is divisible by Θ.

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

$$\overline{\int \frac{1-\bar\Theta(\lambda)\Theta(z)}{\bar\lambda - z} d\mu(z)} = \Theta(\lambda)K\bar\Theta\bar\mu(\lambda) - K\bar\mu(\lambda) \tag{2.55}$$

which implies the statement because the initial integral is zero.

(ii)⇒(i). If η is the measure such that $K\eta = K\bar\mu/\Theta$ in $\mathbb{C}_+$, then η can be chosen as $\bar\Theta\bar\mu$, see for instance [**122**, theorem 3.4]. We can assume that the boundary values of Θ exist μ-a.e. (otherwise Θ can be replaced by a divisor). Then the right hand side of (2.55) 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_Θ we obtain the statement. □

Note that the condition (2.54) implies that $1/E$ is a Cauchy integral of a finite measure μ concentrated on the zero set of E. The pointmass of μ at a zero t of E is equal to $1/|E'(t)|$.

Thus the existence of E like in the statement of theorem 66 is equivalent to the existence of a finite discrete real measure μ 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 μ with all the above properties. Our next lemma shows that a similar statement can be formulated for any inner Θ in place of S^a. Like in de Branges' proof, we use the Krein-Milman theorem on the existence of extreme points in a weak-star closed convex set.

Lemma 9. *Let Θ be an inner function in $\mathbb{C}_+$. Let μ be a finite complex measure whose Cauchy integral $K\mu$ is divisible by Θ (or, equivalently, $\bar{\mu}$ annihilates K_Θ). Then there exists a finite singular complex measure ν such that*

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

2) $K\nu$ is divisible by Θ ($\bar{\nu}$ annihilates K_Θ);

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

4) if Θ is a meromorphic inner function, then ν is concentrated on a discrete set.

Proof. First, let us symmetrize μ. Since together with any $f \in K_\Theta, \Theta\bar{f} \in K_\Theta$, the measure $\bar{\Theta}\mu$, just like $\bar{\mu}$, annihilates K_Θ and $K\Theta\bar{\mu}$ is divisible by Θ. Consider $\eta = \mu + \Theta\bar{\mu}$. Without loss of generality $||\eta|| \le 1$.

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

$$\Theta\bar{\sigma} = \sigma. \tag{2.56}$$

Since $\eta \in A^\Theta_\Sigma$, this set is not empty. It is also weak-star closed and convex. By the Krein-Milman theorem it contains a non-zero extremal point ν. 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 Θ is one-dimensional, and therefore $h = c \in \mathbb{R}$. (This is equivalent to the statement that the closure of K_Θ in $L^1(|\nu|)$ has deficiency 1, i.e. the space of its annihilators is one dimensional)

Let there be a bounded real h such that $Kh\nu$ is divisible by Θ. Without loss of generality $h \ge 0$, since one can add constants, and $||h\nu|| = 1$. Choose $0 < \alpha < 1$ so that $|\alpha h| < 1$. Consider probability measures $\nu_1 = h\nu$ and $\nu_2 = (1-\alpha)^{-1}(\nu - \alpha\nu_1)$. Then both of them belong to A^Θ_Σ and $\nu = \alpha\nu_1 + (1-\alpha)\nu_2$ which contradicts the extremality of ν.

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

$$0 = \int \frac{f_n(z) - f_n(w)}{z-w} d\bar\nu(z) = K f_n \bar\nu(w) - f_n(w) K\bar\nu(w)$$

and therefore

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

Taking the limit,

$$f = \lim f_n = \lim \frac{K f_n \bar\nu}{K\bar\nu} = \frac{K g\bar\nu}{K\bar\nu},$$

where convergence is uniform on compact subsets of $\mathbb{C} \setminus \mathbb{R}$. Since all of 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 $f_n^\pm, f^\pm$ the radial limits of f_n, f on $\mathbb{R}$ taken from $\mathbb{C}_\pm$ correspondingly. Then, since 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 $\mathbb{R}$.

Since the numerator in the representation

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

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

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

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

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

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

by a result from [**122**]. Since $K\nu$ is divisible by Θ, G is divisible by Θ. Let us first show that G/Θ does not have an inner component in the upper half-plane. Suppose that $G = \Theta U H$ for some inner U. Since the measure ν satisfies (2.56), $\bar G = G/\Theta$, $|\nu|$-a.e.

If $F \in K_J$ is the function such that $\bar J G = \bar F$, since $J = 1$, $|\nu|$-a.e., $F = \bar G = G/\Theta$, $|\nu|$-a.e. Since functions in K_J are uniquely determined by their boundary values in $L^2(|\nu|)$ (recall that $|\nu|$ is the Clark measure for J), $F = G/\Theta = UH$. Notice that the function $h = \Theta(1+U)^2 H$ also belongs to K_J:

$$\bar{J}h = \bar{J}\Theta(1+U)^2H = (\bar{J}G)\bar{U}(1+U)^2$$

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

because $\bar{U}(1+U)^2$ is real a.e. on $\mathbb{R}$. Denote by γ 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

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

The Cauchy integral of γ is divisible by Θ because h is divisible by Θ. Since $\bar{U}(1+U)^2$ is real, a constant multiple of γ belongs to A^{Θ}_{Σ}. Since U is non-constant and $|\nu|$ is the Clark measure for J, γ is not a constant multiple of ν. We obtain a contradiction with the property that the space of annihilators is one-dimensional.

Thus $G/\Theta \in K_J$ is outer in C_+. Since $J\bar{G} = F = G/\Theta$, the pseudocontinuation of G does not have an inner factor in $\mathbb{C}_-$ as well. 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 of 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 one-dimensional, since $(x-a)^{-1}$ is bounded and real on the support of ν. 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 66. 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 4. *Let I, Θ 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 $\text{spec}_J \subset \text{spec}_I$ *and the kernel $N[\bar{J}\Theta]$ contains a purely outer function f that does not have any zeros on* $\mathbb{R} \setminus \text{spec}_J$.

If σ_1 is the Clark measure of J then f is also non-zero σ_1-a.e. on spec_J.

If Θ is a meromorphic inner function, then J can be chosen as a meromorphic inner function.

Proof. We will consider the case $\text{spec}_I \subset \mathbb{R}$. 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 σ is the Clark measure corresponding to I. Since $1 - I$ is outer, the Cauchy integral $Kg\sigma$ is divisible by Θ. By lemma 9, there exists a finite singular measure ν,

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

satisfying the 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 Θ. Therefore $l = h/\Theta$ belongs to $N[\bar{J}\Theta]$. Notice that l is purely outer by part 3 of lemma 9. Since $2\pi i\nu = h|\nu|$, l has no zeros a.e. with respect to $|\nu|$, the Clark measure of J.

If Θ is a meromorphic function then, by part 4 of lemma 9, ν can be chosen to be discrete. Then J is meromorphic. □

In chapters 4, 5 and 6 we will return to the de Branges extreme point approach and formulate further versions of theorem 66 [**26**].