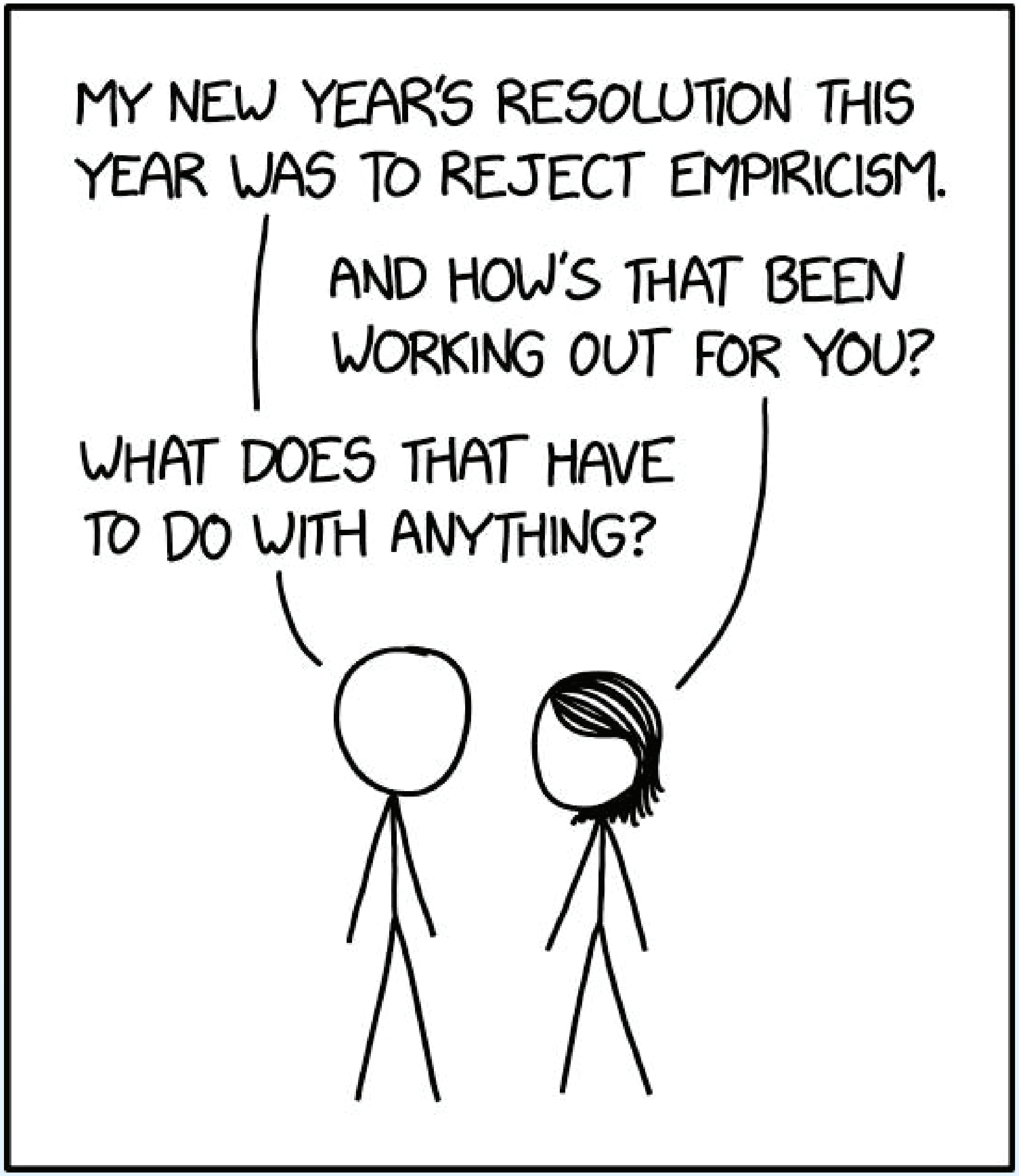
Free Zero Bias and $\boxplus$-Infinite Divisibility

Probabilistic Operator Algebras Seminar

Todd Kemp, UC San Diego
joint work with Larry Goldstein, USC

March 19, 2024

Empiricism



MY NEW YEAR'S RESOLUTION THIS YEAR WAS TO REJECT EMPIRICISM.
AND HOW'S THAT BEEN WORKING OUT FOR YOU?
WHAT DOES THAT HAVE TO DO WITH ANYTHING?

I May be Biased. . .

A Game of Telephone

In the olden days (or nowadays in Canada), people had landline phones. It was common for homes to have several (for parents, for teenage kids, for fax, for dialup internet. . .)

Without access to phone company records, how could a researcher estimate the distribution of number of landlines per household?

A Game of Telephone

In the olden days (or nowadays in Canada), people had landline phones. It was common for homes to have several (for parents, for teenage kids, for fax, for dialup internet. . .)

Without access to phone company records, how could a researcher estimate the distribution of number of landlines per household? A simple sampling method:

- Choose phone numbers randomly from the (local) phonebook.
- Ask each person you call “how many landlines do you have”?
- Assemble a representative sample of such data, and plot a histogram.

A Game of Telephone

In the olden days (or nowadays in Canada), people had landline phones. It was common for homes to have several (for parents, for teenage kids, for fax, for dialup internet. . .)

Without access to phone company records, how could a researcher estimate the distribution of number of landlines per household? A simple sampling method:

- Choose phone numbers randomly from the (local) phonebook.
- Ask each person you call “how many landlines do you have”?
- Assemble a representative sample of such data, and plot a histogram.

Let X be the random variable “$\#$ of landlines per home”. Does the histogram you build up approximate the distribution of X? Even if you called every number in the phonebook?

A Game of Telephone

In the olden days (or nowadays in Canada), people had landline phones. It was common for homes to have several (for parents, for teenage kids, for fax, for dialup internet. . .)

Without access to phone company records, how could a researcher estimate the distribution of number of landlines per household? A simple sampling method:

- Choose phone numbers randomly from the (local) phonebook.
- Ask each person you call “how many landlines do you have”?
- Assemble a representative sample of such data, and plot a histogram.

Let X be the random variable “$\#$ of landlines per home”. Does the histogram you build up approximate the distribution of X? Even if you called every number in the phonebook?

Question: How many people will give you the answer 0?

Size Bias

In the landline sampling scenario, the sample distribution is *not* the distribution of X. Rather, it is the distribution of X^s: the **size bias** of X. (More precisely: if $X \stackrel{d}{=} \mu$, then the sample distribution is μ^s, the **size bias transform** of μ.)

If X is a non-negative random variable with mean $m > 0$, then X^s has distribution

$$\mu^s(dx) = \frac{x}{m} \mathbb{1}_{[0,\infty)}(x)\, \mu(dx).$$

Size Bias

In the landline sampling scenario, the sample distribution is *not* the distribution of X. Rather, it is the distribution of X^s: the **size bias** of X. (More precisely: if $X \stackrel{d}{=} \mu$, then the sample distribution is μ^s, the **size bias transform** of μ.)

If X is a non-negative random variable with mean $m > 0$, then X^s has distribution

$$\mu^s(dx) = \frac{x}{m}\mathbb{1}_{[0,\infty)}(x)\,\mu(dx).$$

This can be understood more effectively as a functional equation: for any nice test function f,

$$\mathbb{E}[Xf(X)] = m\mathbb{E}[f(X^s)] = \mathbb{E}[X]\mathbb{E}[f(X^s)].$$

From here we see that μ can be “recovered” from μ^s...

Size Bias

In the landline sampling scenario, the sample distribution is *not* the distribution of X. Rather, it is the distribution of X^s: the **size bias** of X. (More precisely: if $X \stackrel{d}{=} \mu$, then the sample distribution is μ^s, the **size bias transform** of μ.)

If X is a non-negative random variable with mean $m > 0$, then X^s has distribution

$$\mu^s(dx) = \frac{x}{m} \mathbb{1}_{[0,\infty)}(x)\, \mu(dx).$$

This can be understood more effectively as a functional equation: for any nice test function f,

$$\mathbb{E}[Xf(X)] = m\mathbb{E}[f(X^s)] = \mathbb{E}[X]\mathbb{E}[f(X^s)].$$

From here we see that μ can be "recovered" from μ^s... except for any mass at 0.

Size Biasing Independent Sums

Let $X_1, \ldots, X_n$ be independent non-negative random variables, with $\mathbb{E}[X_i] = m_i > 0$. Let

$$W = \sum_{i=1}^{n} X_i.$$

How does the size bias W^s relate to the X_i^s?

Size Biasing Independent Sums

Let $X_1, \ldots, X_n$ be independent non-negative random variables, with $\mathbb{E}[X_i] = m_i > 0$. Let

$$W = \sum_{i=1}^{n} X_i.$$

How does the size bias W^s relate to the X_i^s?

Let I be a random index chosen from $\{1, \ldots, n\}$, independent from $\{X_1, \ldots, X_n\}$, with $\mathbb{P}(I = i) = m_i / \sum_j^n m_j$. Then

$$W^s \stackrel{d}{=} W - X_I + X_I^s.$$

If the X_i are i.i.d. you can choose any single index uniformly at random, or you can just choose (say) the first one:

$$W^s \stackrel{d}{=} X_1^s + X_2 + \cdots + X_n.$$

Fixed Points of Size Bias

Maybe the size biased distribution doesn't come out so different from the original one; what does this say about the distribution?
Question: Is there a distribution μ for which $\mu^s = \mu$?

$$\mathbb{E}[Xf(X)] = m\mathbb{E}[f(X^s)] = m\mathbb{E}[f(X)]$$

Fixed Points of Size Bias

Maybe the size biased distribution doesn't come out so different from the original one; what does this say about the distribution?

Question: Is there a distribution μ for which $\mu^s = \mu$?

$$\mathbb{E}[Xf(X)] = m\mathbb{E}[f(X^s)] = m\mathbb{E}[f(X)]$$

Answer: Yes. $\mu = \delta_m$, and that's it.

Fixed Points of Size Bias

Maybe the size biased distribution doesn't come out so different from the original one; what does this say about the distribution?
Question: Is there a distribution μ for which $\mu^s = \mu$?

$$\mathbb{E}[Xf(X)] = m\mathbb{E}[f(X^s)] = m\mathbb{E}[f(X)]$$

Answer: Yes. $\mu = \delta_m$, and that's it.

More interesting: what if we allow for a shift as well: $X \mapsto X^s - 1$. Are there fixed points?

Fixed Points of Size Bias

Maybe the size biased distribution doesn't come out so different from the original one; what does this say about the distribution?
Question: Is there a distribution μ for which $\mu^s = \mu$?

$$\mathbb{E}[Xf(X)] = m\mathbb{E}[f(X^s)] = m\mathbb{E}[f(X)]$$

Answer: Yes. $\mu = \delta_m$, and that's it.

More interesting: what if we allow for a shift as well: $X \mapsto X^s - 1$. Are there fixed points?

Exercise: The unique fixed point is $\mathrm{Poisson}(m)$.

Fixed Points of Size Bias

Maybe the size biased distribution doesn't come out so different from the original one; what does this say about the distribution?
Question: Is there a distribution μ for which $\mu^s = \mu$?

$$\mathbb{E}[Xf(X)] = m\mathbb{E}[f(X^s)] = m\mathbb{E}[f(X)]$$

Answer: Yes. $\mu = \delta_m$, and that's it.

More interesting: what if we allow for a shift as well: $X \mapsto X^s - 1$. Are there fixed points?

Exercise: The unique fixed point is $\mathrm{Poisson}(m)$.

$$\mathbb{E}[f(X)] = \mathbb{E}[f(X^s - 1)] = \frac{1}{m}\mathbb{E}[Xf(X-1)]$$

Taking $f(x) = x^k$ sets up a recursion of moments.

Is There a Free Size Bias?

The concept of bias in general is challenging to make sense of in a multivariate context. In the case of a single (selfadjoint) random variable, it is not clear how a “free size bias” should differ from the classical one! (It’s just a transform on probability measures.)

Is There a Free Size Bias?

The concept of bias in general is challenging to make sense of in a multivariate context. In the case of a single (selfadjoint) random variable, it is not clear how a "free size bias" should differ from the classical one! (It's just a transform on probability measures.)

One could ask: does the size bias relate similarly to freely independent sums? I.e. If $X_1, \ldots, X_n$ are f.i.d. is it true that

$$(X_1 + X_2 + \cdots + X_n)^s \stackrel{d}{=} X_1^s + X_2 + \cdots + X_n \text{ ?}$$

Is There a Free Size Bias?

The concept of bias in general is challenging to make sense of in a multivariate context. In the case of a single (selfadjoint) random variable, it is not clear how a "free size bias" should differ from the classical one! (It's just a transform on probability measures.)

One could ask: does the size bias relate similarly to freely independent sums? I.e. If $X_1, \ldots, X_n$ are f.i.d. is it true that

$$(X_1 + X_2 + \cdots + X_n)^s \stackrel{d}{=} X_1^s + X_2 + \cdots + X_n \,?$$

No. Take shifted f.i.d. semicircular random variables.

Perhaps, then, the "free size bias" should be a new transform which *does* have this free sum exchange property? And whose shift-fixed point is a free Poisson?

Is There a Free Size Bias?

The concept of bias in general is challenging to make sense of in a multivariate context. In the case of a single (selfadjoint) random variable, it is not clear how a "free size bias" should differ from the classical one! (It's just a transform on probability measures.)

One could ask: does the size bias relate similarly to freely independent sums? I.e. If $X_1, \ldots, X_n$ are f.i.d. is it true that

$$(X_1 + X_2 + \cdots + X_n)^s \stackrel{d}{=} X_1^s + X_2 + \cdots + X_n \,?$$

No. Take shifted f.i.d. semicircular random variables.

Perhaps, then, the "free size bias" should be a new transform which *does* have this free sum exchange property? And whose shift-fixed point is a free Poisson? We have some thoughts on this, but nothing that can see the light of day just yet.

The Zero Bias

Gaussian Integration by Parts

Let $Z \overset{d}{=} \mathscr{N}(0, \sigma^2)$. A very useful computational tool is the integration by parts formula (aka Stein's formula)

$$\mathbb{E}[Zf(Z)] = \sigma^2\,\mathbb{E}[f'(Z)]$$

which holds for any $f \in C^1(\mathbb{R})$ for which f and f' are sufficiently integrable.

Gaussian Integration by Parts

Let $Z \stackrel{d}{=} \mathcal{N}(0, \sigma^2)$. A very useful computational tool is the integration by parts formula (aka Stein's formula)

$$\mathbb{E}[Zf(Z)] = \sigma^2 \, \mathbb{E}[f'(Z)]$$

which holds for any $f \in C^1(\mathbb{R})$ for which f and f' are sufficiently integrable. In fact, this identity (holding over all $f \in C_c^\infty(\mathbb{R})$ for example) uniquely specifies the centered normal distribution of variance σ^2.

Gaussian Integration by Parts

Let $Z \stackrel{d}{=} \mathscr{N}(0, \sigma^2)$. A very useful computational tool is the integration by parts formula (aka Stein's formula)

$$\mathbb{E}[Zf(Z)] = \sigma^2\,\mathbb{E}[f'(Z)]$$

which holds for any $f \in C^1(\mathbb{R})$ for which f and f' are sufficiently integrable. In fact, this identity (holding over all $f \in C_c^\infty(\mathbb{R})$ for example) uniquely specifies the centered normal distribution of variance σ^2.

The extent to which a distribution fails to satisfy the above equation can be viewed, in multiple ways, as a measure of its distance from a normal distribution. Tools based on this idea are generally called **Stein's Method**, and can produce extremely sharp estimates for normal approximation.

One approach is with *Stein kernels*.

Stein Kernels

Let X be a real-valued random variable. We say that X (or rather its distribution) possesses a **Stein kernel** $A = A_X$ if, for all $f \in C_c^\infty(\mathbb{R})$,

$$\mathbb{E}[Xf(X)] = \mathbb{E}[A(X)f'(X)].$$

If $X \stackrel{d}{=} \mathcal{N}(0, \sigma^2)$, then X possesses the constant Stein kernel $A = \sigma^2$. (So bounds on derivatives of the Stein kernel can measure distance from normality; this leads to *Stein discrepancy* sharply controlling L^2-Wasserstein distance.)

Stein Kernels

Let X be a real-valued random variable. We say that X (or rather its distribution) possesses a **Stein kernel** $A = A_X$ if, for all $f \in C_c^\infty(\mathbb{R})$,

$$\mathbb{E}[X f(X)] = \mathbb{E}[A(X) f'(X)].$$

If $X \stackrel{d}{=} \mathcal{N}(0, \sigma^2)$, then X possesses the constant Stein kernel $A = \sigma^2$. (So bounds on derivatives of the Stein kernel can measure distance from normality; this leads to *Stein discrepancy* sharply controlling L^2-Wasserstein distance.)

Stein kernels are *unique* when they exist; but *not every distribution* μ *has a Stein kernel* A_μ. Characterizing those that do is a difficult problem that is an area of active research. One existence theorem: if μ **has mean** 0 and has a density ρ with connected support, then

$$A_\mu(x) = \frac{1}{\rho(x)} \int_x^\infty y\, \rho(y)\, dy.$$

The Zero Bias

Stein kernels are one way to deform the Stein equation.

$$\mathbb{E}[Xf(X)] = \sigma^2 \cdot \mathbb{E}[f'(X)] \qquad \text{Stein equation}$$

$$\mathbb{E}[Xf(X)] = \mathbb{E}[A(X)f'(X)] \qquad \text{Stein kernel}$$

The Zero Bias

Stein kernels are one way to deform the Stein equation. A different way relates to size bias:

$$\mathbb{E}[Xf(X)] = \sigma^2 \cdot \mathbb{E}[f'(X)] \qquad \text{Stein equation}$$

$$\mathbb{E}[Xf(X)] = \mathbb{E}[A(X)f'(X)] \qquad \text{Stein kernel}$$

$$\mathbb{E}[Xf(X)] = m \cdot \mathbb{E}[f(X^s)] \qquad \text{size bias}$$

The Zero Bias

Stein kernels are one way to deform the Stein equation. A different way relates to size bias:

$$\mathbb{E}[Xf(X)] = \sigma^2 \cdot \mathbb{E}[f'(X)] \qquad \text{Stein equation}$$

$$\mathbb{E}[Xf(X)] = \mathbb{E}[A(X)f'(X)] \qquad \text{Stein kernel}$$

$$\mathbb{E}[Xf(X)] = m \cdot \mathbb{E}[f(X^s)] \qquad \text{size bias}$$

$$\mathbb{E}[Xf(X)] = c \cdot \mathbb{E}[f'(X^*)] \qquad \textbf{zero bias}$$

The Zero Bias

Stein kernels are one way to deform the Stein equation. A different way relates to size bias:

$\mathbb{E}[Xf(X)] = \sigma^2 \cdot \mathbb{E}[f'(X)]$	Stein equation
$\mathbb{E}[Xf(X)] = \mathbb{E}[A(X)f'(X)]$	Stein kernel
$\mathbb{E}[Xf(X)] = m \cdot \mathbb{E}[f(X^s)]$	size bias
$\mathbb{E}[Xf(X)] = c \cdot \mathbb{E}[f'(X^*)]$	**zero bias**

Setting $f \equiv 1$, we see size bias X^* only makes sense if $\mathbb{E}[X] = 0$.

The Zero Bias

Stein kernels are one way to deform the Stein equation. A different way relates to size bias:

$$\mathbb{E}[Xf(X)] = \sigma^2 \cdot \mathbb{E}[f'(X)] \qquad \text{Stein equation}$$
$$\mathbb{E}[Xf(X)] = \mathbb{E}[A(X)f'(X)] \qquad \text{Stein kernel}$$
$$\mathbb{E}[Xf(X)] = m \cdot \mathbb{E}[f(X^s)] \qquad \text{size bias}$$
$$\mathbb{E}[Xf(X)] = c \cdot \mathbb{E}[f'(X^*)] \qquad \textbf{zero bias}$$

Setting $f \equiv 1$, we see size bias X^* only makes sense if $\mathbb{E}[X] = 0$.
Setting $f(x) = x$, we see that $c = \mathbb{E}[X^2] = \mathrm{Var}[X] = \sigma^2$.

The Zero Bias

Stein kernels are one way to deform the Stein equation. A different way relates to size bias:

$\mathbb{E}[Xf(X)] = \sigma^2 \cdot \mathbb{E}[f'(X)]$	Stein equation
$\mathbb{E}[Xf(X)] = \mathbb{E}[A(X)f'(X)]$	Stein kernel
$\mathbb{E}[Xf(X)] = m \cdot \mathbb{E}[f(X^s)]$	size bias
$\mathbb{E}[Xf(X)] = \sigma^2 \cdot \mathbb{E}[f'(X^*)]$	**zero bias**

Setting $f \equiv 1$, we see size bias X^* only makes sense if $\mathbb{E}[X] = 0$
Setting $f(x) = x$, we see that $c = \mathbb{E}[X^2] = \mathrm{Var}[X] = \sigma^2$.

The **zero bias transform** $X \mapsto X^*$ (or more precisely $\mu \mapsto \mu^*$) is well-defined on the space $\mathcal{D}_{0,\sigma^2}$ of probability distributions on $\mathbb{R}$ with mean 0 and variance σ^2. It can be constructed in several different ways (all leading to the same measure μ^*).

The normal distribution $X \stackrel{d}{=} \mathcal{N}(0, t)$ is the unique fixed point.

One Construction of the Zero Bias

For one concrete (probabilistic) construction of the zero bias, we need to get even more biased.

One Construction of the Zero Bias

For one concrete (probabilistic) construction of the zero bias, we need to get even more biased.

Let X be a non-constant L^2 random variable, with distribution μ. The **square bias** of X (or more precisely of μ) is the distribution $\mu^\square$, realized as the distribution of a random variable $X^\square$, defined by

$$\mu^\square(dx) = \frac{1}{\mathbb{E}[X^2]} x^2\, \mu(dx).$$

The associated functional equation is

$$\mathbb{E}[f(X^2 f(X))] = \mathbb{E}[X^2] \cdot \mathbb{E}[f(X^\square)].$$

One Construction of the Zero Bias

For one concrete (probabilistic) construction of the zero bias, we need to get even more biased.

Let X be a non-constant L^2 random variable, with distribution μ. The **square bias** of X (or more precisely of μ) is the distribution $\mu^\square$, realized as the distribution of a random variable $X^\square$, defined by

$$\mu^\square(dx) = \frac{1}{\mathbb{E}[X^2]} x^2 \, \mu(dx).$$

The associated functional equation is

$$\mathbb{E}[f(X^2 f(X))] = \mathbb{E}[X^2] \cdot \mathbb{E}[f(X^\square)].$$

Proposition. If X has mean 0 and finite second moment, then

$$X^* \stackrel{d}{=} U X^\square$$

where $U \stackrel{d}{=} \mathrm{Unif}[0,1]$ is independent from $X^\square$.

Properties of the Zero Bias

- Every mean 0, finite variance random variable has a zero bias.
- For any constant $\alpha \neq 0$, $(\alpha X)^* = \alpha X^*$.
- If $X_n \rightharpoonup X$ and $\mathrm{Var}[X_n] \to \mathrm{Var}[X]$, then $X_n^* \rightharpoonup X^*$.
- The distribution of X^* is always absolutely continuous.
- The support of μ^* is equal to the convex hull of the support of μ.

Properties of the Zero Bias

- Every mean 0, finite variance random variable has a zero bias.
- For any constant $\alpha \neq 0$, $(\alpha X)^* = \alpha X^*$.
- If $X_n \rightharpoonup X$ and $\mathrm{Var}[X_n] \to \mathrm{Var}[X]$, then $X_n^* \rightharpoonup X^*$.
- The distribution of X^* is always absolutely continuous.
- The support of μ^* is equal to the convex hull of the support of μ.

The zero bias has a similar independent sum exchange property to the size bias:

Properties of the Zero Bias

- Every mean 0, finite variance random variable has a zero bias.
- For any constant $\alpha \neq 0$, $(\alpha X)^* = \alpha X^*$.
- If $X_n \rightharpoonup X$ and $\mathrm{Var}[X_n] \to \mathrm{Var}[X]$, then $X_n^* \rightharpoonup X^*$.
- The distribution of X^* is always absolutely continuous.
- The support of μ^* is equal to the convex hull of the support of μ.

The zero bias has a similar independent sum exchange property to the size bias: If $X_1, \ldots, X_n$ are independent mean 0 random variables and $\mathbb{E}[X_i^2] = \sigma_i^2 > 0$, and if I is a random index in $\{1, \ldots, n\}$ independent from $\{X_1, \ldots, X_n\}$ with $\mathbb{P}(I = i) = \sigma_i^2 / \sum_j^n \sigma_j^2$, then

$$\left(\sum_i^n X_i\right)^* \stackrel{d}{=} \sum_i^n X_i - X_I + X_I^*.$$

If the variables are i.i.d. we can just take (say) $I = 1$:

$$(X_1 + X_2 + \cdots + X_n)^* \stackrel{d}{=} X_1^* + X_2 + \cdots + X_n.$$

Connection to Stein Kernels

Let $X \stackrel{d}{=} \mu$. The defining equation for the zero bias is

$$\mathbb{E}[X f(X)] = \sigma^2 \, \mathbb{E}[f'(X^*)]$$

Connection to Stein Kernels

Let $X \stackrel{d}{=} \mu$. The defining equation for the zero bias is

$$\begin{aligned}\mathbb{E}[X f(X)] &= \sigma^2\, \mathbb{E}[f'(X^*)] \\ &= \sigma^2 \int f'(x)\, \mu^*(dx)\end{aligned}$$

Connection to Stein Kernels

Let $X \overset{d}{=} \mu$. The defining equation for the zero bias is

$$\begin{aligned}\mathbb{E}[X f(X)] &= \sigma^2 \, \mathbb{E}[f'(X^*)] \\ &= \sigma^2 \int f'(x)\, \mu^*(dx) = \sigma^2 \int f'(x)\, \rho_{X^*}(x)\, dx.\end{aligned}$$

Connection to Stein Kernels

Let $X \stackrel{d}{=} \mu$. The defining equation for the zero bias is

$$\begin{aligned}\mathbb{E}[X f(X)] &= \sigma^2\, \mathbb{E}[f'(X^*)] \\ &= \sigma^2 \int f'(x)\, \mu^*(dx) = \sigma^2 \int f'(x)\, \rho_{X^*}(x)\, dx.\end{aligned}$$

Now suppose that X has a density ρ_X that is strictly positive on the interior of its support (i.e. $\operatorname{supp}\mu$ is connected). In this case $\operatorname{supp}\mu^* = \operatorname{supp}\mu$, and hence

$$\begin{aligned}\mathbb{E}[X f(X)] &= \sigma^2 \int f'(x)\, \rho_{X^*}(x)\, dx \\ &= \sigma^2 \int \frac{\rho_{X^*}(x)}{\rho_X(x)}\, f'(x)\, \rho_X(x)\, dx\end{aligned}$$

Connection to Stein Kernels

Let $X \stackrel{d}{=} \mu$. The defining equation for the zero bias is

$$\begin{aligned}\mathbb{E}[X f(X)] &= \sigma^2\, \mathbb{E}[f'(X^*)] \\ &= \sigma^2 \int f'(x)\, \mu^*(dx) = \sigma^2 \int f'(x)\, \rho_{X^*}(x)\, dx.\end{aligned}$$

Now suppose that X has a density ρ_X that is strictly positive on the interior of its support (i.e. $\operatorname{supp} \mu$ is connected). In this case $\operatorname{supp} \mu^* = \operatorname{supp} \mu$, and hence

$$\begin{aligned}\mathbb{E}[X f(X)] &= \sigma^2 \int f'(x)\, \rho_{X^*}(x)\, dx \\ &= \sigma^2 \int \frac{\rho_{X^*}(x)}{\rho_X(x)}\, f'(x)\, \rho_X(x)\, dx \\ &= \mathbb{E}[A(X) f'(X)] \qquad \text{where} \qquad A = \sigma^2\, \rho_{X^*}/\rho_X.\end{aligned}$$

In fact μ has a Stein kernel whenever $\mu_X \ll \mu_{X^*}$ (which is equivalent to assuming $\mu_X \approx \mu_{X^*}$).

Surprising Connection to Infinite Divisibility

A distribution (the law of X) is (classically) *infinitely divisible* if, for every n, there are i.i.d. random variables $X_{1,n}, \ldots, X_{n,n}$ with

$$X \stackrel{d}{=} X_{1,n} + \cdots + X_{n,n}.$$

A forthcoming paper by L. Goldstein and U. Schmock proves the following very interesting characterization of infinitely divisible distributions with finite second moment. For this result, we extend the zero bias to the non-centered case by shifting: if $\mathbb{E}[X] = m$, we work with $(X - m)^* + m$.

Surprising Connection to Infinite Divisibility

A distribution (the law of X) is (classically) *infinitely divisible* if, for every n, there are i.i.d. random variables $X_{1,n}, \ldots, X_{n,n}$ with

$$X \stackrel{d}{=} X_{1,n} + \cdots + X_{n,n}.$$

A forthcoming paper by L. Goldstein and U. Schmock proves the following very interesting characterization of infinitely divisible distributions with finite second moment. For this result, we extend the zero bias to the non-centered case by shifting: if $\mathbb{E}[X] = m$, we work with $(X - m)^* + m$.

Theorem. [Goldstein, Schmock, 2023+] X is infinitely divisible if and only if there exist random variables U, Y with $\{U, X, Y\}$ independent, $U \stackrel{d}{=} \mathrm{Unif}[0, 1]$, and

$$(X - m)^* + m \stackrel{d}{=} X + UY.$$

2017 work of Arras and Houdré used non-probabilistic methods to prove a slightly weaker relation to the Kolmogorov formulation:

The Lévy–Khinchine Formula (by Kolmogorov)

An L^2 random variable with variance σ^2 is (classically) infinitely divisible if and only if its cumulant generating function (log Fourier transform) has the following form:

$$C(\xi) = -\frac{\sigma^2}{2}\xi^2\nu(\{0\}) + \sigma^2 \int_{\mathbb{R}\setminus\{0\}} \frac{\exp(i\xi x) - i\xi x - 1}{x^2}\,\nu(dx)$$

for some probability measure ν on $\mathbb{R}$.

The Lévy–Khinchine Formula (by Kolmogorov)

An L^2 random variable with variance σ^2 is (classically) infinitely divisible if and only if its cumulant generating function (log Fourier transform) has the following form:

$$C(\xi) = -\frac{\sigma^2}{2}\xi^2\nu(\{0\}) + \sigma^2\int_{\mathbb{R}\setminus\{0\}} \frac{\exp(i\xi x) - i\xi x - 1}{x^2}\,\nu(dx)$$

for some probability measure ν on $\mathbb{R}$.

Goldstein and Schmock prove directly that an L^2 random variable X has a cumulant generation of the above form if and only if

$$(X-m)^* + m \overset{d}{=} X + UY$$

and moreover ν is the distribution of Y. This yields a concrete meaning for this Lévy–Khinchine measure.

The Free Zero Bias

Free Difference Quotient

In noncommutative probability, we frequently let single-variable functions do double-duty and act on operators by functional calculus. If p is an ordinary polynomial, and $\mathscr{A}$ is a C^* algebra, let $p_{\mathscr{A}}: \mathscr{A} \to \mathscr{A}$ be the associated functional calculus function.

Propositon. $p_{\mathscr{A}} \in C^\infty(\mathscr{A}; \mathscr{A})$, and the Fréchet derivative $Dp_{\mathscr{A}}$ is given by

$$[Dp_{\mathscr{A}}](a)[h] = (\partial p)(a)\#h$$

where $\partial p \colon \mathscr{A} \to \mathscr{A} \otimes \mathscr{A}$ is the **free difference quotient**. Here $(a \otimes b)\#h := ahb$, and ∂p is defined by

$$\partial x^k := \sum_{i=1}^{k} x^{k-i} \otimes x^{i-1}.$$

Equivalently: identifying $\mathbb{C}[x] \otimes \mathbb{C}[x] \approx \mathbb{C}[x, y]$, really

$$(\partial p)(x, y) = \frac{p(x) - p(y)}{x - y}.$$

The Free Stein Equation

The Stein equation (i.e. Gaussian integration by parts) uniquely specific $Z \stackrel{d}{=} \mathcal{N}(0, \sigma^2)$ via the functional equation

$$\mathbb{E}[Zf(Z)] = \sigma^2\, \mathbb{E}[f'(Z)], \qquad f \in C_c^\infty(\mathbb{R}).$$

The Free Stein Equation

The Stein equation (i.e. Gaussian integration by parts) uniquely specific $Z \stackrel{d}{=} \mathscr{N}(0, \sigma^2)$ via the functional equation

$$\mathbb{E}[Zf(Z)] = \sigma^2\,\mathbb{E}[f'(Z)], \qquad f \in C_c^\infty(\mathbb{R}).$$

The natural guess for the free version of this equation is:

$$\mathbb{E}[Sf(S)] = \sigma^2\,\mathbb{E}\otimes\mathbb{E}[\partial f(S)], \qquad f \in C_c^\infty(\mathbb{R}).$$

We should restrict to polynomials f to make sense of this from the definition $\partial f\colon \mathscr{A} \to \mathscr{A}\otimes\mathscr{A}$.

The Free Stein Equation

The Stein equation (i.e. Gaussian integration by parts) uniquely specific $Z \stackrel{d}{=} \mathscr{N}(0, \sigma^2)$ via the functional equation

$$\mathbb{E}[Zf(Z)] = \sigma^2\,\mathbb{E}[f'(Z)], \qquad f \in C_c^\infty(\mathbb{R}).$$

The natural guess for the free version of this equation is:

$$\mathbb{E}[Sf(S)] = \sigma^2\,\mathbb{E}\otimes\mathbb{E}[\partial f(S)], \qquad f \in C_c^\infty(\mathbb{R}).$$

We should restrict to polynomials f to make sense of this from the definition $\partial f\colon \mathscr{A} \to \mathscr{A}\otimes\mathscr{A}$. But if we interpret ∂f as a genuine difference quotient, we can interpret this more directly for any measurable function f as

$$\mathbb{E}[Sf(S)] = \sigma^2\,\mathbb{E}\left[\frac{f(S) - f(S')}{S - S'}\right]$$

where S, S' are two *classically* independent copies of the putative random variable S.

The Semicircle Law and the Free Stein Equation

Proposition. The unique solution (in distribution) to the free Stein equation

$$\mathbb{E}[Sf(S)] = \sigma^2\,\mathbb{E}\otimes\mathbb{E}[\partial f(S)], \qquad f\in\mathbb{C}[x]$$

is the semicircle law S of variance σ^2.

The Semicircle Law and the Free Stein Equation

Proposition. The unique solution (in distribution) to the free Stein equation

$$\mathbb{E}[Sf(S)] = \sigma^2\,\mathbb{E}\otimes\mathbb{E}[\partial f(S)], \qquad f \in \mathbb{C}[x]$$

is the semicircle law S of variance σ^2.

Proof. Renormalize $\sigma^2 = 1$. Take $f(x) = x^k$. Thus

$$\mathbb{E}[S^{k+1}] = \mathbb{E}[S\cdot S^k] = \mathbb{E}\otimes\mathbb{E}\left[\sum_{i=1}^{k} S^{k-i}\otimes S^{i-1}\right] = \sum_{i=1}^{k}\mathbb{E}[S^{k-i}]\mathbb{E}[S^{i-1}].$$

The Semicircle Law and the Free Stein Equation

Proposition. The unique solution (in distribution) to the free Stein equation

$$\mathbb{E}[Sf(S)] = \sigma^2\,\mathbb{E}\otimes\mathbb{E}[\partial f(S)], \qquad f \in \mathbb{C}[x]$$

is the semicircle law S of variance σ^2.

Proof. Renormalize $\sigma^2 = 1$. Take $f(x) = x^k$. Thus

$$\mathbb{E}[S^{k+1}] = \mathbb{E}[S \cdot S^k] = \mathbb{E}\otimes\mathbb{E}\left[\sum_{i=1}^{k} S^{k-i}\otimes S^{i-1}\right] = \sum_{i=1}^{k}\mathbb{E}[S^{k-i}]\mathbb{E}[S^{i-1}].$$

If k is even, see by induction (from $\mathbb{E}[X] = 0$) that all terms are 0; so odd moments of S are 0.

The Semicircle Law and the Free Stein Equation

Proposition. The unique solution (in distribution) to the free Stein equation

$$\mathbb{E}[Sf(S)] = \sigma^2\, \mathbb{E} \otimes \mathbb{E}[\partial f(S)], \qquad f \in \mathbb{C}[x]$$

is the semicircle law S of variance σ^2.

Proof. Renormalize $\sigma^2 = 1$. Take $f(x) = x^k$. Thus

$$\mathbb{E}[S^{k+1}] = \mathbb{E}[S \cdot S^k] = \mathbb{E} \otimes \mathbb{E}\left[\sum_{i=1}^{k} S^{k-i} \otimes S^{i-1}\right] = \sum_{i=1}^{k} \mathbb{E}[S^{k-i}]\mathbb{E}[S^{i-1}].$$

If k is even, see by induction (from $\mathbb{E}[X] = 0$) that all terms are 0; so odd moments of S are 0. . Then taking $k = 2m - 1$,

$$\mathbb{E}[S^{2m}] = \sum_{i=1}^{2m-1} \mathbb{E}[S^{2m-1-i}]\mathbb{E}[S^{i-1}]$$

The Semicircle Law and the Free Stein Equation

Proposition. The unique solution (in distribution) to the free Stein equation

$$\mathbb{E}[Sf(S)] = \sigma^2\, \mathbb{E} \otimes \mathbb{E}[\partial f(S)], \qquad f \in \mathbb{C}[x]$$

is the semicircle law S of variance σ^2.

Proof. Renormalize $\sigma^2 = 1$. Take $f(x) = x^k$. Thus

$$\mathbb{E}[S^{k+1}] = \mathbb{E}[S \cdot S^k] = \mathbb{E} \otimes \mathbb{E}\left[\sum_{i=1}^{k} S^{k-i} \otimes S^{i-1}\right] = \sum_{i=1}^{k} \mathbb{E}[S^{k-i}]\mathbb{E}[S^{i-1}].$$

If k is even, see by induction (from $\mathbb{E}[X] = 0$) that all terms are 0; so odd moments of S are 0. . Then taking $k = 2m - 1$,

$$\mathbb{E}[S^{2m}] = \sum_{i=1}^{2m-1} \mathbb{E}[S^{2m-1-i}]\mathbb{E}[S^{i-1}] = \sum_{j=0}^{m-1} \mathbb{E}[S^{2m-2j}]\mathbb{E}[S^{2j}].$$

$\square$

Defining the Free Zero Bias

Putatively, we define the **free zero bias** X° of a (law of a) centered, variance t random variable X by the functional equation

$$\mathbb{E}[Xf(X)] = \sigma^2\,\mathbb{E}\otimes\mathbb{E}[\partial f(X^\circ)], \qquad f \in C_c^\infty(\mathbb{R})$$

where, to define the right-hand-side beyond polynomials f, we interpret for Y° a classically independent copy of (the putative) X°

$$\mathbb{E}\otimes\mathbb{E}[\partial f(X^\circ)] = \mathbb{E}\left[\frac{f(X^\circ)-f(Y^\circ)}{X^\circ - Y^\circ}\right].$$

Defining the Free Zero Bias

Putatively, we define the **free zero bias** X° of a (law of a) centered, variance t random variable X by the functional equation

$$\mathbb{E}[Xf(X)] = \sigma^2\,\mathbb{E}\otimes\mathbb{E}[\partial f(X^\circ)], \qquad f \in C_c^\infty(\mathbb{R})$$

where, to define the right-hand-side beyond polynomials f, we interpret for Y° a classically independent copy of (the putative) X°

$$\mathbb{E}\otimes\mathbb{E}[\partial f(X^\circ)] = \mathbb{E}\left[\frac{f(X^\circ)-f(Y^\circ)}{X^\circ - Y^\circ}\right].$$

Note: $\frac{f(x)-f(y)}{x-y} = \mathbb{E}[f'(Ux+(1-U)y)]$ where $U \stackrel{d}{=} \mathrm{Unif}[0,1]$, so we could give the definition as

$$\mathbb{E}[Xf(X)] = \sigma^2\,E[f'(UX^\circ + (1-U)Y^\circ)], \qquad f \in C_c^\infty(\mathbb{R}).$$

This means that $X^* \stackrel{d}{=} UX^\circ + (1-U)Y^\circ$.

Existence of the Free Zero Bias

For $z \in \mathbb{C}_+$, taking the resolvent function $f_z(x) = \frac{1}{z-x}$, we calculate that

$$\partial f_z(x,y) = \frac{\frac{1}{z-x} - \frac{1}{z-y}}{x-y}$$

Existence of the Free Zero Bias

For $z \in \mathbb{C}_+$, taking the resolvent function $f_z(x) = \frac{1}{z-x}$, we calculate that

$$\partial f_z(x,y) = \frac{\frac{1}{z-x} - \frac{1}{z-y}}{x-y} = \frac{1}{(z-x)(z-y)} = f_z(x)f_z(y).$$

Plugging this into the defining equation

$$\mathbb{E}[Xf_z(X)] = \sigma^2 \mathbb{E}\otimes\mathbb{E}[\partial f_z(X^\circ)]$$

and simplifying yields the following quadratic equation for Cauchy transforms $G_X(z) = \mathbb{E}[\frac{1}{z-X}]$:

$$\sigma^2 G_{X^\circ}(z)^2 = zG_X(x) - 1.$$

So the question is: given a Cauchy transform $G(z)$ (mean 0, variance t), is there a square root of $\frac{1}{\sigma^2}(zG(z)-1)$ that is a Cauchy transform? **The answer is always yes.**

A Geometric Mean Cauchy Transform

Lemma. Let X, Y be real-valued random variables. Then for $z \in \mathbb{C}_+$,

$$z \mapsto -\sqrt{G_X(z)G_Y(z)} = \sqrt{G_X(z)}\sqrt{G_Y(z)}$$

is a Cauchy transform of a probability measure. We denote the associated measure as the law of a random variable $X \flat Y$.

A Geometric Mean Cauchy Transform

Lemma. Let X, Y be real-valued random variables. Then for $z \in \mathbb{C}_+$,

$$z \mapsto -\sqrt{G_X(z)G_Y(z)} = \sqrt{G_X(z)}\sqrt{G_Y(z)}$$

is a Cauchy transform of a probability measure. We denote the associated measure as the law of a random variable $X \flat Y$.

Example. If $X \equiv 1$ and $Y \equiv -1$, $G_{X \flat Y}(z) = \frac{1}{\sqrt{z^2-1}}$.

A Geometric Mean Cauchy Transform

Lemma. Let X, Y be real-valued random variables. Then for $z \in \mathbb{C}_+$,

$$z \mapsto -\sqrt{G_X(z)G_Y(z)} = \sqrt{G_X(z)}\sqrt{G_Y(z)}$$

is a Cauchy transform of a probability measure. We denote the associated measure as the law of a random variable $X \flat Y$.

Example. If $X \equiv 1$ and $Y \equiv -1$, $G_{X\flat Y}(z) = \frac{1}{\sqrt{z^2-1}}$.

Proposition. If $X_n \rightharpoonup X$ and $Y_n \rightharpoonup Y$ then $X_n \flat Y_n \rightharpoonup X \flat Y$.

A Geometric Mean Cauchy Transform

Lemma. Let X, Y be real-valued random variables. Then for $z \in \mathbb{C}_+$,

$$z \mapsto -\sqrt{G_X(z)G_Y(z)} = \sqrt{G_X(z)}\sqrt{G_Y(z)}$$

is a Cauchy transform of a probability measure. We denote the associated measure as the law of a random variable $X \flat Y$.

Example. If $X \equiv 1$ and $Y \equiv -1$, $G_{X \flat Y}(z) = \frac{1}{\sqrt{z^2-1}}$.

Proposition. If $X_n \rightharpoonup X$ and $Y_n \rightharpoonup Y$ then $X_n \flat Y_n \rightharpoonup X \flat Y$.

We will mostly use this with $Y = 0$. Define $X^\flat := X \flat 0$; i.e.

$$G_{X^\flat}(z) = -\sqrt{\frac{1}{z} G_X(z)}.$$

We call this the 'El Gordo transform'. We will see that it is a free analogue of the map (on distributions) $X \mapsto UX$ where $U \overset{d}{=} \mathrm{Unif}[0,1]$.

Construction of Free Zero Bias Using Square Bias

The square bias $\mathbb{E}[X^2 f(X)] = \mathbb{E}[X^2]\mathbb{E}[f(X^\square)]$ can be identified by its Cauchy transform:

$$G_{X^\square}(z) = \frac{1}{\mathbb{E}[X^2]}(z^2 G_X(z) - \mathbb{E}[X] - z).$$

Construction of Free Zero Bias Using Square Bias

The square bias $\mathbb{E}[X^2 f(X)] = \mathbb{E}[X^2]\mathbb{E}[f(X^\square)]$ can be identified by its Cauchy transform:

$$G_{X^\square}(z) = \frac{1}{\mathbb{E}[X^2]}(z^2 G_X(z) - \mathbb{E}[X] - z).$$

Definition. Given any L^2 random variable X, define (the law of) its **free zero bias** by

$$X^\circ \stackrel{d}{=} (X^\square)^\flat.$$

This is the free version of the zero bias construction $X^* = UX^\square$.

Construction of Free Zero Bias Using Square Bias

The square bias $\mathbb{E}[X^2 f(X)] = \mathbb{E}[X^2]\mathbb{E}[f(X^\square)]$ can be identified by its Cauchy transform:

$$G_{X^\square}(z) = \frac{1}{\mathbb{E}[X^2]}(z^2 G_X(z) - \mathbb{E}[X] - z).$$

Definition. Given any L^2 random variable X, define (the law of) its **free zero bias** by

$$X^\circ \stackrel{d}{=} (X^\square)^\flat.$$

This is the free version of the zero bias construction $X^* = UX^\square$.

In terms of Cauchy transforms:

$$\mathbb{E}[X^2] \cdot G_{X^\circ}(z)^2 = zG_X(x) - \frac{\mathbb{E}[X]}{z} - 1$$

which reduces to the correct equation for the (originally defined) free zero bias when $\mathbb{E}[X] = 0$.

Examples of Free Zero Bias Computations

Example. If X is a (centered) semicircular random variable, $X^\circ \stackrel{d}{=} X$. (If and only if, actually.)

Examples of Free Zero Bias Computations

Example. If X is a (centered) semicircular random variable, $X^\circ \stackrel{d}{=} X$. (If and only if, actually.)

Example. If X is centered with point masses at $-a < 0 < b$, then

$$G_{X^\circ}(z) = \frac{1}{\sqrt{(z+a)(z-b)}}.$$

In particular: if X is Rademacher ($a = b = 1$), X° is arcsine distributed

$$\rho_{X^\circ}(x) = \frac{1}{\pi}\frac{1}{\sqrt{(1-x^2)_+}}.$$

Examples of Free Zero Bias Computations

Example. If X is a (centered) semicircular random variable, $X^\circ \stackrel{d}{=} X$. (If and only if, actually.)

Example. If X is centered with point masses at $-a < 0 < b$, then

$$G_{X^\circ}(z) = \frac{1}{\sqrt{(z+a)(z-b)}}.$$

In particular: if X is Rademacher ($a = b = 1$), X° is arcsine distributed

$$\rho_{X^\circ}(x) = \frac{1}{\pi} \frac{1}{\sqrt{(1-x^2)_+}}.$$

Example. If X is arcsine distributed, then

$$\rho_{X^\circ}(x) = \frac{1}{\pi} \sqrt{1 + \frac{1}{\sqrt{(1-x^2)_+}}}.$$

Properties of the Free Zero Bias

Continuity and Support of the Free Zero Bias

Theorem. Let X be a mean 0, non-constant, L^2 random variable. Let μ denote the distribution of X and let μ° denote the distribution of the free zero bias X°. Then for any compact interval $[a,b] \subset \mathbb{R}$,

$$(\mu^\circ([a,b]))^2 \leq (b-a)\mathbb{E}[|X|].$$

Consequently, μ° is absolutely continuous with respect to Lebesgue measure.

Continuity and Support of the Free Zero Bias

Theorem. Let X be a mean 0, non-constant, L^2 random variable. Let μ denote the distribution of X and let μ° denote the distribution of the free zero bias X°. Then for any compact interval $[a,b] \subset \mathbb{R}$,

$$(\mu^\circ([a,b]))^2 \leq (b-a)\mathbb{E}[|X|].$$

Consequently, μ° is absolutely continuous with respect to Lebesgue measure.

This mirrors a (until now unknown) continuity property of the classical zero bias:

$$(\mu^*([a,b])) \leq (b-a)\mathbb{E}[|X|].$$

Moreover: $\operatorname{supp}\mu^\circ$ is contained in the convex hull of $\operatorname{supp}\mu$.

Continuity and Support of the Free Zero Bias

Theorem. Let X be a mean 0, non-constant, L^2 random variable. Let μ denote the distribution of X and let μ° denote the distribution of the free zero bias X°. Then for any compact interval $[a,b] \subset \mathbb{R}$,

$$(\mu^\circ([a,b]))^2 \le (b-a)\mathbb{E}[|X|].$$

Consequently, μ° is absolutely continuous with respect to Lebesgue measure.

This mirrors a (until now unknown) continuity property of the classical zero bias:

$$(\mu^*([a,b])) \le (b-a)\mathbb{E}[|X|].$$

Moreover: $\operatorname{supp}\mu^\circ$ is contained in the convex hull of $\operatorname{supp}\mu$.

The proofs use the relation $X^* = UX^\circ + (1-U)Y^\circ$, together with several integral representations of the free difference quotient.

Regularity Properties of the Free Zero Bias

Theorem. Let X be a mean 0, non-constant, L^2 random variable with variance σ^2. Then

$$\mathbb{E}[Xf(X)] = \sigma^2\,\mathbb{E}\otimes\mathbb{E}[\partial f(X^\circ)], \qquad f \in C_b^1(\mathbb{R}).$$

Moreover, the following hold.

- For any constant $\alpha \neq 0$, $(\alpha X)^\circ = \alpha X^\circ$.
- If $X_n \rightharpoonup X$ and $\mathrm{Var}[X_n] \to \mathrm{Var}[X] > 0$, then $X_n^\circ \rightharpoonup X^\circ$.
- $X^\circ \overset{d}{=} Y^\circ$ if and only if $X^\square \overset{d}{=} Y^\square$.

Regularity Properties of the Free Zero Bias

Theorem. Let X be a mean 0, non-constant, L^2 random variable with variance σ^2. Then

$$\mathbb{E}[X f(X)] = \sigma^2 \, \mathbb{E} \otimes \mathbb{E}[\partial f(X^\circ)], \qquad f \in C_b^1(\mathbb{R}).$$

Moreover, the following hold.

- For any constant $\alpha \neq 0$, $(\alpha X)^\circ = \alpha X^\circ$.
- If $X_n \rightharpoonup X$ and $\mathrm{Var}[X_n] \to \mathrm{Var}[X] > 0$, then $X_n^\circ \rightharpoonup X^\circ$.
- $X^\circ \overset{d}{=} Y^\circ$ if and only if $X^\Box \overset{d}{=} Y^\Box$.

However, the free sum exchange property (probably?) does not hold: if X, Y, X° are all coupled to be freely independent,

$$(X+Y)^\circ \overset{d}{\neq} X^\circ + Y \quad \text{(???)}$$

But there is an analog.

Cauchy Transforms, Reciprocals, and Inverses

For any probability distribution μ, its Cauchy transform $G_\mu(z) = \int \frac{\mu(dx)}{z-x}$ is analytic in $\mathbb{C}_+$. It is univalent (analytically invertible) in a truncated cone $\{z \in \mathbb{C}_+ \colon |z| > r, \mathrm{Im} z > \alpha |\mathrm{Re} z|\}$, with image contained in a similar truncated cone.

The R**-transform**: $R_\mu(z) = G_\mu^{\langle -1 \rangle}(z) - \frac{1}{z}$.

Also useful: the **reciprocal Cauchy transform** $F_\mu = 1/G_\mu$, and the **Voiculescu transform** $\varphi_\mu(z) = R_\mu(1/z)$. In their terms, $\varphi_\mu(z) = F_\mu^{-1}(z) - z$.

Cauchy Transforms, Reciprocals, and Inverses

For any probability distribution μ, its Cauchy transform $G_\mu(z) = \int \frac{\mu(dx)}{z-x}$ is analytic in $\mathbb{C}_+$. It is univalent (analytically invertible) in a truncated cone $\{z \in \mathbb{C}_+ \colon |z| > r, \mathrm{Im} z > \alpha|\mathrm{Re} z|\}$, with image contained in a similar truncated cone.

The R**-transform**: $R_\mu(z) = G_\mu^{\langle -1 \rangle}(z) - \frac{1}{z}$.

Also useful: the **reciprocal Cauchy transform** $F_\mu = 1/G_\mu$, and the **Voiculescu transform** $\varphi_\mu(z) = R_\mu(1/z)$. In their terms, $\varphi_\mu(z) = F_\mu^{-1}(z) - z$.

Given two probability distributions μ, ν, their **subordinator** is the analytic function

$$\omega_{\mu,\nu} = G_\mu^{\langle -1 \rangle} \circ G_{\mu \boxplus \nu}.$$

I.e. the defining equation is $G_\mu(\omega_{\mu,\nu}(z)) = G_{\mu\boxplus\nu}(z)$; this actually defines $\omega_{\mu,\nu}$ everywhere on $\mathbb{C}_+$.

Fact. If $\mu_n \rightharpoonup \mu$ and $\nu_n \rightharpoonup \nu$, then $\omega_{\mu_n,\nu_n} \to \omega_{\mu,\nu}$ uniformly on compact subsets of $\mathbb{C}_+$. (Folklore known for decades; proof in our paper, using Montel's theorem.)

Free Conditioning, and Free Zero Bias Exchange Property

Biane showed that the subordinator function plays an important role in free conditional expectation. In particular, with $f_z(x) = \frac{1}{z-x}$, if X, Y are freely independent then

$$\mathbb{E}[f_z(X+Y)|X] = \frac{1}{\omega_{\mu,\nu}(z) - X}.$$

Free Conditioning, and Free Zero Bias Exchange Property

Biane showed that the subordinator function plays an important role in free conditional expectation. In particular, with $f_z(x) = \frac{1}{z-x}$, if X, Y are freely independent then

$$\mathbb{E}[f_z(X+Y)|X] = \frac{1}{\omega_{\mu,\nu}(z) - X}.$$

We use this to prove the following.

Proposition. Let $X_1, \ldots, X_n$ be f.i.d. selfadjoint centered L^2 random variables, and let $W = X_1 + \cdots + X_n$. Then

$$G_{W^\circ}(z) = G_{X_1^\circ}(\omega_{X_1, W-X_1}(z)).$$

For comparison: the independent sum exchange property of the (classical) zero bias, in terms of Fourier transforms ψ, says

$$\psi_{W^*}(\xi) = \psi_{X_1^*}(\xi) \cdot \psi_{W-X_1}(\xi).$$

Free Infinite Divisibility

A distribution μ is called $\boxplus$-infinitely divisible if, for all n, there are f.i.d. random variables $X_1, \ldots, X_n$ with $X_1 + \cdots + X_n \stackrel{d}{=} \mu$.

In 1992-1993, in two landmark papers, Bercovici and Voiculescu characterized $\boxplus$-infinitely divisible distributions.

There is an analog to Kolmogorov's Lévy–Khinchine formula in the classical case. If X has mean m and finite variance σ^2, then X is $\boxplus$-infinitely divisible if and only if there is some probability measure ν such that

$$\varphi_X(z) = m + \sigma^2 \int \frac{1}{z - x} \nu(dx) = m + \sigma^2 G_\nu(z).$$

Free Infinite Divisibility

A distribution μ is called ⊞-infinitely divisible if, for all n, there are f.i.d. random variables $X_1, \ldots, X_n$ with $X_1 + \cdots + X_n \stackrel{d}{=} \mu$.

In 1992-1993, in two landmark papers, Bercovici and Voiculescu characterized ⊞-infinitely divisible distributions.

There is an analog to Kolmogorov's Lévy–Khinchine formula in the classical case. If X has mean m and finite variance σ^2, then X is ⊞-infinitely divisible if and only if there is some probability measure ν such that

$$\varphi_X(z) = m + \sigma^2 \int \frac{1}{z-x}\,\nu(dx) = m + \sigma^2 G_\nu(z).$$

Note that this formula gives an analytic continuation for φ_X (and hence R_X) to all of $\mathbb{C}_+$; Bercovici–Voiculescu proved the converse, that the existence of such analytic continuation also implies ⊞-infinite divisibility.

Free Zero Bias and $\boxplus$-Infinite Divisibility

Theorem. Let X be a selfadjoint L^2 random variable with mean m and variance $\sigma^2 > 0$. Then X is $\boxplus$-infinitely divisible if and only if there is a selfadjoint random variable Y such that

$$F_{(X-m)^\circ+m}(z) = F_{Y^\flat}(F_X(z))$$

Free Zero Bias and $\boxplus$-Infinite Divisibility

Theorem. Let X be a selfadjoint L^2 random variable with mean m and variance $\sigma^2 > 0$. Then X is $\boxplus$-infinitely divisible if and only if there is a selfadjoint random variable Y such that

$$F_{(X-m)^\circ+m}(z) = F_{Y^\flat}(F_X(z)) \quad \text{i.e.}$$
$$G_{(X-m)^\circ+m}(z) = G_{Y^\flat}(1/G_X(z)).$$

Moreover, this holds if and only if

$$\varphi_X(z) = m + \sigma^2 \int \frac{1}{z-x}\,\nu(dx)$$

where ν is the distribution if Y; i.e. $\varphi_X(z) = m + \sigma^2 G_Y(z)$.

Free Zero Bias and ⊞-Infinite Divisibility

Theorem. Let X be a selfadjoint L^2 random variable with mean m and variance $\sigma^2 > 0$. Then X is ⊞-infinitely divisible if and only if there is a selfadjoint random variable Y such that

$$F_{(X-m)^\circ+m}(z) = F_{Y^\flat}(F_X(z)) \quad \text{i.e.}$$
$$G_{(X-m)^\circ+m}(z) = G_{Y^\flat}(1/G_X(z)).$$

Moreover, this holds if and only if

$$\varphi_X(z) = m + \sigma^2 \int \frac{1}{z-x}\,\nu(dx)$$

where ν is the distribution if Y; i.e. $\varphi_X(z) = m + \sigma^2 G_Y(z)$.

This is the free version of the Goldstein–Shmock result on classical infinite divisibility equivalent to $(X-m)^* + m \stackrel{d}{=} X + UY$, where the distribution of Y is the measure ν.

What is the Lévy Measure

Our reformulation of the (free) Lévy–Khintchine formula gives some alternative meaning to the Lévy measure, which is the law of the random variable Y for which

$$F_{(X-m)^{\circ}+m}(z) = F_{Y^{\flat}}(F_X(z)).$$

More instructive is the way this arises in our proof.

What is the Lévy Measure

Our reformulation of the (free) Lévy–Khintchine formula gives some alternative meaning to the Lévy measure, which is the law of the random variable Y for which

$$F_{(X-m)^\circ+m}(z) = F_{Y^\flat}(F_X(z)).$$

More instructive is the way this arises in our proof.

Theorem. Let X be (classically or freely) infinitely divisible. For each n, write $X = X_{n,1} + \cdots + X_{n,n}$ for (freely) independent identically distributed $X_{n,j}$. Then

$$X_{n,n}^{\square} \to Y \quad \text{in distribution as} \quad n \to \infty.$$

All Probability Distributions are (Free) Lévy Measures

In the representation $\varphi_X(z) = m + \sigma^2 G_Y(z)$ for $\boxplus$-infinitely divisible X, it was not known exactly which probability measures actually arise as Lévy measures (i.e. which Y's appear this way). In fact, this was not known in the classical case either. Our methods provide the definitive answer, in both cases.

Theorem. Given *any* random variable Y, and any $m \in \mathbb{R}$ and $\sigma^2 > 0$, there is a (unique up to distribution) $\boxplus$-infinitely divisible random variable X satisfying

$$\varphi_X(z) = m + \sigma^2 G_Y(z).$$

All Probability Distributions are (Free) Lévy Measures

In the representation $\varphi_X(z) = m + \sigma^2 G_Y(z)$ for $\boxplus$-infinitely divisible X, it was not known exactly which probability measures actually arise as Lévy measures (i.e. which Y's appear this way). In fact, this was not known in the classical case either. Our methods provide the definitive answer, in both cases.

Theorem. Given *any* random variable Y, and any $m \in \mathbb{R}$ and $\sigma^2 > 0$, there is a (unique up to distribution) $\boxplus$-infinitely divisible random variable X satisfying

$$\varphi_X(z) = m + \sigma^2 G_Y(z).$$

We first show how to achieve this under the condition $\mathbb{E}[Y^{-2}] < \infty$, with X's that are limits of compound free Poisson random variables; then we remove the negative moment condition with a cutoff approximation.

The Compactly-Supported Case

Our proof of the Free Lévy–Khinchine formula is quite different from the Bercovici–Voiculescu proof, and gives new insight into the nature of (free) Lévy measures. Our proof is not *quite* self-contained: in one convergence proof, we need a priori knowledge of the fact that φ_X has an analytic continuous to $\mathbb{C}_+$ when X is $\boxplus$-infinitely divisible.

We can, however, circumvent this argument when X is *compactly-supported*. Here, a compactness argument yields the tightness required for the proof without more advanced analytic techniques. To make this work, we needed to prove a uniformity result which again is probability folklore but doesn't seem to be proved anywhere in writing (until now).

Lemma. Let μ be a $\boxplus$-infinitely divisible random variable, supported in $[-R,R]$, with variance σ^2. For each n, let μ_n be its $\boxplus$ nth root: $\mu_n^{\boxplus n}=\mu$. Then $\operatorname{supp}\mu_n\subseteq[-R-\sigma^2-1,R+\sigma^2+1]$ for all n.

Things I Didn't Mention

As in the classical case, we can use the free zero bias to construct free Stein kernels, under an absolute continuity assumption. Unlike the classical case, free Stein kernels always exist (as shown by Fathi–Nelson, and later by Cébron–Fathi–Mai) and are *never* unique. The free zero bias always exists, and the free Stein kernel so constructed is *different* from any of those found before.

Things I Didn't Mention

As in the classical case, we can use the free zero bias to construct free Stein kernels, under an absolute continuity assumption. Unlike the classical case, free Stein kernels always exist (as shown by Fathi–Nelson, and later by Cébron–Fathi–Mai) and are *never* unique. The free zero bias always exists, and the free Stein kernel so constructed is *different* from any of those found before.

A current goal is to use the free zero bias to prove new sharp estimates on semicircular approximation (i.e. quantitative bounds in free CLTs). The subordination-flavored replacement for the free sum exchange property makes this more challenging than the classical case. *Stay tuned!*