Gaussian Stein envelopes and factors: a kernel calculus and a guide by test class

Ferdinand Rapin* Yvik Swan†

Preprint — September 2026

Abstract

We give a results-first account of pointwise envelopes and uniform Stein factors for the standard Gaussian Stein equation. The bounds are organised by test class: Kolmogorov, bounded measurable, Wasserstein, Fortet–Mourier, and integer-order Zolotarev. Each class is treated at solution derivative orders zero, one, two, and general order, with explicit distinctions between classical, almost-everywhere, and piecewise derivatives. An array of signed kernels yields exact pointwise envelopes for the principal single-seminorm classes and recovers their known optimal uniform factors. For the Fortet–Mourier class we provide convenient bounds under both common norm conventions, together with exact variational envelopes at orders zero and one. Polynomial and localised smooth examples explain why the remaining single-seminorm bounds cannot hold. The proofs retain the boundary and correction terms needed for a valid higher-order calculus.

Keywords: Gaussian approximation; Stein equation; Stein factors; pointwise envelopes; probability metrics; kernel representation.

MSC 2020: 60F05, 60E15.

1 Results at a glance and by test class

1.1 Equation, conventions, and how to use the bounds

Let $N \sim \mathcal{N}(0,1)$, with density φ, distribution function Φ, and survival function $\overline{\Phi} = 1 - \Phi$. For $\mathbb{E}|h(N)| < \infty$, put

$$f_h(x) = \frac{1}{\varphi(x)} \int_{-\infty}^{x} \{h(v) - \mathbb{E}h(N)\}\varphi(v)\, dv. \tag{1.1}$$

This canonical solution satisfies

$$f_h'(x) - x f_h(x) = h(x) - \mathbb{E}h(N) \tag{1.2}$$

almost everywhere, and everywhere if h is continuous. Adding a constant to h does not change f_h. For bounded or Lipschitz h, this is the unique bounded solution. For higher-order classes, f_h need not be bounded; (1.1) remains our convention.

For bounded test functions, $\|h\|_\infty = \sup_x |h(x)|$ and $\operatorname{osc}(h) = \sup_x h(x) - \inf_x h(x)$. For measurable derivatives, the norm is the essential supremum. The pointwise convention for test functions also controls their values at atoms of a compared distribution. In particular, $h^{(0)} = h$:

*Département de mathématiques, Université Libre de Bruxelles.

†Département de mathématiques, Université Libre de Bruxelles; Vrije Universiteit Brussel.

order zero never denotes an oscillation seminorm. For $r \geq 1$, the assumption "$\|h^{(r)}\|_\infty < \infty$" means that $h \in C^{r-1}(\mathbb{R})$, $h^{(r-1)}$ is locally absolutely continuous, and its almost-everywhere derivative is essentially bounded. Thus $h^{(r-1)}$ is globally Lipschitz. Such h has at most polynomial growth and is Gaussian integrable.

An *envelope* $E_n(x)$ gives $|f_h^{(n)}(x)| \leq E_n(x)$ for every test in the indicated normalised class. Its *uniform factor* is a constant C_n such that $\|f_h^{(n)}\|_\infty \leq C_n$. We call an envelope *exact* if it is the pointwise supremum over the class, with limits allowed. Derivative assertions hold everywhere through the continuous orders and almost everywhere at the highest weak derivative. Exactness there can equivalently be tested using functions smooth near the point. For Kolmogorov tests, higher derivatives are explicitly piecewise.

The following functions occur throughout:

$$A(x) = \frac{\Phi(x)\overline{\Phi}(x)}{\varphi(x)}, \qquad a_0 = \sqrt{\frac{\pi}{8}}, \tag{1.3}$$

$$\Phi_0 = \overline{\Phi}_0 = \varphi, \qquad \Phi_j(x) = \int_{-\infty}^{x} \Phi_{j-1}(v)\,dv, \quad \overline{\Phi}_j(x) = \int_{x}^{\infty} \overline{\Phi}_{j-1}(v)\,dv \quad (j \geq 1), \tag{1.4}$$

$$D_n(x) = \frac{2n!\Phi_{n+1}(x)\overline{\Phi}_{n+1}(x)}{\varphi(x)}, \qquad d_n = \frac{\Gamma((n+1)/2)}{\sqrt{2}\,\Gamma(n/2+1)} \quad (n \geq 1). \tag{1.5}$$

In particular,

$$D_1(x) = \frac{2(\varphi(x) + x\Phi(x))(\varphi(x) - x\overline{\Phi}(x))}{\varphi(x)}, \qquad d_1 = \sqrt{2/\pi}, \tag{1.6}$$

$$D_2(x) = \frac{\{(1+x^2)\Phi(x) + x\varphi(x)\}\{(1+x^2)\overline{\Phi}(x) - x\varphi(x)\}}{\varphi(x)}, \qquad d_2 = \sqrt{\pi/8}. \tag{1.7}$$

We have $\sup_x A(x) = A(0) = a_0$ and $\sup_x D_n(x) = D_n(0) = d_n$. Although $D_0 = 2A$ by the same formula, the oscillation-normalised envelope is A, not D_0.

Normalised test class	$\|f_h\|_\infty$	$\|f_h'\|_\infty$	$\|f_h''\|_\infty$
Kolmogorov: $h = \mathbf{1}_{(-\infty,z]}$	$\sqrt{\pi/8}$	1	no finite factor†
Bounded: $\operatorname{osc}(h) \leq 1$	$\sqrt{\pi/8}$	1	no finite factor
Bounded: $\|h\|_\infty \leq 1$	$\sqrt{\pi/2}$	2	no finite factor
Wasserstein: $\operatorname{Lip}(h) \leq 1$	1	$\sqrt{2/\pi}$	2
Fortet–Mourier: $\max(\|h\|_\infty, \operatorname{Lip}(h)) \leq 1$	≤ 1	$\leq \sqrt{2/\pi}$	≤ 2
Zolotarev $r = 2$: $\|h''\|_\infty \leq 1$	no finite factor	1/2	$\sqrt{\pi/8}$

Table 1: Uniform factors at the first three orders. Entries without $\leq$ are optimal. The Fortet–Mourier entries are useful upper bounds; optimality is not asserted. †For Kolmogorov tests the second derivative is considered away from the jump. All-order results and pointwise envelopes follow below.

Each subsection can be used on its own. Choose the test class and the required derivative order, then use the pointwise envelope if location information is available, or the uniform factor otherwise. The joint smoothness bounds in section 1.7 allow several available controls to be combined by taking their minimum. The higher-order uniform factors are known; see Gaunt [1] for their optimality and classification. Our presentation makes the pointwise kernel mechanism and the distinctions between the test classes explicit. All results in this paper concern the standard Gaussian target.

1.2 Kolmogorov tests: exact piecewise envelopes at every order

Let $h_z(v) = \mathbf{1}_{\{v \le z\}}$ and write $f_z = f_{h_z}$. The associated metric is $d_{\mathrm{K}}(X, N) = \sup_z |\Pr(X \le z) - \Phi(z)|$. Define, for every integer $n \ge 0$,

$$Q_n(x) = \frac{n!}{\varphi(x)} \max\{\overline{\Phi}(x)\Phi_{n+1}(x),\ \Phi(x)\overline{\Phi}_{n+1}(x)\}. \tag{1.8}$$

Theorem 1.1 (Kolmogorov envelopes)**.** *For $x \neq z$ and $n \ge 0$,*

$$f_z^{(n)}(x) = \frac{n!}{\varphi(x)} \begin{cases} \overline{\Phi}(z)\Phi_{n+1}(x), & x < z, \\ (-1)^n \Phi(z)\overline{\Phi}_{n+1}(x), & x > z. \end{cases} \tag{1.9}$$

Consequently $\sup_{z \neq x} |f_z^{(n)}(x)| = Q_n(x)$. At $n = 0$ the formula extends continuously to $x = z$ and the supremum is attained there. The first orders and their best uniform factors are

Order	Exact envelope	Best uniform factor
0	$Q_0(x) = A(x)$	$\sqrt{\pi/8}$
1	$Q_1(x) = \max\{\overline{\Phi}(x) + xA(x), \Phi(x) - xA(x)\}$	1
2	$Q_2(x) = (1 + x^2)A(x) + \lvert x\rvert\overline{\Phi}(\lvert x\rvert)$	$+\infty$
$n \ge 2$	$Q_n(x)$ from (1.8)	$+\infty$

For $n \ge 1$, $Q_n(x) \sim |x|^{n-1}$ as $|x| \to \infty$.

The exclusion $x = z$ matters: f_z is continuous, but $f_z'(z+) - f_z'(z-) = -1$. Thus the distributional second derivative contains a Dirac mass. The higher-order entries are bounds on the ordinary derivatives on each side of the jump, not global weak derivatives in L^∞. A finite piecewise envelope at a fixed x does not imply a finite factor uniform in x and z.

1.3 Bounded tests: oscillation and supremum-norm normalisations

For bounded measurable h, put $H_0 = \mathrm{osc}(h)$. The oscillation unit class induces total variation in the convention $d_{\mathrm{TV}}(X, N) = \sup_B |\Pr(X \in B) - \Pr(N \in B)|$. The class $\|h\|_\infty \le 1$ induces $2d_{\mathrm{TV}}$.

Theorem 1.2 (Bounded-test bounds)**.** *The following envelopes and factors are sharp:*

Order	Pointwise bound	Uniform bound
0	$\lvert f_h(x)\rvert \le A(x)H_0$	$\|f_h\|_\infty \le \sqrt{\pi/8}\, H_0$
1	$\lvert f_h'(x)\rvert \le H_0$ a.e.	$\|f_h'\|_\infty \le H_0$
2	no finite class envelope	no finite factor
$n \ge 2$	no finite class envelope	no finite factor

If only $\|h\|_\infty \le B$ is given, replace H_0 by $2B$. This gives the sharp envelopes $2BA(x)$ and $2B$, and the sharp factors $B\sqrt{\pi/2}$ and $2B$.

The negative statements hold even when the class is restricted to smooth bounded tests; they are stronger than the observation that derivatives may fail to exist. Boundedness alone supplies no control of f_h''. If derivatives of h are also controlled, use section 1.7.

1.4 Wasserstein tests: the three finite derivative orders

Let h be globally Lipschitz and set $L = \mathrm{Lip}(h) = \|h'\|_\infty$. The unit class gives the Wasserstein distance $d_{\mathrm{W}}(X, N) = \sup_{\mathrm{Lip}(h)\le 1} |\mathbb{E}h(X) - \mathbb{E}h(N)|$, provided X has a finite first absolute moment.

Theorem 1.3 (Wasserstein bounds)**.** *The solution is continuously differentiable, its first derivative is locally absolutely continuous, and*

Order	Exact envelope multiplied by L	Best uniform factor
0	L	1
1	$D_1(x)L$	$\sqrt{2/\pi}$
2	$2L$ a.e.	2
$n \ge 3$	no finite class envelope	no finite factor

All three finite envelopes and factors are sharp for the Lipschitz class. The negative statements remain true for smooth Lipschitz tests.

The order-zero extremiser is $h(v) = v$. At order one an extremiser at x is $h(v) = -|v - x|$. At order two, tests whose derivative is $+1$ in a shrinking neighbourhood of x and -1 elsewhere approach the envelope. These examples need not be bounded, which is relevant when using the next subsection.

1.5 Fortet–Mourier tests: simultaneous size and Lipschitz control

We first use the rectangular convention

$$\mathcal{H}_{B,L} = \{h : \|h\|_\infty \le B,\ \mathrm{Lip}(h) \le L\}, \qquad B, L \ge 0.$$

The Fortet–Mourier, or bounded-Lipschitz, unit class is $\mathcal{H}_{1,1}$. Its metric is finite without moment assumptions on the compared laws.

Proposition 1.4 (Immediately usable Fortet–Mourier bounds)**.** *For $h \in \mathcal{H}_{B,L}$,*

Order	Pointwise bound	Uniform bound
0	$\min\{2BA(x), L\}$	$\min\{B\sqrt{\pi/2}, L\}$
1	$\min\{2B, LD_1(x)\}$	$\min\{2B, L\sqrt{2/\pi}\}$
2	$2L$ a.e.	$2L$
$n \ge 3$	no finite class envelope if $B, L > 0$	no finite factor

These are valid upper bounds; simultaneous sharpness under the two constraints is not claimed. If $B = 0$ or $L = 0$, all solutions vanish.

Taking the minimum of two sharp bounds for larger classes does not in general give the exact envelope for their intersection. The following formulas exploit both constraints and give exact envelopes at the first two orders. They are also convenient for numerical evaluation.

For $B, L > 0$ and $a \in [-B, B]$, define

$$u_{a,x}(v) = \operatorname{clip}(a + L(x - v), -B, B), \tag{1.10}$$

$$F_0(B, L; x) = \max_{a \in [-B,B]} \left\{ \frac{\overline{\Phi}(x)}{\varphi(x)} \int_{-\infty}^{x} u_{a,x}(v)\varphi(v)\, dv - \frac{\Phi(x)}{\varphi(x)} \int_{x}^{\infty} u_{a,x}(v)\varphi(v)\, dv \right\}, \tag{1.11}$$

$$q_x(v) = \begin{cases} (1 - x\overline{\Phi}(x)/\varphi(x))\varphi(v), & v \le x, \\ (1 + x\Phi(x)/\varphi(x))\varphi(v), & v > x, \end{cases} \tag{1.12}$$

$$F_1(B, L; x) = \int_{\mathbb{R}} \min\{L|v - x|, 2B\}\, q_x(v)\, dv. \tag{1.13}$$

Here $\operatorname{clip}(t, -B, B) = \max(-B, \min(t, B))$, and q_x is a probability density. Then

$$\sup_{h \in \mathcal{H}_{B,L}} |f_h(x)| = F_0(B, L; x), \qquad \sup_{h \in \mathcal{H}_{B,L}} |f_h'(x)| = F_1(B, L; x). \tag{1.14}$$

Thus exact uniform factors, if desired, are $\sup_x F_0(B, L; x)$ and $\sup_x F_1(B, L; x)$. The explicit constants in proposition 1.4 bound them from above. No closed-form optimal constant for this intersection class is asserted here.

The sum-norm convention. Another common definition uses $\|h\|_\infty + \operatorname{Lip}(h) \le 1$. Under that convention the following bounds apply:

Order	Pointwise upper bound	Uniform upper bound
0	$\dfrac{2A(x)}{1 + 2A(x)}$	$\dfrac{\sqrt{\pi/2}}{1 + \sqrt{\pi/2}}$
1	$\dfrac{2D_1(x)}{2 + D_1(x)}$	$\dfrac{2\sqrt{2/\pi}}{2 + \sqrt{2/\pi}}$
2	2 a.e.	2
$n \ge 3$	no finite class envelope	no finite factor

These follow by maximising the preceding minimum bounds over $B + L \le 1$; they are not asserted to be optimal. Exact envelopes at orders zero and one can instead be obtained by taking $\sup_{0 \le B \le 1} F_j(B, 1 - B; x)$, $j = 0, 1$, with the zero cases interpreted by continuity. Always specify which convention is being used.

1.6 Integer-order Zolotarev tests: a complete order classification

For an integer $r \ge 1$, define

$$\mathcal{Z}_r = \{h : h^{(r-1)} \text{ is globally Lipschitz, } \|h^{(r)}\|_\infty \le 1\}, \tag{1.15}$$

with the derivative convention above. Lower derivatives are not bounded or anchored. The associated functional is

$$\zeta_r(X, N) = \sup_{h \in \mathcal{Z}_r} |\mathbb{E}h(X) - \mathbb{E}h(N)|.$$

If X has a finite rth absolute moment, this is finite when its moments of orders $1, \dots, r - 1$ equal those of N. A mismatch at any of these orders makes the supremum infinite. For $r = 1$, this is Wasserstein. This paper uses integer orders; fractional Hölder–Zolotarev classes are not included in the statements below.

Theorem 1.5 (Zolotarev envelopes and factors)**.** *For $h \in \mathcal{Z}_r$, the solution belongs to C^r, its rth derivative is locally absolutely continuous, and*

Solution order n	Exact envelope	Best uniform factor
$0 \le n \le r-2$	no finite class envelope	no finite factor
$r-1$	$1/r$	$1/r$
r	$D_r(x)$	d_r
$r+1$	2 a.e.	2
$n \ge r+2$	no finite class envelope	no finite factor

The first row is empty for $r = 1$. For general $H_r = \|h^{(r)}\|_\infty$, multiply the three finite bounds by H_r.

For quick reference, here are the envelopes / factors at orders zero, one, and two:

Class	Order 0	Order 1	Order 2
$\mathcal{Z}_1$	1 / 1	$D_1(x) \,/\, \sqrt{2/\pi}$	2 / 2
$\mathcal{Z}_2$	none	$\frac{1}{2} \,/\, \frac{1}{2}$	$D_2(x) \,/\, \sqrt{\pi/8}$
$\mathcal{Z}_3$	none	none	$\frac{1}{3} \,/\, \frac{1}{3}$
$\mathcal{Z}_r$, $r \ge 4$	none	none	none

Here "none" means no finite envelope under the stated seminorm alone, not an unproved or omitted bound. For example $h(v) = Mv$ belongs to $\mathcal{Z}_2$ for every M, but $f_h \equiv -M$. Moment matching allows such polynomial terms to cancel in the *metric*; it does not bound the corresponding individual Stein solutions.

1.7 Additional smoothness: combine the three neighbouring orders

Write $H_0 = \operatorname{osc}(h)$ and $H_j = \|h^{(j)}\|_\infty$ for $j \ge 1$. Each term in the next display is available separately under its own regularity assumption; omit a term when that control is not available:

$$|f_h(x)| \le \min\{A(x)H_0, H_1\}, \tag{1.16}$$
$$|f_h'(x)| \le \min\{H_0, D_1(x)H_1, \tfrac{1}{2}H_2\}, \tag{1.17}$$
$$|f_h''(x)| \le \min\{2H_1, D_2(x)H_2, \tfrac{1}{3}H_3\}, \tag{1.18}$$
$$|f_h^{(n)}(x)| \le \min\{2H_{n-1}, D_n(x)H_n, \tfrac{1}{n+1}H_{n+1}\}, \quad n \ge 2. \tag{1.19}$$

To obtain uniform bounds, replace A by a_0 and D_n by d_n. Each single-seminorm envelope is optimal. This does not assert optimality of the minimum for a class imposing several simultaneous restrictions.

The pointwise bounds also give weighted factors without any new proof. For a nonnegative weight w, set $\|g\|_{\infty,w} = \operatorname{ess\,sup}_x w(x)|g(x)|$. For example,

$$\|f_h^{(n)}\|_{\infty,w} \le \big(\operatorname*{ess\,sup}_x w(x)D_n(x)\big)H_n.$$

The diagonal weighted factor is exact whenever this supremum is finite, by pointwise sharpness and continuity near the evaluation point. The same observation applies to the other exact envelopes, with piecewise norms in the Kolmogorov case.

2 The Gaussian kernel calculus

2.1 Iterated tails and derivatives of the Mills ratios

The integrated tails in (1.4) admit the useful formulas

$$\Phi_j(x) = \frac{\mathbb{E}[(x-N)_+^{j-1}]}{(j-1)!}, \qquad \overline{\Phi}_j(x) = \frac{\mathbb{E}[(N-x)_+^{j-1}]}{(j-1)!} \quad (j \geq 1), \tag{2.1}$$

where the zeroth powers mean indicators of strict positivity. In particular, $\Phi_1 = \Phi$, $\overline{\Phi}_1 = \overline{\Phi}$, and

$$\Phi_2(x) = x\Phi(x) + \varphi(x), \qquad \overline{\Phi}_2(x) = -x\overline{\Phi}(x) + \varphi(x), \tag{2.2}$$

$$(j+1)\Phi_{j+2}(x) = x\Phi_{j+1}(x) + \Phi_j(x), \quad (j+1)\overline{\Phi}_{j+2}(x) = -x\overline{\Phi}_{j+1}(x) + \overline{\Phi}_j(x), \quad j \geq 0. \tag{2.3}$$

The minus sign in the right-tail recurrence is essential.

Let $R = \overline{\Phi}/\varphi$ and $S = \Phi/\varphi$. Their integral representations give, for $n \geq 0$,

$$R^{(n)}(x) = (-1)^n \int_0^\infty t^n e^{-xt-t^2/2}\, dt = (-1)^n n! \frac{\overline{\Phi}_{n+1}(x)}{\varphi(x)}, \tag{2.4}$$

$$S^{(n)}(x) = \int_0^\infty t^n e^{xt-t^2/2}\, dt = n! \frac{\Phi_{n+1}(x)}{\varphi(x)}. \tag{2.5}$$

These follow by substituting $v = x \pm t$ into the Gaussian tails and differentiating under an integrable Gaussian majorant on compact sets of x. In particular $(-1)^n R^{(n)} > 0$ and $S^{(n)} > 0$ on $\mathbb{R}$.

2.2 Definition, signs, and masses of the kernels

For integers $i, j \geq 0$, define

$$K^{i,j}(x,v) = \frac{1}{\varphi(v)} \left\{ (-1)^j R^{(i)}(x) \Phi_j(v) \mathbf{1}_{\{v \leq x\}} - S^{(i)}(x) \overline{\Phi}_j(v) \mathbf{1}_{\{v > x\}} \right\}. \tag{2.6}$$

The indices i, j refer respectively to the solution derivative and the test-function derivative. Superscripts in parentheses always denote ordinary differentiation.

Lemma 2.1 (Kernel masses)**.** *For all $i, j \geq 0$,*

$$\mathbb{E}K^{i,j}(x,N) = \frac{i!}{\varphi(x)} \{ (-1)^{i+j} \overline{\Phi}_{i+1}(x) \Phi_{j+1}(x) - \Phi_{i+1}(x) \overline{\Phi}_{j+1}(x) \}, \tag{2.7}$$

$$\mathbb{E}|K^{i,j}(x,N)| = \frac{i!}{\varphi(x)} \{ \overline{\Phi}_{i+1}(x) \Phi_{j+1}(x) + \Phi_{i+1}(x) \overline{\Phi}_{j+1}(x) \}. \tag{2.8}$$

If $i+j$ is even, the kernel is positive to the left of x and negative to the right. If $i+j$ is odd, it is negative on both sides. Moreover,

$$\mathbb{E}K^{n,n}(x,N) = 0, \qquad \mathbb{E}|K^{n,n}(x,N)| = D_n(x), \tag{2.9}$$

$$\mathbb{E}K^{n,n-1}(x,N) = -1, \qquad \mathbb{E}|K^{n,n-1}(x,N)| = 1 \quad (n \geq 1), \tag{2.10}$$

$$\mathbb{E}K^{n,n+1}(x,N) = -\frac{1}{n+1}, \qquad \mathbb{E}|K^{n,n+1}(x,N)| = \frac{1}{n+1} \quad (n \geq 0). \tag{2.11}$$

Proof. Insert (2.4)–(2.5) into (2.6), multiply by $\varphi(v)$, and integrate each half-line. The signs follow at once. For the neighbouring masses, the recurrence (2.3) yields

$$n! \{ \overline{\Phi}_{n+1} \Phi_n + \Phi_{n+1} \overline{\Phi}_n \} = \varphi, \qquad n \geq 1.$$

Indeed the expression on the left at index n equals that at index $n-1$; the base case follows from (2.2). This proves (2.10); replacing n by $n+1$ proves (2.11). The diagonal identities are immediate from the two general formulas. □

2.3 Representations and correction terms

Theorem 2.2 (Three neighbouring representations)**.** *For bounded measurable h,*

$$f_h(x) = \mathbb{E}[K^{0,0}(x,N)h(N)], \tag{2.12}$$

$$f_h'(x) = \mathbb{E}[K^{1,0}(x,N)\{h(N) - h(x)\}] \quad a.e. \tag{2.13}$$

For $r \geq 1$ and $\|h^{(r)}\|_\infty < \infty$ in the convention of section 1,

$$f_h^{(r-1)}(x) = \mathbb{E}[K^{r-1,r}(x,N)h^{(r)}(N)], \tag{2.14}$$

$$f_h^{(r)}(x) = \mathbb{E}[K^{r,r}(x,N)h^{(r)}(N)], \tag{2.15}$$

$$f_h^{(r+1)}(x) = \mathbb{E}[K^{r+1,r}(x,N)\{h^{(r)}(N) - h^{(r)}(x)\}]. \tag{2.16}$$

The first two equalities hold everywhere; the last holds almost everywhere, in particular at Lebesgue points of $h^{(r)}$.

Proof. Splitting (1.1) at x gives (2.12). The identities $R' = xR - 1$ and $S' = xS + 1$, together with (1.2) and $\mathbb{E}K^{1,0} = -1$, give (2.13).

For smooth tests of polynomial growth, integrate by parts separately on the two half-lines. Since $\Phi_{j+1}' = \Phi_j$ and $\overline{\Phi}_{j+1}' = -\overline{\Phi}_j$, this gives

$$\mathbb{E}[K^{i,j}(x,N)h^{(j)}(N)] = \mathbb{E}[K^{i,j+1}(x,N)h^{(j+1)}(N)] + \mathbb{E}[K^{i,j}(x,N)]h^{(j)}(x). \tag{2.17}$$

The terms at infinity vanish because the appropriate integrated tail decays Gaussianly and the test derivatives grow at most polynomially. On the diagonal, the mean in the correction term is zero. Differentiating a diagonal representation produces the boundary term $h^{(i)}(x)$:

$$f_h^{(i+1)}(x) = h^{(i)}(x) + \mathbb{E}[K^{i+1,i}(x,N)h^{(i)}(N)]. \tag{2.18}$$

The coefficient of this boundary term is one, by the identity used in (2.10) (and directly for $i = 0$). Applying (2.17) with $(i,j) = (i+1,i)$ cancels that term and produces the next diagonal representation. Induction proves the diagonal formulas. One further integration by parts gives (2.14); (2.18) and the mass -1 give (2.16).

For the stated regularity, convolve h with a smooth compactly supported approximate identity. The rth derivatives remain bounded by $\|h^{(r)}\|_\infty$, and the lower derivatives admit a common polynomial majorant. Dominated convergence proves the first two formulas. The differentiated Stein equation gives the last formula almost everywhere. The same argument, or local regularity for (1.2), shows that $f_h \in C^r$ and $f_h^{(r)}$ is locally absolutely continuous. □

Corollary 2.3 (Representations farther from the diagonal)**.** *If $\|h^{(k)}\|_\infty < \infty$, $k \geq 1$, and $0 \leq n \leq j \leq k$, then*

$$f_h^{(n)}(x) = \mathbb{E}[K^{n,j}(x,N)h^{(j)}(N)] + \sum_{\ell=n}^{j-1} \mathbb{E}[K^{n,\ell}(x,N)]h^{(\ell)}(x). \tag{2.19}$$

The sum is empty for $j = n$ and zero for $j = n+1$. In general it must be retained for $j \geq n+2$.

Proof. Iterate (2.17) from the diagonal representation. Gaussian tail decay justifies every boundary limit. □

For example, $h(v) = v$ has $h'' = 0$ but $f_h = -1$; the correction term with $K^{0,1}$ supplies precisely this missing constant. A representation involving only h'' would be false.

3 Proofs of the envelopes, factors, and sharpness

3.1 Uniform diagonal factors

Lemma 3.1. *The functions A and D_n satisfy $A(x) \le A(0) = \sqrt{\pi/8}$ and $D_n(x) \le D_n(0) = d_n$ for $n \ge 1$.*

Proof. Here is a proof that also explains the gamma quotient. Let $P_t h(x) = \mathbb{E}h(e^{-t}x + \sqrt{1-e^{-2t}}N)$ be the Gaussian semigroup. For a smooth test of polynomial growth, integration of its generator identity shows that $u = -\int_0^\infty (P_t h - \mathbb{E}h(N))\,dt$ solves $u'' - xu' = h - \mathbb{E}h(N)$ and $f_h = u'$. Differentiation and Gaussian integration by parts give

$$f_h^{(n)}(x) = -\int_0^\infty \frac{e^{-(n+1)t}}{\sqrt{1-e^{-2t}}} \mathbb{E}\left[N h^{(n)}(e^{-t}x + \sqrt{1-e^{-2t}}N)\right] dt. \tag{3.1}$$

For $n \ge 1$, the absolute value is at most

$$\|h^{(n)}\|_\infty \sqrt{\frac{2}{\pi}} \int_0^\infty \frac{e^{-(n+1)t}}{\sqrt{1-e^{-2t}}}\,dt = d_n \|h^{(n)}\|_\infty.$$

The substitution $u = e^{-2t}$ evaluates the integral as a beta integral. Mollification extends the bound to our stated test classes.

For each fixed x, take $h^{(n)}(v) = \mathbf{1}_{\{v<x\}} - \mathbf{1}_{\{v>x\}}$, and take successive antiderivatives to define h. Formula (2.15) then gives $f_h^{(n)}(x) = D_n(x)$. Hence $D_n(x) \le d_n$. Using (2.1) at zero and the gamma duplication formula gives $D_n(0) = d_n$.

For $n = 0$ in (3.1), subtract the midpoint of the range of a bounded h. The same calculation gives $\|f_h\|_\infty \le \sqrt{\pi/2}\,\operatorname{osc}(h)/2$. The test $h_x = \mathbf{1}_{(-\infty,x]}$ has $f_{h_x}(x) = A(x)$; therefore $A(x) \le \sqrt{\pi/8}$, with equality at zero. □

3.2 Kolmogorov and bounded tests

Proof of theorem 1.1. Directly from (1.1),

$$f_z(x) = \begin{cases} \overline{\Phi}(z)S(x), & x \le z, \\ \Phi(z)R(x), & x > z. \end{cases}$$

Differentiating on each open half-line and using (2.4)–(2.5) gives (1.9). At fixed x, the largest values on the two branches are approached as z tends to x from the corresponding side. This proves (1.8) and its exactness. The formulas for Q_1, Q_2 follow from (2.3). The order-one upper bound 1 follows from theorem 1.2, proved just below. Finally, for $x \to +\infty$,

$$\frac{\overline{\Phi}_{n+1}(x)}{\varphi(x)} \sim x^{-n-1}, \qquad \Phi_{n+1}(x) \sim \frac{x^n}{n!}, \qquad \frac{\overline{\Phi}(x)}{\varphi(x)} \sim x^{-1}.$$

These follow from the Laplace integrals (2.4) and from (2.1). They give $Q_n(x) \sim x^{n-1}$ for $n \ge 1$. Reflection gives the negative tail. Thus the factor 1 is sharp at order one, and no finite uniform factor exists at higher orders. □

Proof of the finite bounds in theorem 1.2. The centred kernel $K^{0,0}$ has absolute mean $2A(x)$. Subtracting the midpoint of the essential range of h in (2.12) gives $|f_h(x)| \le A(x)\operatorname{osc}(h)$. Since $K^{1,0} \le 0$ has absolute mean one, (2.13) gives $|f_h'(x)| \le \operatorname{osc}(h)$ almost everywhere. The uniform factors follow from lemma 3.1.

The test $h_x = \mathbf{1}_{(-\infty,x]}$ attains the order-zero envelope. For sharpness of the constant envelope at order one, use $h_{x,\varepsilon} = \mathbf{1}_{[x-\varepsilon,x+\varepsilon]}$. It is constant near x and has oscillation one, and

$$f'_{h_{x,\varepsilon}}(x) = 1 + \mathbb{E}[K^{1,0}(x,N)h_{x,\varepsilon}(N)] \longrightarrow 1.$$

The derivative is continuous near x, so the same construction proves sharpness for the essential supremum. Smooth plateaux give the same limit. Affine rescaling from $[0,1]$ to $[-B,B]$ proves the supremum-norm assertions. The higher-order impossibility is proved in section 4. □

3.3 The three smoothness bands and their extremisers

Proof of the finite bounds in theorem 1.5. Apply lemma 2.1 to theorem 2.2. The lower representation gives $|f_h^{(r-1)}| \le H_r/r$. The diagonal gives $|f_h^{(r)}(x)| \le D_r(x)H_r$. The upper representation gives $|f_h^{(r+1)}| \le 2H_r$ almost everywhere. The diagonal uniform factor follows from lemma 3.1.

All are exact pointwise. For order $r-1$, take $h(v) = v^r/r!$; then $h^{(r)} = 1$ and $f_h^{(r-1)} = -1/r$. For order r, take $h^{(r)}(v) = \mathbf{1}_{\{v<x\}} - \mathbf{1}_{\{v>x\}}$, as in the proof of lemma 3.1. For order $r+1$, prescribe

$$h_{x,\varepsilon}^{(r)}(v) = \mathbf{1}_{\{|v-x|\le\varepsilon\}} - \mathbf{1}_{\{|v-x|>\varepsilon\}}.$$

Its successive antiderivatives belong to the required class. At x,

$$f_{h_{x,\varepsilon}}^{(r+1)}(x) = -2\mathbb{E}[K^{r+1,r}(x,N)\mathbf{1}_{\{|N-x|>\varepsilon\}}] \longrightarrow 2.$$

Dominated convergence applies because the kernel has absolute mean one. The function is smooth near x, so the argument also proves sharpness for the essential supremum. Smooth approximations preserving the derivative bound give the same limiting values. Uniform sharpness follows by evaluating the diagonal at zero and using the other two constructions at any fixed point. □

theorem 1.3 is the case $r = 1$. Combining the available bounds proves (1.16)–(1.19). For the moment statement about ζ_r, a polynomial of degree less than r has zero rth derivative and can be multiplied by any constant. This proves divergence when a lower moment differs. Conversely, subtract the Taylor polynomial at zero of degree $r-1$ from h; the remainder is at most $|x|^r/r!$ in absolute value. Under moment matching its expectations have the same difference as those of h, proving finiteness.

3.4 Fortet–Mourier bounds and exact first-order refinements

Proof of proposition 1.4 and (1.14). The simple bounds follow from $\operatorname{osc}(h) \le 2B$, $\|h'\|_\infty \le L$, and theorems 1.2 and 1.3. To prove the exact order-zero formula, fix $h(x) = a$. On $v \le x$, $h(v) \le \min(B, a + L(x-v))$, and on $v > x$, $h(v) \ge \max(-B, a - L(v-x))$. The signs of $K^{0,0}$ therefore imply $f_h(x) \le f_{u_{a,x}}(x)$. The function $u_{a,x}$ itself belongs to $\mathcal{H}_{B,L}$ and has value a at x. Maximise over a; continuity and compactness give a maximum. Symmetry under $h \mapsto -h$ shows that this is the supremum of the absolute value as well.

For order one, (2.13) reads $f'_h(x) = \int (h(x) - h(v))q_x(v)\,dv$. The kernel sign and mass show that q_x in (1.12) is a probability density. Every test satisfies

$$|h(x) - h(v)| \le \min\{L|v-x|, 2B\}.$$

Equality in the resulting upper bound is attained by $h(v) = B - \min\{L|v-x|, 2B\}$. This function is bounded by B and is L-Lipschitz, proving (1.13).

For the sum-norm convention, maximising $\min\{aB, bL\}$ over $B, L \ge 0$, $B + L \le 1$, gives $ab/(a+b)$ for $a, b > 0$. Apply this with $(a,b) = (2A(x), 1)$ and $(2, D_1(x))$. The second-derivative bound follows from $L \le 1$. □

4 Why the missing bounds cannot hold

4.1 Too few solution derivatives: polynomial freedom

Fix $r \geq 2$ and $0 \leq n \leq r-2$. Let H_j be the probabilists' Hermite polynomials, so $H_n' - xH_n = -H_{n+1}$. For $h = MH_{n+1}$, we have $h^{(r)} = 0$, while the canonical solution is $f_h = -MH_n$. Thus $f_h^{(n)} = -Mn!$ at every point. Letting $|M| \to \infty$ proves the first negative row of theorem 1.5, including failure of any finite pointwise envelope.

4.2 Too many solution derivatives: a smooth local construction

Fix $x_0 \in \mathbb{R}$, integers $j \geq 0$ and $n > j+1$, and choose $\psi \in C_c^\infty(\mathbb{R})$ with $\psi^{(n)}(0) \neq 0$. For $0 < \delta \leq 1$, set

$$g_\delta(x) = \delta^{j+1}\psi((x-x_0)/\delta), \qquad h_\delta(x) = g_\delta'(x) - xg_\delta(x). \tag{4.1}$$

Gaussian integration by parts gives $\mathbb{E}h_\delta(N) = 0$, and integration of $(\varphi g_\delta)' = \varphi h_\delta$ shows that $f_{h_\delta} = g_\delta$. The support is localised near x_0 and

$$\sup_{0<\delta\leq 1} \|h_\delta^{(j)}\|_\infty < \infty, \qquad f_{h_\delta}^{(n)}(x_0) = \delta^{j+1-n}\psi^{(n)}(0).$$

Rescale all tests by a fixed constant to make the controlled norm at most one. The displayed derivative diverges as $\delta \downarrow 0$. With $j = 0$, this proves the negative bounded-test entries. With $j = 1$, it proves the negative Wasserstein entries; in this case $\|h_\delta\|_\infty = O(\delta)$ and $\mathrm{Lip}(h_\delta) = O(1)$, so a fixed rescaling also places the tests in any $\mathcal{H}_{B,L}$ with $B, L > 0$, or in the sum-norm unit ball. With $j = r$, it proves the last negative row of theorem 1.5. All counterexamples are smooth.

4.3 What the kernel asymptotics do, and do not, prove

The general absolute mass in (2.8) satisfies, for $x \to +\infty$,

$$\mathbb{E}|K^{i,j}(x,N)| \sim \frac{i!}{j!}x^{j-i-1} + x^{i-j-1}. \tag{4.2}$$

Consequently it is uniformly bounded precisely in the three bands $|i-j| \leq 1$. Notice the placement of the coefficient $i!/j!$. This explains the special role of the three neighbouring representations, but divergence of a kernel mass by itself does not rule out every other representation or estimate. The preceding polynomial and localised constructions supply the required impossibility proofs for the solution operator itself.

5 Using and evaluating the bounds

When the Stein equation holds at X almost surely and the terms are integrable,

$$\mathbb{E}h(X) - \mathbb{E}h(N) = \mathbb{E}\{f_h'(X) - Xf_h(X)\}. \tag{5.1}$$

For a discontinuous test and an X with atoms, the classical derivative may not be defined at every value taken by X. In that case use the Stein-equation version $f_{h,\mathrm{S}}'(x) = xf_h(x) + h(x) - \mathbb{E}h(N)$ in (5.1). It agrees with f_h' almost everywhere and obeys the bounded-test order-one bound everywhere; for h_z at z it is the left derivative. This convention does not create classical higher derivatives at a jump.

The tables identify the derivatives available when bounding the right side, or a coupling expansion of it. If an argument produces a term $\mathbb{E}[|T|\,|f_h^{(n)}(Y)|]$, the envelope gives $\mathbb{E}[|T|E_n(Y)]$,

while a uniform factor gives $C_n\mathbb{E}|T|$. The former can retain useful information about the location of Y.

For instance, if a Wasserstein calculation requires f_h'', the sharp factor is 2. A third derivative cannot be bounded from the Lipschitz condition alone. If a Zolotarev-2 calculation uses f_h', its factor is $1/2$, even though the solutions themselves have no class-uniform bound. For Kolmogorov tests, a Taylor expansion crossing z must account for the jump in f_z'; piecewise higher-derivative formulas do not remove that term.

For evaluation at moderate x, the recurrences (2.3) produce every envelope using only Φ and φ. In extreme tails the subtraction $\varphi(x) - x\overline{\Phi}(x)$ may lose precision. Use the positive integral

$$\frac{\overline{\Phi}_j(x)}{\varphi(x)} = \frac{1}{(j-1)!}\int_0^\infty t^{j-1}e^{-xt-t^2/2}\,dt, \qquad j \geq 1,$$

or a stable tail routine there, and use reflection $\Phi_j(-x) = \overline{\Phi}_j(x)$. The Fortet–Mourier refinements need only Gaussian integrals of clipped affine functions and a compact one-dimensional maximisation.

Finally, Gaussian location and scale can be restored without redoing the analysis. For $Z \sim \mathcal{N}(\mu, \sigma^2)$, $\sigma > 0$, use

$$\sigma^2 f'(x) - (x-\mu)f(x) = h(x) - \mathbb{E}h(Z).$$

Writing $y = (x-\mu)/\sigma$ and $g(y) = h(\mu + \sigma y)$, its canonical solution satisfies

$$f(x) = \sigma^{-1} f_g(y), \qquad f^{(n)}(x) = \sigma^{-n-1} f_g^{(n)}(y), \qquad \|g^{(r)}\|_\infty = \sigma^r \|h^{(r)}\|_\infty.$$

Thus a standardised (n, r) derivative bound acquires the factor σ^{r-n-1}; oscillation is unchanged. For Fortet–Mourier tests, the parameters become $(B, \sigma L)$ before multiplying the solution bound by σ^{-n-1}.

References

[1] R. E. Gaunt (2025). On Stein factors in Stein's method for normal approximation. *Statistics & Probability Letters* **219**, 110339. doi:10.1016/j.spl.2024.110339. Preprint: arXiv:2410.05475.