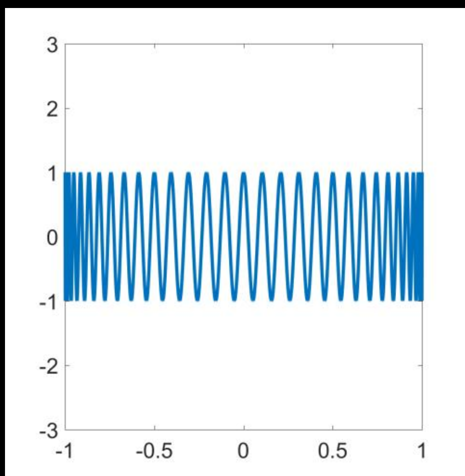
Three Perspectives on Orthogonal Polynomials

Paul Valiant

Brown University

(Based on joint work with Gregory Valiant, mostly STOC'11: "Estimating the Unseen: An n/log(n)-sample Estimator for Entropy and Support Size, Shown Optimal via New CLTs")

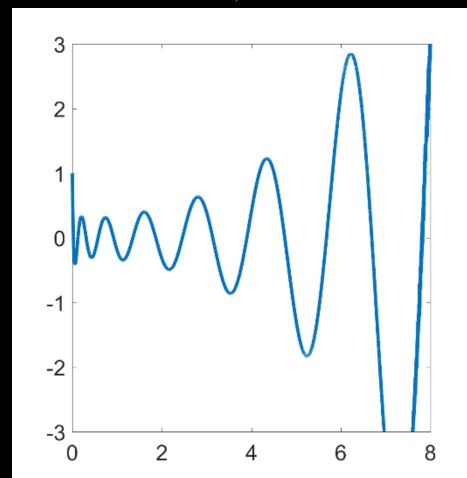
Structure of this talk:
3 polynomial challenges…
and solutions



Chebyshev

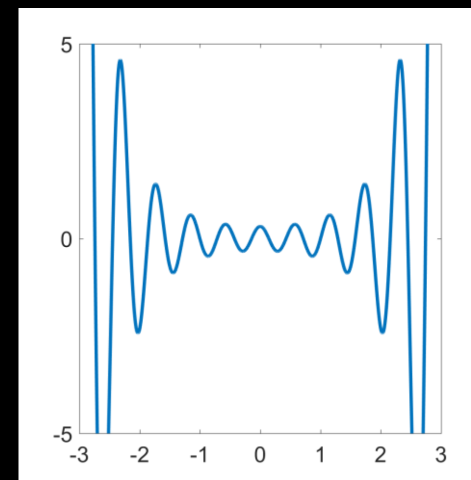
Seems like:

“cosine”



Laguerre

“cosine times exponential”



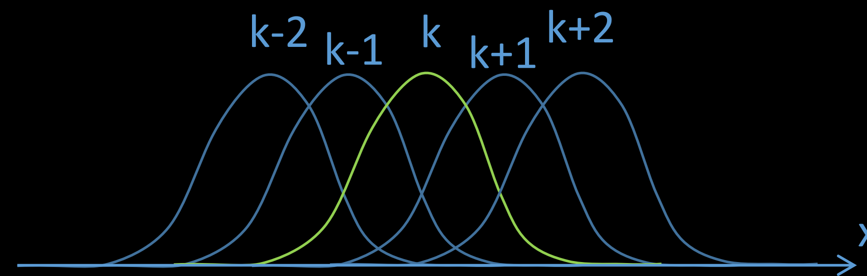
Hermite

“cosine over Gaussian”

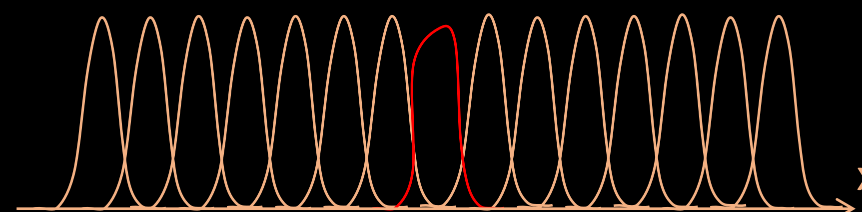
“Seems,” madam? Nay, *it is*. I know not “seems.” – Hamlet

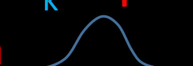
Challenge 1: Poisson bumps → thinnest bumps

$$poi(\lambda, k) = \frac{\lambda^k e^{-k}}{k!}$$



Linear transform ↓ ?

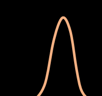


Motivation: Given an event with probability p, $poi(pn, k)$ captures the probability of it occurring exactly k times in Poi(n) samples. Let $\mathcal{F}_k$ be total number of events that were observed k times. $\mathcal{F}_k$ captures probabilities from . Is there a linear combination of $\mathcal{F}_k$ that captures ?

x-resolution

1. Thin as possible

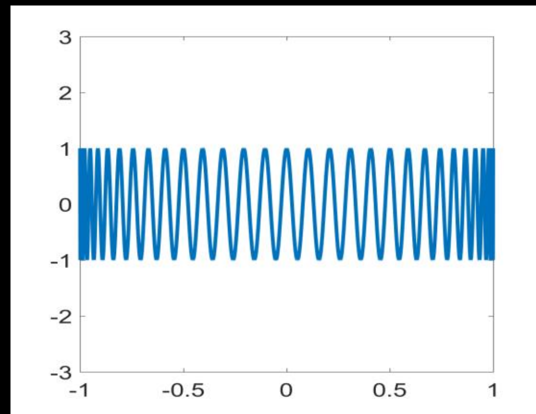
(with bounded coeffs)

2. $\sum$  ≈ 1

y-resolution

Thm: general log n factor improvement in resolution, #samples

Chebyshev Polynomials



$$T_j(\cos x) = \cos(jx)$$

Chebyshev is exactly like cosine,
except on distorted x-axis

x-resolution

1. Thin as possible

(with bounded ~~coeffs~~)

j

Both unchanged under
x-axis distortion!

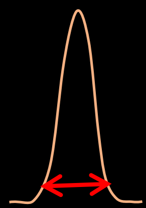
2. $\sum$ =1

y-resolution

New question: thin cosine bumps

Thinnest Cosine Bumps

Thinnest linear combination of $\cos(jx)$ for $j < b$:



$\approx 1/b$

(Intuition: Fourier transform of degree b gives resolution 1/b)

x-resolution

1. Thin as possible

(with bounded ~~coeffs~~)

j

Sum of all possible x-translated bumps is constant

2. $\sum$ =1

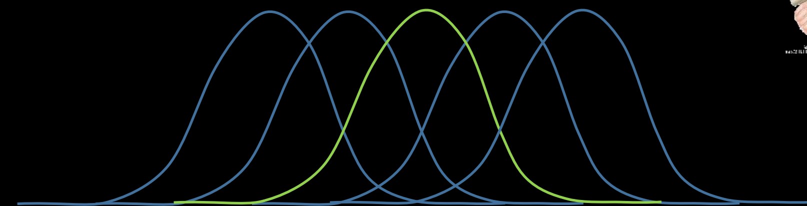
y-resolution

(Trig functions are well-behaved under x-translation)

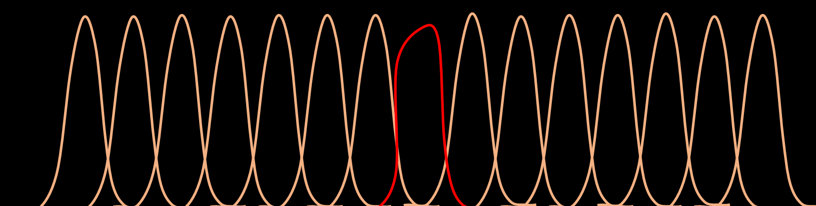
Chebyshev Takeaways:

(Modulo x-axis distortion) "polynomials are cosines"

$$poi(\lambda, k) = \frac{\lambda^k e^{-k}}{k!}$$



Linear transform



Motivation: Given an event with probability p, $poi(pn, k)$ captures the probability of it occurring exactly k times in Poi(n) samples. Let $\mathcal{F}_k$ be total number of events that were observed k times. $\mathcal{F}_k$ captures probabilities from . Is there a linear combination of $\mathcal{F}_k$ that captures ?

Thm: general log n factor improvement in resolution, #samples

x-resolution

1. Thin as possible

b times thinner

(with coeffs << n)

2. $\sum$ $\approx$1

exp(b)

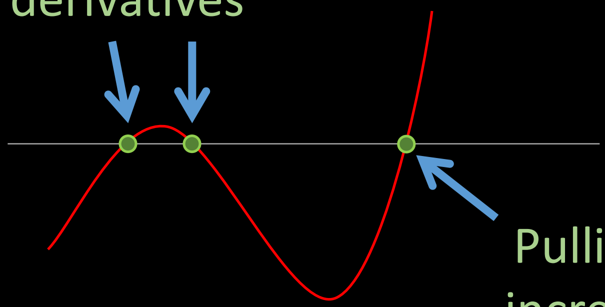
y-resolution

Thus: $b = \theta(\log n)$

Challenge 2: Exponentially Growing Derivatives

Find: Degree j polynomial with roots at ε, 2ε; and all remaining roots have much larger derivative, growing exponentially with x

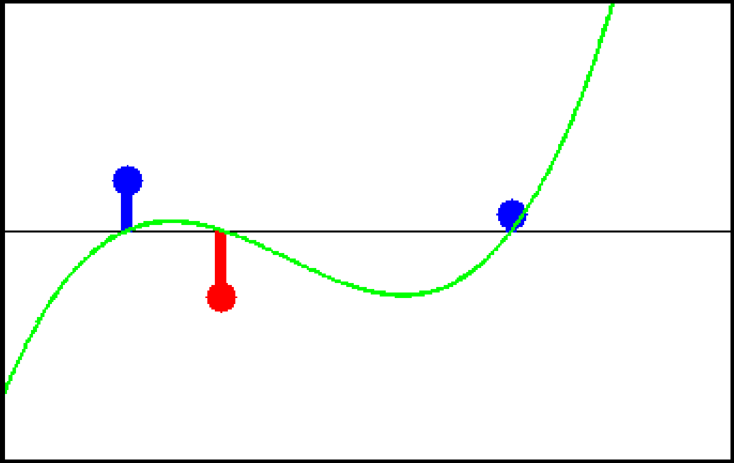
Roots close together have small derivatives



Pulling a root farther away increases its derivative, but only polynomially

Success requires a delicate balancing act!

Orthogonal to Polynomials



Motivation: Want to construct a pair of distributions g^+, g^- that are, respectively, close to the uniform distributions on T and 2T elements, but where for each (small) k, the expected number of domain elements seen k times from Poi(n) samples is identical for g^+, g^-.

Fact: If P is a degree j polynomial with distinct real roots $\{x_i\}$, then the signed measure h_P having point mass $1/P'(x_i)$ at each root x_i is orthogonal to all polynomials of degree ≤j-2

Essentially: find a signed measure g(x) that is

1) Orthogonal to $poi(x,k) = \frac{x^k e^{-k}}{k!}$ for each small k,
2) Has most of its positive mass at 1/T and most of its negative mass at 1/(2T)

$$g(x) \triangleq e^x h(x)$$

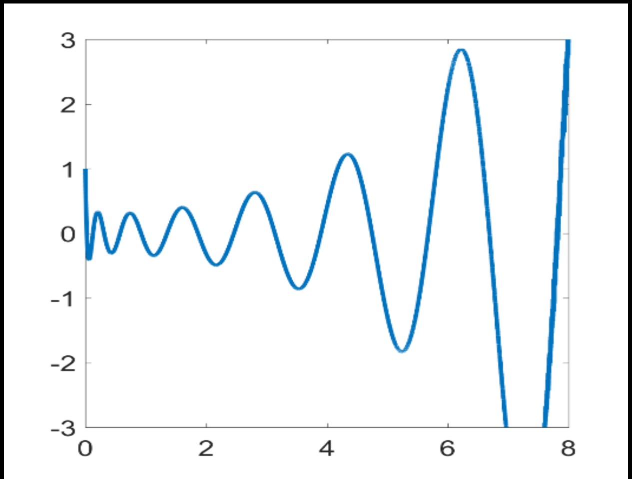
Essentially: find a signed measure $h(x)$ that is

1) Orthogonal to all degree ≤k polynomials
2) Has most of its positive mass at 1/T and most of its negative mass at 1/(2T)
   - and otherwise decays $\ll e^{-x}$

Task: find P such that $P'(x_i)$ grows exponentially in x_i

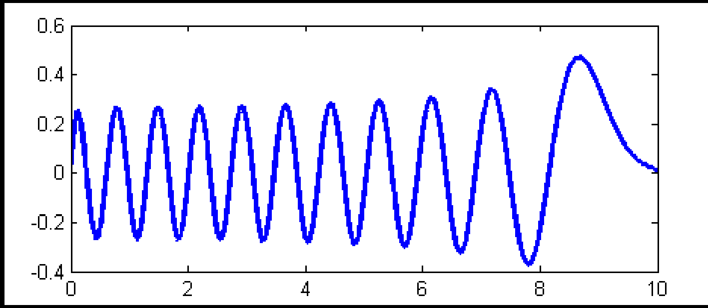
Laguerre Polynomials

Defined by $L_n(x) = e^x \frac{d^n}{dx^n} \frac{e^{-x} x^n}{n!}$ and
orthogonal as: $\int_0^\infty L_n(x) L_m(x) e^{-x} dx = [m = n]$



Why should the derivative be so nicely behaved at its roots, in particular, growing exponentially?

Transform the Laguerre: $v = e^{-x^2/2} \sqrt{x} \cdot L_n(x^2)$



Many differential equations, including

$$v'' + \left(4n + 2 - x^2 + \frac{1}{4x^2}\right) v = 0$$

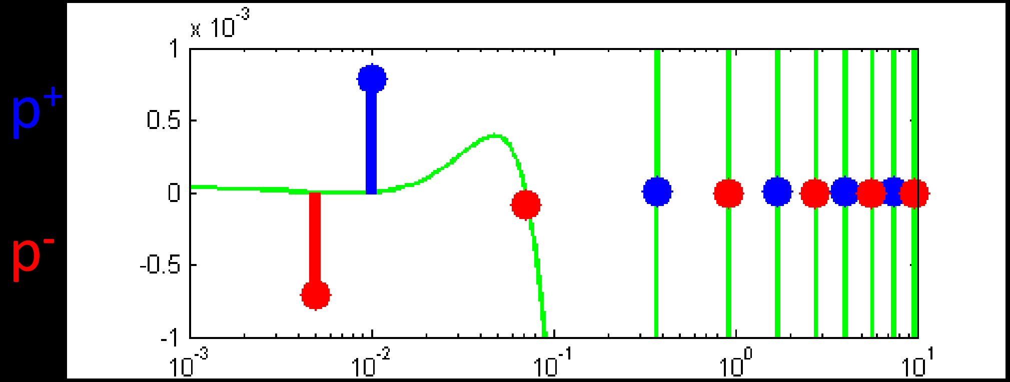
Almost harmonic motion, v→sine

Nicely spaced zeros, and max derivative at the zeros

The Construction

Recall: We want a *signed measure* g on the positive reals that:

- Is orthogonal to low degree polynomials
- Decays exponentially fast
- Its positive portion has most of its mass at 2ϵ
- Its negative portion has most of its mass at ϵ



Theorem: p^+ is "close" to $U_{n/2}$, and p^- is "close" to U_n , and p^+ and p^- are indistinguishable via *cn/log n* samples

(Modulo diff-eq distortion) "polynomials are $e^x \sin(x)$"

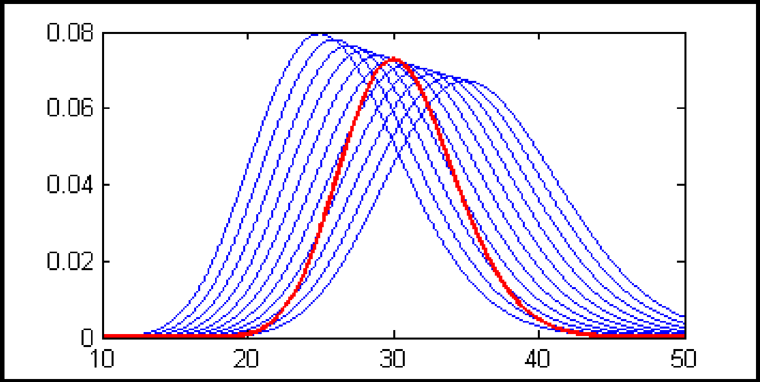
Challenge 3: exponentially good bump approximations

Motivation: Previously, constructed lowerbound distributions g^+, g^- where expectation of every measurement matched. Lower bound? No… until we show variances match too.

Aim: show that variances can be approximated as linear combinations of expectations, with moderate coefficients; thus matching means implies matching variances. Since means come from poi(j,x), second moments come from $poi(j,x)^2$.

Find a linear combination over j of poi(x,j) that approximates $poi(x,k)^2$ to within ε, using coefficients $\leq 1/\varepsilon$

Think of $\varepsilon = 1/\exp(j)$



These look like Gaussians!

1) What's the answer for Gaussians?

2) Analyze via Hermite polynomials instead

Approximating "Thin" Gaussians as Linear Combinations of Gaussians

What do we convolve a Gaussian with to approximate a *thinner* Gaussian?

(Other direction is easy, since convolving Gaussians adds their variances)

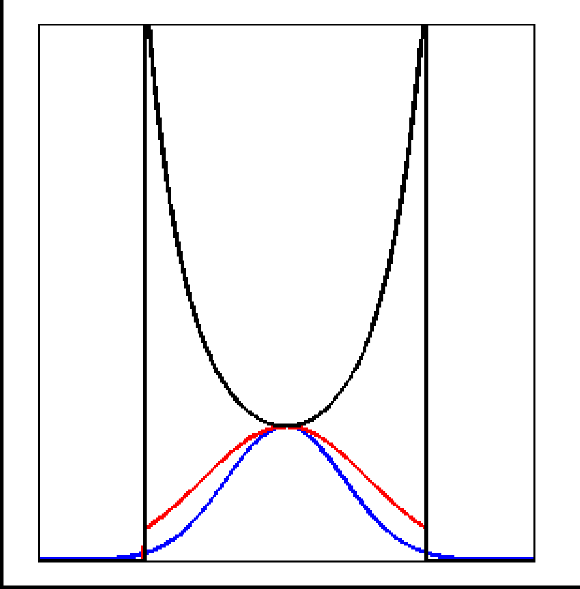
"Blurring is easy, unblurring is hard" → can only do it approximately

How to analyze? Fourier transform! Convolution becomes multiplication

Now: what do we *multiply* a Gaussian with to approximate a *fatter* Gaussian?

$$e^{-x^2} \cdot ??? = e^{-x^2/2}$$

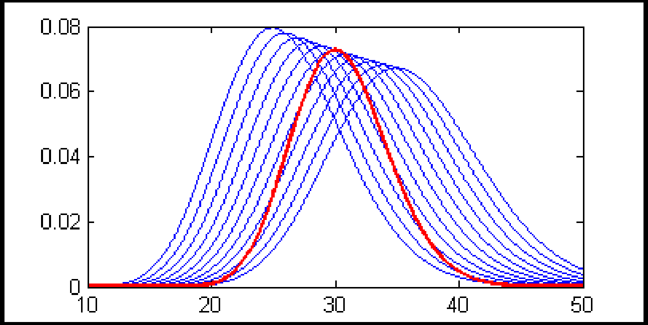
$e^{x^2/2}$ Problem: blows up



Answer: if we want to approximate to within ϵ, we only need to approximate out to where $e^{-x^2/2} = \epsilon$. How big is $e^{x^2/2}$ here? $1/\epsilon$

Result: Can approximate to within ϵ using coefficients no bigger than $1/\epsilon$

Hermite Polynomials



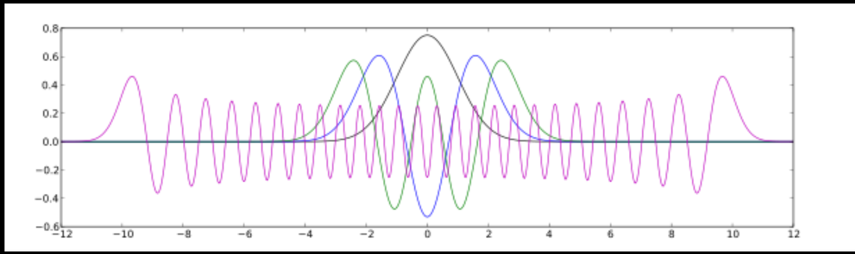
Poissons seem a lot like Gaussians

(I was stuck here for about a month)

Idea: $x \rightarrow y^2$

$$\frac{x^j e^{-x}}{j!} = \frac{y^{2j} e^{-y^2}}{j!} \xrightarrow{fourier} \frac{1}{j!} \frac{d^{2j}}{dw^{2j}} e^{-w^2} = \frac{1}{j!} H_{2j}(w) e^{-w^2}$$

Hermite Polynomials! Sequence of orthogonal polynomials, from which we take the even ones: orthogonal basis for even functions

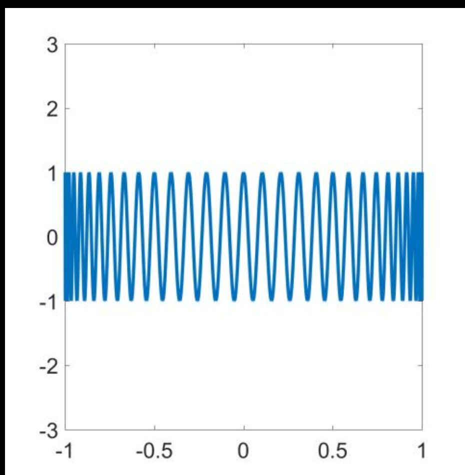


To express any function in this basis, just compute each coefficient as an inner product

Which function? Fourier transform of "thin" Poisson, cut off at ϵ

Proposition: Can approximate $\Pr[Poi(2\lambda) = k]$ to within ϵ as a linear combination $\sum_j \alpha_{k,j} \Pr[Poi(\lambda) = j]$ with coefficients that sum to $\sum_j |\alpha_{k,j}| \leq \frac{1}{\epsilon} 200 \max\{\sqrt[4]{k}, 24 \log^{\frac{3}{2}} \frac{1}{\epsilon}\}$

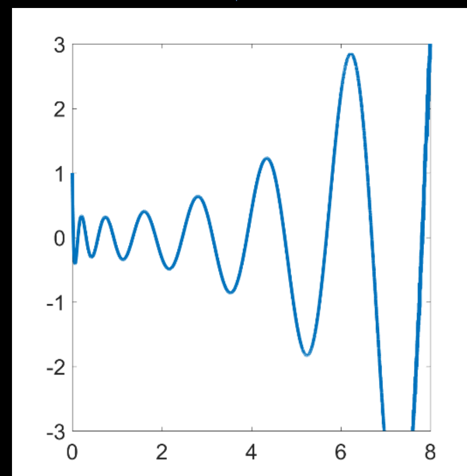
Structure of this talk:
3 polynomial challenges…
and solutions



Chebyshev

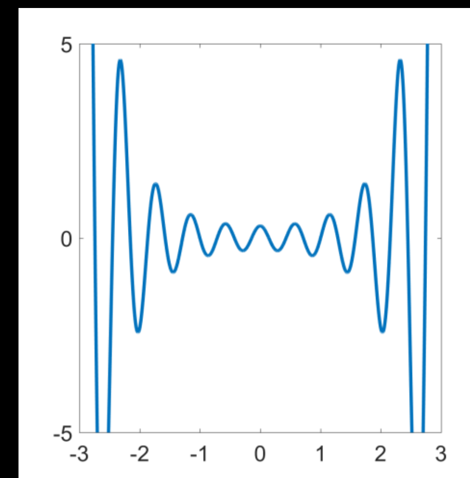
Seems like:

“cosine”



Laguerre

“cosine times exponential”



Hermite

“cosine over Gaussian”

“Seems,” madam? Nay, *it is*. I know not “seems.” — Hamlet