FREE PROBABILITY AND RANDOM MATRICES

LECTURE 4: FREE HARMONIC ANALYSIS, OCTOBER 4, 2007

The Cauchy Transform. Let $\mathbb{C}^+ = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$ denote the complex upper half plane, and $\mathbb{C}^- = \{z \mid \mathrm{Im}(z) < 0\}$ denote the lower half plane. Let ν be a probability measure on $\mathbb{R}$ and for $z \notin \mathbb{R}$ let

$$G(z) = \int_{\mathbb{R}} \frac{1}{z-t}\, d\nu(t)$$

G is the *Cauchy transform* of the measure ν. Let us briefly check that the integral converges to an analytic function on $\mathbb{C}^+$.

Lemma. G is an analytic function on $\mathbb{C}^+$ with range contained in $\mathbb{C}^-$.

Since $|z-t|^{-1} \leq |\mathrm{Im}(z)|^{-1}$ and ν is a probability measure the integral is always convergent. If $\mathrm{Im}(w) \neq 0$ and $|z-w| < |\mathrm{Im}(w)|/2$ then for $t \in \mathbb{R}$ we have

$$\left|\frac{z-w}{t-w}\right| < \frac{|\mathrm{Im}(w)|}{2} \cdot \frac{1}{|\mathrm{Im}(w)|} = \frac{1}{2}$$

so the series $\sum_{n=0}^{\infty}(\frac{z-w}{t-w})^n$ converges uniformly to $\frac{t-w}{t-z}$ on $|z-w| < |\mathrm{Im}(w)|/2$. Thus $(z-t)^{-1} = -\sum_{n=0}^{\infty}(t-w)^{-(n+1)}(z-w)^n$ on $|z-w| < |\mathrm{Im}(w)|/2$. Hence

$$G(z) = -\sum_{n=0}^{\infty}\left[\int_{\mathbb{R}} (t-w)^{-(n+1)} d\nu(t)\right](z-w)^n$$

is analytic on $|z-w| < |\mathrm{Im}(w)|/2$.

Finally note that for $\mathrm{Im}(z) > 0$, we have for $t \in \mathbb{R}$, $\mathrm{Im}((z-t)^{-1}) < 0$, and hence $\mathrm{Im}(G(z)) < 0$. Thus G maps $\mathbb{C}^+$ into $\mathbb{C}^-$. □

Lemma. (i) $\lim\limits_{y\to\infty} y\, G(iy) = -i$ and (ii) $\sup\limits_{y\geq 0, x\in\mathbb{R}} y\, |G(x+iy)| = 1$

Proof. (i)

$$\begin{aligned} y\,\mathrm{Im}(G(iy)) &= \int_{\mathbb{R}} y\,\mathrm{Im}\Big(\frac{1}{iy-t}\Big)\, d\nu(t) = \int_{\mathbb{R}} \frac{-y^2}{y^2+t^2}\, d\nu(t) \\ &= -\int_{\mathbb{R}} \frac{1}{1+(t/y)^2}\, d\nu(t) \to -\int_{\mathbb{R}} d\nu(t) \end{aligned}$$

where $y \to \infty$ and since $(1+(t/y)^2)^{-1} \leq 1$ we may apply Lebesgue's Dominated Convergence Theorem.

We have $y\,\mathrm{Re}(G(iy)) = \displaystyle\int_{\mathbb{R}} \frac{-yt}{y^2+t^2}\,d\nu(t)$. But for all $y > 0$ and for all t

$$\left|\frac{yt}{y^2+t^2}\right| \leq \frac{1}{2}$$

and $|yt/(y^2+t^2)|$ converges to 0 as $y \to \infty$. Therefore $y\,\mathrm{Re}(G(iy) \to 0$ as $y \to \infty$, again by the Dominated Convergence Theorem.

(ii) For $y > 0$ and $z = x + iy$,

$$y\,|G(z)| \leq \int_{\mathbb{R}} \frac{y}{|z-t|}\,d\nu(t) = \int_{\mathbb{R}} \frac{y}{\sqrt{(x-t)^2+y^2}}\,d\nu(t) \leq 1$$

Thus $\sup_{y\geq 0, x\in\mathbb{R}} y\,|G(x+iy)| \leq 1$. By (i) however, the supremum is 1. □

Theorem. *Suppose $a < b$. Then*

$$\lim_{y\to 0^+} \frac{-1}{\pi}\int_a^b \mathrm{Im}(G(x+iy))\,dx = \nu((a,b)) + \frac{1}{2}\nu(\{a,b\})$$

If ν_1 and ν_2 are probability measures with $G_{\nu_1} = G_{\nu_2}$, then $\nu_1 = \nu_2$.

Proof. We have

$$\mathrm{Im}(G(x+iy)) = \int_{\mathbb{R}} \mathrm{Im}\Big(\frac{1}{x-t+iy}\Big)\,d\nu(t) = \int_{\mathbb{R}} \frac{-y}{(x-t)^2+y^2}d\nu(t)$$

Thus

$$\begin{aligned}\int_a^b \mathrm{Im}(G(x+iy))\,dx &= \int_{\mathbb{R}}\int_a^b \frac{-y}{(x-t)^2+y^2}\,dx\,d\nu(t)\\ &= -\int_{\mathbb{R}}\int_{(a-t)/y}^{(b-t)/y} \frac{1}{1+\tilde{x}^2}\,d\tilde{x}\,d\nu(t)\\ &= -\int_{\mathbb{R}} \tan^{-1}\Big(\frac{b-t}{y}\Big) - \tan^{-1}\Big(\frac{a-t}{y}\Big)\,d\nu(t)\end{aligned}$$

where we have let $\tilde{x} = (x-t)/y$.

So let $f(y,t) = \tan^{-1}((b-t)/y) - \tan^{-1}((a-t)/y)$ and

$$f(t) = \begin{cases} 0 & t \notin [a,b] \\ \pi/2 & t \in \{a,b\} \\ \pi & t \in (a,b) \end{cases}$$

Then $\lim_{y\to 0^+} f(y,t) = f(t)$, and, for all $y > 0$ and for all t, we have $|f(y,t)| \leq \pi$. So by Lebesgue's Dominated Convergence Theorem

$$\begin{aligned}\lim_{y\to 0^+}\int_a^b \mathrm{Im}(G(x+iy))\,dx &= -\lim_{y\to 0^+}\int_{\mathbb{R}} f(y,t)\,d\nu(t)\\ &= -\int_{\mathbb{R}} f(t)\,d\nu(t) = -\pi\Big(\nu((a,b)) + \frac{1}{2}\nu(\{a,b\})\Big)\end{aligned}$$

This proves the first claim. The first claim shows that if $G_{\nu_1} = G_{\nu_2}$ then ν_1 and nu_2 agree on all intervals and thus are equal. □

Remark. The proof of the next theorem depends on a fundamental result of R. Nevalinna which provides an integral representation for an analytic function from $\mathbb{C}^+$ to $\mathbb{C}^+$. Suppose that $\phi : \mathbb{C}^+ \to \mathbb{C}^+$ is analytic, then there is a unique Borel measure σ on $\mathbb{R}$ and real numbers α and β, with $\beta \geq 0$ such that for $z \in \mathbb{C}^+$

$$\phi(z) = \alpha + \beta z + \int_{\mathbb{R}} \frac{1}{t-z}\,d\sigma(t)$$

This integral representation is achieved by letting $\xi = \frac{iz+1}{iz-1}$, and thus $z = i\frac{1+\xi}{1-\xi}$. Then one defines ψ by $\psi(\xi) = i\phi(z)$ and obtains an analytic function ψ mapping the open unit disc, $\mathbb{D}$, into the complex right half plane. Since ψ the real and imaginary parts of ψ are harmonic conjugates, it is enough to work with the real part. Now $\mathrm{Re}(\psi)$ is a positive harmonic function on $\mathbb{D}$; when $\mathrm{Re}(\psi)$ is harmonic on an open set containing $\overline{\mathbb{D}}$ one gets the measure from Cauchy's Integral Theorem; in general one dilates $\mathrm{Re}(\psi)$ to get a harmonic function on $\overline{\mathbb{D}}$, and then takes a limit as the dilation shrinks to $\mathrm{Re}(\psi)$. This integral representation is usually attributed to G. Herglotz (1911). The details can be found in the book of Akhiezer and Glazman[1].

Theorem (R. Nevanlinna)**.** *Suppose* $G : \mathbb{C}^+ \to \mathbb{C}^-$ *is analytic and* $\sup y\,|G(x+iy)| = 1$ *. Then there is a unique probability measure on* $\mathbb{R}$ *such that* $G(z) = \int_{\mathbb{R}} \frac{1}{z-t}\,d\nu(t)$.

Proof. By the remark above there is a unique finite measure on $\mathbb{R}$ such that $G(z) = \alpha + \beta z + \int \frac{1+tz}{z-t}\,d\sigma(t)$ with $\alpha \in \mathbb{R}$ and $\beta \leq 0$.

Next we will use the condition

$$\sup_{y>0,\,x\in\mathbb{R}} y\,|G(x+iy)| = 1 \tag{1}$$

Considering first the imaginary part we get that $\beta = 0$ and for all N, $\int_{-N}^{N} \frac{y^2}{t^2+y^2}(1+t^2)\,d\sigma(t) \leq 1$. Thus $\int_{\mathbb{R}}(1+t^2)\,d\sigma(t) \leq 1$. Let

[1]N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space* (1961), vol. 2, Ch. VI, §58

$\nu(E) = \int_E 1 + t^2 d\sigma(t)$. From the real part of (1) we get that for all $y > 0$

$$y \left| \alpha + \int_{\mathbb{R}} \frac{t(y^2-1)}{t^2+y^2} d\nu(t) \right| \leq 1$$

Therefore

$$\begin{aligned} \alpha &= \lim_{y\to\infty} \int_{\mathbb{R}} \frac{t(1-y^2)}{t^2+y^2}\, d\sigma(t) \\ &= -\lim_{y\to\infty} \int_{\mathbb{R}} t \frac{1-y^{-2}}{1+(t/y)^2}\, d\sigma(t) = -\int t d\sigma(t) \end{aligned}$$

Hence

$$\begin{aligned} G(z) &= \int -t + \frac{1+tz}{z-t} d\sigma(t) = \int \frac{-tz+t^2+1+tz}{z-t} d\sigma(t) \\ &= \int_{\mathbb{R}} \frac{1}{z-t}(1+t^2) d\sigma(t) = \int_{\mathbb{R}} \frac{1}{z-t} d\nu(t) \end{aligned}$$

□

Remark. If $\{\nu_n\}_n$ is a sequence of finite Borel measures on $\mathbb{R}$ we say that $\{\nu_n\}_n$ *converges in distribution* to the measure ν if for every $f \in C_b(\mathbb{R})$ (the continuous bounded functions on $\mathbb{R}$) we have $\lim_n \int f(t)\, d\nu_n(t) \to \int f(t)\, d\nu(t)$. We say that $\{\nu_n\}_n$ *converges vaguely* to ν if for every $f \in C_0(\mathbb{R})$ (the continuous functions on $\mathbb{R}$ vanishing at infinity) we have $\lim_n \int f(t)\, d\nu_n(t) \to \int f(t)\, d\nu(t)$. Convergence in distribution implies vague convergence but not conversely. However if ν is a probability measure then $\{\nu_n\}_n$ converges vaguely to ν does imply that $\{\nu_n\}_n$ converges to ν in distribution[2].

Theorem. *Suppose that $(\nu_n)_n$ is a sequence of probability measures on $\mathbb{R}$ with G_n the Cauchy transform of ν_n. Suppose $\{G_n\}_n$ converges pointwise to G on $\mathbb{C}^+$. If $\lim_{y\to\infty} y\, G(iy) = -i$ then there is a unique probability measure ν on $\mathbb{R}$ such that $\nu_n \to \nu$ in distribution, and $G(z) = \int \frac{1}{z-t} d\nu(t)$.*

Proof. $\{G_n\}_n$ is uniformly bounded on compact subsets of $\mathbb{C}^+$ (as $|G(z)| \leq |\mathrm{Im}(z)|^{-1}$ for the Cauchy transform of any probability measure), so by Montel's theorem $\{G_n\}_n$ is relatively compact in the topology of uniform convergence on compact subsets of $\mathbb{C}^+$, thus, in particular, $\{G_n\}_n$ has a subsequence which converges uniformly on compact subsets of $\mathbb{C}^+$ to an analytic function, which must be G. Thus G is analytic, as it is the uniform limit of analytic functions. Now for $z \in \mathbb{C}^+$, $G(z) \in \overline{\mathbb{C}^-}$. Also for each $n \in \mathbb{N}$, $x \in \mathbb{R}$ and $y \geq 0$ $y\,|G_n(x+iy)| \leq 1$.

[2]see Kai Lai Chung *Probability*, 2^{nd} ed. (1984) §3.3 and 3.4.

Thus $\forall x \in \mathbb{R}$, $\forall y \geq 0$, $y\,|G(x+iy)| \leq 1$. So in particular, G is non-constant. If for some $z \in \mathbb{C}^+$, $\mathrm{Im}(G(z)) = 0$ then by the minimum modulus principle G would be constant. Thus G maps $\mathbb{C}^+$ into $\mathbb{C}^-$. Hence by Nevanlinna's theorem there is a unique probability measure ν such that $G(z) = \int_{\mathbb{R}} \frac{1}{z-t} d\nu(t)$.

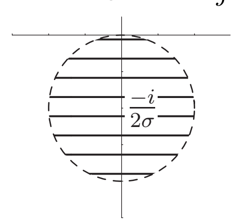
Now by the Helley selection theorem[3] there is a subsequence $\{\nu_{n_k}\}_k$ converging vaguely to some measure $\tilde{\nu}$. For fixed z the function $t \mapsto 1/(z-t)$ is in $C_0(\mathbb{R})$. Thus for $\mathrm{Im}(z) > 0$, $G_{n_k}(z) = \int_{\mathbb{R}} \frac{1}{z-t}\, d\nu_{n_k}(t) \to \int_{\mathbb{R}} \frac{1}{z-t}\, d\tilde{\nu}(t)$. Therefore $G(z) = \int \frac{1}{z-t}\, d\nu(t)$ i.e. $\nu = \tilde{\nu}$. Thus $\{\nu_{n_k}\}_k$ converges vaguely to ν. Since ν is a probability measure $\{\nu_{n_k}\}_k$ converges vaguely to ν. Thus $\{\nu_n\}$ converges in distribution to ν. □

Corollary. If $\{\nu_n\}_n$ is a sequence of probability measures on $\mathbb{R}$ with $G_n \to G$ pointwise on $\mathbb{C}^+$. Then $\{\nu_n\}$ converges vaguely to a finite measure ν.

The analytic relation between moments and free cumulants involves finding a functional inverse for the Cauchy transform of a probability measure, ν. One cannot do this in general without making some assumptions about ν. A fairly general condition was found by Hans Maassen in 1992[4]

Theorem (H. Maassen). *Let ν be a probability measure on $\mathbb{R}$ with $\sigma^2 = \int_{-\mathbb{R}} t^2 d\nu(t) < \infty$ and $\int t\, d\nu(t) = 0$. Then G is a bijection between*

$\{z \mid \mathrm{Im}(z) > \sigma\}$ *and* $\{\frac{1}{z} \mid \mathrm{Im}(z) > 2\sigma\}$



Compactly Supported Measures. When the probability measure ν has compact support then one can find a region in $\mathbb{C}^+$ upon which the Cauchy transform is univalent.

Suppose ν is a probability measure on $\mathbb{R}$ and $\mathrm{supp}(\nu) \subseteq [-r, r]$. Then ν has moments $m_n = \int t^n d\nu(t)$ of all orders and $|m_n| \leq r^n$. Let $M(z) = 1 + m_1 z + m_2 z + \cdots$ be the moment generating function. If

[3]E. Lukacs, *Characteristic Functions*, 2^{nd} ed., Griffin, 1972

[4]Addition of Freely Independent Random Variables, *J. Funct. Anal.*, **106** (1992), 409-438.

$|z| < r^{-1}$ the series converges. If $|z| > r$ then $|z^{-1}| < r^{-1}$ and

$$\begin{aligned} G(z) &= \int_{\mathbb{R}} \frac{1}{z-t}\, d\nu(t) = \frac{1}{z} \int_{\mathbb{R}} \sum_{n\geq 0} z^{-n} t^n d\nu(t) \\ &= \frac{1}{z} \sum_{n\geq 0} z^{-n} \int_{\mathbb{R}} t^n d\nu(t) = \frac{1}{z} \sum_{n\geq 0} z^{-n} m_n = \frac{1}{z} M\Big(\frac{1}{z}\Big) \end{aligned}$$

Let $f(z) = zM(z) = G(\frac{1}{z}) = z + m_1 z^2 + m_2 z^3 + \cdots$. Then $f(0) = 0$ and $f'(0) = 1$. Suppose $|z_1|, |z_2| \leq (4r)^{-1}$ then

$$\begin{aligned} \left|\frac{f(z_1) - f(z_2)}{z_1 - z_2}\right| &\geq \mathrm{Re}\Big(\frac{f(z_1) - f(z_2)}{z_1 - z_2}\Big) \\ &= \mathrm{Re} \int_0^1 \frac{d}{dt} \left[\frac{f(z_1 + t(z_2 - z_1))}{z_2 - z_1}\right] dt \\ &= \int_0^1 \mathrm{Re}(f'(z_1 + t(z_2 - z_1)))\, dt \end{aligned}$$

Now

$$\begin{aligned} \mathrm{Re}(f'(z)) &= \mathrm{Re}(1 + 2zm_1 + 3x^2 m_2 + \cdots) \\ &\geq 1 - 2\,|z|\,r - 3\,|z|^2\,r^2 - \cdots \\ &= 2 - (1 + 2(|z|\,r) + 3(|z|\,r)^2 + \cdots) = 2 - \frac{1}{(1 - |z|\,r)^2} \end{aligned}$$

So for $|z| < (4r)^{-1}$ we have

$$\mathrm{Re}(f'(z)) \geq 2 - \frac{1}{(1 - 1/4)^2} = \frac{2}{9}$$

Hence for $|z_1|, |z_2| < (4r)^{-1}$ we have $|f(z_1) - f(z_2)| \geq \frac{2}{9}|z_1 - z_2|$. In particular f is one-to-one on $\{z \mid |z| < (4r)^{-1}\}$. Also

$$\begin{aligned} |f(z)| &= |z|\,|1 + m_1 z + m_2 z^2 + \cdots| \\ &\geq |z|\,(2 - (1 + r|z| + r^2|z|^2 + \cdots)) = |z|\Big(2 - \frac{1}{1 - r\,|z|}\Big) \\ &\geq |z|\,\Big(2 - \frac{1}{1 - 1/4}\Big) = \frac{2}{3}|z| \end{aligned}$$

Thus for $|z| = (4r)^{-1}$ we have $|f(z)| \geq (6r)^{-1}$. Moreover, since f is one-to-one on $\{z \mid |z| \geq (4r)^{-1}\}$, the curve $\gamma = \{f(e^{it}/(4r)) \mid 0 \leq t \leq 2\pi\}$ only wraps once around the origin.

Thus f maps $\{z \mid |z| < (4r)^{-1}\}$ to the interior of the curve γ. Hence

$$\{z \mid |z| < (6r)^{-1}\} \subseteq \{f(z) \mid |z| < (4r)^{-1}\}$$

We will denote the inverse of f under composition by $f^{\langle -1\rangle}$. Thus $f^{\langle -1\rangle}$ is analytic on $\{z \mid |z| < (6r)^{-1}\}$ and has a simple zero at $z = 0$. Thus

$$K(z) = \frac{1}{f^{\langle -1\rangle}(z)}$$

is meromorphic on $\{z \mid |z| < (6r)^{-1}\}$ and has a simple pole at $z = 0$. Now

$$K(G(z)) = 1/(f^{\langle -1\rangle}(G(z))) = 1/(f^{\langle -1\rangle}(f(z^{-1}))) = z$$

and $G(K(z)) = f(f^{\langle -1\rangle}(z)) = z$.

Theorem. *Suppose* $\mathrm{supp}(\nu) \subset [-r, r]$, *then* G_ν *is univalent on the open disc with centre 0 and radius* $(6r)^{-1}$

Let $R(z) = K(z) - \frac{1}{z}$. Then R is analytic on $\{z \mid |z| < (6r)^{-1}\}$, and

$$G\Big(\frac{1}{z} + R(z)\Big) = z = R(G(z)) + \frac{1}{G(z)}$$

which is the same relation as was found combinatorially.

Bercovici and Voiculescu found a theorem that relates convergence in distribution of a sequence of probability measures with support in $[-r, r]$ to the uniform convergence of the corresponding R-transforms on a disc.

Theorem (Bercovici & Voiculescu)**.** *Let* $\{\nu_n\}_n$ *be a sequence of probability measures on* $\mathbb{R}$ *with compact support, let* R_n *be the corresponding sequence of R-transforms. Then*
i) $\exists\, r$ *s.t.* $\mathrm{supp}(\nu_n) \subseteq [-r, r]$ *and* ν *with* $\mathrm{supp}(\nu) \subseteq [-r, r]$ *such that* $\nu_n \xrightarrow{\mathcal{D}} \nu$

if and only if

ii) $\exists\, s > 0$ *such that* $\{R_n\}_n$ *converges uniformly on* $\{z \mid |z| < s\}$ *to an analytic function* R *on* $\{z \mid |z| < s\}$.

Example: The Semi-circle Law. As an example of Stieltjes inversion let us take a familiar example and calculate its Cauchy transform and then using only the Cauchy transform find the density by using Stieltjes inversion. The density of the semi-circle law is given by

$$\frac{d\nu}{dt} = \frac{\sqrt{4-t^2}}{2\pi} s\, dt \text{ on } [-2, 2]$$

the moments are given by

$$m_n = \int_{-2}^{2} t^n d\nu(t) = \begin{cases} 0 & n \text{ odd} \\ c_{n/2} & n \text{ even} \end{cases}$$

where the c_n's are the Catalan numbers:

$$c_n = \frac{1}{n+1}\binom{2n}{n}$$

Now let $M(z)$ be the moment generating function

$$M(z) = 1 + c_1 z^2 + c_2 z^4 + \cdots$$

then

$$M(z)^2 = \sum_{m,n\geq 0} c_m c_n z^{2(m+n)} = \sum_{k\geq 0}(\sum_{m+n=k} c_m c_n) z^{2k}$$

Now we saw in lecture three that

$$\sum_{m+n=k} c_m c_n = c_{k+1}$$

so

$$M(z)^2 = \sum_{k\geq 0} c_{k+1} z^{2k} = \frac{1}{z^2}\sum_{k\geq 0} c_{k+1} z^{2(k+1)}$$

therefore

$$z^2 M(z)^2 = M(z) - 1 \text{ or } M(z) = 1 + z^2 M(z)^2$$

Therefore $z^2M(z)^2 - M(z) + 1 = 0$. Solving the quadratic equation we have

$$M(z) = \frac{1 \pm \sqrt{1-4z^2}}{2z^2} \text{ and } M(0) = 1$$

Now

$$\frac{1+\sqrt{1-4t^2}}{2t^2} \to \infty \text{ as } t \to 0,$$

whereas

$$\frac{1-\sqrt{1-4t^2}}{2t^2} = \frac{2}{1+\sqrt{1-4t^2}} \to 1 \text{ as } t \to 0$$

Therefore we choose the minus sign

$$M(z) = \frac{1-\sqrt{1-4z^2}}{2z^2}$$

Now

$$G(z) = \frac{1}{z} M\Big(\frac{1}{z}\Big) = \frac{1}{z}\cdot\frac{1-\sqrt{1-4z^{-2}}}{2z^{-2}} = \frac{z-\sqrt{z^2-4}}{2}$$

Also $1 + z^2M(z)^2 = M(z)$ implies

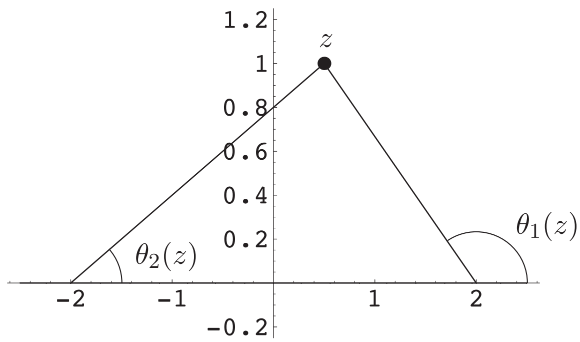
$$1 + \Big(\frac{1}{z}M\Big(\frac{1}{z}\Big)\Big)^2 = z\Big(\frac{1}{z}M(z)\Big)$$

i.e.

$$1 + G(z)^2 = zG(z) \text{ or } z = G(z) + \frac{1}{G(z)}$$

Thus $K(z) = z + \frac{1}{z}$ and $R(z) = z$ i.e. all so all cumulants of the semi-circle law are 0 except κ_2, which equals 1.

Now let us apply Stieltjes inversion to $G(z)$. To do this we have to choose a branch of $\sqrt{z^2-4}$. To achieve this we choose a branch for each of $\sqrt{z-2}$ and $\sqrt{z+2}$. Now $\sqrt{z-2} = |z-2|^{1/2}e^{i\theta/2}$ where θ is the argument of $z-2$. Choosing a branch means connecting 2 to ∞ with a curve along which θ will have a jump discontinuity.



We choose θ_1, the argument of $z-2$, to be such that $0 \le \theta_1 < 2\pi$. Similarly we choose θ_2, the argument of $z+2$, such that $0 \le \theta_2 < 2\pi$. Thus $\theta_1 + \theta_2$ is continuous on $\mathbb{C} \setminus [-2, \infty)$. However $e^{i(\theta_1+\theta_2)/2}$ is continuous on $\mathbb{C} \setminus [-2, 2]$ because $e^{i(0+0)/2} = 1 = e^{i(2\pi+2\pi)/2}$, so there is no jump as the half line $[2, \infty)$ is crossed.

Now

$$\sqrt{z^2-4} = \sqrt{z-2}\sqrt{z+2} = \left|z^2-4\right|^{1/2} e^{i\theta_1/2}e^{i\theta_2/2} = \left|z^2-4\right|^{1/2} e^{i(\theta_1+\theta_2)/2}$$

$$\mathrm{Im}(\sqrt{(x+iy)^2-4}) = \left|(x+iy)^2-4\right|^{1/2} \sin((\theta_1+\theta_2)/2)$$

$$\begin{aligned}\lim_{y\to 0} \mathrm{Im}(\sqrt{(x+iy)^2-4}) &= |x^2-4|^{1/2} \begin{cases} 0 & |x|>2 \\ 1 & |x|<2 \end{cases} \\ &= \begin{cases} 0 & |x|>0 \\ \sqrt{4-x^2} & |x|<2 \end{cases}\end{aligned}$$

$$\begin{aligned}\lim_{y\to 0^+} \mathrm{Im}(G(x+iy)) &= \lim_{y\to 0^+} \mathrm{Im}\Big(\frac{x+iy-\sqrt{(x+iy)^2-4}}{2}\Big) \\ &= \begin{cases} 0 & |x|>2 \\ \frac{-\sqrt{4-x^2}}{2} & |x|<2 \end{cases}\end{aligned}$$

Therefore

$$\lim_{y\to 0^+} \frac{-1}{\pi}\mathrm{Im}(G(x+iy)) = \begin{cases} 0 & |x|>2 \\ \dfrac{\sqrt{4-x^2}}{2\pi} & |x|<2 \end{cases}$$

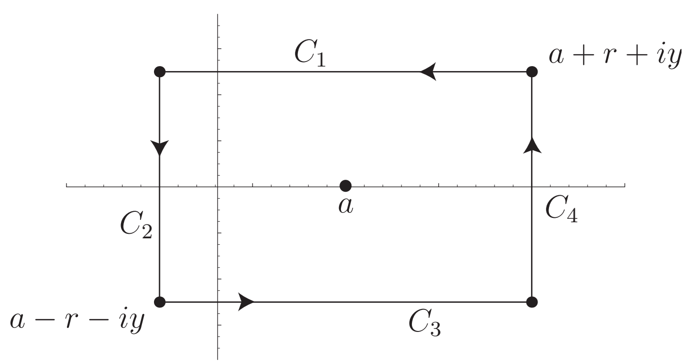
Hence we recover our original density.

Stieltjes Inversion at an Atom. When the Cauchy transform has a pole at $a \in \mathbb{R}$ the limit calculation using the dominated convergence theorem used above will not apply. A situation that frequently occurs is that $G(z)$ has a simple pole at a, i.e. $G(z) = f(a)/(z-a)$ with f analytic on a neighbourhood of a, then ν has an atom at a with weight $f(a)$.

Theorem. *Suppose $G(z) = f(z)/(z-a)$ with a a real number and f analytic on a neighbourhood, $B(a, 2r)$, of a. Then*

$$\lim_{y\to 0^+} \frac{-1}{\pi}\int_{a-r}^{a+r} \mathrm{Im}(G(x+iy))\, dx = f(a)$$

Proof. Let C be the boundary of the rectangle with vertices at $a+y+ir$, $a-y+ir$, $a-y-ir$, and $a+y-ir$. We will parameterize the four sides as follows.



C_1: the top side; $\gamma_1(t) = t+iy$ where t decreases from $a+r$ to $a-r$;
C_2: the left side; $\gamma_2 = a-r+it$ where t decreases from y to $-y$;
C_3: the bottom side; $\gamma_3(t) = t-iy$ where t increases from $a-r$ to $a+r$
C_4: the right side; $\gamma_4(t) = a+r+it$ where t increases from $-y$ to y

Then by Cauchy's integral theorem

$$\frac{1}{2\pi i}\int_C \frac{f(z)}{z-a}\, dz = f(a)$$

So let us calculate separately each integral over C_i. Now

$$\int_{C_1} G(z)\,dz = \int_{a+r}^{a-r} G(x+iy) \text{ and } \int_{C_3} G(z)\,dz = \int_{a-r}^{a+r} G(x-iy)$$

Thus

$$\begin{aligned} \int_{C_1+C_3} G(z)\,dz &= \int_{a-r}^{a+r} G(x-i\delta) - G(x+iy)\,dx \\ &= -\int_{a-r}^{a+r} 2i\,\mathrm{Im}(G(x+iy))\,dx \end{aligned}$$

Thus

$$\int_{a-r}^{a+r} \frac{-1}{\pi}\mathrm{Im}(G(x+iy))\,dx = \frac{1}{2\pi i}\int_{C_1+C_3} G(z)\,dz$$

So we only need to prove that for $i = 2$ or 4

$$\lim_{y\to 0^+} \int_{C_i} G(z)\,dz = 0$$

we shall only do the case $i = 2$. Since f is analytic on $B(0,r)$ there is $M > 0$ such that $|f(z)| \le M$ for $z \in C_2$. Thus

$$\left|\int_{C_2} G(z)\,dz\right| \le \int_{-y}^{y} \frac{2M}{|r-it|}\,dt \le \frac{2My}{r}$$

Thus for $i = 2$ and 4

$$\lim_{y\to 0^+} \int_{C_i} G(z)\,dz = 0$$

□