

SZEGÖ LIMIT THEOREMS FOR THE HARMONIC OSCILLATOR

BY

A. J. E. M. JANSSEN AND STEVEN ZELDITCH

ABSTRACT. Let $H = -\frac{1}{2}d^2/dx^2 + \frac{1}{2}x^2$ be the harmonic oscillator Hamiltonian on $L^2(\mathbf{R})$, and let A be a selfadjoint DO of order O in the Beals-Fefferman class with weights $\varphi = 1$, $\Phi(x, \xi) = (1 + |\xi|^2 + |x|^2)^{1/2}$. Form the measure $\mu(f) = \lim_{\lambda \rightarrow \infty} (1/\text{rank } \pi_\lambda) \text{tr } f(\pi_\lambda A \pi_\lambda)$ where $\pi_\lambda A \pi_\lambda$ is the compression of A onto the span of the Hermite functions with eigenvalue less than or equal to λ . Then one has the following Szegő limit theorem:

$$\mu(f) = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H(x, \xi) \leq T} f(a(x, \xi)) dx d\xi \quad \text{for } f \in C(\mathbf{R}).$$

For the special case where $f(x) = x$, this will be proved for a considerably wider class of operators by employing the Weyl correspondence. Moreover, by using estimates on Wigner functions of Hermite functions we are able to prove the full Szegő theorem for a fairly general class of multiplication operators.

0. Introduction. Pseudo-differential operator theory provides a natural framework in which to discuss correspondence principles between classical and quantum mechanics. Indeed, one of the origins of the theory was the “Weyl correspondence” [We] between functions $a(x, \xi)$ on phase space $T^*\mathbf{R}$ and operators $a^w(x, D)$ on $L^2(\mathbf{R})$ by means of the formula

$$(0.1) \quad a^w(x, D)u(x) = \frac{1}{2\pi} \int a\left(\frac{1}{2}(x+y), \xi\right) u(y) e^{i\langle x-y, \xi \rangle} dy d\xi.$$

Moreover, the requirement that operator brackets $[a^w, b^w]$ should correspond (as much as possible) to the poisson bracket of the symbols $\{a, b\}$ leads to the selection of classes of symbols to which orders may be assigned, so that $[a^w, b^w] = \{a, b\}^w +$ lower order. What “lower order” should be depends upon the application; frequently in applications to quantum mechanics it means a higher order dependence on Planck’s constant h [V].

In this paper, we will use the pseudo-differential and Weyl correspondence to relate various averages of operators to averages of the symbols. Whether using conventional ψDO theorem (as in Part I) or other kinds of analysis (as in Part II), a basic principle is that lower order terms should vanish in the averaging.

Let us now state the basic results. First, fix $H = -\frac{1}{2}d^2/dx^2 + \frac{1}{2}x^2$, the harmonic oscillator Hamiltonian on $L^2(\mathbf{R})$ and let π_λ be orthogonal projection onto the span of

Received by the editors December 2, 1981 and, in revised form, September 16, 1982.
 1980 *Mathematics Subject Classification*. Primary 35S99; Secondary 46N05, 81C99.

©1983 American Mathematical Society
 0002-9947/83 \$1.00 + \$.25 per page

the Hermite functions with eigenvalue less than λ . Correspondingly, let $H(x, \xi) = \frac{1}{2}\xi^2 + \frac{1}{2}x^2$ be the classical Hamiltonian and let B_λ be the energy disc $\{(x, \xi) | H(x, \xi) \leq \lambda\}$ of energy less than or equal to λ . Then for certain corresponding pairs (a, A) we will show:

(i) AVERAGING THEOREM.

$$\frac{1}{2\pi\lambda} \int_{B_\lambda} a(x, \xi) dx d\xi \sim \frac{1}{\text{rank } \pi_\lambda} \text{tr}(\pi_\lambda A \pi_\lambda)$$

as $\lambda \rightarrow \infty$. Intuitively the compression of A to $\{\varphi \in L^2(\mathbf{R}) | \langle H_\varphi, \varphi \rangle / \langle \varphi, \varphi \rangle \leq \lambda\}$ has average value asymptotic to the average value of the compression of its symbol to $\{(x, \xi) \in T^*\mathbf{R} | H(x, \xi) \leq \lambda\}$.

(ii) SZEGÖ THEOREM.

$$\frac{1}{2\pi\lambda} \int_{B_\lambda} f(a(x, \xi)) dx d\xi \sim \frac{1}{\text{rank } \pi_\lambda} \text{tr} f(\pi_\lambda A \pi_\lambda)$$

as $\lambda \rightarrow \infty$, and for $f \in C(\mathbf{R})$. This defines a limiting measure μ of the sequence $\mu_\lambda / \text{rank } \pi_\lambda$, with $\mu_\lambda(f) = \text{tr} f(\pi_\lambda A \pi_\lambda)$. Intuitively, letting f be approximately the characteristic function of an interval, this says that the distribution of values of the compressed a is asymptotic to the distribution of values of the compressed A .

In Part I of this paper the “certain corresponding pairs” will consist of symbols a of order O in a special case of the Beals-Fefferman calculus (see [Be, Hô]). We could also use the more special, and for applications less realistic, calculi of pseudo-Toeplitz operators of Grossman, Loupias and Stein (see [GLS, H]). The great utility of these calculi is that (as we mentioned above) symbols of products are products of symbols plus terms which vanish in the averaging, and, in general, lower order terms will never show up in the limits. We should also add that some extra conditions on the symbols are needed to insure the existence of the limits.

In Part II, we will prove the averaging theorem for operators A corresponding, under the Weyl correspondence, to symbols $a \in L(T^*\mathbf{R})$ which have mean values on $T^*\mathbf{R}$. Here the symbols do not belong to symbol calculi, and so one must work with explicit formulae for symbols and kernels, rather than with ψDO approximations. In the ensuing analysis, however, one gets more precise error estimates than is possible in the ψDO theory. This is related to the fact that Weyl’s correspondence can be described in terms of Wigner functions (symbols of projection operators) and the fact that we are able to get good estimates on the Wigner functions of Hermite functions. In this part we also prove the Szegö theorem for potentials $V(x)$, although some indications for the general case are given at the end of §4. The assumptions on V are not of “symbol” type, so that we can free ourselves of the ψDO approach. (One advantage of working with potentials is that symbols of powers are the powers of symbols, so no lower terms enter in. In general, all symbol manipulations are simplified in using potentials.)

The similarity of these results to the work of Widom should be clear to anyone familiar with it (see [Wi, Gu]). Some of the proofs are borrowed with little change from these papers, as will be noted below. We would also like to thank Alan Weinstein for conversations and encouragement.

PART I. THE ψDO APPROACH

In this part we will prove the averaging and Szegö theorems for symbols $a(x, \xi)$ in a special Beals-Fefferman class S_Φ^0 .

0. Preliminaries.

DEFINITION 0.1. Let $S_\Phi^m = \{a(x, \xi) \mid a \in C^\infty(\mathbf{R}_x \times \mathbf{R}_\xi), \mid \partial_x^\alpha \partial_\xi^\beta a \mid \leq C_{\alpha\beta} \Phi^{m-|\beta|}\}$ where $\Phi(x, \xi) = (1 + |x|^2 + |\xi|^2)^{1/2}$.

In the language of [Be], S_Φ^m is the space of symbols of order m relative to the weights $\varphi = 1$ and Φ above.

DEFINITION 0.2. Let $\mathcal{L}_\Phi^m = \{A = a(x, D) \mid a \in S_\Phi^m\}$. Here we will use the ‘‘Weyl correspondence’’ definition

$$a(x, D)u(x) = \frac{1}{2\pi} \int a\left(\frac{1}{2}(x+y), \xi\right) u(y) e^{i\xi \cdot (x-y)} dy d\xi,$$

although our conclusions are valid for the usual ‘‘derivatives first’’ definition as well.

We also write $a(x, D) = Op_{1/2}(a(x, \xi))$ and $a(x, \xi) = \sigma_A^w(x, \xi)$.

The class of ψDO ’s in this calculus is closed under composition, and one has the following

(0.3) *Composition rule.* If $A = a(x, D) \in \mathcal{L}_\Phi^m$, and $B = b(x, D) \in \mathcal{L}_\Phi^n$ and $a \circ b = \sigma_{A \cdot B}^w$ then $A \circ B \in \mathcal{L}_\Phi^{m+n}$ and

$$\begin{aligned} a \circ b &= a \exp\left(\frac{i}{2} \left(\frac{\bar{\partial}}{\partial x} \frac{\bar{\partial}}{\partial \xi} - \frac{\bar{\partial}}{\partial x} \frac{\bar{\partial}}{\partial \xi}\right)_b\right) \quad (\text{see [V, p.114]}) \\ &\sim \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{i}{2}\right)^j a \left(\frac{\bar{\partial}}{\partial x} \frac{\bar{\partial}}{\partial \xi} - \frac{\bar{\partial}}{\partial x} \frac{\bar{\partial}}{\partial \xi}\right)^j b \end{aligned}$$

where asymptotics are taken relative to the symbol grading, i.e.

$$a \circ b - \sum_{j=1}^k \frac{1}{j!} \left(\frac{i}{2}\right)^j a \left(\frac{\bar{\partial}}{\partial x} \frac{\bar{\partial}}{\partial \xi} - \frac{\bar{\partial}}{\partial x} \frac{\bar{\partial}}{\partial \xi}\right)^j b \in S_\Phi^{m+n-k}.$$

PROOF. Essentially the same as [Be, Theorem 4.1].

We will need the following well-known

(0.4) *Trace rule.* If A is a trace class ψDO with smooth kernel $k_A(x, y)$, then $\text{tr } A = \int \sigma_A^w(x, \xi) dx d\xi$.

PROOF.

$$k_A(x, y) = \frac{1}{2\pi} \int a\left(\frac{1}{2}(x+y), \xi\right) e^{i\xi \cdot (x-y)} d\xi$$

so

$$\text{tr } A = \int k_A(x, x) dx = \frac{1}{2\pi} \int a(x, \xi) dx d\xi \quad \text{where } a = \sigma_A^w.$$

Finally, we recall these basic facts about the Hermite operator (cf. [Sz, U]):

$$\begin{aligned}
 & \text{(i)} \quad H \in \mathcal{L}_\Phi^2, \\
 & \text{(ii)} \quad \text{spec } H = \{n + \frac{1}{2}\}, \\
 & \text{(iii)} \quad e^{-tH} \text{ is a trace class } \psi DO \text{ with Weyl symbol} \\
 (0.5) \quad & \sigma_{e^{-tH}}^w(x, \xi) = \sigma(t, x, \xi) = \frac{1}{\cosh t/2} e^{-\tanh(t/2)(x^2 + \xi^2)}, \\
 & \text{(iv)} \quad \text{tr } e^{-tH} = \frac{1}{2 \sinh t/2}.
 \end{aligned}$$

1. Computing traces. One advantage in using the Weyl calculus is that we get a very simple formula for the trace $\text{tr } Ae^{-tH}$. Indeed, for $A \in \mathcal{L}_\Phi^0$,

$$\text{tr } Ae^{-Ht} = \frac{1}{2\pi} \int \sigma_{Ae^{-Ht}}^w(x, \xi) dx d\xi = \frac{1}{2\pi} \int a(x, \xi) \circ \sigma(t, x, \xi) dx d\xi.$$

In the Weyl calculus, we have for symbols a and b of arbitrary order, and with $\omega(z, \zeta; t, \tau) = \zeta t - z\tau$,

$$a \circ b(x, \xi) = \int a(x + z, \xi + \zeta) b(x + t, \xi + \tau) e^{2i\omega(z, \zeta; t, \tau)} dz d\zeta dt d\tau$$

(see [Hö, p. 374]). Writing $x + z = \bar{z}$, $\xi + \zeta = \bar{\zeta}$, $x + t = \bar{t}$, $\xi + \tau = \bar{\tau}$, one gets

$$\begin{aligned}
 & \int a \circ b(x, \xi) dx d\xi \\
 & = \int a(\bar{z}, \bar{\zeta}) b(\bar{t}, \bar{\tau}) \exp(2i\{x(\bar{\zeta} - \bar{\tau}) + \xi(\bar{t} - \bar{z}) \\
 & \qquad \qquad \qquad + \omega(\bar{z}, \bar{\zeta}; \bar{t}, \bar{\tau})\}) dx d\xi d\bar{z} d\bar{\zeta} d\bar{t} d\bar{\tau} \\
 & = \int a(t, \tau) b(t, \tau) d\tau dt.
 \end{aligned}$$

Therefore

$$\begin{aligned}
 (1.3) \quad \text{tr } Ae^{-tH} & = \int a(x, \xi) \sigma(t, x, \xi) dx d\xi \\
 & = \frac{1}{2\pi \cosh t/2} \int a(x, \xi) \exp(-\tanh t/2(x^2 + \xi^2)) dx d\xi.
 \end{aligned}$$

Changing to polar variables $z = x + i\xi = re^{i\theta}$ we get

$$(1.4) \quad \frac{1}{2\pi \cosh t/2} \int_0^{2\pi} \int_0^\infty a(re^{i\theta}) \exp(-\tanh t/2r^2) r dr d\theta,$$

whence letting $\rho = r^2$ we get

$$(1.5) \quad \frac{1}{2 \cosh(t/2)} \int_0^\infty e^{-\tanh t/2\rho} a_{\text{ave}}(\sqrt{\rho}) d\rho,$$

where $a_{\text{ave}}(\rho) = (1/2\pi) \int_0^{2\pi} a(\rho e^{i\theta}) d\theta$.

Finally,

$$\operatorname{tr} e^{-tH} = \frac{1}{2\pi \cosh(t/2)} \int \exp(-\tanh(t/2)(x^2 + \xi^2)) dx d\xi = \frac{1}{2 \sinh(t/2)},$$

so that

$$(1.6) \quad \frac{1}{\operatorname{tr} e^{-tH}} \operatorname{tr} A e^{-tH} = \frac{\sinh(t/2)}{\cosh(t/2)} \int_0^\infty e^{-\tanh(t/2)r} a_{\text{ave}}(\sqrt{r}) dr$$

or letting $w = e^{-2t}$

$$(1.7) \quad = \frac{1-w}{1+w} \int_0^\infty e^{-r(1-w)/(1+w)} a_{\text{ave}}(\sqrt{r}) dr.$$

Summing up, we have, for $A \in \mathcal{L}_\Phi^0$,

PROPOSITION 1.8. *If one of the following limits exists, then so does the other and they assume the same value:*

- (i) $\lim_{t \downarrow 0} (1/\operatorname{tr} e^{-tH}) \operatorname{tr} (A e^{-tH})$.
- (ii) $\lim_{\varepsilon \rightarrow 0} \varepsilon \int_0^\infty e^{-\varepsilon r} a_{\text{ave}}(\sqrt{r}) dr$.

2. Abelian-Tauberian theorem and the averaging theorem. The limits (i) and (ii) may be replaced by limits of other averaging procedures via Karamata's Tauberian theorem. Indeed, we have

PROPOSITION 2.1.

$$\lim_{t \downarrow 0} \frac{1}{\operatorname{tr} e^{-tH}} \operatorname{tr} A e^{-tH} = \lim_{n \rightarrow \infty} \frac{1}{n} \operatorname{tr} (\pi_n A \pi_n)$$

where $\pi_n =$ projection onto the span of the first n Hermite functions, $\{\psi_j\}$ and where equality means that if one side exists, then so does the other and assumes the same value.

PROOF. Let $\varphi(\lambda) = \operatorname{tr} \pi_\lambda A \pi_\lambda = \sum_{j+1/2 \leq \lambda} \langle A \psi_j, \psi_j \rangle$. Then

$$\frac{d\varphi}{d\lambda} = \sum_{j=1}^\infty \langle A \psi_j, \psi_j \rangle \delta\left(\lambda - \left(j + \frac{1}{2}\right)\right)$$

so the Laplace transform

$$\mathcal{L}\left(\frac{d\varphi}{d\lambda}\right)(t) = \sum_{j=1}^\infty \langle A \psi_j, \psi_j \rangle e^{-(j+1/2)t} = \operatorname{tr} A e^{-Ht}.$$

But Karamata's theorem says $\lim_{t \downarrow 0} t \mathcal{L}(d\varphi/d\lambda)(t) = \lim_{n \rightarrow \infty} \varphi[0, n]/n$ which is equivalent to the stated proposition, since $\operatorname{tr} e^{-tH} \sim 1/t$.

PROPOSITION 2.2.

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \int_0^\infty e^{-\varepsilon r} a_{\text{ave}}(\sqrt{r}) dr = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T a_{\text{ave}}(\sqrt{r}) dt$$

where equality again means one limit exists if and only if the other does.

PROOF. This is again an application of Karamata's Tauberian theorem (see e.g. [Si]).

Summing up §§1 and 2 we may state the averaging theorem for $A \in \mathcal{L}_\Phi^0$:

THEOREM 1. *If one of the following limits exists, then so does the other and assumes the same value:*

- (i) $\lim_{n \rightarrow \infty} (1/n) \text{tr}(\pi_n A \pi_n)$.
- (ii) $\lim_{T \rightarrow \infty} (1/2\pi T) \int_{H(x, \xi) \leq T} a(x, \xi) dx d\xi$.

PROOF.

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T a_{\text{ave}}(\sqrt{r}) dr &= \lim_{T \rightarrow \infty} \frac{1}{\pi T^2} \int_0^{2\pi} \int_0^T a(re^{i\theta}) r dr d\theta \\ &= \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H(x, \xi) \leq T} a(x, \xi) dx d\xi \end{aligned}$$

and the rest follows from Propositions 1.8, 2.1 and 2.2.

REMARK 1. One may paraphrase the averaging theorem as follows: The average value of A , compressed to that part of quantum phase space where the energy is less than n , is asymptotic to the average value of a compressed to that part of phase space where the energy is less than n . Here we interpret $2\pi T = \text{area}(H \leq T)$ and $n = \text{rank } \pi_n$, and compare volumes to dimensions as usual. In this form, the theorem is valid in great generality, as we will show in a later paper.

REMARK 2. A simple sufficient condition for the existence of these limits is that a_{ave} has a limit as $r \rightarrow \infty$, or that a has radial limits. These are strong assumptions, but if they hold for a , then they do as well for $f(a)$ ($f \in C(\mathbf{R})$), so that the Szegő theorem of the next section is not vacuous.

3. The Szegő theorem. In this section, we prove the following Szegő type theorem.

THEOREM 2. *Let $\mu_\lambda(f) = \text{tr } f(\pi_\lambda A \pi_\lambda)$, $f \in C(\mathbf{R})$. Then the sequence of measures $(1/\lambda)\mu_\lambda$ has a weak limit μ given by*

$$\mu(f) = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H \leq T} f(a(x, \xi)) dx d\xi.$$

The main step in the proof, following Widom, is to replace $(1/\lambda)\text{tr } f(\pi_\lambda A \pi_\lambda)$ by $(1/\lambda)\text{tr } \pi_\lambda f(A) \pi_\lambda$ for polynomial f , and to apply the averaging theorem to the latter sequence. We may say this in another way. By the spectral theorem, one has for each vector $\psi \in L^2(\mathbf{R})$ the spectral measure $\mu_\psi(f) = \langle \psi, f(A)\psi \rangle$. In particular, one has the measures μ_{ψ_j} , ψ_j being the j th Hermite function, and one may form the spectral measures

$$\nu_\lambda = \sum_{j+1/2 \leq \lambda} \mu_{\psi_j}.$$

One then needs to show $\lim_{\lambda \rightarrow \infty} (1/\lambda)\mu_\lambda = \lim_{\lambda \rightarrow \infty} (1/\lambda)\nu_\lambda$ and that

$$\lim_{\lambda \rightarrow \infty} (1/\lambda)\nu_\lambda(f) = \lim_{T \rightarrow \infty} (1/2\pi T) \int_{H \leq T} f(a(x, \xi)) dx d\xi.$$

To exploit Widom's proof, we will have to pass from H to \sqrt{H} . The eigenvalues of \sqrt{H} are $\lambda_n = \sqrt{n + 1/2}$; let P_n be the projection onto the eigenspace of eigenvalue $\sqrt{n + 1/2}$ and let $P_\lambda = \sum_{\sqrt{n+1/2} \leq \lambda} P_n (= \pi_{\lambda^2})$.

PROPOSITION 3.1. $\sqrt{H} = OP_{1/2}(\sqrt{\xi^2/2 + x^2/2}) + B$ where B is bounded on L^2 .

PROOF. This is a familiar fact in ψDO theory, so we only briefly recall the proof (see [T, p. 289]).

Let

$$H_{1/2} = OP_{1/2}(\sqrt{(\xi^2 + x^2)/2}).$$

Then

$$\begin{aligned} \sigma_{H_{1/2}, H_{1/2}} &= \frac{1}{2} \sqrt{\xi^2 + x^2} \circ \sqrt{\xi^2 + x^2} \\ &= \frac{1}{2} \sqrt{\xi^2 + x^2} \exp\left(\frac{i}{2} \left(\frac{\bar{\partial}}{\partial x} \frac{\bar{\partial}}{\partial \xi} - \frac{\bar{\partial}}{\partial x} \frac{\bar{\partial}}{\partial \xi}\right)\right) \sqrt{\xi^2 + x^2} \\ &= \frac{1}{2} (\xi^2 + x^2) + r(x, \xi) \quad \text{where } r \in S_\Phi^{-1}. \end{aligned}$$

Thus $(H_{1/2})^2 = H + R$, where $R \in \mathcal{L}_\Phi^{-1}$ hence bounded (see [T or H]). Now let Γ be a curve in the complex plane enclosing both $\text{spec } H$ and $\text{spec}(H + R)$. Then

$$\begin{aligned} H_{1/2} &= \frac{1}{2\pi i} \int_\Gamma \lambda^{1/2} ((H_{1/2})^2 - \lambda)^{-1} d\lambda \\ &= \frac{1}{2\pi i} \int_\Gamma \lambda^{1/2} (H + R - \lambda)^{-1} d\lambda \\ &= \frac{1}{2\pi i} \int_\Gamma \lambda^{1/2} (H - \lambda)^{-1} [I + (H - \lambda)^{-1} R]^{-1} d\lambda. \end{aligned}$$

Using $(1 + A)^{-1} = 1 - (1 + A)^{-1}A$ we get

$$\begin{aligned} &= \frac{1}{2\pi i} \int_\Gamma \lambda^{1/2} (H - \lambda)^{-1} d\lambda \\ &\quad - \frac{1}{2\pi i} \int_\Gamma \lambda^{1/2} (H - \lambda)^{-1} (I + (H - \lambda)^{-1} R)^{-1} (H - \lambda)^{-1} R d\lambda. \end{aligned}$$

It is easy to see the second term is bounded, so we have $H_{1/2} = \sqrt{H} + B$, as desired.

We now derive the basic lemma required in replacing μ_λ by ν_λ .

LEMMA 3.2. If $A: L^2(\mathbf{R}) \rightarrow L^2(\mathbf{R})$ is in \mathcal{L}_Φ^0 and P_λ are the projections for \sqrt{H} as above, then

$$\|P_\lambda A(1 - P_\lambda)\|_2 = o(\sqrt{\text{rank } P_\lambda}) = o(\lambda).$$

Here $\|\cdot\|_2$ is the Hilbert-Schmidt norm (compare [Gu, p. 244]).

PROOF.

$$P_m A P_n = \frac{1}{\lambda_m - \lambda_n} P_m [\sqrt{H}, A] P_n.$$

For any $\lambda, \delta > 0$, $\|P_\lambda A(1 - P_\lambda)\|_2^2 = \|P_\lambda A(P_{\lambda+\delta} - P_\lambda)\|_2^2 + \|P_\lambda A(1 - P_{\lambda+\delta})\|_2^2$. But $[\sqrt{H}, A] = [H_{1/2}, A] + [B, A]$ is bounded by Lemma 3.2 and the composition rules of the ψDO calculus. So

$$\|P_\lambda A(1 - P_\lambda)\|_2^2 \leq \|A\|^2 \text{rank}(P_{\lambda+\delta} - P_\lambda) + \delta^{-2} \|[H, A]\| \text{rank } P_\lambda.$$

But $\text{rank}(P_{\lambda+\delta} - P_\lambda) = 2\lambda\delta + \delta^2$, $\text{rank } P_\lambda = \lambda^2$. So

$$\lambda^{-2} \|P_\lambda A(1 - P_\lambda)\|_2^2 \leq C \left\{ \frac{2\delta}{\lambda} + \frac{\delta^2}{\lambda^2} + \delta^{-2} \right\} \rightarrow \frac{C}{\delta^2} \quad \text{as } \lambda \rightarrow \infty.$$

But δ is arbitrarily large, so $\|P_\lambda A(1 - P_\lambda)\|_2 = o(\lambda)$, as desired.

Recalling that $P_{\lambda^2} = \pi_\lambda$, we get

COROLLARY 3.3. $\|\pi_\lambda A \pi_\lambda^\perp\|_2 = o(\sqrt{\lambda})$.

It now follows, as in Widom's proof, that $\lambda^{-1}(\nu_\lambda(f) - \mu_\lambda(f)) = o(1)$ for polynomial f . Indeed, one writes $\pi^\perp = 1 - \pi$, $\pi_\lambda A^k \pi_\lambda = \pi_\lambda A(\pi_\lambda + \pi_\lambda^\perp) A \cdots A \pi_\lambda = (\pi_\lambda A \pi_\lambda)^k +$ terms with at least one factor of $\pi_\lambda A \pi_\lambda^\perp$. But we have just seen that $\|\pi_\lambda A \pi_\lambda^\perp\|_1 \leq \|\pi_\lambda\|_2 \|\pi_\lambda A \pi_\lambda\|_2 = o(\lambda)$, so $\|\pi_\lambda A^k \pi_\lambda - (\pi_\lambda A \pi_\lambda)^k\|_1 = o(\lambda)$. Consequently, $\|\pi_\lambda f(A) \pi_\lambda - f(\pi_\lambda A \pi_\lambda)\|_1 = o(\lambda)$ for any polynomial f .

Thus $\lim_{\lambda \rightarrow \infty} (\nu_\lambda(f)/\lambda) = \lim_{\lambda \rightarrow \infty} (\mu_\lambda(f)/\lambda)$ for polynomials f .

LEMMA 3.4.

$$\lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \nu_\lambda(f) = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H \leq T} f(a(x, \xi)) \, dx \, d\xi$$

for polynomials f .

PROOF. For any polynomial f , $f(A) \in \mathcal{L}_\Phi^0$. By the averaging theorem,

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \nu_\lambda(f) &= \lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \text{tr } \pi_\lambda f(A) \pi_\lambda \\ &= \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H \leq T} \sigma_{f(A)}(x, \xi) \, dx \, d\xi, \end{aligned}$$

assuming either of the limits exists.

By the composition rule, $\sigma_{f(A)}(x, \xi) = f(a(x, \xi)) + r(x, \xi)$ with $r(x, \xi) \in S_\Phi^{-1}$. Since any $r \in S_\Phi^{-1}$ satisfies $|r(x, \xi)| \leq C \cdot (1 + |x|^2 + |\xi|^2)^{-1/2}$, it is clear that its mean value is 0. Thus

$$\lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \nu_\lambda(f) = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H \leq T} f(a(x, \xi)) \, dx \, d\xi,$$

assuming either limit exists.

(As previously mentioned, the assumption that a has radial limits implies that these means exist for all continuous f .)

COROLLARY 3.5 (SZEGÖ THEOREM).

$$\lim_{\lambda \rightarrow \infty} \lambda^{-1} \mu_\lambda(f) = \mu(f) = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H \leq T} f(a(x, \xi)) \, dx \, d\xi$$

for all $f \in C(\mathbf{R})$.

PROOF. The eigenvalues a_i^λ of $\pi_\lambda A \pi_\lambda$ are bounded by $\|A\|$ for all λ, i ; the values of $a(x, \xi)$ are bounded by some constant C . Any continuous $g \in C(\mathbf{R})$ can be approximated uniformly on $I = \{|x| \leq \max(\|A\|, C)\}$ by a polynomial f_ϵ so $\|f_\epsilon - g\|_{\infty, I} < \epsilon$. Then

$$\begin{aligned} \left| \frac{1}{\lambda} (\mu_\lambda(g) - \mu_\lambda(f_\epsilon)) \right| &= \left| \frac{1}{\lambda} \sum_i g(a_i^\lambda) - f_\epsilon(a_i^\lambda) \right| < \epsilon \\ \Rightarrow \lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \mu_\lambda(g) &= \lim_{\lambda \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \frac{1}{\lambda} \mu_\lambda(f_\epsilon) \\ &= \lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H \leq T} f_\epsilon(a(x, \xi)) dx d\xi \\ &= \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{H \leq T} g(a(x, \xi)) dx d\xi. \end{aligned}$$

So the theorem holds for all $g \in C(\mathbf{R})$.

PART II. THE SPECIAL FUNCTION APPROACH

0. Introduction. In this part we take $H = -(1/4\pi^2)(d^2/dx^2) + x^2$ (Hermite operator; the normalization chosen here is more convenient for our purposes than that of Part I) and consider trace and Szegő limit theorems for bounded linear operators A of $L^2(\mathbf{R})$ outside the context of ψDO -theory. The simple commutator argument of Part I will not be available here to estimate effects of off-diagonal terms of A . However, the framework of operators and their Weyl symbols will be maintained; it is the calculus of symbols that is dropped.

We again use the Weyl symbols of operators since these give rise to explicit formulas, in terms of Wigner functions, for the matrix elements of bounded operators A relative to the basis $(\psi_k)_k$ of Hermite functions. These formulas will allow us to deal in the trace theorem with fairly general A 's, and in the Szegő theorem with fairly general multiplication operators M_V . (The restriction to multiplication operators in the Szegő theorem is natural since we need operators A for which the Weyl symbol σ_A satisfies $f(\sigma_A) = \sigma_{f(A)}$ if f is a continuous function.) In particular, smoothness assumptions may be dropped.

More precisely, we will show the following theorems.

THEOREM II.1. *Let A be a bounded operator of $L^2(\mathbf{R})$ whose Weyl symbol a is bounded. Let π_n be the orthogonal projection onto the span of the first n Hermite functions. If one of the two limits*

$$\lim_{T \rightarrow \infty} \frac{1}{\pi T} \int_{x^2 + \xi^2 \leq T} a(x, \xi) dx d\xi$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \text{trace}(\pi_n A \pi_n)$$

exists, then so does the other and assumes the same value. If $A = M_V$ (multiplication by $V \in L^\infty(\mathbf{R})$) the first limit may be replaced by $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T V(x) dx$.

For multiplication operators M_V where $V: \mathbf{R} \rightarrow \mathbf{R}$ belongs to a fairly large subset of $L^\infty(\mathbf{R})$ including all almost periodic functions and all V 's for which $\lim_{x \rightarrow \pm \infty} V(x)$ exists, we have the following Szegő limit theorem.

THEOREM II.2. *If $f: \mathbf{R} \rightarrow \mathbf{C}$ is continuous, then*

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \text{trace } f(\pi_n M_V \pi_n) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(V(x)) dx.$$

1. The Weyl correspondence and Wigner functions. We begin by reviewing the Weyl correspondence between functions $a(x, \xi)$ on \mathbf{R}^2 and operators acting on $L^2(\mathbf{R})$. Relevant references here are [GLS, Hö, dB, D and H], but since our normalizations are slightly different we prefer to give some details.

Let $a(x, \xi)$ be in Schwartz' class $S'(\mathbf{R}^2)$. Then we define the Weyl operator $a^W = a^W(x, D)$ associated with a as follows. For any two functions f and g in $S(\mathbf{R})$ we define the Wigner function $W(f, g)$ or $W(x, \xi; f, g)$ of f and g as (an overhead bar denotes complex conjugation)

$$W(x, \xi; f, g) = \int e^{-2\pi i t \xi} f\left(x + \frac{1}{2}t\right) \bar{g}\left(x - \frac{1}{2}t\right) dt.$$

We have $W(f, g) \in S(\mathbf{R}^2)$ for all f and g in $S(\mathbf{R})$. Now we define a^W by requiring that $(a^W f, g) = (W(f, g), \bar{a})$ for all f and g in $S(\mathbf{R})$. The operator a^W so defined maps $S(\mathbf{R})$ continuously into $S'(\mathbf{R})$; its Schwartz kernel $A^W \in S'(\mathbf{R}^2)$ is related to a by the formula

$$a(x, \xi) = \int A^W\left(x + \frac{1}{2}t, x - \frac{1}{2}t\right) e^{-2\pi i t \xi} dt$$

(this identity is to be interpreted in the distributional sense). In particular, if we take $a = W(\varphi, \psi)$, a^W is the operator $\varphi \otimes \bar{\psi}$, i.e. $a^W f = (f, \psi)\varphi$.

The attention paid below to Wigner functions is consistent with the pseudo-differential theme of treating operators via their symbols rather than their kernels. Much of it is devoted to getting exact results on full symbols, rather than on leading order terms.

It is known that if $a \in L^p(\mathbf{R}^2)$, $1 \leq p \leq 2$, then a^W is of Schatten p' -class where $p' = p(p-1)^{-1}$ (cf. [H, §3.5]). On the other hand, if A is a bounded operator on $L^2(\mathbf{R})$, then there is an $a \in [S^5]'$ such that $(Af, g) = (W(f, g), \bar{a})$ for all $f, g \in S(\mathbf{R})$. Here $[S^5]'$ is the dual of

$$S^5 = \left\{ f \mid f \in C^5(\mathbf{R}^2), \max_{|m| \leq 5} \sup |(1+x^2+\xi^2)^{(5-|m|)/2} \partial_{x,\xi}^m f| < \infty \right\}$$

(see [D] for details). We call a the Weyl symbol of A . Hence, if A is a bounded linear operator of $L^2(\mathbf{R})$, then we can write

$$\text{trace}(\pi_n A \pi_n) = \left(\sum_{k=0}^n W(\psi_k, \psi_k), \bar{a} \right)$$

where the ψ_k 's are the Hermite functions. In general the matrix elements of A (with respect to the ψ_k 's) are given by $(A\psi_k, \psi_l) = (W(\psi_k, \psi_l), \bar{a})$.

In §2 we shall give a number of estimates and formulas about the $W(\psi_k, \psi_1)$'s. The explicitness of these show clearly the advantage of the Weyl symbol over the usual one for our purposes; in particular, one has in many cases a good idea how fast $(1/n + 1) \text{trace}(\pi_n A \pi_n)$ approaches its limit if $n \rightarrow \infty$. It is, furthermore, important to note that A is selfadjoint if a is real.

The following gives an alternative description of the Weyl correspondence (this is Weyl's original set up). For $(q, p) \in \mathbf{R}^2$ we define the unitary operators $T_q, R_p, W_{q,p}$ by

$$(T_q f)(x) = f(x + q), \quad (R_p f)(x) = e^{-2\pi i p x} f(x),$$

$$W_{q,p} f = e^{-\pi i q p} R_p T_q f.$$

If $a \in S'(\mathbf{R}^2)$, we define the symplectic Fourier transform $\tilde{a} \in S'(\mathbf{R}^2)$ of a by

$$\tilde{a}(q, p) = \int e^{2\pi i(xp - \xi q)} a(x, \xi) dx d\xi.$$

It can be shown that for all $f \in S(\mathbf{R})$

$$a^w f = \int W_{q,p} f \tilde{a}(q, p) dq dp;$$

that is, in the inversion formula

$$a(x, \xi) = \int e^{2\pi i(q\xi - px)} \tilde{a}(q, p) dq dp,$$

we replace the exponential $\exp(2\pi i(q\xi - px))$ by its "quantization" $W_{q,p}$.

2. Hermite functions and Wigner functions. The following objects and observations are basic to the Hermite theory in this paper.

- (a) The Hermite operator $H = -(1/4\pi^2)(d^2/dx^2) + x^2$.
- (b) The Hermite function $\psi_n(x) = e^{-\pi x^2} H_n(x\sqrt{2\pi})(2^{n-1/2}n!)^{1/2}$, where $H_n(x) = (-1)^n e^{x^2} (d/dx)^n e^{-x^2}$ is the n th Hermite polynomial, $n = 0, 1, \dots$
- (c) $H\psi_n = (n + \frac{1}{2})\pi^{-1}\psi_n$, $n = 0, 1, \dots$
- (d) $\exp(\pi x^2 - 2\pi(x-w)^2) = \sum_{n=0}^{\infty} c_n w^n \psi_n(x)$ where $c_n = (n!)^{-1/2} 2^{-1/4} (4\pi)^{n/2}$.
- (e) Mehler's formula for the kernel $e^{-\pi t H}$ (compare [dB, §27.6.3]) may be written as

$$\sum_{n=0}^{\infty} e^{-(n+1/2)t} \psi_n(x) \psi_n(y) = \left(\frac{1}{\sinh t} \right)^{1/2} \exp\left(-\frac{\pi}{\sinh t} ((x^2 + y^2) \cosh t + 2xy) \right)$$

for $\text{Re } t > 0, x, y \in \mathbf{C}$. Alternatively,

$$\sum_{n=0}^{\infty} w^n \psi_n(x) \psi_n(y) = \left(\frac{2}{1-w^2} \right)^{1/2} \exp\left(-\pi \frac{1+w^2}{1-w^2} (x^2 + y^2) + \frac{4\pi xyw}{1-w^2} \right).$$

(f) The Wigner function $W(x, \xi; \psi_m, \psi_m)$ is given by

$$W(x, \xi; \psi_m, \psi_m) = 2 \cdot (-1)^m e^{-2\pi(x^2 + \xi^2)} L_m(4\pi(x^2 + \xi^2)),$$

where L_m is the m th Laguerre polynomial of order 0 (cf. [Ja, Gr]). Note that $W(x, \xi; \psi_m, \psi_m)$ is the Weyl symbol of the projection operator $\psi_m \otimes \psi_m$ onto the span of ψ_m .

(g) If we write $x + i\xi = re^{i\theta}$, then

$$W(x, \xi; \psi_{m+k}, \psi_m) = 2 \cdot (-1)^m \sqrt{\frac{m!}{(m+k)!}} e^{-ik\theta} (4\pi r^2)^{k/2} e^{-2\pi r^2} L_m^{(k)}(4\pi r^2)$$

where $L_m^{(k)}$ is the m th Laguerre polynomial of order k , for which the generating function is given by $\sum_{m=0}^{\infty} w^m L_m^{(k)}(x) = (1-w)^{-k-1} \exp(-xw(1-w)^{-1})$ for $|w| < 1$, $x \geq 0$. Consequently, $W(re^{i\theta}; \psi_{m+k}, \psi_m)$ has the form $e^{-ik\theta} \Phi_{m,k}(r)$, as already observed in [Gr]. The above follows from the formula

$$c_k c_l W(x, \xi; \psi_k, \psi_l) = 2^{1/2} \exp(-2\pi r^2) \sum_{j=1}^{\min(k,l)} \frac{(-4\pi)^j}{j!} \frac{(4\pi r)^{k+l-2j}}{(k-j)!} \frac{e^{i(l-k)\theta}}{(l-j)!}$$

(the c_k 's are as in (d)). This formula can be found in [Ja, Gr].

(h) π_n , the orthogonal projection onto $\text{span}\{\psi_k | k \leq n\}$, itself has the Weyl symbol

$$\sigma_{\pi_n} = \sum_{k=0}^n W(x, \xi; \psi_k, \psi_k) = 2 \cdot \sum_{k=0}^n (-1)^k e^{-2\pi(x^2 + \xi^2)} L_k(4\pi(x^2 + \xi^2)).$$

Since

$$\text{tr } \pi_n A \pi_n = \int \sigma_{\pi_n}(x, \xi) a(x, \xi) dx d\xi,$$

one interpretation of the trace theorem (as observed to us by Alan Weinstein) is that σ_{π_n} is approximately the characteristic function of the n th energy disc. One of us (A. J. E. M. Janssen) intends to elaborate this point further. At present we can say the following: The sums $\sum_{k=0}^n W(x, \xi; \psi_k, \psi_k) =: S_n(x, \xi)$ are nonnegative, uniformly bounded, very small if $\pi(x^2 + \xi^2) \geq n(1 + \epsilon)$ if $\epsilon > 0$ is fixed and n is large, and $S_n(x, \xi) \rightarrow 1$ uniformly on any compact set of (x, ξ) 's not containing $(0, 0)$. This information, though useful, is not sufficient to prove the trace theorem since the behavior of $S_n(x, \xi)$ should also be known in the region $\frac{1}{2}n \leq \pi(x^2 + \xi^2) \leq n$, say.

In the remainder of this section, we shall derive estimates for $W(re^{i\theta}; \psi_k, \psi_l)$. We have, of course, $|W(re^{i\theta}; \psi_k, \psi_l)| \leq 2$. More precise information can be obtained from [Sz, Theorem 8.91.2] which gives the asymptotic behavior of

$$\max_x e^{-x/2} x^\lambda |L_n^{(\alpha)}(x)| \quad (n \rightarrow \infty)$$

if α is fixed and $\lambda \in \mathbf{R}$. Unfortunately, we need estimates for $W(re^{i\theta}; \psi_k, \psi_l)$ where $k, l \rightarrow \infty$ independently. The result we prove below is not the best possible, but it suffices for the analysis of the multiplication operators and certain operators of more general type. We are especially interested in an estimate that shows that $W(re^{i\theta}; \psi_k, \psi_l)$ is small if $l - k$ is sufficiently large. For that gives us an estimate of the ‘‘off-diagonal’’ part of the matrix of an operator A (relative to the basis $(\psi_k)_k$) whose matrix elements are given by $(A\psi_k, \psi_l) = (W(\psi_k, \psi_l), \bar{a})$. Such an estimate is, of course, not necessary if we are only interested in trace theorems, but if we want to prove Szegő limit theorems, we definitely cannot do without it.

The basic result is contained in the following lemma.

LEMMA II. 2.1. *We have for $l > k, r \geq 0, \theta \in \mathbf{R}$*

$$|W(re^{i\theta}; \psi_k, \psi_l)| \leq \frac{1}{2} \left(\frac{l!}{k!} \right)^{1/2} \left(\frac{2e\pi^{1/2}r}{l-k} \right)^{l-k} \left(1 + \frac{4\pi r^2}{l-k} \right)^k e^{-2\pi r^2}.$$

PROOF. As observed before $|W(re^{i\theta}; \psi_{m+k}, \psi_m)|$ is a function of r alone. Hence it is sufficient to investigate $(T_a\psi_k, \psi_l) = 2(-1)^l W(\frac{1}{2}a, 0; \psi_k, \psi_l)$ (recall that $(T_a f)(t) = f(t+a)$). We get a convenient expression for $(T_a\psi_k, \psi_l)$ by using the Bargmann transform (cf. [Jb]). The Bargmann transform Bf of an $f \in L^2(\mathbf{R})$ is defined as

$$\begin{aligned} (Bf)(z) &= 2^{1/4} \int_{-\infty}^{\infty} \exp\left(\frac{1}{2}\pi z^2 - \pi(t-z)^2\right) f(t) dt \\ &= \sum_{n=0}^{\infty} \frac{(f, \psi_n)}{\sqrt{n!}} (\pi^{1/2}z)^n \end{aligned}$$

for $z \in \mathbf{C}$. We have the formulas (cf. [Jb or Ba])

$$\begin{aligned} (f, g) &= (Bf, Bg)_C = \int_C \exp(-\pi|z|^2) (Bf)(z) \overline{(Bg)(z)} dz, \\ (BT_a f)(z) &= \exp\left(-\frac{1}{2}\pi a^2 - \pi az\right) (Bf)(z+a), \\ (B\psi_n)(z) &= \frac{(\pi^{1/2}z)^n}{\sqrt{n!}}. \end{aligned}$$

Here $f, g \in L^2(\mathbf{R}), a \in \mathbf{R}, n = 0, 1, \dots$. Therefore we get

$$(T_a\psi_k, \psi_l) = \frac{\pi^{(k+l)/2}}{\sqrt{k!l!}} \int_C \exp\left(-\frac{1}{2}\pi a^2 - \pi az\right) (z+a)^k \bar{z}^l \exp(-\pi|z|^2) dz.$$

Using polar coordinates $z = re^{i\theta}$ in the integral

$$c_{kl}(a) := \int_C \exp(-\pi|z|^2 - \pi az) (z+a)^k \bar{z}^l dz$$

we get

$$\begin{aligned} c_{kl}(a) &= \int_0^{\infty} \int_0^{2\pi} \exp(-\pi r^2 - \pi a r e^{i\theta}) (re^{i\theta} + a)^k r^{l+1} e^{-il\theta} d\theta \\ &= \int_0^{\infty} r^{l+1} \exp(-\pi r^2) \left\{ \int_0^{2\pi} \exp(-\pi a r e^{i\theta}) (re^{i\theta} + a)^k e^{-il\theta} d\theta \right\} dr. \end{aligned}$$

Substituting $w = e^{i\theta}$ ($\theta \in [0, 2\pi]$) in the integral $I(r, a, l, k)$ between $\{ \}$ we get

$$\begin{aligned} I(r, a, l, k) &= \frac{1}{i} \int_{|w|=1} e^{-\pi a r w} (r w + a)^k w^{-l-1} dw \\ &= \frac{1}{i} \int_{|w|=R} e^{-\pi a r w} (r w + a)^k w^{-l-1} dw \end{aligned}$$

for any $R > 0$. Take $R = (l - k)/\pi|a|r$. On the circle $|w| = R$ we estimate

$$\begin{aligned} |\exp(-\pi awr)(rw + a)^k w^{-l-1}| &\leq e^{l-k} \left(\frac{l-k}{\pi|a|} + |a| \right)^k \left(\frac{\pi|a|r}{l-k} \right)^{l+1} \\ &= e^{l-k} (\pi|a|)^{l-k+1} (l-k)^{-(l-k)-1} \left(1 + \frac{\pi|a|^2}{l-k} \right)^k r^{l+1}. \end{aligned}$$

Here we have assumed that $l > k$. We therefore get

$$\begin{aligned} |c_{kl}(a)| &\leq 2\pi e^{l-k} (\pi|a|)^{l-k} (l-k)^{-(l-k)} \left(1 + \frac{\pi|a|^2}{l-k} \right)^k \int_0^\infty r^{2l+1} e^{-\pi r^2} dr \\ &= \left(\frac{e|a|}{l-k} \right)^{l-k} \left(1 + \frac{\pi|a|^2}{l-k} \right)^k \pi^{-k} l!. \end{aligned}$$

Hence, if $a \in \mathbf{R}$ and $l > k$,

$$|(T_a \psi_k, \psi_l)| \leq \left(\frac{l!}{k!} \right)^{1/2} \left(\frac{e\pi^{1/2}|a|}{l-k} \right)^{l-k} \left(1 + \frac{\pi|a|^2}{l-k} \right)^k \exp\left(-\frac{1}{2}\pi a^2\right).$$

This concludes the proof.

This estimate on $(T_a \psi_k, \psi_l)$ is fairly good, for it shows that $(T_a \psi_k, \psi_l)$ has a zero of order $l - k$ at $a = 0$ and that $(T_a \psi_k, \psi_l) = O(|a|^{l+k} \exp(-\frac{1}{2}\pi a^2))$ for large a .

COROLLARY. For any $C > 0$ there is a $\sigma > 0$ such that

$$|(T_a \psi_k, \psi_l)| \leq C^{l-k} \exp\left(-\frac{1}{2}\pi a^2\right)$$

for all $a \in \mathbf{R}$ and all k, l with $k \geq a^2$, $l - k \geq |a|\sqrt{k}/\sigma$.

PROOF. Let $\varepsilon > 0$. We consider $(T_a \psi_k, \psi_l)$ for $k \geq a^2$, $l - k \geq \sqrt{k}/\varepsilon$. We have (since $l!/k! \leq l^{l-k}$, $k/(l-k) \leq \varepsilon^2(l-k)$)

$$\begin{aligned} |(T_a \psi_k, \psi_l)| &\leq \left(\frac{l!}{k!} \right)^{1/2} \left(\frac{e\pi^{1/2}|a|}{l-k} \right)^{l-k} \exp\left(\frac{\pi k a^2}{l-k}\right) e^{-\frac{1}{2}\pi a^2} \\ &\leq \left\{ \frac{e\pi^{1/2}|a|l^{1/2}}{l-k} \exp(\pi a^2 \varepsilon^2) \right\}^{l-k} e^{-\frac{1}{2}\pi a^2}. \end{aligned}$$

Now $l^{1/2}/(l-k) \leq \varepsilon(1 + \varepsilon^{-1}k^{-1/2})^{1/2} \leq \varepsilon(1 + (|a|\varepsilon)^{-1})^{1/2}$, so

$$|(T_a \psi_k, \psi_l)| \leq \left\{ e\pi^{1/2}|a|\varepsilon(1 + (|a|\varepsilon)^{-1})^{1/2} \exp(\pi|a|^2\varepsilon^2) \right\}^{l-k} e^{-\frac{1}{2}\pi a^2}.$$

Hence, if σ is the solution of the equation $e\pi^{1/2}\sigma(1 + \sigma^{-1})^{1/2} \exp(\pi\sigma^2) = C$, and $\varepsilon = \sigma/|a|$, then $|(T_a \psi_k, \psi_l)| \leq C^{l-k} \exp(-\frac{1}{2}\pi a^2)$ for $k \geq a^2$, $l - k \geq \sqrt{k}/\varepsilon = |a|\sqrt{k}/\sigma$. This completes the proof.

REMARK. This lemma says that for $k \geq a^2$ the size of the set of all $l \geq k$ for which $(T_a \psi_k, \psi_l)$ cannot be neglected is $O(|a|\sqrt{k})$.

3. The proof of Theorem II.1. The proof of the first half of Theorem II.1 uses similar arguments as the proof of Theorem 1 of Part I; to keep the two parts independent we give the following quick proof with the aid of Wigner functions. Let

A be a bounded operator of $L^2(\mathbf{R})$ whose Weyl symbol a is bounded. As observed in §1 of Part II we have

$$\text{trace}(\pi_n A \pi_n) = \sum_{k=0}^n (A\psi_k, \psi_k) = \sum_{k=0}^n (W(\psi_k, \psi_k), \bar{a}).$$

Consider $F(w) := \sum_{k=0}^\infty w^k (A\psi_k, \psi_k)$ for $|w| < 1$. The numbers $(A\psi_k, \psi_k)$ are bounded since A is bounded. By a simple Tauberian theorem we have: if one of the limits $\lim_{n \rightarrow \infty} (\text{trace}(\pi_n A \pi_n)) / (n + 1)$ and $\lim_{w \rightarrow 1^-} (1 - w)F(w)$ exists then so does the other and assumes the same value. By §2(f), (g) we can write $F(w)$ as $(G(\cdot, \cdot; w), \bar{a})$, where $x, \xi \in \mathbf{R}, |w| \leq 1$, and

$$\begin{aligned} G(x, \xi; w) &:= \sum_{k=0}^\infty w^k W(x, \xi; \psi_k, \psi_k) \\ &= 2(1 + w)^{-1} \exp\left(-2\pi(x^2 + \xi^2) \frac{1 - w}{1 + w}\right). \end{aligned}$$

Hence, it is sufficient to investigate when

$$\lim_{\epsilon \rightarrow 0^+} 2\epsilon \int \exp(-2\pi(x^2 + \xi^2)\epsilon) a(x, \xi) dx d\xi$$

exists. Changing to polar coordinates $(x, \xi) = r(\cos \theta, \sin \theta)$ and performing the substitution $u = 2\pi r^2$ we can write the above limit as $\lim_{\epsilon \rightarrow 0^+} \epsilon \int_0^\infty e^{-\epsilon u} a_{\text{ave}}(u) du$, where

$$a_{\text{ave}}(u) = \frac{1}{2\pi} \int_0^{2\pi} a(\sqrt{u/2\pi}(\cos \theta, \sin \theta)) d\theta.$$

From the Tauberian theorem in [Si, pp. 107–108] applied with

$$\mu([0, T]) = \int_0^T a_{\text{ave}}(u) du, \quad \gamma = 1,$$

we know that if one of the limits

$$\lim_{\epsilon \rightarrow 0^+} \epsilon \int_0^\infty \exp(-\epsilon u) a_{\text{ave}}(u) du \quad \text{and} \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T a_{\text{ave}}(u) du$$

exists then so does the other and assumes the same value. Now since

$$\frac{1}{T} \int_0^T a_{\text{ave}}(u) du = \frac{1}{\pi T} \int_{x^2 + \xi^2 \leq T} a(x, \xi) dx d\xi$$

the proof of the first part of Theorem II.1 is complete.

Note. A sufficient Tauberian condition on a to apply the theorem just quoted is: there is a $z \in \mathbf{C}$ and $\theta \in \mathbf{R}$ such that all values of $a_{\text{ave}}(u)$ lie in the set $\{z + re^{i\alpha} \mid r \in [0, \infty), \alpha \in [\theta, \theta + \pi/2]\}$. This is definitely weaker than the boundedness assumption on a . Since the Tauberian theorem is applied on a_{ave} , and not on a itself, we can even allow certain distributions for a . We shall not pursue this point.

The second half of the proof of Theorem II.1 deals with the case that $A = M_V$, the multiplication operator by $V \in L^\infty(\mathbf{R})$ whose Weyl symbol a is given by $a(x, \xi) = V(x)$. That is, we want to replace

$$(*) \quad \lim_{T \rightarrow \infty} \frac{1}{\pi T} \int_{x^2 + \xi^2 \leq T} V(x) dx d\xi = \lim_{A \rightarrow \infty} \frac{2}{\pi A} \int_{-A}^A V(x) \sqrt{1 - (x/A)^2} dx$$

by

$$(**) \quad \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V(x) dx.$$

Although this can be done directly by using Wiener’s Tauberian theorem, the following argument is perhaps more instructive. As before we can consider $\lim_{w \rightarrow 1-} (1-w)F(w)$ instead of (*), where $F(w) = \sum_{n=0}^{\infty} w^n (M_V \psi_n, \psi_n)$. By Mehler’s formula (§2(g))

$$F(w) = \left(\frac{2}{1-w^2} \right)^{1/2} \int_{-\infty}^{\infty} V(x) \exp\left(-2\pi \frac{1-w}{1+w} x^2\right) dx.$$

Hence studying $(1-w)F(w)$ as $w \rightarrow 1-$ amounts to studying

$$\varepsilon^{1/2} \int_{-\infty}^{\infty} V(x) \exp(-\pi \varepsilon x^2) dx \quad \text{as } \varepsilon \rightarrow 0+.$$

We need the following Tauberian theorem.

PROPOSITION II.3.1. *Let $f: [0, \infty) \rightarrow \mathbf{C}$ be bounded and measurable. If one of the limits $\lim_{T \rightarrow \infty} (1/T) \int_0^T f(x) dx$ and $\lim_{\varepsilon \rightarrow 0+} (2\varepsilon)^{1/2} \int_0^{\infty} f(x) \exp(-\pi \varepsilon x^2) dx$ exists, then so does the other and assumes the same value.*

PROOF. Let $K_1(t) = e^{-t} \chi_{[0, \infty)}(t)$, $K_2(t) = e^{-t} \exp(-\pi e^{-2t})$ for $t \in \mathbf{R}$. If we take $T = e^y$, $\varepsilon = e^{-2y}$ and substitute $x = e^u$, we can write $(1/T) \int_0^T f(x) dx$ and $\varepsilon^{1/2} \int_0^{\infty} f(x) \exp(-\pi \varepsilon x^2) dx$ as $\int_{-\infty}^{\infty} f(e^u) K_1(y-u) du$ and $\int_{-\infty}^{\infty} f(e^u) K_2(y-u) du$ respectively. The Fourier transforms of K_1 and K_2 are given respectively by

$$\int_{-\infty}^{\infty} e^{-2\pi i \lambda t} K_1(t) dt = \frac{1}{1 + 2\pi i \lambda},$$

$$\int_{-\infty}^{\infty} e^{-2\pi i \lambda t} K_2(t) dt = 2\pi^{-1/2} \Gamma\left(\frac{1}{2} + \pi i \lambda\right)$$

for $\lambda \in \mathbf{R}$. Since these Fourier transforms have no real zeros, the proposition follows from Wiener’s Tauberian theorem.

Note. A slightly more general theorem is proved in [Jc, §4.2].

By applying Proposition II.3.1 to $f(x) = V(x) + V(-x)$ ($x \geq 0$) we see at once that if one of the limits

$$\lim_{n \rightarrow \infty} (\text{trace}(\pi_n M_V \pi_n)) / (n + 1) \quad \text{and} \quad \lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T V(x) dx$$

exists, then so does the other and assumes the same value. The proof of Theorem II.1 is complete.

4. The Szegő limit theorem for multiplication operators. We restrict ourselves in this section mainly to multiplication operators $A = M_V$ where $V \in L^\infty(\mathbf{R})$; we shall give some indications how to treat more general operators at the end of this section. We shall often write V instead of M_V .

We want to find a sufficiently large class of functions $V \in L^\infty(\mathbf{R})$ for which the following holds: if f is a continuous function defined on \mathbf{R} , then $\lim_{n \rightarrow \infty} (f(\pi_n M_V \pi_n)) / (n + 1)$ exists and equals $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T f(V(x)) dx$. To do this we first introduce the class $\mathfrak{C}(\mathfrak{L})$.

DEFINITION II.4.1. \mathfrak{W} is the set of all $V \in L^\infty(\mathbf{R})$ for which $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T V^k(x) dx$ exists for all $k = 0, 1, \dots$.

For the time being we allow V to be complex-valued, but later on we shall have to restrict ourselves to real-valued V 's. We note that for any $n = 1, 2, \dots$ there is a $V \in L^\infty(\mathbf{R})$ such that $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T V^k(x) dx$ exists for $k = 0, 1, \dots, n - 1$, and such that $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T V^n(x) dx$ does not exist. The class \mathfrak{W} is not linear.

We give some simple properties of the class \mathfrak{W} .

LEMMA II.4.1. $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T f(V(x)) dx$ exists if $f: \mathbf{R} \rightarrow \mathbf{C}$ is continuous and $V \in \mathfrak{W}$.

PROOF. Since f is continuous and $V \in L^\infty(\mathbf{R})$ we can find a sequence $(P_n)_n$ of polynomials such that $P_n(V(x)) \rightarrow f(V(x))$ uniformly in $x \in \mathbf{R}$. The lemma easily follows.

DEFINITION II.4.2. For $V \in L^\infty(\mathbf{R})$ we put

$$\rho(V) := \limsup_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |V(x)| dx.$$

LEMMA II.4.2. Let $V_n \in L^\infty(\mathbf{R})$, $\|V_n\|_\infty \leq M$, $V \in L^\infty(\mathbf{R})$, $V_n \in \mathfrak{W}$, $\rho(V_n - V) \rightarrow 0$ ($n \rightarrow \infty$). Then $V \in \mathfrak{W}$.

PROOF. Let $k = 0, 1, \dots, \varepsilon > 0$. We have for $n = 0, 1, \dots, T > 0$

$$\left| \frac{1}{2T} \int_{-T}^T (V_n^k(x) - V^k(x)) dx \right| \leq \frac{A}{2T} \int_{-T}^T |V_n(x) - V(x)| dx,$$

where $A := k \max(M^{k-1}, \|V\|_\infty^{k-1})$. Let

$$I_n^{(k)} := \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V_n^k(x) dx.$$

Take n so large that $\rho(V_n - V) < \varepsilon/2A$, and then take T_0 so large that

$$\left| \frac{1}{2T} \int_{-T}^T V_n^k(x) dx - I_n^{(k)} \right| < \varepsilon,$$

$$\frac{1}{2T} \int_{-T}^T |V_n(x) - V(x)| dx \leq 2\rho(V_n - V)$$

if $T > T_0$. Then we have for $T_1, T_2 > T_0$ that

$$\left| \frac{1}{2T_1} \int_{-T_1}^{T_1} V^k(x) dx - \frac{1}{2T_2} \int_{-T_2}^{T_2} V^k(x) dx \right| \leq 4\varepsilon.$$

From this it follows that $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T V^k(x) dx$ exists for all k . Hence $V \in \mathfrak{W}$.

LEMMA II.4.3. (i) \mathfrak{W} contains all $V \in L^\infty(\mathbf{R})$ for which

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |V(x) - L_+| dx = 0 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T}^0 |V(x) - L_-| dx$$

for some $L_+, L_- \in \mathbf{C}$.

(ii) \mathfrak{W} contains all V which are almost periodic in the following sense: there exists a bounded sequence of trigonometric polynomials V_n such that $\rho(V_n - V) \rightarrow 0$ ($n \rightarrow \infty$).

(iii) $\mathcal{O}\mathcal{B}$ contains all $V \in L^\infty(\mathbf{R})$ for which there exists a $g: \mathbf{R} \rightarrow \mathbf{C}$, locally of bounded variation, such that $\rho(V_A - V) \rightarrow 0$ if $A \rightarrow \infty$. Here $V_A(x) = \int_{-A}^A e^{-2\pi i \lambda x} dg(\lambda)$ is assumed to be uniformly bounded.

PROOF. (i) Define $V_0(x) = L_+$ or L_- according as $x \geq 0$ or $x < 0$. Then $V_0 \in \mathcal{O}\mathcal{B}$ and $\rho(V - V_0) = 0$. Hence $V \in \mathcal{O}\mathcal{B}$ by Lemma II.4.2.

(ii) We only need to check that trigonometric polynomials are in $\mathcal{O}\mathcal{B}$. That is easy.

(iii) Let $A > 0$ be such that $\pm A$ are continuity points of g . Then $V_A(x) = \int_{-A}^A e^{-2\pi i \lambda x} dg(\lambda + 0)$ for $x \in \mathbf{R}$ (we have $g(\lambda + 0) = \lim_{\eta \downarrow \lambda} g(\eta)$). Now put $g_A(\lambda) = g(\lambda + 0)$ if $|\lambda| \leq A$, $g_A(\lambda) = g(A)$ if $\lambda \geq A$ and $g_A(\lambda) = g(-A)$ if $\lambda \leq -A$. It follows from the theory of characteristic functions of probability distributions that

$$(V_A(x))^k = \left(\int e^{-2\pi i \lambda x} dg_A(\lambda) \right)^k = \int e^{-2\pi i \lambda x} dg_{k,A}(\lambda),$$

where $dg_{k,A}$ is the k -fold convolution of dg_A with itself. Now, by Fubini's theorem and dominated convergence,

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \left\{ \int e^{-2\pi i \lambda x} dg_{k,A}(\lambda) \right\} dx = (dg_{k,A})\{0\}.$$

Hence $V_A \in \mathcal{O}\mathcal{B}$. From this the lemma easily follows.

As in Part I we follow Widom's approach as presented in Guillemin's survey paper. That is, we let $k = 1, 2, \dots$, and we want to show that $\|\pi_n V^k \pi_n - (\pi_n V \pi_n)^k\|_1 = o(n)$ ($n \rightarrow \infty$). Here $\|\cdot\|_1$ is the trace class norm. It will be sufficient to show that $\|\pi_n V(I - \pi_n)\|_2 = o(n^{1/2})$ ($n \rightarrow \infty$). Here $\|\cdot\|_2$ denotes the Hilbert-Schmidt norm.

LEMMA II.4.4. Denote $V_\lambda(x) := e^{-2\pi i \lambda x}$ ($x \in \mathbf{R}$) for $\lambda \in \mathbf{R}$. Then

$$\|\pi_n M_{V_\lambda}(I - \pi_n)\|_2^2 = O(\lambda^2 + |\lambda| \sqrt{n}).$$

PROOF. We have $|(V_\lambda \psi_k, \psi_l)| = |(T_\lambda \psi_k, \psi_l)|$ for all $\lambda \in \mathbf{R}$, $k = 0, 1, \dots$, $l = 0, 1, \dots$. This is so because $\mathfrak{F} \psi_k = (-i)^k \psi_k$ (\mathfrak{F} denotes the Fourier transform) for $k = 0, 1, \dots$ and $(\mathfrak{F} V_\lambda f)(x) = (\mathfrak{F} f)(x + \lambda) = (T_\lambda \mathfrak{F} f)(x)$ for $\lambda, x \in \mathbf{R}$, $f \in L^2(\mathbf{R})$. Let $C < 1$. By the corollary at the end of §2 we can find $\sigma > 0$ such that

$$|(T_\lambda \psi_k, \psi_l)| \leq C^{l-k} \exp(-\frac{1}{2} \pi \lambda^2)$$

for $\lambda \in \mathbf{R}$, $k \geq \lambda^2$, $l - k \geq |\lambda| \sqrt{k} / \sigma$. Now

$$\|\pi_n V_\lambda(I - \pi_n)\|_2^2 = \sum_{k=0}^n \sum_{l=n+1}^\infty |(V_\lambda \psi_k, \psi_l)|^2.$$

For $k < \lambda^2$ we estimate $\sum_{l=n+1}^\infty |(V_\lambda \psi_k, \psi_l)|^2$ by 1. For $\lambda^2 \leq k \leq n$ we have

$$\sum_{l=n+1}^\infty |(V_\lambda \psi_k, \psi_l)|^2 = \sum_{n+1 \leq l < n + |\lambda| \sqrt{n} / \sigma} |(V_\lambda \psi_k, \psi_l)|^2 + \sum_{l \geq n + |\lambda| \sqrt{n} / \sigma} |(V_\lambda \psi_k, \psi_l)|^2.$$

The second one of these sums can be estimated by

$$\exp(-\pi \lambda^2) (1 - C^2)^{-1} C^{2(n-k + |\lambda| \sqrt{n} / \sigma)},$$

and if we sum this over all k we get a number that is at most $(1 - C^2)^{-2} \exp(-\pi \lambda^2)$. And the sum of $|(V_\lambda \psi_k, \psi_l)|^2$ over all k and all l with $n + 1 \leq l < n + |\lambda| \sqrt{n} / \sigma$

yields a number $\leq |\lambda| \sqrt{n} / \sigma$. Putting the estimates together we get

$$\|\pi_n V_\lambda (I - \pi_n)\|_2^2 \leq \lambda^2 + |\lambda| \sqrt{n} / \sigma + (1 - C^2)^{-2} \exp(-\pi \lambda^2)$$

as desired.

Lemma II.4.4 is used to prove the following proposition.

PROPOSITION II.4.1. *If V is a trigonometric polynomial and f is a polynomial, then $\|f(\pi_n V \pi_n) - \pi_n f \circ V \pi_n\|_1 = O(n^{3/4})$.*

PROOF. Let $V(x) = \sum_{l=0}^m c_l V_{\lambda_l}(x)$. We may assume, by linearity and the triangle inequality for $\|\cdot\|_1$, that $f(x) = x^k$ for some $k = 0, 1, \dots$. We have

$$\pi_n \left(\sum_{l=0}^m c_l V_{\lambda_l} \right)^k \pi_n = \sum_{l_1=0}^m \cdots \sum_{l_k=0}^m c_{l_1} \cdots c_{l_k} \pi_n V_{\lambda_{l_1}} \cdots V_{\lambda_{l_k}} \pi_n.$$

We write

$$\pi_n V_{\lambda_{l_1}} \cdots V_{\lambda_{l_k}} \pi_n = \pi_n (V_{\lambda_{l_1}}(\pi_n + (I - \pi_n)) V_{\lambda_{l_2}}(\pi_n + (I - \pi_n)) \cdots V_{\lambda_{l_k}}(\pi_n + (I - \pi_n))),$$

and we expand the product at the right-hand side. We get $\pi_n V_{\lambda_{l_1}} \pi_n \cdots \pi_n V_{\lambda_{l_k}} \pi_n +$ sum of products each of which involves at least one factor $\pi_n V_{\lambda_{l_j}} (I - \pi_n)$. The trace norm of each of the terms in this sum can be estimated (as $\pi_n, I - \pi_n, V_\lambda$ all have operator norm ≤ 1) by $\|\pi_n V_\lambda (I - \pi_n)\|_1 \leq \|\pi_n\|_2 \|\pi_n V_\lambda (I - \pi_n)\|_2 = (n + 1)^{1/2} O(n^{1/4})$. Here λ is one of the λ_j 's. Hence we get

$$\left\| \pi_n \left(\sum_{l=0}^m c_l V_{\lambda_l} \right)^k \pi_n - \sum_{l_1=0}^m \cdots \sum_{l_k=0}^m \prod_{j=1}^k c_{l_j} \pi_n V_{\lambda_{l_j}} \pi_n \right\|_1 = O(n^{3/4}).$$

The multiple sum here equals $[\pi_n (\sum_{l=0}^m c_l V_{\lambda_l}) \pi_n]^k = (\pi_n V \pi_n)^k$, and the proof is complete.

Note. With a similar proof one can show the following. Let g be locally of bounded variation, and let f be a polynomial. Then we have for all $A > 0$ that

$$\left\| \pi_n f \left(\int_{-A}^A V_\lambda dg(\lambda) \right) \pi_n - f \left(\pi_n \int_{-A}^A V_\lambda dg(\lambda) \pi_n \right) \right\|_1 = O(n^{3/4}).$$

We assume from now on that V is real-valued. This is mainly done to be able to include general continuous functions f .

LEMMA II.4.5. *Let $g: \mathbf{R} \rightarrow \mathbf{C}$ be continuous and $V \in L^\infty(\mathbf{R})$ real-valued. Then $|\text{trace } g(\pi_n V \pi_n)| \leq (n + 1) \sup\{|g(x)| \mid |x| \leq \|V\|_\infty\}$.*

PROOF. The operator $\pi_n V \pi_n$ has, as a finite rank operator, a real symmetric matrix $((V \psi_k, \psi_l))_{k,l=0,\dots,n}$ which is diagonalizable as $\text{diag}(a_0, \dots, a_n)$ where $|a_k| \leq \|V\|_\infty$. Now $\text{trace } g(\pi_n V \pi_n) = \text{trace } \text{diag}(g(a_0), \dots, g(a_n))$, and the lemma follows.

The following lemma is important, for it expresses that the trace norm of $\pi_n V \pi_n$ can be estimated by an expression related to $\rho(V)$.

LEMMA II.4.6. *There is a constant $a > 0$ with the following property. For any $V, V_0 \in L^\infty(\mathbf{R}), k = 0, 1, \dots$ we have*

$$\|(\pi_n V \pi_n)^k - (\pi_n V_0 \pi_n)^k\|_1 = O\left(n \sup_{T \geq a\sqrt{n}} \frac{1}{2T} \int_{-T}^T |V(x) - V_0(x)| dx\right).$$

PROOF. For any two bounded operators A, B of $L^2(\mathbf{R})$ we have that $A^k - B^k = ((A - B) + B)^k - B^k =$ sum in which each term has at least one factor $A - B$. Hence $\|(\pi_n V \pi_n)^k - (\pi_n V_0 \pi_n)^k\|_1 \leq C \|\pi_n(V - V_0)\pi_n\|_1$ where C is independent of n . Here we use that π_n, V, V_0 are bounded in operator norm. We observe further that we can write $\pi_n(V - V_0)\pi_n = \pi_n(V - V_0)_+ \pi_n - \pi_n(V - V_0)_- \pi_n$, where $(V - V_0)_+ = \max(V - V_0, 0), (V - V_0)_- = \max(V_0 - V, 0)$. The operators $\pi_n(V - V_0)_\pm \pi_n$ are positive, whence we have that $\|\pi_n(V - V_0)_\pm \pi_n\|_1 = \text{trace}(\pi_n(V - V_0)_\pm \pi_n)$. And

$$\begin{aligned} \|\pi_n(V - V_0)\pi_n\|_1 &\leq \|\pi_n(V - V_0)_+ \pi_n\|_1 + \|\pi_n(V - V_0)_- \pi_n\|_1 \\ &= \int_{-\infty}^{\infty} (V - V_0)_+(x) \sum_{k=0}^n \psi_k^2(x) dx + \int_{-\infty}^{\infty} (V - V_0)_-(x) \sum_{k=0}^n \psi_k^2(x) dx \\ &= \int_{-\infty}^{\infty} |V(x) - V_0(x)| \sum_{k=0}^n \psi_k^2(x) dx. \end{aligned}$$

For $n = 1, 2, \dots$ let $t_n = 2^{-1/n}$. We have by Mehler's formula

$$\begin{aligned} \|\pi_n(V - V_0)\pi_n\|_1 &\leq 2 \int_{-\infty}^{\infty} |V(x) - V_0(x)| \sum_{k=0}^{\infty} t_n^k \psi_k^2(x) dx \\ &= 2^{3/2} \int_{-\infty}^{\infty} |V(x) - V_0(x)| (1 - t_n^2)^{-1/2} \exp\left(-2\pi x^2 \frac{1 - t_n}{1 + t_n}\right) dx. \end{aligned}$$

If $\epsilon > 0$, we have

$$\sqrt{\epsilon} \exp(-\pi \epsilon x^2) \leq \sum_{l=1}^{\infty} \sqrt{\epsilon} 2^{-l+1} \chi_{[-a_l, a_l]},$$

where $a_l = (l \log 2 / \pi \epsilon)^{1/2}$ for $l = 1, 2, \dots$. Hence if $C = 4(\pi^{-1} \log 2)^{1/2}$ we have

$$\sqrt{\epsilon} \exp(-\pi \epsilon x^2) \leq C \sum_{l=1}^{\infty} 2^{-l/2} \frac{1}{2a_l} \chi_{[-a_l, a_l]}.$$

We get

$$\sqrt{\epsilon} \int_{-\infty}^{\infty} |V(x) - V_0(x)| \exp(-\pi \epsilon x^2) dx \leq C \sum_{l=1}^{\infty} 2^{-l/2} \frac{1}{2a_l} \int_{-a_l}^{a_l} |V(x) - V_0(x)| dx.$$

Observing that $a_l \geq (\log 2 / \pi)^{1/2} \epsilon^{-1/2}$ for $l = 1, 2, \dots$ and that $1 - t_n = (1/n) \log 2 + O(1/n^2)$ we easily get the result by taking $\epsilon = 2(1 - t_n)(1 + t_n)^{-1}$.

COROLLARY. *There are $a, b > 0$ such that for $V \in L^\infty(\mathbf{R})$*

$$\frac{1}{n} \|\pi_n V \pi_n\|_1 \leq b \sup_{T \geq a\sqrt{n}} \frac{1}{2T} \int_{-T}^T |V(x)| dx.$$

With all these lemmas available now we can prove the main results of this section. We give one final definition.

DEFINITION II.4.3. Let \mathfrak{W}_1 be the set of all real-valued $V \in \mathfrak{W}$ for which $\lim_{n \rightarrow \infty} (\text{trace } f(\pi_n V \pi_n)) / (n + 1)$ exists and equals $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T f(V(x)) dx$ if $f: \mathbf{R} \rightarrow \mathbf{C}$ is continuous.

THEOREM II.2a. Let $A > 0$, and let $g: \mathbf{R} \rightarrow \mathbf{R}$ be of bounded variation on $[-A, A]$. Assume that $V(x) := \int_{-A}^A e^{-2\pi i \lambda x} dg(\lambda) \in \mathbf{R}$ for $x \in \mathbf{R}$. Then $V \in \mathfrak{W}_1$.

PROOF. We know that $V \in \mathfrak{W}$. Let $k = 0, 1, \dots$. By the note following Proposition II.4.1 we know that

$$|\text{trace}(\pi_n V \pi_n)^k - \text{trace}(\pi_n V^k \pi_n)| = O(n^{3/4}).$$

By Theorem II.1 we know that $\lim_{n \rightarrow \infty} (\text{trace}(\pi_n V^k \pi_n)) / (n + 1) = \lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T V^k(x) dx$. Hence $\lim_{n \rightarrow \infty} (\text{trace}(\pi_n V \pi_n)^k) / (n + 1)$ exists and has the required value. If $f: \mathbf{R} \rightarrow \mathbf{C}$ is continuous, then we can approximate f uniformly on the set $\{x \mid |x| \leq \|V\|_\infty\}$ by polynomials. Now Lemma II.4.5 gives the result.

Note. All trigonometric polynomials are in \mathfrak{W}_1 .

THEOREM II.2b. Let $V \in L^\infty(\mathbf{R})$ be real-valued, and let (V_l) be a sequence in \mathfrak{W}_1 with $\|V_l\|_\infty \leq M$, $\rho(V - V_l) \rightarrow 0$. Then $V \in \mathfrak{W}_1$.

PROOF. By Lemma II.4.2 we know that $V \in \mathfrak{W}$. As in the proof of Theorem II.2a it is enough to show that

$$\lim_{n \rightarrow \infty} \frac{1}{n + 1} \text{trace}(\pi_n V \pi_n)^k = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V^k(x) dx \quad \text{for } k = 0, 1, \dots$$

There are $M_k, a > 0$ such that

$$\begin{aligned} \frac{1}{n + 1} |\text{trace}(\pi_n V \pi_n)^k - \text{trace}(\pi_n V_l \pi_n)^k| &\leq \frac{1}{n + 1} \|(\pi_n V \pi_n)^k - (\pi_n V_l \pi_n)^k\|_1 \\ &\leq M_k \sup_{T \geq a\sqrt{n}} \frac{1}{2T} \int_{-T}^T |V(x) - V_l(x)| dx \end{aligned}$$

by Lemma II.4.6 (the proof of that lemma shows that we can take M_k independent of l). Take $\epsilon > 0$, and let l be such that $\rho(V - V_l) < \epsilon/2M_k$. Next let N be so large that

$$\frac{1}{2T} \int_{-T}^T |V(x) - V_l(x)| dx < 2\rho(V - V_l) < \epsilon/M_k \quad (T \geq a\sqrt{N}).$$

We then have that $|((\text{trace}(\pi_n V \pi_n)^k) / (n + 1) - \text{trace}(\pi_n V_l \pi_n)^k)| < \epsilon$ if $n \geq N$. Since $V_l \in \mathfrak{W}_1$ we can take $N_l \geq N$ so big that

$$\left| \frac{1}{n + 1} \text{trace}(\pi_n V_l \pi_n)^k - \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V_l^k(x) dx \right| < \epsilon$$

if $n \geq N_1$. Then for $n \geq N_1$

$$\begin{aligned} & \left| \frac{1}{n+1} \text{trace}(\pi_n V \pi_n)^k - \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V^k(x) dx \right| \\ & \leq \left| \frac{1}{n+1} \text{trace}(\pi_n V \pi_n)^k - \frac{1}{n+1} \text{trace}(\pi_n V_l \pi_n)^k \right| \\ & \quad + \left| \frac{1}{n+1} \text{trace}(\pi_n V_l \pi_n)^k - \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V_l^k(x) dx \right| \\ & \quad + \left| \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (V_l^k(x) - V^k(x)) dx \right| \\ & \leq 2\varepsilon + k \max(M^k, \|V\|_\infty^k) \varepsilon / 2M_k. \end{aligned}$$

The proof is complete.

Consequences. Almost periodic functions are elements of \mathfrak{W}_1 . Here we mean by an almost periodic function a real-valued function $V \in L^\infty(\mathbf{R})$ for which there exists a bounded sequence $(V_n)_n$ of trigonometric polynomials such that $\rho(V - V_n) \rightarrow 0$ ($n \rightarrow \infty$). Also, any $V \in L^\infty(\mathbf{R})$, for which there exists a g which is locally of bounded variation such that $\rho(V_A - V) \rightarrow 0$ ($A \rightarrow \infty$) where $V_A(x) := \int_{-A}^A e^{-2i\lambda x} dg(\lambda)$ is uniformly bounded, is an element of \mathfrak{W}_1 .

THEOREM II.2c. Let $V \in L^\infty(\mathbf{R})$, and assume that $\lim_{x \rightarrow \pm\infty} V(x) = \pm 1$. Then $V \in \mathfrak{W}_1$.

PROOF. Let $A > 0$, and define $V_A(x) := (1/\pi) \int_{-A}^A ((\sin tx)/t) dt$. We claim that $V_A \in \mathfrak{W}_1$. Unfortunately we cannot apply one of our theorems immediately; instead we have to check certain steps in our proofs. We observe that $\lim_{T \rightarrow \infty} (1/2T) \int_{-T}^T V_A^k(x) dx$ exists for all k , since $V_A(x) \rightarrow \text{sgn}(x)$ if $|x| \rightarrow \infty$. In accordance with the general theme we shall show that $\|\pi_n V_A(I - \pi_n)\|_2 = o(n^{1/2})$. Therefore we consider for $k = 0, 1, \dots, l = k + 1, \dots$

$$(V_A \psi_k, \psi_l) = \frac{1}{\pi} \int_{-A}^A \left(\frac{\sin tx}{t} \psi_k(x), \psi_l(x) \right) dt.$$

We have for $|t| \leq A$ (recall that $(R_b f)(x) = \exp(-2\pi i b x) f(x)$)

$$\left(\frac{\sin tx}{t} \psi_k(x), \psi_l(x) \right) = \frac{1}{2it} \left[(R_{-t/2\pi} \psi_k, \psi_l) - (R_{t/2\pi} \psi_k, \psi_l) \right].$$

The terms $(R_{\pm t/2\pi} \psi_k, \psi_l)$ can be estimated by using the results of §2. Note that the fact that $(T_a \psi_k, \psi_l)$ has a zero of order $l - k$ at $a = 0$ comes in quite handy here. We easily get that $t^{-1}(\sin tx \psi_k(x), \psi_l(x)) = O(C^{l-k})$ ($k \geq t^2/4\pi^2$, $l \geq k + \varepsilon\sqrt{k}$) for some $C \in (0, 1)$, $\varepsilon > 0$; here 0 holds uniformly in $t \in [-A, A]$. We conclude as before that $\|\pi_n V_A(I - \pi_n)\|_2 = O(n^{1/4})$. Hence $V_A \in \mathfrak{W}_1$.

Now $V_A(x) = (1/\pi) \int_{-A|x|}^{A|x|} ((\sin t)/t) dt \text{sgn}(x) \rightarrow \text{sgn}(x)$ uniformly in any set not containing 0 if $A \rightarrow \infty$, and the V_A 's are uniformly bounded. Hence $\text{sgn} \in \mathfrak{W}_1$, and obviously $\rho(V - \text{sgn}) = 0$, so that by Theorem II.2b we see that $V \in \mathfrak{W}_1$, and the proof is complete.

We conclude Part II with some remarks. Our explicit formulas enable us to investigate in certain cases how fast $(\text{trace } f(\pi_n V \pi_n))/(n + 1)$ converges to its limit. Since the exponentials played a special role in this part we first assume that $V \equiv V_\lambda(x) = \exp(-2\pi i \lambda x)$, and that $f(x) = x$. We have the formula

$$\begin{aligned} \text{trace}(\pi_n V_\lambda \pi_n) &= \sum_{k=0}^n (V_\lambda \psi_k, \psi_k) = \frac{1}{2} \sum_{k=0}^n (-1)^k W\left(0, \frac{1}{2}\lambda; \psi_k, \psi_k\right) \\ &= \sum_{k=0}^n e^{-\pi \lambda^2 / 2} L_k(\pi \lambda^2) = e^{-\pi \lambda^2 / 2} L_n^{(1)}(\pi \lambda^2) \end{aligned}$$

(here $L_n^{(1)}$ is the n th Laguerre polynomial of first order; cf. §2(g) of Part II). We have used here [Sz, (5.1.13)]. Now one has very precise knowledge of the behavior of the functions $\exp(-x/2)L_n^{(1)}(x)$ (cf. e.g. [Sz, Chapter 8 or AW, table on p. 699]). As a consequence one can show that

$$\text{trace}(\pi_n V_\lambda \pi_n) = \begin{cases} n + 1, & \lambda = 0, \\ O(n), & 0 \leq \pi \lambda^2 \leq 1/n, \\ O(n^{1/4}(\pi \lambda^2)^{-3/4}), & 1/n \leq \pi \lambda^2 \leq \delta n, \\ O(n^{1/6}/\lambda), & \pi \lambda^2 \geq \delta n, \\ O(\exp(-\xi \lambda^2)), & \pi \lambda^2 \geq 4(1 + \eta)n. \end{cases}$$

Here $\xi > 0$, and δ, η are positive and sufficiently small.

We have e.g.: if $\lambda \neq 0$, then $(\text{trace}(\pi_n V_\lambda \pi_n))/(n + 1) = O(n^{-3/4})$. One might have expected here $O(n^{-1})$ since $(1/2T) \int_{-T}^T e^{-2\pi i \lambda x} dx = O(1/T)$.

For a more general V we must know how well V can be approximated by linear combinations of exponentials; here Lemma II.4.6 is relevant. And if e.g. V is of the form $\int_{-\infty}^{\infty} e^{-2\pi i \lambda x} dg(\lambda)$, then

$$\text{trace}(\pi_n V \pi_n) = \int_{-\infty}^{\infty} \exp(-\pi \lambda^2 / 2) L_n^{(1)}(\pi \lambda^2) dg(\lambda),$$

and we can read off from the estimates just given how fast $1(\text{trace}(\pi_n V \pi_n))/(n + 1)$ converges to $(dg)\{0\}$.

If V is a trigonometric polynomial again, and f is a polynomial, then II.4.1 together with the remarks made above can tell us how fast $(\text{trace } f(\pi_n V \pi_n))/(n + 1)$ approaches its limit if $n \rightarrow \infty$. In the general case we must, of course, know how well f and V can be approximated by polynomials and trigonometric polynomials respectively.

We finally indicate how the results of this section can be generalized to operators which are not multiplication operators. For any complex measure μ defined on \mathbf{R}^2 with $|\mu|(\mathbf{R}^2) < \infty$ we let A_μ be the operator whose Weyl symbol a is given by

$$a(x, \xi) = \int \exp(2\pi i(q\xi - px)) d\mu(q, p)$$

(cf. §1 of Part II; it is easy to check that A_μ is a bounded operator of $L^2(\mathbf{R})$), i.e. d_μ is the symplectic Fourier transform of a . For this A_μ the trace theorem holds, for we have (cf. the formulas in §1)

$$\begin{aligned} \frac{1}{n+1} \text{trace}(\pi_n A_\mu \pi_n) &= \frac{1}{n+1} \sum_{k=0}^n (A_\mu \psi_k, \psi_k) \\ &= \frac{1}{n+1} \int \sum_{k=0}^n (W_{q,p} \psi_k, \psi_k) d\mu(q, p) \\ &= \frac{1}{n+1} \int \exp(-\pi(q^2 + p^2)/2) L_n^{(1)}(\pi(q^2 + p^2)) d\mu(q, p), \end{aligned}$$

and this tends to $(d\mu)\{(0,0)\} = \lim_{T \rightarrow \infty} (1/\pi T) \int_{x^2 + \xi^2 \leq T} a(x, \xi) dx d\xi$ (for the latter relation one uses Fubini's theorem and dominated convergence).

It can be shown that for A_μ the Szegő theorem holds as well if μ decays sufficiently fast at ∞ , provided that A_μ is selfadjoint. To sketch the proof, let μ_1 and μ_2 be as μ in the previous paragraph. It is a well-known fact of the Weyl calculus that $A_{\mu_1} A_{\mu_2} = A_{\mu_3}$, where μ_3 is the twisted convolution product of μ_1 and μ_2 , given by

$$\int f d\mu_3 = \int \left\{ \int \exp(-\pi i p' q + \pi i p q') f(q', p') d\mu_2(q' - q, p' - p) \right\} d\mu_1(q, p)$$

for all continuous $f: \mathbf{R}^2 \rightarrow \mathbf{C}$ of compact support. Observe that μ_2 is a complex measure with $|\mu_3|(\mathbf{R}^2) < \infty$. We conclude that $\lim_{n \rightarrow \infty} (\text{trace}(\pi_n A_\mu^k \pi_n)) / (n+1)$ exists and has the right value for $k = 0, 1, \dots$. We finally need an estimate on $\|\pi_n A_\mu (I - \pi_n)\|_2$ ($o(n^{1/2})$ suffices). Therefore one proceeds as in Lemma II.4.4 and estimates the sum $\sum_{k=0}^n \sum_{l=n}^{\infty} |(A_\mu \psi_k, \psi_l)|^2$. Since

$$(A_\mu \psi_k, \psi_l) = \frac{1}{2} (-1)^l \int W\left(\frac{1}{2} p, \frac{1}{2} q; \psi_k, \psi_l\right) d\mu(q, p)$$

one can use the estimates on the Wigner functions of §2. It is not hard to formulate sufficient conditions on μ so that $\|\pi_n A_\mu (I - \pi_n)\|_2 = o(n^{1/2})$ (e.g. μ has compact support suffices).

REFERENCES

- [AW] R. Askey and S. Wainger, *Mean convergence of expansions in Laguerre and Hermite series*, Amer. J. Math. **87** (1965), 695–708.
- [Ba] V. Bargmann, *On a Hilbert space of analytic functions and an associated integral transform*. Part I, Comm. Pure Appl. Math. **14** (1961), 187–214.
- [Be] R. Beals, *A general calculus of pseudodifferential operators*, Duke Math. J. **42** (1975), 1–42.
- [dB] N. G. de Bruijn, *A theory of generalized functions, with applications to Wigner distribution and Weyl correspondence*, Nieuw Arch. Wisk. **21** (1973), 205–280.
- [D] I. Daubechies, *On the distributions corresponding to bounded operators in the Weyl quantization*, Comm. Math. Phys. **75** (1980), 229–238.
- [Gr] H. J. Groenewold, *On the principle of elementary quantum mechanics*, Physica **21** (1946), 405–460.
- [Gu] V. Guillemin, *Some classical theorems in spectral theory revisited*, Seminar and Singularities of Solutions of Linear Partial Differential Equations (L. Hörmander, editor), Princeton Univ. Press, Princeton, N. J., 1979.
- [GLS] A. Grossman, G. Loupias and E. M. Stein, *An algebra of pseudo-differential operators and quantum mechanics in phase space*, Ann. Inst. Fourier (Grenoble) **18** (1969), 363–368.
- [H] R. Howe, *Quantum mechanics and partial differential equations*, J. Funct. Anal. **38** (1980), 188–254.
- [Hö] L. Hörmander, *The Weyl calculus of pseudo-differential operators*, Comm. Pure Appl. Math. **32** (1979), 359–443.

- [Ja] A. J. E. M. Janssen, *Positivity of weighted Wigner distributions*, SIAM J. Math. Anal. **12** (1981), 752–758.
- [Jb] ———, *Bargmann transform, Zak transform and coherent states*, J. Math. Phys. **23** (1982), 720–731.
- [Jc] ———, *Application of the Wigner distribution to harmonic analysis of generalized stochastic processes*, MC-tract 114, Amsterdam, 1979.
- [M] G. Mauceri, *The Weyl transform and bounded operators on $L^p(\mathbf{R}^n)$* , J. Funct. Anal. **39** (1980), 408–429.
- [P] J. Peetre, *The Weyl transform and Laguerre polynomials*, Matematiche (Catania) **27** (1972), 301–323.
- [Ru] W. Rudin, *Functional analysis*, McGraw-Hill, New York, 1973.
- [Si] B. Simon, *Functional integration and quantum physics*, Academic Press, New York, 1979.
- [Sz] G. Szegö, *Orthogonal polynomials*, 4th ed., Amer. Math. Soc. Colloq. Publ., vol. 23, Amer. Math. Soc., Providence, R. I., 1975.
- [T] M. Taylor, *Pseudodifferential operators*, Princeton Univ. Press, Princeton, N. J., 1981.
- [U] A. Unterberger, *Oscillateur harmonique et operateurs pseudo-differentiels*, Ann. Inst. Fourier (Grenoble) **29** (1979), 201–221.
- [V] A. Voros, *An algebra of pseudodifferential operators and the asymptotics of quantum mechanics*, J. Funct. Anal. **29** (1978), 104–132.
- [We] H. Weyl, *The theory of groups and quantum mechanics*, Dover, New York, 1950.
- [Wi] H. Widom, *Eigenvalue distribution theorems for certain homogeneous spaces*, J. Funct. Anal. **32** (1979), 139–147.

DEPARTMENT OF MATHEMATICS, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA 91125

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY, NEW YORK, NEW YORK 10027 (Current address of Steven Zelditch)

Current address (A. J. E. M. Janssen): Philips Research Laboratories, 5600 JA Eindhoven, The Netherlands