Exponential Rigidity and a Log-Value Normal Form for Imbalanced Market Making

Peter Cotton*

August 28, 2026

Abstract

Let $G(K) = \sup_m (m - K)\bar{F}(m)$ be the optimized value of an enquiry with break-even strike K, against a best competing quote with survival curve $\bar{F}$. For every positive, strictly decreasing enquiry-value curve, the coordinate $\Phi = -\log(G/G_\star)$ converts buy–sell imbalance into an exact translation of two log-value coordinates. The translation is induced by state-independent shifts of the break-even strikes exactly for $G = \kappa + Be^{-hK}$ on the active strike interval; state-independent shifts of the submitted quotes, or global rationalizability by a finite-mean markup distribution, force $\kappa = 0$. For general G the canonical sidewise gauge need not satisfy the edge compatibility that lets both strike ladders derive from one inventory potential; we quantify the defect. On a finite inventory lattice, under nonsingularity of the anchor linearization, any smooth deformation path yields a perturbation hierarchy whose every coefficient uses one linear operator, with distributional perturbations entering through an envelope formula and local rationalizability conditions. Reflection symmetry makes the skew odd and the convexity response even in the log-odds of the flow. All formulas and asymptotic orders are checked numerically.

Keywords: market making, order imbalance, skew, bid–ask spread, normal form, perturbation theory, request for quote.
MSC codes: 91G15, 93E20, 41A58.

1 Introduction

Cotton (2026a) considers a market maker who acquires and disposes of inventory only by responding to sealed-bid enquiries, arriving as a Poisson stream in which a seller calls with probability q and a buyer with probability $1 - q$. Under the assumption that the best competing response is exponentially distributed beyond fair value with mean $w = 1/h$, the steady state of the imbalanced problem compresses exactly onto the balanced one, and the compression has two equivalent representations. One subtracts $\delta = \frac{1}{2h} \log \frac{q}{1-q}$ from the break-even skew and adds $\gamma = \frac{1}{h} \log \frac{1}{2\sqrt{q(1-q)}}$ to the overhead; the other tilts the inventory potential by δx and solves the balanced problem with carrying cost multiplied by $M(q) = e^{h\gamma} = 1/(2\sqrt{q(1-q)})$. These are one transformation written in two frames, not simultaneous corrections: the imbalanced

*The exact steady-state equivalence under exponential competing quotes is proved in Cotton (2026a), and the finite-horizon case is treated in Cotton (2026b). An accompanying script, `verify_normal_form.py`, checks every formula and asymptotic order numerically. Comments welcome.

dealer's quotes coincide with those of the balanced dealer at cost $M(q)c$ shifted by δ, so her physical width response to imbalance is that problem's second-order convexity response, not γ.

The exponential assumption invites the obvious question, and Cotton (2026a) gave it a local answer: only the hazard at the quotes actually made enters, and if the hazard varies slowly across the visited strikes the equivalence holds approximately with the local width, with error of the order of the hazard's relative variation. That is a robustness statement. The present paper replaces it with a structure: an exponential *normal form*, in which general win curves are deformations of an integrable anchor and the corrections are computable order by order along a deformation path.

The organizing observation is that after the dealer optimizes her own markup, the distribution of the best competing quote enters the steady state through one function only,

$$G(K) \;=\; \sup_{m}\,(m-K)\,\bar F(m), \qquad \bar F = 1-F,$$

the optimized value of the next enquiry as a function of its strike K; quoting against the survival curve of the best competing bid in this way descends from Friedman (1956). Throughout, *win curve* means the survival function $\bar F$ and G is called the *enquiry-value curve*; the two are related by the envelope calculus of Section 2 and Proposition 2. The right object to deform is not $\bar F$ but the *effective log-value*

$$\Phi(K) \;=\; -\log\frac{G(K)}{G_\star},$$

with $G_\star$ an arbitrary normalization. By the envelope identity of Cotton (2026a), $-\frac{d}{dK}\log G(K)$ equals the hazard rate of the win curve evaluated at the optimal markup, so Φ is the integral of the effective hazard sampled along the optimal response; exponential competing quotes correspond exactly to Φ affine, and Φ'' is the natural local measure of non-exponentiality.

The results separate five levels at which imbalance can be removed, and the levels do not coincide:

- translation in the coordinates built from Φ — every positive enquiry-value curve;
- sidewise gauge rigidity, each side's gauge a translation — the exponential only;
- rigid translation of the break-even strikes, balancing the two-sided problem — the affine-exponential family $\kappa + Be^{-hK}$;
- rigid translation of the submitted quotes, the optimal markups $m^*(K)$ of Section 2 — the exponential only;
- global rigidity for a win curve with finite mean markup — the exponential only.

The vocabulary is from control theory. A *normal form* is a canonical coordinate system together with a classification of which members of a family become trivial there, as in feedback linearization; *rigidity* is meant as in the memorylessness characterization of the exponential distribution: a weak hypothesis forces a short list. Exponentiality plays the role here that linearizability plays in control, and the results away from it measure the obstruction and compute the correction.

Five results organize the theory.

1. *Normal form* (Theorem 1). For every smooth positive enquiry-value curve, in the nonlinear coordinates (U, V) built from Φ, imbalance is exactly a translation: by $\log M(q)$ in the width-like coordinate and by the log-odds $\ell = \frac{1}{2}\log\frac{q}{1-q}$ in the skew-like coordinate. Exponentiality is not what creates the symmetry; it is what makes the coordinates affine, so that the translation is a rigid motion of the strike coordinates — the physical quotes then shift by the flow term δ alone, widths unchanged.

2. *Rigidity* (Theorems 2 and 3, Corollary 2). The sidewise gauge maps are translations only for exponential G. With the convexity every enquiry-value curve has (Lemma 1), rigid strike translations balance the two-sided Hamiltonian exactly for the affine-exponential family $G = \kappa + Be^{-hK}$; the constant cancels from the consistency equation, so the compression of Cotton (2026a) extends verbatim to that family. The pure exponential is recovered in two independent ways: rigidity of the submitted quotes, and a finite-mean tail.

3. *Integrability defect* (Theorem 4). Outside the rigid class the pointwise removal survives but generically fails to respect the constraint that both strike ladders derive from a single inventory potential; the exact criterion is that a defect functional be constant on the visited strike set. The defect's variation begins at first order in the curvature, with an explicit oscillation bound, and its *constant part* at the affine point is exactly 2γ: the gauged-overhead displacement γ of Cotton (2026a) — an overhead-frame object, not a physical widening — is the value of the defect functional at the integrable point.

4. *Perturbation hierarchy* (Theorem 5), *envelope transfer* (Proposition 1), and *admissibility* (Proposition 2). On a finite inventory lattice, and under nonsingularity of the exponential-anchor linearization, writing $\Phi_\eta = a + hK + \eta\psi(K)$, where the fixed function ψ carries the shape of the departure from exponentiality and η its size (Section 2), the inventory potential expands as $\nu_\eta = \nu_0 + \eta\nu_1 + \cdots$ where ν_0 is the exact exponential solution and every ν_j solves a linear system with one fixed operator, tridiagonal in the interior with two boundary rows and a rank-one centering correction. First-order deformations of standard families enter through one envelope evaluation at the exponential optimum — for the Weibull family with shape $1 + \eta$, $\psi(K) = (1 + hK)\log(1 + hK)$ — with higher orders using higher derivatives of the log-survival curve and the same linear operator; the theory is a functional expansion of the log-survival curve rather than one calculation per named distribution, and the admissibility proposition gives local sufficient conditions for the deformations traversed to correspond to genuine markup distributions. Corollary 4 turns the value correction into the quote correction itself.

5. *Parity* (Theorem 6). Distribution-free: with a common win curve on both sides and even carrying cost, the skew contains only odd powers of ℓ and the width response only even powers. Under nondegeneracy of the leading coefficients the familiar reading follows: skew first order, width second order. The exponential normal-form translations $\delta = \ell/h$ and $\gamma = \frac{1}{h}\log\cosh\ell$ exhibit the same odd/even pattern, but γ is not in general the endogenous convexity response, which depends on the carrying cost.

The formulas and their predicted asymptotic orders are independently checked in the accompanying script `verify_normal_form.py`, at machine precision for the exact identities and by Richardson ratio tests for the expansions (Section 9).

2 Setup: the effective log-value

The model is that of Cotton (2026a). Enquiries arrive with mean gap τ; each is a sell with probability q, a buy with probability $1-q$; trades have size s; every won trade concedes ϵ to adverse selection; holding x costs $c(x)$ per unit time with $c(0)=0$. The policy is characterized by an inventory indifference cost $\nu(x)$[1] with slope and convexity

$$S(x) = \frac{\nu(x+s)-\nu(x-s)}{2s}, \qquad C(x) = \frac{\nu(x+s)-2\nu(x)+\nu(x-s)}{2s},$$

and break-even strikes

$$K^{\downarrow}(x) = \epsilon + \frac{\nu(x+s)-\nu(x)}{s} = A(x)+S(x), \qquad K^{\uparrow}(x) = \epsilon + \frac{\nu(x-s)-\nu(x)}{s} = A(x)-S(x), \tag{1}$$

where $A = \epsilon + C$ collects the overhead and convexity. In words: $\nu(x)$ is the haircut at which the dealer would clear her book, so $K^{\downarrow}(x)$ is the all-in break-even markdown for buying one more lot — adverse selection plus the marginal inventory cost — and $K^{\uparrow}(x)$ is its selling counterpart. The two strikes sit symmetrically about $A(x)$ at distance $S(x)$: the slope of the inventory cost is the skew and its convexity the discretionary width, the identities of Cotton (2026a). The best competing response beyond fair value has survival function $\bar{F}$, assumed smooth and positive with $G(K) = \sup_m (m-K)\bar{F}(m)$ finite, positive, and attained at an interior optimum $m^*(K)$ on the strike range of interest; $G(K)$ is what the next enquiry is worth to a dealer whose all-in break-even is K, and it falls as K rises.

Lemma 1 (Indirect-value properties)**.** *Where $\bar{F} > 0$ and the optimum is interior and unique, G is convex and strictly decreasing, with $-1 \le G'(K) = -\bar{F}(m^*(K)) < 0$.*

Proof. G is a supremum of the affine functions $K \mapsto (m-K)\bar{F}(m)$, each with slope $-\bar{F}(m) \in [-1, 0)$, hence convex with slopes in the same range; the envelope theorem identifies the slope at K as $-\bar{F}(m^*(K))$. □

The steady-state consistency equation is, for all x,

$$\frac{\tau\, c(x)}{s} \;=\; H_q\big(A(x), S(x)\big) - H_q\big(A(0), S(0)\big), \qquad H_q(A,S) = q\,G(A+S) + (1-q)\,G(A-S). \tag{2}$$

The left side of (2) is the direct cost of carrying x to the next enquiry; the right side is the extra option value that enquiry has at inventory x versus at zero, a loaded dealer valuing the chance to get out. Indifference requires the two to balance state by state. For exponential $\bar{F}(m) = e^{-hm}$ one has $G(K) = \frac{1}{h}e^{-1-hK}$ and (2) is the equation solved in closed correspondence by Cotton (2026a).

Define the effective log-value $\Phi(K) = -\log(G(K)/G_\star)$. The optimizer's first-order condition $\bar{F}(m^*) + (m^*-K)\bar{F}'(m^*) = 0$ and the envelope slope $G'(K) = -\bar{F}(m^*(K))$ of Lemma 1 combine into $-G'(K)/G(K) = 1/(m^*(K)-K) = h(m^*(K))$, the envelope identity of Cotton (2026a):

$$\Phi'(K) \;=\; h\big(m^*(K)\big) \;>\; 0, \tag{3}$$

[1] The sign of ν is opposite to the usual Bellman convention: in the average-reward dynamic programming formulation the differential value function is $-\nu$. The flip makes ν a *cost* — the non-myopic price of holding inventory, netting direct carrying charges against the option value of trading out through future enquiries — which is the object a dealer quotes from.

where $h(m) = -\frac{d}{dm}\log \bar{F}(m)$ is the hazard rate of the win curve; for the exponential win curve $h(m) \equiv h$, so the two uses of the symbol agree. The identity says how fast the enquiry loses value as the break-even worsens: a small increase in K costs the dealer only in the states where she wins, and at the optimum that sensitivity, per unit of value, is the hazard at the quote actually made. Thus Φ is strictly increasing and invertible on its range: it is the integral of the effective hazard sampled at the optimal response. Three consequences of (3) frame everything below: exponential competing markups are exactly the case $\Phi(K) = a + hK$; departures from exponentiality are exactly departures of Φ from affinity; and Φ'' is the natural local measure of non-exponentiality. Accordingly we write deformations of the anchor as

$$\Phi_\eta(K) \;=\; a + hK + \eta\,\psi(K), \tag{4}$$

where ψ is a fixed smooth function, the *deformation*, carrying the shape of the departure from exponentiality, and the scalar η carries its size. Constant and affine components of ψ are absorbed into a and h, and one may normalize ψ and ψ' to vanish at a chosen anchor strike, so that the perturbation genuinely begins with curvature. Section 7 supplies ψ from named distribution families; until then it is simply given.

One point governs the interpretation of everything built on (4). The coefficients of the perturbation hierarchy are attached to a parameterized *path* of curves, not to its endpoint: the first coefficient depends only on the initial tangent $\psi = \partial_\eta \Phi_\eta|_{\eta=0}$, the j-th on the path jet through order j, and different paths to the same endpoint can differ already at first order because their tangents differ. The centered consistency equation has null directions that make this concrete. It is invariant under $G \mapsto G + \kappa$, so the additive path $G_\eta = G_0 + \eta\kappa$, whose tangent is $\psi_{\text{add}} = -\kappa/G_0$, produces forcing that the centering annihilates: the potential never moves. The direction is null only for the centered potential equation — the gain and the submitted quotes do move along it, the latter through the direct term of Corollary 4. The straight line (4) to the same endpoint has tangent $-\log(1+\kappa/G_0)$ and moves the potential already at first order. A path convention must therefore be chosen and reported; (4), the log-geometric interpolation $G_\eta = G_0^{1-\eta} G_1^{\eta}$ with target $G_1 = G_0 e^{-\psi}$, is this paper's default, and for a named parametric family the family's own parameter path is the natural alternative. Equivalently, the hierarchy computes directional Taylor derivatives of the solution map $\Phi \mapsto \nu[\Phi]$ in the affine chart that Lemma 2 distinguishes; η is only the device for taking them, and parametric families enter through the chain rule along their own paths. Bluntly: the log-value coordinate is canonical for removing multiplicative flow weights (Lemma 2); the straight path in that coordinate is a declared interpolation convention; and the coordinate adapted to the full rigid manifold $\kappa + Be^{-hK}$ is $\Xi = -\log(-G')$ (Remark 7).

3 The normal form

The coordinates that follow average and difference the log-values of the dealer's two strikes: U measures how expensive her two-sided quote is overall, a width in log-value units, and V measures its asymmetry, a skew. The theorem says the flow translates the width-like coordinate by $\log M(q)$ and the skew-like coordinate by $-\ell$, both constants. The logarithm is not merely convenient:

Lemma 2 (The log-value coordinate is forced)**.** *Let $\chi : (0,\infty) \to \mathbb{R}$ be C^1 and suppose $\chi(\lambda z) - \chi(z)$ is independent of z for every λ in a neighborhood of one. Then $\chi(z) = a + b\log z$.*

Up to affine rescaling, $\Phi = -\log(G/G_\star)$ is therefore the unique C^1 coordinate on enquiry values under which the multiplicative imbalance weights act as translations.

Proof. Differentiate in z: $\lambda\chi'(\lambda z) = \chi'(z)$ for all λ near one, so $z\chi'(z)$ is constant; integrate. □

Theorem 1 (Imbalance is a translation in the Φ-coordinates)**.** *Let G be any positive function with $\Phi = -\log(G/G_\star)$ strictly increasing, and define*

$$U(A,S) = \frac{\Phi(A+S)+\Phi(A-S)}{2}, \qquad V(A,S) = \frac{\Phi(A+S)-\Phi(A-S)}{2}, \qquad \ell = \frac{1}{2}\log\frac{q}{1-q}.$$

Then, exactly,

$$H_q(A,S) \;=\; 2\sqrt{q(1-q)}\, G_\star\, e^{-U(A,S)} \cosh\big(V(A,S)-\ell\big), \tag{5}$$

and, writing $\mathcal{H}_q(U,V)$ for H_q expressed through the coordinates (U,V),

$$\mathcal{H}_q(U,V) \;=\; \mathcal{H}_{1/2}\big(U+\log M(q),\; V-\ell\big), \qquad M(q) = \frac{1}{2\sqrt{q(1-q)}} = \cosh\ell. \tag{6}$$

Proof. $\Phi(K^{\downarrow}) = U+V$ and $\Phi(K^{\uparrow}) = U-V$, so $H_q = G_\star e^{-U}\big(qe^{-V}+(1-q)e^{V}\big)$, and the identity of Cotton (2026a), $qe^{-V}+(1-q)e^{V} = 2\sqrt{q(1-q)}\cosh(V-\ell)$, gives (5). Since $\mathcal{H}_{1/2}(U,V) = G_\star e^{-U}\cosh V$ and $2\sqrt{q(1-q)} = e^{-\log M(q)}$, (6) follows. That $M(q) = \cosh\ell$ is the elementary computation $\cosh\ell = \frac{q+(1-q)}{2\sqrt{q(1-q)}}$. □

Corollary 1 (The theorem of Cotton (2026a), and why it is rigid there)**.** *For exponential competing quotes, $\Phi(K) = a+hK$, so $U = a+hA$ and $V = hS$ are affine and the translation (6) is a rigid motion of the strike coordinates themselves, the balancing map*

$$S \longmapsto S - \frac{\ell}{h} = S-\delta, \qquad A \longmapsto A + \frac{\log M(q)}{h} = A+\gamma;$$

equivalently, the imbalanced solution is the balanced one with $S_q = S_{\mathrm{bal}}+\delta$. Moreover $e^{-h(A+\gamma)} = e^{-h\gamma}e^{-hA}$ factors the overhead displacement through the equation, so it may be moved to the cost side of (2) as the multiplier $M(q) = e^{h\gamma}$: for exponential G, widened overhead and inflated carrying cost are the same statement, at exchange rate h. Outside the affine-exponential family of Theorem 3 no such factorization exists, and the overhead displacement — not the multiplier — is the invariant form of the correction. Neither representation widens the dealer's physical quotes by γ: her quotes are those of the balanced dealer at cost $M(q)c$, shifted by δ, so the physical width response to imbalance is that problem's convexity response, second order and cost-dependent; the accompanying script demonstrates the distinction.

Remark 1 (What the theorem does and the coordinates hide)**.** *Equation (6) holds for every smooth positive enquiry-value curve, so the imbalance symmetry as such costs nothing. (On a strike interval I the translated coordinates correspond to physical strikes only where $U \pm V$, shifted, stay inside $\Phi(I)$; the identity itself is algebraic and unrestricted.) What it does not by itself deliver is a solution map: the consistency equation couples the (U,V) values at different inventories through the requirement that all strikes derive from one potential ν, and the translation in (U,V) moves each inventory's strike pair by a different physical amount unless Φ is affine. Sections 4 and 5 make that obstruction precise.*

4 Two rigidity theorems

Pointwise, imbalance can always be removed. Define the gauge maps, for $\lambda > 0$,

$$T_\lambda(K) \;=\; \Phi^{-1}\big(\Phi(K) - \log\lambda\big), \qquad \text{so that} \qquad G\big(T_\lambda(K)\big) = \lambda\, G(K). \tag{7}$$

Then

$$H_q\big(A(x), S(x)\big) \;=\; \tfrac{1}{2}\, G\big(T_{2q}(K^{\downarrow}(x))\big) + \tfrac{1}{2}\, G\big(T_{2(1-q)}(K^{\uparrow}(x))\big) : \tag{8}$$

every imbalanced Hamiltonian is a balanced one evaluated at gauged strikes. On an interval I the gauge is defined on $I_\lambda = \{K \in I : \Phi(K) - \log\lambda \in \Phi(I)\}$, and every gauge statement below is on these domains. In words, T_λ moves a strike to where the enquiry is λ times as valuable: probability weight on a side is traded for a slide of its strike. For exponential Φ the gauge is the constant translation $T_\lambda(K) = K - \frac{1}{h}\log\lambda$; in general $T_\lambda(K) - K$ depends on K, and the exponential case is the only one in which it does not.

Theorem 2 (Sidewise gauge rigidity)**.** *Let G be positive and strictly decreasing with Φ continuously differentiable on an interval I, and suppose that for every λ in some open interval around 1 the gauge T_λ restricted to I is a translation: $T_\lambda(K) = K - t(\lambda)$ for a constant $t(\lambda)$, for all $K \in I$ with $T_\lambda(K) \in I$. Then Φ' is constant on I: G is exponential there.*

Proof. Fix an interior K_0. The shift $t(\lambda) = K_0 - \Phi^{-1}\big(\Phi(K_0) - \log\lambda\big)$ is continuous in λ with $t(1) = 0$, so the available shifts fill an interval $(-\tau_0, \tau_0)$. For each such shift, $G(K - t) = \lambda\, G(K)$, i.e. $\Phi(K) - \Phi(K - t) = \log\lambda$ with the left side independent of K, wherever K and $K - t$ both lie in I; differentiating, $\Phi'(K) = \Phi'(K - t)$ there. A continuous function on a connected interval that is invariant under every sufficiently small shift is constant, since any two points are joined by steps shorter than τ_0. □

Theorem 2 is a statement about the two sidewise gauges separately. It does not by itself characterize the curves for which the *summed* Hamiltonian can be balanced by rigid translations of the strikes. The family

$$G(K) \;=\; \kappa + B\, e^{-hK}, \qquad B > 0,\ \kappa > 0,$$

has curved Φ, yet

$$\begin{aligned} H_q(A, S) &= \kappa + Be^{-hA}\big(qe^{-hS} + (1-q)e^{hS}\big) \\ &= \kappa + 2B\sqrt{q(1-q)}\, e^{-hA}\cosh(hS - \ell) \;=\; H_{1/2}(A + \gamma,\, S - \delta) : \end{aligned}$$

the constant passes through both sides and the exponential part translates rigidly.

Moreover κ cancels from the consistency equation (2), which sees only differences of H_q, so the compression of Cotton (2026a) — skew shift δ with the carrying cost multiplied by $M(q)$ — extends to this family *at the level of the two-sided centered equation*; and the family is locally realizable by a genuine markup distribution, with optimizer $m^*(K) = K + \frac{1}{h} + \frac{\kappa}{hB}e^{hK}$ and survival $\bar{F}(m^*(K)) = -G'(K) = hBe^{-hK}$ on any interval where the latter is at most one. The extension is an interior statement: on the finite lattice with one-sided boundary rows the additive constant does not scale — the one-sided identity requires $q\kappa = \frac{D(q)}{2}\kappa$, which fails unless $q = \frac{1}{2}$ or $\kappa = 0$ — so the boundary rows of the corrected finite-state model single out the pure exponential, and the affine extension holds only under a two-sided closure. Among enquiry-value curves, the affine-exponential family is the exact boundary of rigid behavior:

Theorem 3 (Strike rigidity)**.** *Let G be twice continuously differentiable, strictly decreasing, and convex on an open interval I — the properties Lemma 1 grants every enquiry-value curve — and suppose that for each q in a nondegenerate interval of imbalances there exist constants $\gamma(q)$ and $\delta(q)$ with*

$$H_q(A,S) \;=\; H_{1/2}\big(A+\gamma(q),\, S-\delta(q)\big)$$

for all (A,S) in an open set whose strikes $A \pm S$ fill a common interval I, the translated strikes $A+\gamma(q) \pm (S-\delta(q))$ remaining in I as well. Then $-G'$ satisfies the multiplicative translation property of Theorem 2, hence is a decaying exponential, and

$$G(K) \;=\; \kappa + B\,e^{-hK} \qquad \text{on } I$$

for some constants $\kappa \in \mathbb{R}$, $B>0$, $h>0$, and the translations are necessarily the standard ones, $\gamma(q) = \log M(q)/h$ and $\delta(q) = \ell/h$. If moreover I contains a right half-line and $G(K) \to 0$ along it, then $\kappa = 0$ and G is exponential.

Proof. Differentiate the identity with respect to S and with respect to A, and add:

$$2q\,G'(A+S) \;=\; G'\big(A+\gamma+S-\delta\big),$$

i.e. with $K = A+S$: $2q\,G'(K) = G'(K+t(q))$ for all K, where $t(q) = \gamma(q) - \delta(q)$; subtracting instead gives the sell-side relation $2(1-q)\,G'(K') = G'(K'+\gamma(q)+\delta(q))$. Write $f = -G'$. It is positive — at an interior zero of G', convexity would make G nondecreasing to the right, against strict decrease — and nonincreasing, by convexity. From $f(K_0+t(q)) = 2q\,f(K_0)$ the map $q \mapsto t(q)$ is strictly decreasing, hence continuous off a countable set; pick a continuity point q_0 and $q_n \to q_0$, so the nonzero shifts $\Delta_n = t(q_n) - t(q_0) \to 0$ satisfy, by composing the q_n and q_0 relations, $f(K+\Delta_n) = (q_n/q_0)\,f(K)$ for all relevant K. The difference quotient $\big[\log f(K+\Delta_n) - \log f(K)\big]/\Delta_n = \log(q_n/q_0)/\Delta_n$ is independent of K, and $\log f$ is differentiable since G is C^2; letting $n \to \infty$ makes $(\log f)'$ constant, so $f(K) = A\,e^{-hK}$ for some $A>0$ and real h. Since f is nonincreasing, $h \ge 0$; and $h=0$ would make f constant, forcing $2q = 1$, against the nondegenerate range. So $h>0$; writing $A = hB$ and integrating gives $G = \kappa + Be^{-hK}$. Substituting back into the two translation relations gives $e^{-h(\gamma-\delta)} = 2q$ and $e^{-h(\gamma+\delta)} = 2(1-q)$, hence $\gamma = \log M(q)/h$ and $\delta = \ell/h$. On a right half-line the tail condition forces $\kappa = 0$. ∎

Remark 2 (The concave impostor)**.** *Convexity is load-bearing. The decreasing, twice differentiable, concave family $G(K) = \kappa - Be^{hK}$ also balances rigidly, with the opposite-signed translations $(A,S) \mapsto (A-\gamma, S+\delta)$: a narrowing rather than a widening. Lemma 1 rules it out — no enquiry-value curve is concave — and the accompanying script checks its rigid balancing at machine precision alongside its exclusion.*

Corollary 2 (Quote rigidity)**.** *Suppose additionally that G is locally rationalized as an enquiry-value curve with a unique interior optimizer — as in the realization preceding Theorem 3, or Proposition 2 — so that $m^*(K) - K = G(K)/(-G'(K))$ by the envelope identity. Within the family of Theorem 3 this gives $m^*(K) = K + \frac{1}{h} + \frac{\kappa}{hB}e^{hK}$, so*

$$m^*(K+t) - m^*(K) \;=\; t + \tfrac{\kappa}{hB}\,e^{hK}\big(e^{ht}-1\big),$$

which, for $t \neq 0$, is independent of K only when $\kappa = 0$. Rigid translation of the break-even strikes therefore characterizes the affine-exponential class, while rigid translation of the submitted quotes*, over a nondegenerate range of imbalances, characterizes the pure exponential, with no tail condition.*

Remark 3 (The constant is a borderline tail)**.** *Along the family, $m^*(K)\,\bar{F}(m^*(K)) \to \kappa$ as K grows: a positive constant is exactly a $1/m$ survival tail, an infinite-mean-markup regime. A finite mean markup gives $m\bar{F}(m) \to 0$, since for nonincreasing $\bar{F}$ one has $\frac{m}{2}\bar{F}(m) \le \int_{m/2}^{m} \bar{F}(u)\,du \to 0$; hence $G(K) \le \sup_{m>K} m\bar{F}(m) \to 0$ and $\kappa = 0$. The pure exponential is the finite-mean member of the rigid class.*

Remark 4 (What each theorem governs)**.** *Theorem 2 is the right tool for the defect functional of the next section, which is built from the sidewise gauges. Theorem 3 governs the break-even strikes: rigid strike translations remove imbalance exactly when G is affine-exponential. Corollary 2 and Remark 3 then recover the pure exponential at the level of the submitted quotes and of global tails respectively. The defect functional $\mathcal{D}$ below measures the failure of the* chosen termwise gauge *to lift to a potential; Theorem 3 closes the remaining gap by ruling out every other rigid balancing of the strikes.*

Remark 5 (Rigidity run backwards is identification)**.** *A dealer who finds that sliding her* submitted quotes *by a constant absorbs the flow has, by Corollary 2, identified her market's rejection curve as exponential on the strike range her book visits, and learned its width there. If only her break-even strikes translate rigidly, she has identified the affine-exponential class, two constants short of the curve. Either way an observed sufficiency of rigid skewing becomes a statement about the shape of the competition.*

5 The integrability defect

The strikes (1) are not free sequences: adjacent inventory states share the increment of ν, which telescopes into the edge identity

$$K^{\downarrow}(x) + K^{\uparrow}(x+s) \;=\; 2\epsilon \qquad \text{for all } x. \tag{9}$$

In words: a buy at inventory x followed by a sell at $x+s$ is a round trip, the two marginal inventory costs cancel, and what remains is twice the overhead.

Lemma 3 (Reconstruction)**.** *A pair of sequences $\big(\tilde{K}^{\downarrow}(x), \tilde{K}^{\uparrow}(x)\big)$ derives from a single potential $\tilde{\nu}$ and overhead $\tilde{\epsilon}$ via (1) if and only if $\tilde{K}^{\downarrow}(x) + \tilde{K}^{\uparrow}(x+s) = 2\tilde{\epsilon}$ for all x, in which case $\tilde{\nu}$ is recovered, up to an additive constant, by summing $\tilde{\nu}(x+s) - \tilde{\nu}(x) = s\,(\tilde{K}^{\downarrow}(x) - \tilde{\epsilon})$.*

Proof. Necessity is the computation (9); sufficiency is the stated summation, whose consistency with the up-strikes is precisely the edge identity. □

After the pointwise gauge (8), the transformed strikes are $\tilde{K}^{\downarrow}(x) = T_{2q}(K^{\downarrow}(x))$ and $\tilde{K}^{\uparrow}(x) = T_{2(1-q)}(K^{\uparrow}(x))$, and by (9) the reconstruction condition becomes: the *defect functional*

$$\mathcal{D}(K) \;=\; T_{2q}(K) + T_{2(1-q)}(2\epsilon - K) \tag{10}$$

must be constant in K across the visited strike set. That equivalence is exact and governs everything in this section: lifting occurs if and only if $\mathcal{D}$ is constant on the strikes actually visited. Affinity of Φ guarantees it globally; curvature destroys it generically but not universally — a curved G can make $\mathcal{D}$ accidentally constant on a particular finite ladder, and at $q = \frac{1}{2}$ the obstruction vanishes trivially since both gauges are the identity.

Theorem 4 (The defect measures curvature; its constant part is the gauged overhead)**.** (i) *For affine* $\Phi = a + hK$,

$$\mathcal{D}(K) \;\equiv\; 2\epsilon - \frac{\log\big(4q(1-q)\big)}{h} \;=\; 2\epsilon + 2\gamma :$$

the defect is constant, the gauged problem is a balanced problem with potential $\tilde\nu(x) = \nu(x) - \delta x$ *and overhead* $\tilde\epsilon = \epsilon + \gamma$, *and the gauged-overhead displacement* γ *of Cotton (2026a) — an overhead-frame object, not a physical widening — is the value of the defect functional at the integrable point.* (ii) *For the deformation (4), with* ψ *twice continuously differentiable on a neighborhood of the shifted and reflected copies of the strike interval and the* $O(\eta^2)$ *remainders uniform there, and with* $a_1 = \frac{1}{h}\log(2q)$ *and* $a_2 = \frac{1}{h}\log\big(2(1-q)\big)$ *(so* $a_1 + a_2 = -2\gamma$*),*

$$T_\lambda(K) \;=\; K - \frac{\log\lambda}{h} + \frac{\eta}{h}\Big[\psi(K) - \psi\Big(K - \frac{\log\lambda}{h}\Big)\Big] + O(\eta^2), \tag{11}$$

and

$$\mathcal{D}(K) \;=\; 2(\epsilon+\gamma) + \frac{\eta}{h}\Big[\psi(K) - \psi(K-a_1) + \psi(2\epsilon - K) - \psi(2\epsilon - K - a_2)\Big] + O(\eta^2). \tag{12}$$

Affine components of ψ *contribute a constant to (12) and no* K*-dependence: they renormalize the exponential width and overhead and create no genuine defect. The variation of* $\mathcal{D}$ *therefore begins, at first order in* η, *with a second-difference functional of* ψ *— non-exponential curvature is the integrability defect of the local imbalance gauge.*

Proof. (i) Substitute $T_\lambda(K) = K - \frac{1}{h}\log\lambda$ into (10); the K's cancel and $\log(2q)+\log(2(1-q)) = \log(4q(1-q)) = -2h\gamma$. That $\tilde\nu = \nu - \delta x$ matches: $\tilde K^{\downarrow} = K^{\downarrow} - \frac{1}{h}\log 2q$ and $\tilde\epsilon + \mathrm{D}^+\tilde\nu = \epsilon + \gamma + \mathrm{D}^+\nu - \delta$ agree because $\gamma - \delta = \frac{1}{h}\log\frac{1}{2q}$. (ii) Write $\Phi_\eta(T) = \Phi_\eta(K) - \log\lambda$ and expand $T = T^{(0)} + \eta T^{(1)} + O(\eta^2)$: at order zero $h(T^{(0)} - K) = -\log\lambda$; at order one $hT^{(1)} + \psi(T^{(0)}) - \psi(K) = 0$, giving (11); summing the two gauges per (10) gives (12). For affine $\psi(K) = uK + v$ the bracket in (12) is $u(a_1 + a_2) = -2u\gamma$, a constant. ☐

Corollary 3 (A quantitative curvature bound)**.** *Write* $B_\psi(K)$ *for the bracket in (12), and let* $\|\psi''\|_\infty$ *be taken over the reflected and shifted copies of the visited strike interval* I *at which (12) evaluates* ψ. *Then* $|B'_\psi(K)| \le (|a_1| + |a_2|)\,\|\psi''\|_\infty$, *and on* I,

$$\operatorname*{osc}_{K\in I}\mathcal{D}(K) \;\le\; \frac{|\eta|}{h}\,\mathrm{diam}(I)\,\big(|a_1| + |a_2|\big)\,\|\psi''\|_\infty \;+\; O(\eta^2).$$

Proof. $B'_\psi(K) = \big[\psi'(K) - \psi'(K-a_1)\big] - \big[\psi'(2\epsilon - K) - \psi'(2\epsilon - K - a_2)\big]$, and each bracket is bounded by $|a_i|\,\|\psi''\|_\infty$ by the mean value theorem; integrate over I. ☐

The bound controls the failure of the canonical sidewise gauge to reconstruct a common potential. It does not by itself bound the true policy correction, because the centered equation has null directions — notably additive constants in G, for which the defect is nonconstant while the potential never moves. After a deformation path and a normalization transverse to those null directions are fixed, curvature controls the policy correction directly through the forcing: that is Corollary 5 below, the rigorous successor of the slowly-varying-hazard remark of Cotton (2026a).

6 The perturbation hierarchy

A nonconstant defect obstructs compression by the canonical gauge, not a computation — and, by the null directions of the centered equation, not necessarily the policy. Work on the inventory lattice $x \in \{-Ns, \dots, Ns\}$ with the normalization $\nu(0) = 0$, write $G_\eta = G_\star e^{-\Phi_\eta}$ with (4), and define the residual map

$$\mathcal{F}_x(\nu, \eta) \;=\; q\, G_\eta\big(\epsilon + \mathrm{D}^+\nu(x)\big) + (1-q)\, G_\eta\big(\epsilon + \mathrm{D}^-\nu(x)\big) - \big[\, x \to 0 \,\big] - \frac{\tau c(x)}{s}, \tag{13}$$

where $\mathrm{D}^\pm\nu(x) = (\nu(x \pm s) - \nu(x))/s$ and $[\, x \to 0\,]$ denotes the same expression at $x = 0$, together with a boundary closure at the lattice ends; the closure used here, as in the script accompanying Cotton (2026a), sets the third difference to zero at each end,

$$\nu(\pm Ns) - 3\nu(\pm(N-1)s) + 3\nu(\pm(N-2)s) - \nu(\pm(N-3)s) \;=\; 0,$$

extrapolating the convexity of the potential through the boundary. At $\eta = 0$ the exact solution is known from Cotton (2026a): $\nu_0 = \nu_{\mathrm{bal}} + \delta x$, the tilt of the balanced solve at cost $M(q)\, c$. The hierarchy applies to any smooth path $\eta \mapsto G_\eta$ through the anchor, with $\psi = \partial_\eta \Phi_\eta|_{\eta=0}$ its first-order deformation; the straight line (4) is the default path.

Theorem 5 (One linear operator computes every order)**.** *Let $K_0^\downarrow, K_0^\uparrow$ denote the strikes of ν_0, write $G_0 = G_{\eta=0}$, and suppose the linearization*

$$(L_0 u)(x) \;=\; h\, \Delta_0\Big[\, q\, G_0(K_0^\downarrow)\, \mathrm{D}^+ u + (1-q)\, G_0(K_0^\uparrow)\, \mathrm{D}^- u\,\Big](x), \qquad \Delta_0 f(x) = f(x) - f(0), \tag{14}$$

closed with the boundary rows and the pinning $u(0) = 0$, is nonsingular. If G_η is C^r jointly in (K, η) on the relevant domain, then for small η the consistency equation has a unique C^r branch ν_η, the expansion $\nu_\eta = \nu_0 + \eta\nu_1 + \cdots$ with $\nu_j = \frac{1}{j!}\partial_\eta^j \nu_\eta|_{\eta=0}$ is valid through order r, and

$$L_0\, \nu_1 = -R_\psi, \qquad R_\psi(x) = \Delta_0\Big[\, q\, G_0(K_0^\downarrow)\, \psi(K_0^\downarrow) + (1-q)\, G_0(K_0^\uparrow)\, \psi(K_0^\uparrow)\,\Big](x), \tag{15}$$

with, for every $2 \le j \le r$, $L_0\, \nu_j = R_j$, where R_j is an explicit forcing constructed from $\nu_1, \dots, \nu_{j-1}$ and the path jet of Φ_η through order j — for the straight line (4) the higher η-derivatives of Φ_η vanish and R_j uses ψ alone; in particular

$$L_0 \nu_2 \;=\; \lim_{\eta \to 0} \frac{1}{\eta^2}\, \mathcal{F}\big(\nu_0 + \eta\nu_1,\ \eta\big).$$

If G_η is real analytic in (K, η), the branch is analytic and the series converges on a neighborhood of $\eta = 0$.

Proof. $\mathcal{F}(\nu_0, 0) = 0$ by the theorem of Cotton (2026a). The partial $\partial_\nu \mathcal{F}(\nu_0, 0)$ is $-L_0$, since $\mathcal{F}$ carries $+H_q$ and $\partial_K G_0 = -h\, G_0$; the partial $\partial_\eta \mathcal{F}(\nu_0, 0)$ is $-R_\psi$, since $\partial_\eta G_\eta(K)\big|_{\eta=0} = -\psi(K)\, G_0(K)$. The implicit function theorem in the stated regularity class gives the branch, and $\nu_1 = -(\partial_\nu \mathcal{F})^{-1} \partial_\eta \mathcal{F}$ is (15). The displayed formula for ν_2 is the order-η^2 term of $0 = \mathcal{F}(\nu_\eta, \eta) = \mathcal{F}(\nu_0 + \eta\nu_1, \eta) - \eta^2 L_0 \nu_2 + O(\eta^3)$, and the same rearrangement produces each higher R_j from residual evaluations along the already-computed truncation. $\square$

Remark 6 (Structure and cost of the hierarchy)**.** *Before centering, the operator inside* (14) *is a birth–death tridiagonal operator with positive weights* $qG_0(K_0^{\downarrow})$ *and* $(1-q)G_0(K_0^{\uparrow})$*; the centering* Δ_0 *subtracts one fixed linear functional of* u *from every row, so* L_0 *is tridiagonal in the interior, carries two narrow-band boundary rows from the closure, and adds a rank-one centering correction; retaining the ergodic constant as an unknown instead of eliminating it through* Δ_0 *gives a banded block with one border column. After the pinning* $\nu(0) = 0$ *there are* $2N$ *unknowns, matched by the* $2N-2$ *centered interior equations and the two boundary rows. Nonsingularity is not a technicality: in the zero-carry limit the anchor operator genuinely develops null modes, so the hypothesis encodes the economics of a binding carrying cost. Either way the matrix is factored once, for the fixed anchor problem (fixed* q*, cost, grid, and overhead), and every order of the expansion and every basis direction of Section 7 is then one back-substitution; nothing nonlinear is ever re-solved. The accompanying script factors* L_0 *once and verifies, on a Weibull-born deformation,* $\|\nu_\eta - \nu_0 - \eta\nu_1\|_\infty = O(\eta^2)$ *and* $\|\nu_\eta - \nu_0 - \eta\nu_1 - \eta^2\nu_2\|_\infty = O(\eta^3)$*, with observed error ratios* $3.9 \approx 4$ *and* $8.1 \approx 8$ *under halving of* η *(Section 9).*

Remark 7 (The rigid manifold and its coordinate)**.** *A fully normal-form treatment would expand transversely to the exactly reducible manifold* $\kappa(\eta) + B(\eta)e^{-h(\eta)K}$ *rather than from the pure exponential alone, treating exact directions exactly and only transverse curvature perturbatively. The coordinate adapted to that split is* $\Xi = -\log(-G')$*: by Theorem 3, affine* Ξ *is exactly the rigid class, and* Ξ *removes the null constant automatically, while* Φ *remains the coordinate of the pointwise translation — the two geometries serve different theorems. The osculating rigid anchor at a reference strike* K_0 *with* $G''(K_0) > 0$ *matches* G*,* G'*,* G''*:*

$$h_0 = \frac{G''(K_0)}{-G'(K_0)}, \qquad B_0 e^{-h_0 K_0} = \frac{[G'(K_0)]^2}{G''(K_0)}, \qquad \kappa_0 = G(K_0) - \frac{[G'(K_0)]^2}{G''(K_0)};$$

equivalently, fit an affine tangent to Ξ *and recover* κ_0 *from the value of* G*. At the pure exponential point the tangent space of the rigid manifold in the* Φ*-chart is* $\mathrm{span}\{1, K, e^{hK}\}$*: the deformation* (4) *absorbs the first two directions into* a *and* h*, and the third is the additive-*κ *null direction, which is why a curved* ψ *can carry a gauge defect while leaving the potential unmoved. We leave the transverse expansion to future work and use the pure-exponential anchor with a reported path.*

The value correction becomes a quote correction by one more envelope evaluation. Combining (3) with the first-order condition $m^* - K = 1/h(m^*)$ of Cotton (2026a) gives $m^*_\eta(K) = K + 1/\Phi'_\eta(K)$, and the exact strikes expand as $K^{\downarrow}_\eta(x) = K^{\downarrow}_0(x) + \eta\,\mathrm{D}^+\nu_1(x) + O(\eta^2)$, and likewise on the sell side with D^-, so:

Corollary 4 (The tangent policy)**.** *On interior states, the optimal markdown and markup of the deformed problem satisfy*

$$m^{*\downarrow}_\eta(x) \;=\; K^{\downarrow}_0(x) + \frac{1}{h} \;+\; \eta\left[\mathrm{D}^+\nu_1(x) \;-\; \frac{\psi'\big(K^{\downarrow}_0(x)\big)}{h^2}\right] \;+\; O(\eta^2),$$
$$m^{*\uparrow}_\eta(x) \;=\; K^{\uparrow}_0(x) + \frac{1}{h} \;+\; \eta\left[\mathrm{D}^-\nu_1(x) \;-\; \frac{\psi'\big(K^{\uparrow}_0(x)\big)}{h^2}\right] \;+\; O(\eta^2).$$

The first-order quote correction decomposes into the inventory-value response $\mathrm{D}^{\pm}\nu_1$ *and the direct markup-distribution response* $-\psi'/h^2$*; the two are computable separately, from one linear solve and one derivative evaluation.*

Remark 8 (The curvature amplification factor)**.** *From* (15), $\|\nu_1\| \le \|L_0^{-1}\|\,\|R_\psi\|$*. The forcing is bounded by the curvature through Corollary 3-type estimates, but the resolvent norm belongs to the anchor. In a weighted-*ℓ^2 *realization of the interior birth–death block it would be controlled by a spectral gap; we do not develop that here, and use* $\|L_0^{-1}\|$ *only as the conditioning amplification factor of the anchor problem, which grows as the anchor approaches a low-carry or weakly mean-reverting regime. Small uncertainty in win-curve curvature therefore need not imply small uncertainty in policy when the inventory dynamics are near instability. Since the submitted quote also responds directly through* ψ'*, the factor amplifies the* inventory-feedback *component of the policy rather than the whole of it, and it is the quantitative bridge between the present paper and robust policy learning built on it.*

Corollary 5 (Sensitivity bounds: inventory potential and submitted quotes)**.** *Normalize* $\psi(K_0) = \psi'(K_0) = 0$ *at a reference strike, let* R *be the largest distance* $|K - K_0|$ *over the two strike ladders of* ν_0*, and let* $G_{\max}$ *be the largest value of* G_0 *there; take* $\|\psi''\|_\infty$ *over the convex hull of* K_0 *and the visited strikes. Then*

$$\|R_\psi\|_\infty \;\le\; G_{\max}\, R^2\, \|\psi''\|_\infty, \qquad \|\nu_1\|_\infty \;\le\; \|L_0^{-1}\|_\infty\, G_{\max}\, R^2\, \|\psi''\|_\infty,$$

and the first-order submitted-quote correction $m_1^\pm = \mathrm{D}^\pm \nu_1 - \psi'(K_0^\pm)/h^2$ *of Corollary 4 satisfies*

$$\|m_1^\pm\|_\infty \;\le\; \Big(\|\mathrm{D}^\pm L_0^{-1}\|_\infty\, G_{\max}\, R^2 + \frac{R}{h^2}\Big)\|\psi''\|_\infty.$$

The first term is the amplified inventory-feedback component and the second the direct quote-map component; the potential can be unmoved while the quotes move, as along the additive null direction.

Proof. By Taylor's theorem with the normalization, $|\psi(K)| \le \frac{1}{2}|K-K_0|^2\|\psi''\|_\infty \le \frac{1}{2}R^2\|\psi''\|_\infty$ and $|\psi'(K)| \le R\,\|\psi''\|_\infty$. Each Hamiltonian term in R_ψ is a q-weighted average of values $G_0\psi$, bounded by $G_{\max}\|\psi\|_\infty$, and the centering difference at most doubles the bound; apply $\|L_0^{-1}\|$, and for the quotes apply $\mathrm{D}^\pm L_0^{-1}$ to the forcing and bound the direct term by $R\|\psi''\|_\infty/h^2$. $\square$

Remark 9 (The robustness experiment of Cotton (2026a), reinterpreted)**.** *Cotton (2026a) reported numerically that the error in the zero-inventory skew under a slowly varying hazard is about a quarter of the hazard's relative variation across the visited strikes, shrinking linearly. In the present language that experiment perturbs* Φ *by a specific near-quadratic* ψ *and reads one component of* $\eta\nu_1$*; the observed constant is a functional of* ψ *computable from* (15)*. The robustness statement is thereby upgraded from an error bound to the leading term of a convergent expansion.*

7 From distributions to deformations

The hierarchy consumes deformations of the *effective log-value*; named distribution families are specified by deformations of the *log-survival curve*. The envelope theorem converts one to the other without recomputing the optimal markup.

Proposition 1 (Envelope transfer)**.** *Let* $-\log \bar{F}_\theta(m) = hm + \sum_j \theta_j\, r_j(m)$ *with each* r_j *three times continuously differentiable, and the optimizer of* $(m-K)\bar{F}_\theta(m)$ *unique, interior, and* *nondegenerate, uniformly over the compact strike interval considered. Then, with* $m_0(K) = K + \frac{1}{h}$ *the exponential optimum,*

$$\Phi_\theta(K) \;=\; \Phi_0(K) + \sum_j \theta_j\, r_j\big(m_0(K)\big) \;-\; \frac{1}{2h^2}\sum_{j,k}\theta_j\theta_k\, r_j'\big(m_0(K)\big)\, r_k'\big(m_0(K)\big) + O(\|\theta\|^3). \quad (16)$$

To first order the movement of the optimizer drops out: the deformation of the log-value is the deformation of the log-survival curve evaluated at the exponential optimum. If the family is itself nonlinear in θ*,* $-\log \bar F_\theta = hm + \theta r(m) + \theta^2 r_2(m) + \cdots$*, the family's own curvature adds* $\theta^2 r_2(m_0(K))$ *to* (16) *at second order.*

Proof. Write $g(m;\theta) = \log(m-K) - hm - \sum_j \theta_j r_j(m)$, so $\log G_\theta(K) = g(m^*(\theta);\theta)$. By the envelope theorem $\partial_{\theta_j} \log G_\theta = -r_j(m^*)$, which at $\theta = 0$ is $-r_j(m_0)$. For the second order, differentiate the first-order condition $\frac{1}{m-K} - h - \sum \theta_j r_j'(m) = 0$: with $\partial_m^2 g\big|_0 = -h^2$, one gets $\partial_{\theta_k} m^*\big|_0 = -r_k'(m_0)/h^2$, hence $\partial^2_{\theta_j\theta_k} \log G\big|_0 = -r_j'(m_0)\,\partial_{\theta_k} m^*\big|_0 = r_j'(m_0) r_k'(m_0)/h^2$. Change sign for $\Phi = -\log(G/G_\star)$. □

Corollary 6 (Weibull example)**.** *For* $\bar F_k(m) = e^{-(hm)^k}$ *with* $k = 1+\eta$*, the expansion* $(hm)^{1+\eta} = hm + \eta\, hm \log(hm) + \frac{\eta^2}{2} hm \log^2(hm) + \cdots$ *gives* $r(m) = hm\log(hm)$ *and hence the effective deformation*

$$\psi(K) \;=\; (1+hK)\,\log(1+hK),$$

ready for insertion into (15)*, on the domain* $1 + hK > 0$*, where the anchor optimum* $m_0(K) = K + 1/h$ *is a positive markup. Note* $\psi'(0) = h$*: the raw tangent combines genuine curvature with a first-order recalibration of the local exponential rate, and for the fixed-scale Weibull family the raw* ψ *is the actual path tangent. The transverse part* $\psi_\perp(K) = (1+hK)\log(1+hK) - hK$ *corresponds to simultaneously recalibrating the exponential rate, and the removed affine response must then be carried by the anchor parameters.*

The hierarchy perturbs through effective log-values, and one should ask when a proposed Φ corresponds to any markup distribution at all. The question is a rationalizability question of the kind familiar from producer theory, where indirect profit functions determine supply through the envelope; the modern converse direction is Sinander (2022). The envelope identities answer it locally here: the proposition constructs a survival curve on $m^*(I)$ and proves optimality over that interval, a local sufficiency result rather than a global rationalization.

Proposition 2 (Local admissibility)**.** *Let* G *be* C^2 *on an interval* I *with*

$$G > 0, \qquad -1 < G' < 0, \qquad G'' > 0$$

on I *— equivalently, in the* Φ*-coordinates used by the hierarchy,* $\Phi' > 0$*,* $\Phi'' < \Phi'^2$*, and* $G\Phi' < 1$*. Define* $m^*(K) = K - G(K)/G'(K) = K + 1/\Phi'(K)$ *and* $\bar F(m^*(K)) = -G'(K) = G(K)\,\Phi'(K)$*. Then* m^* *is strictly increasing on* I*,* G *is strictly convex,* $\bar F$ *is a well-defined strictly decreasing function with values in* $(0,1)$ *on* $m^*(I)$*, and for every strike* $\widetilde K$ *with* $m^*(\widetilde K)$ *interior to* $m^*(I)$ *the program* $\sup_m (m - \widetilde K)\,\bar F(m)$ *over* $m^*(I)$ *has the unique interior optimum* $m^*(\widetilde K)$ *with value* $G(\widetilde K)$*. In particular, sufficiently small* C^2 *deformations of an affine anchor that satisfies the three inequalities with a uniform margin on a compact strike interval remain admissible there.*

Proof. $dm^*/dK = 1 - \Phi''/\Phi'^2 > 0$ and, from $G' = -G\Phi'$, $G'' = G(\Phi'^2 - \Phi'') > 0$; also $\frac{d}{dK}[G\Phi'] = G(\Phi'' - \Phi'^2) < 0$, so $\bar{F}$ is strictly decreasing along the strictly increasing curve m^*, with values in $(0, 1)$ by hypothesis. The hazard of $\bar{F}$ along the curve is

$$h_{\bar{F}}\big(m^*(K)\big) = -\frac{d}{dm}\log \bar{F} = \frac{-\frac{d}{dK}\log\big[G(K)\Phi'(K)\big]}{dm^*/dK} = \frac{\Phi' - \Phi''/\Phi'}{1 - \Phi''/\Phi'^2} = \Phi'(K),$$

so the first-order condition $1/(m - \widetilde{K}) = h_{\bar{F}}(m)$ holds at $m = m^*(\widetilde{K})$, with value $(1/\Phi') \cdot G\Phi' = G(\widetilde{K})$. Uniqueness and global optimality are immediate: every $m \in m^*(I)$ is $m^*(K)$ for a unique K, with $h_{\bar{F}}(m^*(K)) = \Phi'(K) = 1/(m^*(K) - K)$, so the m-derivative of the log-objective along the curve is

$$\frac{1}{m^*(K) - \widetilde{K}} - \frac{1}{m^*(K) - K} \;=\; \frac{\widetilde{K} - K}{\big(m^*(K) - \widetilde{K}\big)\big(m^*(K) - K\big)},$$

positive for $K < \widetilde{K}$ and negative for $K > \widetilde{K}$ wherever $m^*(K) > \widetilde{K}$, while the objective is nonpositive where $m^*(K) \le \widetilde{K}$: the objective rises until $K = \widetilde{K}$ and falls after. The final statement holds because the three inequalities are open under C^2 perturbation wherever they hold with a uniform margin. □

Remark 10 (A response library, not one calculation per family)**.** *Expand a log-survival deformation on a basis $\{r_j\}$ — splines, Chebyshev polynomials, or hazard bumps, on the compact strike interval the policy actually visits — and precompute the responses $u_j = -L_0^{-1} R_{\psi_j}$ once, at the exponential anchor. Then*

$$\nu(\theta) \;=\; \nu_0 + \sum_j \theta_j\, u_j + \tfrac{1}{2}\sum_{j,k} \theta_j \theta_k\, u_{jk} + \cdots$$

reprices Weibull shape perturbations, Gompertz hazards, gamma and lognormal displacement families, log-logistic and bounded-support curves, mixtures, and empirically estimated survival curves (Fermanian et al., 2017) from one factorization of one matrix, for the fixed anchor problem.

Two qualifications keep the claim precise. Some of these families contain the exponential as a native parameter limit (Weibull, gamma, Gompertz) and are reached along their own parameter homotopy; others are reached through a basis approximation of their log-survival difference on the visited interval, with the approximation error entering alongside the truncation error. And Proposition 1 is local, conditional on a unique interior optimizer throughout the range traversed. Within those bounds the theory is a functional expansion of the log-survival curve; the named families merely select coordinates θ. If the splines are C^2 rather than analytic, Theorem 5 applies in its C^r form, through second order.

8 The distribution-free parity theorem

Theorem 6 (Parity)**.** *Suppose the buy and sell sides face the same enquiry-value curve G, the carrying cost is even, $c(-x) = c(x)$, and the consistency equation (2) (with any reflection-symmetric boundary closure) has a locally unique solution branch depending on ℓ in class C^r. Then*

$$\nu_{1-q}(x) \;=\; \nu_q(-x) \qquad \textit{(after the normalization } \nu(0) = 0\textit{)},$$

so $S_q(0)$ is odd and $C_q(0)$ is even in ℓ: every derivative of the wrong parity vanishes at balance, and the Taylor jets through order r contain only the stated powers,

$$S_q(0) = d_1\ell + d_3\ell^3 + \cdots, \qquad C_q(0) - C_{1/2}(0) = g_2\ell^2 + g_4\ell^4 + \cdots,$$

with $o(\ell^r)$ remainders; if the branch is real analytic the displayed series converge.

Proof. Let ν solve the problem at imbalance q and set $\tilde{\nu}(x) = \nu(-x)$. Then $\mathrm{D}^{\pm}\tilde{\nu}(x) = \mathrm{D}^{\mp}\nu(-x)$, so the strikes swap sides at the reflected inventory, and since c is even, $\tilde{\nu}$ satisfies (2) with q replaced by $1-q$. Local uniqueness identifies $\tilde{\nu}$ with ν_{1-q}. Then $S_{1-q}(0) = -S_q(0)$ and $C_{1-q}(0) = C_q(0)$; since $q \mapsto 1-q$ is $\ell \mapsto -\ell$, oddness and evenness follow. A C^r odd (even) function of ℓ has only odd (even) powers in its Taylor jet, and analyticity upgrades the jet to a convergent series. □

Corollary 7 (Parity of the submitted quotes)**.** *From $\nu_{1-q}(x) = \nu_q(-x)$, the strikes reflect: $K^{\downarrow}_{1-q}(x) = K^{\uparrow}_q(-x)$ and $K^{\uparrow}_{1-q}(x) = K^{\downarrow}_q(-x)$. The same quote map $m^*(\cdot)$ serves both sides, so at zero inventory the submitted quotes swap under $\ell \mapsto -\ell$: the submitted-quote skew is odd and the submitted-quote half-width even in the log-odds of the flow. The theorem's parity of S and C concerns the break-even geometry; this corollary is its observable counterpart.*

Remark 11 (Parity versus leading order)**.** *Symmetry alone does not make d_1 or g_2 nonzero, nor sign the width response: in principle a particular curve could have cubic-leading skew or quartic-leading widening. What the theorem establishes is the parity; the familiar reading — skew first order in imbalance, width second order — holds under the nondegeneracy $d_1 \neq 0$, $g_2 \neq 0$. The exponential family satisfies it with zero-inventory skew $S_q(0) = \ell/h$ exactly, at all orders. The width response is a different object from the displacement $\gamma = \frac{1}{h}\log\cosh\ell$ of the normal form: γ is the exogenous translation of the transformed overhead coordinate, while $C_q(0) - C_{1/2}(0)$ is the endogenous convexity response, which depends on the carrying cost through the multiplier $M(q)$; the script's $g_2 \approx 0.019$ measures the latter. The strongly curved test curve of Section 9 exhibits both nondegeneracies, with $d_1 \approx 0.797/h$ against the exponential $1/h$. Exponentiality supplies coefficients; parity is universal; nondegeneracy sits in between, generic but not automatic.*

9 Numerical checks

A single seventeen-state experiment checks instances, not theorems. What the accompanying script `verify_normal_form.py` provides is a reproducibility and diagnostic check of the formulas and their predicted asymptotic orders, on one fully specified configuration in the lattice formulation of the corresponding script of Cotton (2026a): $N = 8$ (states $x \in \{-8, \ldots, 8\}$), $c(x) = 0.002x^2$, $q = 0.6$, $h = 1$, $\epsilon = 0.3$, $s = \tau = 1$, the third-difference closure of Section 6, the Weibull-born deformation $\psi(K) = (1 + hK)\log(1 + hK)$, and η-ladders halving from 0.1 or 0.04; the script and its output live with the paper's source. Expansion checks compare a formula against an independent nonlinear solve or direct optimization rather than reevaluating the same expression, and the reported balancing gaps are relative, normalized by $|H_q|$. It verifies:

- the normal form (5) and translation form (6) on randomized (A, S, q) at relative error $\sim 10^{-16}$, for a deformation far from exponential ($\eta = 0.35$), and the gauge property $G(T_\lambda(K)) = \lambda\, G(K)$ at $\sim 10^{-15}$;

- the affine defect $\mathcal{D} \equiv 2\epsilon + 2\gamma$ at $\sim 10^{-16}$, and for curved Φ the expansions (11)–(12) with $O(\eta^2)$ remainders (ratios 3.6–3.8 ≈ 4 under halving), with the measured oscillation inside the bound of Corollary 3 by a factor of about two;

- the hierarchy of Theorem 5: $\|\nu_\eta - \nu_0 - \eta\nu_1\|_\infty$ shrinks at ratio $3.9 \approx 4$ and $\|\nu_\eta - \nu_0 - \eta\nu_1 - \eta^2\nu_2\|_\infty$ at ratio $8.1 \approx 8$; the resolvent inequality of Remark 8 holds with amplification factor $\|L_0^{-1}\|_\infty \approx 73$ on the test anchor;

- the envelope transfer (16) against direct optimization of the true Weibull families: first order at ratio ≈ 4, second order (including the family's own curvature term) at ratio ≈ 8;

- the rigidity boundary of Theorem 3 in both directions: the affine-exponential family balances rigidly, and satisfies $2q\,G'(K) = G'(K + \gamma - \delta)$, at $\sim 10^{-16}$, while the curved test curve's balancing gap, *minimized over all translation pairs*, stays above 10^{-2}; the concave impostor balances rigidly at $\sim 10^{-16}$ with the opposite-signed translations, and $m^*\,\bar{F}(m^*) \to \kappa$ numerically along the affine family;

- the admissibility construction of Proposition 2, with the reconstructed survival curve re-optimizing to G at $\sim 10^{-16}$, and the tangent policy of Corollary 4 at ratio $3.9 \approx 4$;

- the parity of Theorem 6: $|\nu_{1-q}(x) - \nu_q(-x)| \sim 10^{-15}$, with coefficients $d_1 \to 0.797$ and $g_2 \to 0.019$ stable under $\ell \to \ell/2$;

- first-order path dependence: the additive path's tangent produces exactly zero forcing and zero ν_1, while the Φ-line to the same endpoint produces $\|\nu_1\| \approx 0.06$; both sensitivity chains of Corollary 5, for the potential and for the submitted quotes, hold on the test problem;

- the two-representations statement of Corollary 1: the imbalanced dealer's half-width equals the balanced-at-cost-$M(q)c$ dealer's to machine precision, while exceeding the balanced-at-c dealer's by the convexity response only (6.5×10^{-4} in this configuration, an order of magnitude below $\gamma \approx 0.020$).

10 Relation to earlier work and the program

Cotton (2026a) proved the exact compression under exponentiality and remarked that the assumption is needed only locally; the present paper develops that remark into a perturbation theory, with the exponential case as the integrable anchor of a normal form.

Cotton (2026b) plays the analogous role in the transient Avellaneda–Stoikov line (Avellaneda and Stoikov, 2008; Guéant et al., 2012), where the exponential intensity curve is what linearizes the Hamilton–Jacobi–Bellman equation in the first place; its concluding problem — a finite-horizon robustness statement for locally exponential intensity curves — is now posed correctly as the finite-horizon instance of this paper's hierarchy, with the tilted terminal condition joining the forcing terms.

The nearest existing literature treats general intensity curves and asymmetric sides numerically or through general Hamiltonians without structural reductions. Guéant and Lehalle (2015) is the canonical treatment of general intensity shapes, on the execution side; Guéant (2017) allows fully general asymmetric intensities and characterizes quotes through an implicit hazard-type condition solved through the HJB system; Bergault and Guéant (2021) does the same at scale for OTC market makers; Campi and Zabaljauregui (2020) handles general intensities under partial information via viscosity solutions. To our knowledge, none states a structural reduction of the asymmetric problem, which is what the normal form and its boundary provide.

The perturbation genre around solvable market-making models expands in scalar parameters (Bergault et al., 2021); the hierarchy here deforms the win curve as a function. Among contemporaneous work, Barzykin and Ciceri (2026) works with a general fill-probability curve and exact quote-map inversion but without symmetry structure, and Barzykin et al. (2025) obtains first-order corrections in effect-strength parameters for adverse selection and price reading.

The algebraic ancestry is the same as in those papers: the identity behind Theorem 1 is the Cramér–Chernoff exponential tilt of an asymmetric walk (Chernoff, 1952; Feller, 1968), and the affine case's rigid translation is the birth–death symmetrization known since Ledermann and Reuter (1954) and Karlin and McGregor (1957). The envelope argument of Proposition 1 is the standard machinery of Milgrom and Segal (2002); the win-curve estimation that would supply empirical θ coordinates is the econometrics of Fermanian et al. (2017).

Two further ancestries deserve naming. The core of Theorem 2 is the multiplicative Cauchy equation, and its conclusion is the lack-of-memory characterization of the exponential law: the systematic reference is Galambos and Kotz (1978), the regularity machinery Aczél (1966), and Marsaglia and Tubilla (1975) show how far the continuum-of-shifts hypothesis can be weakened; what is not in that literature is the two-sided statement, Theorem 3 lifting the characterization from the distribution to the pricing Hamiltonian, where the memoryless family acquires its affine extension. And conditions under which per-edge weights lift to a potential are the territory of reversibility theory (Kelly, 1979), where the obstruction lives on cycles; the inventory lattice is a line, so the obstruction of Section 5 enters instead through the pairing of up and down strikes across each bond, a pricing-theoretic member of that family of discrete integrability conditions rather than a transplant of the cycle criterion.

What is new here is the geometry above that algebra. For arbitrary enquiry-value curves the tilt acts as an exact translation in the log-value coordinates, and rigid strike translations work exactly on the affine-exponential family, with the pure exponential recovered from quote rigidity or a finite-mean tail (Theorems 2 and 3, Corollary 2). The failure to lift the pointwise gauge to a potential is a curvature functional with the gauged-overhead displacement of Cotton (2026a) as its integrable value and an explicit oscillation bound (Theorem 4, Corollary 3). The correction, to the potential and to the quotes, is a hierarchy driven by one bordered-tridiagonal operator with envelope-transferred, admissibility-checked forcing (Theorem 5, Corollary 4, Propositions 1 and 2). The parity theorem, finally, marks the boundary between what imbalance does to a dealer because of arithmetic — oddness and evenness, which no model in the class can evade — and what it does because the market's rejection curve is memoryless.

Two objects built here are the intended interface to sequels: the tangent policy of Corollary 4, which tells a learning system what the first-order-correct quotes are rather than only what the corrected value function is, and the curvature amplification factor of Remark 8, which prices how win-curve uncertainty propagates to policy uncertainty through the anchor's conditioning. A robust controller assembled from the exponential anchor, the tangent responses, and a learned residual is developed separately.

References

Aczél, J. (1966). *Lectures on Functional Equations and Their Applications*. Academic Press.

Avellaneda, M., Stoikov, S. (2008). High-frequency trading in a limit order book. *Quantitative Finance* 8(3), 217–224.

Barzykin, A., Ciceri, C. (2026). Bond market making with a hit-ratio target. arXiv:2604.20406.

Barzykin, A., Bergault, P., Guéant, O., Lemmel, M. (2025). Optimal quoting under adverse selection and price reading. arXiv:2508.20225.

Bergault, P., Guéant, O. (2021). Size matters for OTC market makers: general results and dimensionality reduction techniques. *Mathematical Finance* 31(1), 279–322.

Bergault, P., Evangelista, D., Guéant, O., Vieira, D. (2021). Closed-form approximations in multi-asset market making. *Applied Mathematical Finance* 28(2), 101–142.

Campi, L., Zabaljauregui, D. (2020). Optimal market making under partial information with general intensities. *Applied Mathematical Finance* 27(1–2), 1–45.

Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics* 23(4), 493–507.

Cotton, P. (2026a). On a simple relationship between order imbalance, skew and width in over-the-counter trading. Working paper, https://inventory.microprediction.org.

Cotton, P. (2026b). Finite-horizon market making with imbalanced flow: an exact symmetrization and its term structure. Working paper, https://inventory.microprediction.org.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed., Ch. XIV. Wiley.

Fermanian, J.-D., Guéant, O., Pu, J. (2017). The behavior of dealers and clients on the European corporate bond market: the case of multi-dealer-to-client platforms. *Market Microstructure and Liquidity* 2(3–4), 1750004.

Friedman, L. (1956). A competitive-bidding strategy. *Operations Research* 4(1), 104–112.

Galambos, J., Kotz, S. (1978). *Characterizations of Probability Distributions*. Lecture Notes in Mathematics 675, Springer.

Guéant, O. (2017). Optimal market making. *Applied Mathematical Finance* 24(2), 112–154.

Guéant, O., Lehalle, C.-A. (2015). General intensity shapes in optimal liquidation. *Mathematical Finance* 25(3), 457–495.

Guéant, O., Lehalle, C.-A., Fernandez-Tapia, J. (2012). Dealing with the inventory risk: a solution to the market making problem. *Mathematics and Financial Economics* 7(4), 477–507.

Karlin, S., McGregor, J. (1957). The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Transactions of the American Mathematical Society* 85(2), 489–546.

Kelly, F. P. (1979). *Reversibility and Stochastic Networks*. Wiley.

Ledermann, W., Reuter, G. E. H. (1954). Spectral theory for the differential equations of simple birth and death processes. *Philosophical Transactions of the Royal Society A* 246(914), 321–369.

Marsaglia, G., Tubilla, A. (1975). A note on the "lack of memory" property of the exponential distribution. *Annals of Probability* 3(2), 353–354.

Milgrom, P., Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica* 70(2), 583–601.

Sinander, L. (2022). The converse envelope theorem. *Econometrica* 90(6), 2795–2819.