Stabilité des familles exponentielles par convolution gaussienne : caractérisation quadratique, fermeture de Cole–Hopf, et équivalence de Hamilton–Jacobi

Emmanuel Sérié

17 avril 2026

Résumé

We characterise sufficient statistics $\varphi : \mathbb{R}^d \to \mathbb{R}^r$ for which the exponential family $\{p_\theta(x) \propto e^{-\theta^\top \varphi(x)}\}$ is closed under Gaussian convolution $p_\theta * \gamma_\sigma$. Our main theorem treats the quadratic statistic $\varphi(x) = (\mathrm{vec}(xx^\top), x)$: the family is stable, the reparametrisation is the closed form $(\widetilde{M}, \widetilde{b}) = (M(I_d + \sigma^2 M)^{-1}, (I_d + \sigma^2 M)^{-1} b)$, and the induced flow in heat time is the matrix Riccati equation $\dot{M} = -M^2$. Using a Cole–Hopf closure condition ($G_{ij} := \langle \nabla\varphi_i, \nabla\varphi_j \rangle$ and $H_i := \Delta\varphi_i$ belong to the affine tube spanned by φ and $\mathbf{1}$), we prove that among minimal polynomial statistics the stable ones have degree at most two; the cubic case $\varphi(x) = x^4$ fails because the closure produces the direction x^6 orthogonal to the tube. Etienne's equivalence between stability and a Hamilton–Jacobi path equation is established under a *regular stability* hypothesis (C^2 regularity of φ, minimality, C^1 path, positivity-restricted heat uniqueness of Widder 1944), and the three hidden jumps are made explicit. The main obstruction arguments carry to harmonic, exponential, trigonometric, radial, and entire statistics of polynomial growth; the Gaussian Riccati class is therefore the unique stable polynomial skeleton on $(\mathbb{R}^d, \Delta, \gamma_\sigma)$. A Lean 4 formalisation companion (Paper B, `docs/paper-formal/paper-formal-v1.tex`) is developed in parallel: the theorem statements for Q2 and Q3 together with an adversarial corpus are established as fidelity anchors, while the mechanical verification of the proof bodies remains work-in-progress.

1 Introduction

Soit $\varphi : \mathbb{R}^d \to \mathbb{R}^r$ une statistique mesurable. À chaque $\theta \in \Theta^\star$ (domaine naturel du problème (1) de [Bro86]), on associe la densité

$$p_\theta(x) \;=\; Z_\theta^{-1}\, e^{-\theta^\top \varphi(x)}, \qquad Z_\theta = \int_{\mathbb{R}^d} e^{-\theta^\top \varphi(x)}\, dx.$$

La convolution $\tilde{p}_\theta := p_\theta * \gamma_\sigma$ avec le noyau gaussien isotrope $\gamma_\sigma(x) = (2\pi\sigma^2)^{-d/2} \exp(-\|x\|^2/(2\sigma^2))$ lisse la loi sans changer sa signification informationnelle. La *stabilité* est la propriété suivante : pour tout $\theta \in \Theta^\star$ et tout $\sigma > 0$, il existe $\tilde{\theta} \in \Theta^\star$ tel que $\tilde{p}_\theta \propto e^{-\tilde{\theta}^\top \varphi}$, à φ fixé.

Motivation. Cette condition rassemble trois questions indépendantes qui admettent pourtant une même réponse quand φ est quadratique : (i) inférence et filtrage (Kalman [Kal60]) – quelles familles conjuguées au bruit blanc gaussien existent ? (ii) géométrie de l'information [AN00] – quelles e-familles sont invariantes sous le semigroupe de la chaleur et donc, par [JKO98], totalement m-géodésiques dans l'espace de Wasserstein $\mathcal{P}_2(\mathbb{R}^d)$? (iii) méthodes MaxEnt bruitées [Jay57 ; Sta59] – à quelles contraintes de moments correspond une image stable de la convolution gaussienne ?

Contributions.

1. **Théorème principal (§3).** Pour $\varphi(x) = (\mathrm{vec}(xx^\top), x)$, on prouve la stabilité, on explicite la reparamétrisation $(M, b) \mapsto (\widetilde{M}, \widetilde{b})$, on établit la stabilité du cône SPD, l'injectivité de l'application et sa non-surjectivité, et on identifie le flot temporel avec l'équation de Riccati matricielle.
2. **Fermeture algébrique** (†) **(§4).** La stabilité régulière équivaut à l'appartenance des quantités G_{ij} et H_i au tube affine $\mathcal{A} = \mathrm{Span}\{\varphi_1, \ldots, \varphi_r\} \oplus \mathbb{R} \cdot \mathbf{1}$, uniformément en θ. Cela entraîne $\deg \varphi \leq 2$ dans la classe polynomiale minimale.
3. **Contre-exemple rigoureux (§5).** La statistique $\varphi(x) = x^4$ échoue à (†) car $\mathcal{L}(\theta x^4) = 8\theta^2 x^6 - 6\theta x^2$ produit $x^6 \notin \mathcal{A}$. L'argument se généralise à $\deg \varphi \geq 3$.
4. **Équivalence Hamilton–Jacobi (§6).** L'équivalence entre stabilité et existence d'un chemin C^1 $t \mapsto \theta_t$ vérifiant l'équation (6) de [Bro86, §3] (restituée en (14)) est établie sous *stabilité régulière* (Définition 6.1), dont les trois sauts sont identifiés.
5. **Rigidité exotique (§7).** Harmonique, exponentielle, trigonométrique, radiale, entière à croissance polynomiale : dans chaque classe naturelle non polynomiale, la stabilité force polynomial de degré ≤ 2.

Plan. La §2 fixe la notation et rappelle Cole–Hopf. La §3 énonce et prouve le théorème principal. La §4 dérive la fermeture algébrique et la borne de degré. La §5 étudie le bord du domaine et les contre-exemples. La §6 établit l'équivalence Hamilton–Jacobi. La §7 passe en revue les classes exotiques. La §8 conclut.

2 Notation et cadre

Convention de temps. Tout au long du papier, $\sigma^2 = 2t$ où t est le temps de l'équation $\partial_t u = \frac{1}{2}\Delta u$. Nous paramétrons le flot par $s := \sigma^2 \in (0, \infty)$, ce qui absorbe les facteurs 2 dans l'équation de Riccati (Annexe §A).

Tube affine et opérateurs. Étant donnée la statistique φ avec composantes $\varphi_1, \ldots, \varphi_r$ de classe C^2, on pose

$$\mathcal{A} \ := \ \mathrm{Span}\{\varphi_1, \ldots, \varphi_r\} \oplus \mathbb{R} \cdot \mathbf{1} \ \subset \ C^2(\mathbb{R}^d, \mathbb{R}),$$
$$G_{ij}(x) \ := \ \langle \nabla \varphi_i(x), \nabla \varphi_j(x) \rangle, \qquad H_i(x) \ := \ \Delta \varphi_i(x).$$

L'opérateur de Cole–Hopf (générateur de Hamilton–Jacobi normalisé) est

$$\mathcal{L} f \ := \ \tfrac{1}{2}\|\nabla f\|^2 - \tfrac{1}{2}\Delta f. \tag{1}$$

Le carré du champ associé au demi-laplacien est $\Gamma(f, g) = \frac{1}{2}\langle \nabla f, \nabla g \rangle$ [BGL14].

Paramétrisation quadratique. Pour $\varphi(x) = (\mathrm{vec}(xx^\top), x) \in \mathbb{R}^{d^2+d}$, on travaille en coordonnées canoniques minimales $(M, b) \in \mathrm{Sym}_d \times \mathbb{R}^d$ via $\theta^\top \varphi(x) = \frac{1}{2} x^\top M x + b^\top x$, avec $M \succ 0$ pour garantir $Z_\theta < \infty$.

Lemme 2.1 (Cole–Hopf, formulation demi-laplacien)**.** *Soit $u \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ strictement positive et $f := -\log u$. Alors*

$$\partial_t u \ = \ \tfrac{1}{2}\Delta u \quad \Longleftrightarrow \quad \partial_t f \ = \ -\mathcal{L} f \ = \ \tfrac{1}{2}\Delta f - \tfrac{1}{2}\|\nabla f\|^2.$$

Hypothèses. $u > 0$ de classe $C^{1,2}$ sur $\mathbb{R}_+ \times \mathbb{R}^d$.

Démonstration. Idée : On substitue $u = e^{-f}$ dans l'équation de la chaleur et on identifie les dérivées en t et x grâce à la règle de la chaîne. On a $\partial_t u = -u\,\partial_t f$, $\nabla u = -u\nabla f$, $\Delta u = u(\|\nabla f\|^2 - \Delta f)$. L'équation $\partial_t u = \frac{1}{2}\Delta u$ devient $-u\,\partial_t f = \frac{u}{2}(\|\nabla f\|^2 - \Delta f)$. La positivité de u permet de diviser, d'où $\partial_t f = -\frac{1}{2}\|\nabla f\|^2 + \frac{1}{2}\Delta f$, soit $\partial_t f = -\mathcal{L}f$. La réciproque est immédiate par le même calcul. Voir [Hop50 ; Col51] pour l'origine historique de la transformation. □

Remarque 2.2 (Dimension effective de la statistique quadratique)**.** La statistique $\varphi(x) = (\mathrm{vec}(xx^\top), x)$ n'est pas minimale : le noyau de $T \mapsto T + T^\top$ est l'espace Asym_d de dimension $d(d-1)/2$. En coordonnées $(M, b) \in \mathrm{Sym}_d \times \mathbb{R}^d$, la dimension effective est $d(d+1)/2 + d$ et la famille est identifiable. Toutes les assertions d'unicité formulées ci-dessous sont relatives à cette paramétrisation minimale ; leur relevé à $\mathbb{R}^{d^2+d}$ se fait modulo Asym_d.

3 Théorème principal : cas quadratique unifié

Théorème 3.1 (Caractérisation Q2 unifiée)**.** *Soit $d \geq 1$ et la statistique quadratique complète $\varphi(x) = (\mathrm{vec}(xx^\top), x) \in \mathbb{R}^{d^2+d}$ paramétrée par $(M, b) \in \mathrm{SPD}_d \times \mathbb{R}^d$ via $\theta^\top\varphi(x) = \frac{1}{2}x^\top Mx + b^\top x$. Alors :*

1. ***(Stabilité)** Pour tout $(M, b) \in \Theta^\star$ et tout $\sigma > 0$, il existe un unique $(\widetilde{M}, \widetilde{b}) \in \Theta^\star$ tel que $p_\theta * \gamma_\sigma \propto e^{-\tilde{\theta}^\top\varphi}$.*
2. ***(Reparamétrisation explicite)***
$$\widetilde{M} \;=\; M\,(I_d + \sigma^2 M)^{-1}, \qquad \widetilde{b} \;=\; (I_d + \sigma^2 M)^{-1}\,b. \tag{2}$$
3. ***(Stabilité du cône)** $M \succ 0 \;\Rightarrow\; \widetilde{M} \succ 0$.*
4. ***(Injectivité non surjective)** L'application $(M, b) \mapsto (\widetilde{M}, \widetilde{b})$ est injective sur $\Theta^\star$; son image est le sous-cône strict $\{\widetilde{M} \prec \sigma^{-2}I_d\}$.*
5. ***(Flot de Riccati)** En temps $s := \sigma^2$, la trajectoire (M_s, b_s) satisfait*
$$\dot{M}_s = -M_s^2, \qquad \dot{b}_s = -M_s b_s, \tag{3}$$
avec solution fermée $M_s = M_0(I_d + sM_0)^{-1}$, $b_s = (I_d + sM_0)^{-1}b_0$.
6. ***(Cohérence avec (†))** Pour $f = \frac{1}{2}x^\top Mx + b^\top x$, on a $\mathcal{L}f \in \mathcal{A}$: la fermeture algébrique (§4) est vérifiée.*

Hypothèses. $M \in \mathrm{SPD}_d$ (intégrabilité), $b \in \mathbb{R}^d$, $\sigma > 0$. Domaine minimal $\mathrm{Sym}_d \times \mathbb{R}^d$ (voir Remarque 2.2).

Démonstration. Idée : On passe en Fourier, on complète le carré, on identifie l'exponentielle gaussienne résultante comme une nouvelle densité quadratique, puis on utilise l'identité de Woodbury en version commutative pour lire la reparamétrisation. Les propriétés structurelles suivent d'arguments matriciels élémentaires.

Étape 1 (Fourier). Pour $M \succ 0$, la densité $p_\theta(x) \propto e^{-\frac{1}{2}x^\top Mx - b^\top x}$ s'obtient par complétion de carré : p_θ est la gaussienne de moyenne $-M^{-1}b$ et covariance M^{-1}. Sa transformée de Fourier est $\hat{p_\theta}(\xi) = \exp(-i\xi^\top M^{-1}b - \frac{1}{2}\xi^\top M^{-1}\xi)$. Celle de γ_σ est $\hat{\gamma_\sigma}(\xi) = \exp(-\frac{1}{2}\sigma^2\|\xi\|^2)$. Par conséquent,
$$\widehat{p_\theta * \gamma_\sigma}(\xi) \;=\; \hat{p_\theta}(\xi)\hat{\gamma_\sigma}(\xi) \;=\; \exp\Big(-i\xi^\top M^{-1}b - \tfrac{1}{2}\xi^\top(M^{-1} + \sigma^2 I_d)\xi\Big)\,. \tag{4}$$

Étape 2 (Inversion gaussienne). L'expression (4) est la transformée de Fourier d'une gaussienne de moyenne $-M^{-1}b$ et covariance $M^{-1} + \sigma^2 I_d$. Par unicité du noyau gaussien multivarié, $\tilde{p}_\theta$ est la gaussienne de mêmes paramètres. Sa densité s'écrit $\tilde{p}_\theta(x) \propto \exp(-\frac{1}{2}x^\top\widetilde{M}x - \widetilde{b}^\top x)$ avec
$$\widetilde{M} = (M^{-1} + \sigma^2 I_d)^{-1}, \qquad \widetilde{b} = \widetilde{M}\cdot M^{-1}b.$$

Étape 3 (Woodbury commutatif). Pour $M \succ 0$ et $\sigma^2 > 0$, l'identité

$$(M^{-1} + \sigma^2 I_d)^{-1} \;=\; M(I_d + \sigma^2 M)^{-1} \;=\; (I_d + \sigma^2 M)^{-1} M \tag{5}$$

se vérifie par multiplication directe : $M(I_d + \sigma^2 M)^{-1} \cdot (M^{-1} + \sigma^2 I_d) = MM^{-1} = I_d$ après simplification (possible par commutativité de M avec $(I_d + \sigma^2 M)^{-1}$, qui est un polynôme de M). On substitue (5) dans l'expression de $\widetilde{M}$ et on obtient (2). Pour $\widetilde{b}$: $\widetilde{b} = M(I_d + \sigma^2 M)^{-1} M^{-1} b = (I_d + \sigma^2 M)^{-1} b$ par commutativité.

Étape 4 (Stabilité du cône). $M \succ 0$ et $I_d + \sigma^2 M \succ I_d \succ 0$ commutent ; leur produit $M(I_d + \sigma^2 M)^{-1}$ est donc symétrique et à valeurs propres positives (image des valeurs propres λ_i de M par $\lambda \mapsto \lambda/(1 + \sigma^2\lambda) > 0$).

Étape 5 (Injectivité / image). Soient $M_1, M_2 \succ 0$ avec $\widetilde{M_1} = \widetilde{M_2}$, soit $M_1(I_d + \sigma^2 M_1)^{-1} = M_2(I_d + \sigma^2 M_2)^{-1}$. En multipliant à droite par $(I_d + \sigma^2 M_2)$ et à gauche par $(I_d + \sigma^2 M_1)$ et en simplifiant, on aboutit à $M_1 - M_2 = \sigma^2[M_2, M_1]$: la différence symétrique est égale à un commutateur antisymétrique, donc nulle. Pour l'image : $\widetilde{M} = M(I_d + \sigma^2 M)^{-1}$ a valeurs propres $\lambda/(1 + \sigma^2\lambda) < \sigma^{-2}$ pour tout $\lambda > 0$, d'où $\widetilde{M} \prec \sigma^{-2} I_d$ strictement. Réciproquement, tout $\widetilde{M}$ dans ce sous-cône strict admet un antécédent $M = \widetilde{M}(I_d - \sigma^2\widetilde{M})^{-1} \succ 0$ (cf. [Kal60], inversion du filtre de Kalman).

Étape 6 (Riccati). Posons $U_s := I_d + sM_0$, d'où $M_s = M_0 U_s^{-1}$. Par dérivation et l'identité $\frac{d}{ds}U_s^{-1} = -U_s^{-1} M_0 U_s^{-1}$: $\dot{M}_s = -M_0 U_s^{-1} M_0 U_s^{-1} = -M_s^2$ par commutativité. Idem pour $b_s = U_s^{-1} b_0$: $\dot{b}_s = -U_s^{-1} M_0 U_s^{-1} b_0 = -M_s b_s$.

Étape 7 (Cohérence). Calcul direct pour $f(x) = \frac{1}{2}x^\top Mx + b^\top x$: $\nabla f = Mx + b$, $\Delta f = \operatorname{tr}(M)$. Donc $\mathcal{L}f = \frac{1}{2}(Mx+b)^\top(Mx+b) - \frac{1}{2}\operatorname{tr}(M)$, polynôme de degré 2 en x à coefficients quadratiques dans (M, b). Les termes $\frac{1}{2}x^\top M^2 x$, $b^\top Mx$, $\frac{1}{2}\|b\|^2$, $\frac{1}{2}\operatorname{tr}(M)$ sont tous dans $\mathcal{A}$ puisque $x^\top Ax$ et $b^\top x$ sont des combinaisons linéaires de composantes de φ. $\square$

Remarque 3.2 (Cas scalaire $d = 1$, test pédagogique)**.** Pour $d = 1$, $M = m > 0$ et $b = \beta$: on a $\tilde{m} = m/(1 + \sigma^2 m)$, $\tilde{\beta} = \beta/(1 + \sigma^2 m)$; Riccati devient $\dot{m} = -m^2$, de solution $m_s = m_0/(1 + sm_0)$; le paramètre $\tilde{m}$ ne peut excéder σ^{-2}, ce qui correspond à $\tilde{\sigma}^2 \geq \sigma^2$, soit : la variance de la loi lissée est au moins la variance ajoutée.

Remarque 3.3 (Lecture information-théorique)**.** Le flot $(\Sigma_s := M_s^{-1}) = \Sigma_0 + sI_d$ ajoute linéairement la covariance du bruit gaussien ; la non-surjectivité exprime la contraction du canal AWGN (Stam–Blachman [Sta59 ; Bla65]). Le flot (Σ_s) est le flot de gradient Wasserstein de l'entropie relative à la mesure de Lebesgue, restreint à la classe gaussienne [JKO98 ; Ott01].

4 Fermeture algébrique de Cole–Hopf

Nous isolons la condition locale sur φ qui encode la stabilité au niveau pointwise, sans supposer l'existence préalable d'un chemin $t \mapsto \theta_t$.

Définition 4.1 (Minimalité)**.** La statistique φ est *minimale* en $\theta \in \operatorname{int}(\Theta^\star)$ si $\{\varphi_1, \ldots, \varphi_r, \mathbf{1}\}$ sont linéairement indépendantes comme fonctions de x. De manière équivalente, la matrice de Fisher $I(\theta) = \operatorname{Cov}_\theta(\varphi)$ est définie positive.

Théorème 4.2 (Fermeture (†) de Cole–Hopf)**.** *Soit $\varphi \in C^2(\mathbb{R}^d, \mathbb{R}^r)$ minimale, et fixons $\theta \in \operatorname{int}(\Theta^\star)$. Les assertions suivantes sont équivalentes :*

(a) La famille $\{p_\theta\}$ est régulièrement stable *au sens de la Définition 6.1 de la §6.*

(b) Il existe, uniformément en $\theta \in \operatorname{int}(\Theta^\star)$, des identités polynomiales en θ

$$G_{ij} \in \mathcal{A} \qquad et \qquad H_i \in \mathcal{A}, \qquad \forall i, j \in \{1, \ldots, r\}.$$

Hypothèses. $\varphi \in C^2$ (nécessaire pour G_{ij} et H_i), minimalité (identifiabilité de la reparamétrisation), intégrabilité de p_θ et $\tilde{p}_\theta$ (croissance adéquate, fournie par $M \succ 0$ dans le cas quadratique).

Démonstration. Idée : (b) ⇒ (a) construit le chemin par intégration Cauchy–Lipschitz d'un champ polynomial lu dans les coefficients de $\mathcal{L}$, puis reconstruit la densité lissée via Cole–Hopf inverse et l'unicité positive de Widder. La direction (a) ⇒ (b) dérive (4) en $t = 0^+$, identifie coefficient par coefficient dans le polynôme de θ, et conclut par minimalité.

(b) ⇒ (a). Pour $f(x,\theta) = \theta^\top \varphi(x)$, on développe $\mathcal{L}f(x,\theta) = \frac{1}{2}\sum_{i,j}\theta_i\theta_j G_{ij}(x) - \frac{1}{2}\sum_i \theta_i H_i(x)$. Par (b), $G_{ij}(x) = \sum_k a_k^{ij}\varphi_k(x) + a_0^{ij}$ et $H_i(x) = \sum_k c_k^i \varphi_k(x) + c_0^i$. Donc $\mathcal{L}f(x,\theta) = A(\theta)^\top\varphi(x) + B(\theta)$ avec $A(\theta) \in \mathbb{R}^r$, $B(\theta) \in \mathbb{R}$ polynomiaux de degré 2 en θ. Posons $g_t(x) := \theta_t^\top\varphi(x) + \zeta_t$, avec $\theta_0 = \theta$ et (θ_t, ζ_t) solution de l'EDO couplée $\dot\theta_t = -A(\theta_t)$, $\dot\zeta_t = -B(\theta_t)$. Cauchy–Lipschitz local donne (θ_t, ζ_t) sur un intervalle $[0,T]$; on choisit T maximal tel que $\theta_t \in \operatorname{int}(\Theta^\star)$ (prolongeable autant que le champ reste régulier, voir §5 pour l'existence globale). Par construction $\partial_t g_t = \dot\theta_t^\top\varphi + \dot\zeta_t = -\mathcal{L}g_t$. Par le Lemme 2.1, $u_t := e^{-g_t}$ résout $\partial_t u_t = \frac{1}{2}\Delta u_t$ avec donnée initiale $u_0 = p_\theta Z_\theta$ (à renormalisation près). Le noyau gaussien donne $u_t = (p_\theta Z_\theta) * \gamma_{\sqrt{2t}}$ via l'unicité positive de Widder [Wid44] (les deux fonctions sont positives, résolvent la même équation avec même donnée initiale, et croissance sous-gaussienne uniforme assurée par intégrabilité de p_θ). Donc $\tilde p_\theta \propto e^{-\theta_t^\top\varphi}$ pour $s = 2t$, ce qui donne la stabilité régulière avec $\tilde\theta(\sigma) = \theta_{\sigma^2/2}$, de classe C^1 puisque A, B le sont.

(a) ⇒ (b). *Notation à deux variables.* Par (H3) de la Définition 6.1, pour chaque condition initiale $\theta_0 \in \operatorname{int}(\Theta^\star)$ il existe un chemin C^1 $\sigma \mapsto \tilde\theta(\sigma;\theta_0) \in \operatorname{int}(\Theta^\star)$ avec $\tilde\theta(0^+;\theta_0) = \theta_0$. Posant $t := \sigma^2/2$, nous notons

$$\theta_t : (t,\theta_0) \longmapsto \theta_t(t,\theta_0) := \tilde\theta(\sqrt{2t};\theta_0), \qquad \zeta_t : (t,\theta_0) \longmapsto \zeta_t(t,\theta_0) := \log \tilde Z_\theta,$$

vues comme fonctions des deux variables indépendantes : le temps de chaleur $t \geq 0$ et la condition initiale θ_0. Par construction $\theta_t(0,\theta_0) = \theta_0$ et, pour chaque θ_0 fixé, $t \mapsto \theta_t(t,\theta_0)$ est C^1 au voisinage droit de $t = 0$ (H3 reparamétrée via $\sigma = \sqrt{2t}$). Nous utilisons ∂_t (resp. ∂_{θ_0}) pour la dérivation en t (resp. en θ_0) et réservons le point (« $\dot{}$ ») aux identités a posteriori, une fois le flot ODE établi.

Identité différentielle à condition initiale fixée. Fixons $\theta_0 \in \operatorname{int}(\Theta^\star)$ et posons

$$u_t(x) \;:=\; Z_{\theta_0}\,(p_\theta[\theta_0] * \gamma_{\sqrt{2t}})(x), \qquad t \geq 0,$$

i.e. u_t est la convolution gaussienne (non-normalisée) de $e^{-\theta_0^\top\varphi}$ avec le noyau de la chaleur à l'instant t. Par la propriété de semigroupe de la chaleur (solution fondamentale), u_t satisfait l'équation de la chaleur

$$\partial_t u_t \;=\; \tfrac{1}{2}\Delta u_t \qquad \text{sur } (0,\infty)\times\mathbb{R}^d, \qquad u_0 = e^{-\theta_0^\top\varphi}, \tag{6}$$

au sens classique ($C^{1,2}$) grâce à la régularisation gaussienne, avec croissance sous-gaussienne uniforme sur les compacts en t (H4). Par ailleurs, la stabilité régulière (H1) fournit, pour $\sigma = \sqrt{2t}$, la reparamétrisation $\tilde p_\theta[\theta_0] \propto e^{-\tilde\theta(\sigma;\theta_0)^\top\varphi}$; en notation à deux variables,

$$u_t(x) \;=\; e^{-g_t(x)}, \qquad g_t(x) \;:=\; \theta_t(t,\theta_0)^\top\varphi(x) + \zeta_t(t,\theta_0), \tag{7}$$

où $\zeta_t(t,\theta_0)$ regroupe les constantes de normalisation (cf. Définition 6.1). Comme $u_t > 0$ et $g_t \in C^{1,2}((0,\infty)\times\mathbb{R}^d)$ (H0, H3, $\varphi \in C^2$), le Lemme 2.1 s'applique point-par-point et

$$\partial_t g_t \;=\; -\mathcal{L}g_t \qquad \text{sur } (0,\infty)\times\mathbb{R}^d. \tag{8}$$

Expansion explicite et évaluation en $t = 0^+$. Dérivant (7) en t par la règle de la chaîne, $\partial_t g_t(x) = \partial_t\theta_t(t,\theta_0)^\top\varphi(x) + \partial_t\zeta_t(t,\theta_0)$; en développant $-\mathcal{L}g_t$ à partir de la définition (1) de $\mathcal{L}$ et des définitions de G_{ij} et H_i :

$$-\mathcal{L}g_t(x) \;=\; \tfrac{1}{2}\sum_{i,j}\theta_t(t,\theta_0)_i\,\theta_t(t,\theta_0)_j\,G_{ij}(x) \;-\; \tfrac{1}{2}\sum_i \theta_t(t,\theta_0)_i\,H_i(x).$$

Par (H3), $t \mapsto \theta_t(t, \theta_0)$ et $t \mapsto \zeta_t(t, \theta_0)$ sont C^1 au voisinage droit de $t = 0$ à θ_0 fixé, donc $\partial_t\theta_t(0^+, \theta_0) \in \mathbb{R}^r$ et $\partial_t\zeta_t(0^+, \theta_0) \in \mathbb{R}$ existent ; faisant $t \downarrow 0$ dans (8) (avec $\theta_t(0, \theta_0) = \theta_0$, continuité de φ, G_{ij}, H_i), on obtient, pour tout $x \in \mathbb{R}^d$:

$$\partial_t\theta_t(0^+, \theta_0)^\top \varphi(x) + \partial_t\zeta_t(0^+, \theta_0) \;=\; \tfrac{1}{2}\theta_0^\top \mathbf{G}(x)\, \theta_0 \;-\; \tfrac{1}{2}\theta_0^\top \mathbf{H}(x), \tag{9}$$

où $\mathbf{G}(x) := (G_{ij}(x))_{i,j} \in \mathbb{R}^{r\times r}$ et $\mathbf{H}(x) := (H_i(x))_i \in \mathbb{R}^r$ sont indépendants de θ_0.

Identification par variation de la condition initiale. Le point-clef est que (9) est valide, avec le *même* $(\mathbf{G}, \mathbf{H})$, pour *tout* $\theta_0 \in \mathrm{int}(\Theta^\star)$: (H3) fournit un chemin C^1 à partir de *chaque* point intérieur. Le membre de droite est polynomial de degré ≤ 2 en θ_0, à coefficients ne dépendant que de x. Nous allons en déduire que le membre de gauche est, lui aussi, polynomial en θ_0 — de manière explicite via une étape d'extraction linéaire par évaluation en $r + 1$ points (étape 2 ci-dessous) —, puis identifier les coefficients monôme par monôme.

Étape 1 : identifiabilité de $\theta \mapsto p_\theta$ sous minimalité (argument auto-contenu). Rappelons $Z(\theta) := \int_{\mathbb{R}^d} e^{-\theta^\top\varphi(x)}\, dx < \infty$ pour $\theta \in \mathrm{int}(\Theta^\star)$ (domaine naturel, ouvert convexe). Par dérivation sous le signe intégral (légitime sur l'intérieur du domaine naturel, classique pour les familles exponentielles [Bar78 ; Bro86 ; WJ08]), $\nabla \log Z(\theta) = -\mathbb{E}_\theta[\varphi]$ et

$$\nabla^2 \log Z(\theta) \;=\; \mathrm{Cov}_\theta(\varphi) \;=\; I(\theta), \tag{10}$$

la matrice de Fisher. Sous (H2) (minimalité, Définition 4.1), $\{\varphi_1, \ldots, \varphi_r\}$ est linéairement indépendante modulo les constantes comme famille de fonctions de x, donc $I(\theta) \succ 0$ partout sur $\mathrm{int}(\Theta^\star)$; en particulier $\log Z$ est *strictement convexe* sur cet ouvert convexe. Par stricte convexité, $\theta \mapsto \nabla \log Z(\theta)$ est strictement monotone (donc injective) ; comme p_θ détermine le moment $\mathbb{E}_\theta[\varphi] = -\nabla \log Z(\theta)$, l'application $\theta \mapsto p_\theta$ est injective sur $\mathrm{int}(\Theta^\star)$. Cf. [WJ08, Prop. 3.1] pour une rédaction équivalente, et [Bro86 ; Bar78 ; Gey] pour des traitements classiques.

Étape 2 : extraction linéaire de $(a(\theta_0), b(\theta_0))$. Notons $a(\theta_0) := \partial_t\theta_t(0^+, \theta_0) \in \mathbb{R}^r$ et $b(\theta_0) := \partial_t\zeta_t(0^+, \theta_0) \in \mathbb{R}$, de sorte que (9) s'écrit

$$a(\theta_0)^\top \varphi(x) + b(\theta_0) \;=\; p_x(\theta_0) \;:=\; \tfrac{1}{2}\theta_0^\top \mathbf{G}(x)\, \theta_0 \;-\; \tfrac{1}{2}\theta_0^\top \mathbf{H}(x), \tag{11}$$

pour tout $x \in \mathbb{R}^d$ et tout $\theta_0 \in \mathrm{int}(\Theta^\star)$. Par minimalité (H2, Définition 4.1), $\{\varphi_1, \ldots, \varphi_r, \mathbf{1}\}$ est linéairement indépendante comme famille de $r + 1$ fonctions de x sur $\mathbb{R}^d$. Il s'ensuit, par un résultat élémentaire d'algèbre linéaire (par récurrence sur r : toute famille libre de n fonctions sur un ensemble S admet n points d'évaluation rendant la matrice des valeurs inversible), qu'il existe $r + 1$ points $x_1, \ldots, x_{r+1} \in \mathbb{R}^d$ tels que la matrice

$$\Phi \;:=\; \big(\varphi_1(x_k),\; \ldots,\; \varphi_r(x_k),\; 1\big)_{k=1,\ldots,r+1} \;\in\; \mathbb{R}^{(r+1)\times(r+1)}$$

soit inversible. Évaluant (11) en chaque x_k, on obtient le système linéaire, à coefficients *constants* (indépendants de θ_0),

$$\Phi \begin{pmatrix} a(\theta_0) \\ b(\theta_0) \end{pmatrix} \;=\; \big(p_{x_k}(\theta_0)\big)_{k=1}^{r+1} \;\in\; \mathbb{R}^{r+1}. \tag{12}$$

Chaque composante $p_{x_k}(\theta_0)$ est, par définition, un polynôme de degré ≤ 2 en $\theta_0 \in \mathbb{R}^r$, à coefficients constants $\frac{1}{2}\mathbf{G}(x_k)$ et $-\frac{1}{2}\mathbf{H}(x_k)$ ne dépendant que de x_k (pas de θ_0). Multipliant (12) par Φ^{-1} — matrice *constante* (indépendante de θ_0) —,

$$\begin{pmatrix} a(\theta_0) \\ b(\theta_0) \end{pmatrix} \;=\; \Phi^{-1}\big(p_{x_k}(\theta_0)\big)_{k=1}^{r+1}. \tag{13}$$

Chaque composante de $(a(\theta_0), b(\theta_0))$ est donc une combinaison linéaire à coefficients *constants* de polynômes de degré ≤ 2 en θ_0. Conclusion : *les applications $\theta_0 \mapsto a_k(\theta_0)$ ($k = 1, \ldots, r$) et $\theta_0 \mapsto b(\theta_0)$ sont polynomiales de degré ≤ 2 sur $\mathbb{R}^r$ tout entier* (composante par composante), leurs coefficients étant des combinaisons linéaires — à coefficients les entrées de Φ^{-1} — des scalaires $\{G_{ij}(x_k), H_i(x_k)\}_{k=1,\ldots,r+1}$. L'identité (11) est ainsi, à chaque $x \in \mathbb{R}^d$ fixé, une *égalité polynomiale en* θ_0 entre deux polynômes de degré ≤ 2 à coefficients dépendant de x, ce qui légitime l'identification des monômes ci-dessous.

Remarque. La construction précédente subsume l'unicité de la décomposition dans $\mathcal{A}$ (*a fortiori* : la liberté évaluation-par-points est plus forte que la liberté fonctionnelle) et fournit en outre la polynomialité *vectorielle* de (a, b), point dont l'étape suivante a crucialement besoin.

Étape 3 : identification des monômes en θ_0. Fixons $x \in \mathbb{R}^d$. Les deux membres de (11) sont, par l'étape 2, des polynômes en $\theta_0 \in \mathbb{R}^r$ de degré ≤ 2, à coefficients dépendant uniquement de x :

— à gauche, $q_x(\theta_0) := a(\theta_0)^\top \varphi(x) + b(\theta_0)$ est polynomial en θ_0 via (13), ses coefficients étant des combinaisons linéaires — via Φ^{-1} et $(\varphi(x), 1)$ — des entrées de $\{\mathbf{G}(x_k), \mathbf{H}(x_k)\}_{k=1}^{r+1}$;

— à droite, $p_x(\theta_0) = \frac{1}{2}\theta_0^\top \mathbf{G}(x)\theta_0 - \frac{1}{2}\theta_0^\top \mathbf{H}(x)$.

L'égalité $q_x(\theta_0) = p_x(\theta_0)$ vaut pour tout $\theta_0 \in \mathrm{int}(\Theta^\star)$, ouvert non vide de $\mathbb{R}^r$; or deux polynômes de $\mathbb{R}[\theta_{0,1}, \ldots, \theta_{0,r}]$ qui coïncident sur un ouvert non vide sont identiques (l'anneau est intègre ; un polynôme qui s'annule sur un ouvert a tous ses coefficients nuls). Les monômes $\{1,\ \theta_{0,i},\ \theta_{0,i}\theta_{0,j} : i \leq j\}$ formant une base de $\mathbb{R}[\theta_0]_{\leq 2}$ (polynômes de degré ≤ 2 en θ_0), l'identification coefficient par coefficient — légitime car les deux membres sont désormais polynomiaux — fournit, pour chaque $x \in \mathbb{R}^d$ et tous $i, j \in \{1, \ldots, r\}$:

— coefficient de $\theta_{0,i}\theta_{0,j}$ $(i \leq j)$: une combinaison linéaire à coefficients *constants* de $\{\varphi_1(x), \ldots, \varphi_r(x), 1\}$ est égale à $\frac{1}{2}G_{ij}(x)$ (symétrisée si $i < j$) ; donc $G_{ij}(x) \in \mathcal{A}$;

— coefficient de $\theta_{0,i}$: une combinaison linéaire à coefficients constants de $\{\varphi_1(x), \ldots, \varphi_r(x), 1\}$ est égale à $-\frac{1}{2}H_i(x)$; donc $H_i(x) \in \mathcal{A}$.

Ceci étant vrai pour tout $x \in \mathbb{R}^d$, on conclut $G_{ij}, H_i \in \mathcal{A}$ pour tous $i, j \in \{1, \ldots, r\}$. C'est précisément la condition (b).

Bilan sur la rigueur. L'argument combine (i) la stricte convexité de $\log Z$ sous minimalité, classique en théorie des familles exponentielles [Bar78 ; Bro86 ; WJ08 ; Gey] et (ii) l'indépendance linéaire, algébrique, des monômes de degré ≤ 2 sur un ouvert de $\mathbb{R}^r$. Aucune régularité C^2 jointe du flot $(t, \theta_0) \mapsto \theta_t(t, \theta_0)$ n'est nécessaire : seule l'existence, en chaque θ_0, d'une dérivée en temps à l'origine (H3) est utilisée. La régularité $\varphi \in C^2$ garantit que G_{ij} et H_i sont bien définis comme fonctions continues.

Remarque (régularité a posteriori). La polynomialité (donc la régularité C^∞) des champs $A : \theta_0 \mapsto -a(\theta_0) = -\partial_t\theta_t(0^+, \theta_0)$ et $B : \theta_0 \mapsto -b(\theta_0) = -\partial_t\zeta_t(0^+, \theta_0)$ est déjà acquise à l'*étape 2* par extraction linéaire ((13)), *indépendamment* de l'établissement préalable de (b). C'est précisément cette polynomialité en θ_0 qui rend légitime l'identification des monômes menant à (b) : Cauchy–Lipschitz n'intervient *pas* pour obtenir la régularité à $t = 0$. Il sert en revanche à étendre la régularité au flot complet pour $t > 0$: une fois (b) établi et le système autonome $\dot{\theta} = -A(\theta)$, $\dot{\zeta} = -B(\theta)$ posé avec A, B polynomiaux (donc C^∞), Cauchy–Lipschitz avec dépendance lisse des conditions initiales garantit que $(t, \theta_0) \mapsto (\theta_t(t, \theta_0), \zeta_t(t, \theta_0))$ est C^∞ au voisinage de tout $(0, \theta_0^\star)$ avec $\theta_0^\star \in \mathrm{int}(\Theta^\star)$. La notation « $\dot{\theta}_t$ » employée dans la direction (b) $\Rightarrow$ (a) est cohérente avec ce flot.

Non-circularité. La direction (b) $\Rightarrow$ (a) n'utilise aucun chemin préexistant : le champ $(\theta, \zeta) \mapsto (-A(\theta), -B(\theta))$ est construit à partir des coefficients de (†), et l'existence du chemin en découle par Cauchy–Lipschitz. La direction (a) $\Rightarrow$ (b) part d'un chemin C^1 donné par la stabilité régulière et dérive (†) comme identité polynomiale. □

Théorème 4.3 (Borne polynomiale $D \leq 2$)**.** *Soit $\varphi : \mathbb{R}^d \to \mathbb{R}^r$ polynomial, minimal, avec $D := \max_i \deg(\varphi_i) \geq 1$. Si la famille est régulièrement stable, alors $D \leq 2$. Hypothèses. φ polynomial (composantes polynomiales en x), minimal, stable régulièrement (Définition 6.1).*

Démonstration. Idée : Sous (†), le gram G_{ij} vit dans $\mathcal{A}$, de degré maximal D en x. Mais G_{ij} a degré $2D - 2$ si un des φ_i est de degré D. L'inégalité $2D - 2 \leq D$ force $D \leq 2$.

Soit $k \in \{1, \ldots, r\}$ tel que $\deg \varphi_k = D \geq 1$ et $\varphi_k^{(D)}$ désigne son homogène dominant, non nul par définition. Le gradient $\nabla\varphi_k$ est de degré exactement $D - 1$; par construction $G_{kk}(x) = \|\nabla\varphi_k(x)\|^2$ est polynomial de degré $2(D-1) = 2D-2$, avec terme dominant $\|\nabla\varphi_k^{(D)}\|^2$ non identiquement nul (car $\varphi_k^{(D)}$ est un polynôme homogène non constant). Par le Théorème 4.2 (direction (a) $\Rightarrow$ (b)), $G_{kk} \in \mathcal{A}$, d'où $\deg(G_{kk}) \leq D$. Combinant : $2D - 2 \leq D$, soit $D \leq 2$. $\square$

Corollaire 4.4 (Q2 est l'unique cas polynomial)**.** *Parmi les φ polynomiaux minimaux, les seules statistiques qui rendent la famille $\{p_\theta\}$ régulièrement stable sont celles de degré ≤ 2. La classe maximale est, modulo action affine, la statistique quadratique complète du Théorème 3.1.*

Démonstration. Idée : Combiner la borne polynomiale du Théorème 4.3 avec la stabilité du cas quadratique (Théorème 3.1). Par le Théorème 4.3, $D \leq 2$. Par le Théorème 3.1, la statistique quadratique complète vérifie effectivement la stabilité. Toute sous-famille (gaussienne centrée $b = 0$, isotropique $M = \lambda I_d$, translation pure $M = I_d$) satisfait (†) trivialement par restriction à un sous-espace de $\mathcal{A}$. $\square$

5 Bord du domaine et contre-exemples

La borne polynomiale du Théorème 4.3 est *effective* : il existe un contre-exemple de degré 3 qui fait échouer (†) de manière explicite et quantitative.

Théorème 5.1 (Contre-exemple $\varphi(x) = x^4$)**.** *Soit $d = 1$ et $\varphi(x) = x^4$. La famille $\{p_\theta(x) \propto e^{-\theta x^4}\}_{\theta>0}$ n'est* pas *stable par convolution gaussienne : pour tout $\theta > 0$ et tout $\sigma > 0$, il n'existe pas de $\tilde{\theta} \in \mathbb{R}$ tel que $p_\theta * \gamma_\sigma \propto e^{-\tilde{\theta}x^4}$. Hypothèses. $\theta > 0$ pour que $Z_\theta < \infty$ ($\int_{\mathbb{R}} e^{-\theta x^4}\,dx < \infty$ exactement quand $\theta > 0$).*

Démonstration. Idée : On calcule $\mathcal{L}(\theta x^4)$, on identifie une composante x^6 qui n'est pas dans l'espace affine $\mathcal{A} = \mathrm{Span}\{1, x^4\}$, et on conclut par le Théorème 4.2.

Calcul direct : avec $f(x) = \theta x^4$, $f'(x) = 4\theta x^3$, $f''(x) = 12\theta x^2$, donc

$$\mathcal{L}f(x) \;=\; \tfrac{1}{2}(4\theta x^3)^2 - \tfrac{1}{2}(12\theta x^2) \;=\; 8\theta^2 x^6 - 6\theta x^2.$$

Le tube affine est $\mathcal{A} = \mathrm{Span}\{1, x^4\}$, donc $\mathcal{A}$ ne contient que des polynômes de degrés $\in \{0, 4\}$. Or $\mathcal{L}f$ contient un monôme x^6 à coefficient $8\theta^2 \neq 0$. Donc $\mathcal{L}f \notin \mathcal{A}$. Par Théorème 4.2 (contraposée), la famille n'est pas régulièrement stable. Tout $\tilde{\theta}(\sigma)$ éventuel est C^1 par le théorème des fonctions implicites : la jacobienne est $I(\theta) = \mathrm{Var}_\theta(x^4) = 1/(4\theta^2) > 0$ (calcul direct depuis $\log Z_\theta = -\frac{1}{4}\log\theta + \text{const}$, ou [Bro86, Thm. 1.13] pour la log-concavité générale de $\log Z$ sur $\mathrm{int}(\Theta^\star)$). Stabilité ponctuelle et stabilité régulière coïncident donc ici, et l'échec de (†) exclut toute reparamétrisation. $\square$

Théorème 5.2 (Exclusion polynomiale $D \geq 3$)**.** *Soit $\varphi : \mathbb{R}^d \to \mathbb{R}^r$ polynomial, minimal, avec $\max_i \deg(\varphi_i) \geq 3$. La famille $\{p_\theta\}$ n'est pas régulièrement stable. Hypothèses. φ polynomial minimal, au moins une composante de degré ≥ 3, $\Theta^\star$ non vide et d'intérieur non vide (nécessaire pour $Z_\theta < \infty$, qui impose une coercivité de $\theta^\top\varphi$).*

Démonstration. Idée : C'est la contraposée du Théorème 4.3. Par le Théorème 4.3, la stabilité régulière force $D \leq 2$. Par hypothèse $D \geq 3$, donc stabilité régulière exclue. $\square$

Remarque 5.3 (Limites $\sigma \to 0, \infty$ et non-surjectivité)**.** Pour φ quadratique, $\sigma \to 0$ redonne $\widetilde{M} \to M$ (identité) ; $\sigma \to \infty$ envoie $\widetilde{M} \to 0$ (bord de $\Theta^\star$). Le bord est atteint asymptotiquement mais jamais en temps fini ; le flot Riccati (3) reste strictement dans SPD_d sur $[0, \infty)$. La non-surjectivité ($\mathrm{Im} = \{\widetilde{M} \prec \sigma^{-2} I_d\}$) exprime l'irréversibilité du canal gaussien [DCT91].

6 Équivalence Hamilton–Jacobi sous stabilité régulière

Nous établissons la version rigoureuse de l'équivalence entre stabilité et chemin Hamilton–Jacobi. Trois hypothèses cachées doivent être explicitées ; leur faisceau forme la *stabilité régulière.*

Définition 6.1 (Stabilité régulière)**.** La famille $\{p_\theta\}$ est *régulièrement stable* en $\theta \in \mathrm{int}(\Theta^\star)$ si elle vérifie :

- (H0) $\varphi \in C^2(\mathbb{R}^d, \mathbb{R}^r)$.
- (H1) Stabilité ponctuelle : pour tout $\sigma > 0$, il existe $\tilde\theta(\sigma) \in \mathrm{int}(\Theta^\star)$ avec $p_\theta * \gamma_\sigma \propto e^{-\tilde\theta(\sigma)^\top \varphi}$.
- (H2) Minimalité (Définition 4.1).
- (H3) L'application $\sigma \mapsto \tilde\theta(\sigma)$ est C^1 sur $(0, \infty)$ avec $\tilde\theta(0^+) = \theta$.
- (H4) Croissance sous-gaussienne uniforme sur les compacts en t (redondante dans le cas densitaire par Widder 1944).

Les trois hypothèses (H2)–(H4) correspondent précisément aux trois sauts cachés identifiés dans l'énoncé d'Etienne : minimalité (identifiabilité de $\dot\theta_t$), régularité C^1 (existence de la dérivée temporelle), unicité du Cauchy chaleur (réciproque). Le paragraphe suivant montre qu'elles sont non redondantes.

Théorème 6.2 (Q1c + Q3 — équivalence)**.** *Soit $\varphi \in C^2(\mathbb{R}^d, \mathbb{R}^r)$ minimale et $\theta \in \mathrm{int}(\Theta^\star)$. Sont équivalentes :*

(a) La famille est régulièrement stable au point θ (Définition 6.1).

(b) Il existe un chemin C^1 $t \mapsto \theta_t \in \mathrm{int}(\Theta^\star)$ et $c : \mathbb{R}_+ \to \mathbb{R}$ tels que $\theta_0 = \theta$, $c(t) = \frac{d}{dt} \log Z_{\theta_t}$, et

$$\dot\theta_t^\top \varphi(x) + c(t) \;=\; -\tfrac{1}{2}\|(\nabla\varphi(x))\theta_t\|^2 + \tfrac{1}{2}\theta_t^\top \Delta\varphi(x), \qquad \forall x \in \mathbb{R}^d,\ t > 0. \tag{14}$$

Hypothèses. $\varphi \in C^2$ (H0), minimalité (H2) ; θ intérieur au domaine (H1 automatique sur un voisinage) ; unicité de la chaleur garantie par positivité via Widder dans le cas densitaire (H4 redondante).

Démonstration. Idée : (a)⇒(b) dérive l'équation de la chaleur pour $q_t := p_\theta * \gamma_{\sqrt{2t}}$ et applique Cole–Hopf en reconnaissant que $-\log q_t$ est dans l'affine engendré par $(\varphi, \mathbf{1})$. (b)⇒(a) inverse l'argument : (†) déduit de (14) par minimalité permet de reconstruire $\tilde p_\theta \propto e^{-\theta_t^\top \varphi}$ via Cole–Hopf + unicité positive de Widder.

(a) ⇒ (b). Soit $q_t := p_\theta * \gamma_{\sqrt{2t}}$. Alors $\partial_t q_t = \frac{1}{2}\Delta q_t$ sur $\mathbb{R}_+ \times \mathbb{R}^d$. Par (H1), $q_t(x) = \widetilde Z_t^{-1} e^{-\theta_t^\top \varphi(x)}$ avec $\theta_t = \tilde\theta(\sqrt{2t}) \in C^1((0,\infty))$ par (H3). Posons $g_t(x) := -\log q_t(x) = \theta_t^\top \varphi(x) + \log \widetilde Z_t$. Par le Lemme 2.1, $\partial_t g_t + \mathcal{L} g_t = 0$ partout. Développement : $\partial_t g_t = \dot\theta_t^\top \varphi(x) + \frac{d}{dt}\log \widetilde Z_t$; $\mathcal{L} g_t = \frac{1}{2}\|\nabla\varphi(x)\theta_t\|^2 - \frac{1}{2}\theta_t^\top \Delta\varphi(x)$ (le terme $\log \widetilde Z_t$ ne contribue pas car constant en x). Posant $c(t) := \frac{d}{dt}\log \widetilde Z_t$, on aboutit à (14).

(b) ⇒ (a). Soit $(\theta_t, c(t))$ comme en (b). Définissons $g_t(x) := \theta_t^\top \varphi(x) + \int_0^t c(s)\,ds + \log Z_\theta$ et $u_t := e^{-g_t}$. Par construction $u_0 = p_\theta$. Par (14) et linéarité, $\partial_t g_t + \mathcal{L} g_t = 0$, donc par Lemme 2.1 u_t résout $\partial_t u_t = \frac{1}{2}\Delta u_t$. Comme $u_t > 0$ partout (densité), l'unicité de Widder [Wid44] s'applique : $u_t = p_\theta * \gamma_{\sqrt{2t}}$. Donc $\tilde p_\theta = Z_0^{-1} u_{\sigma^2/2}$ est proportionnelle à $e^{-\theta_t^\top \varphi}$ pour $t = \sigma^2/2$. La régularité C^1 de $\sigma \mapsto \tilde\theta$ découle de (b) par composition. La minimalité a déjà été supposée. □

Remarque 6.3 (Nécessité de chaque hypothèse)**.** *(H2)* sans minimalité, $\dot\theta_t$ est non identifiable : l'exemple $\varphi = (x, 2x)$ admet une fibre de reparamétrisations indiscernables et (6) devient une famille à un degré de liberté. *(H3)* sans C^1, $\dot\theta_t$ n'existe pas et (14) est vide de sens (une fonction simplement continue suffirait pour (H1) mais pas pour (H3)). *(H4)* dans le cas densitaire, les q_t sont strictement positives ; Widder (positivity-restricted uniqueness) remplace Tychonoff sans condition de sous-gaussianité uniforme supplémentaire. Sans positivité, on retombe sur Tychonoff [Tyc35].

7 Rigidité exotique et cadre invariant

Les théorèmes précédents concernent la classe polynomiale. Nous étendons maintenant la conclusion à toutes les classes naturelles non polynomiales, en utilisant systématiquement le Théorème 4.2.

Proposition 7.1 (Exclusion dans les classes naturelles)**.** *Soit $\varphi : \mathbb{R}^d \to \mathbb{R}^r$ de classe C^2, minimale, régulièrement stable. Dans chacune des classes suivantes, φ est au plus quadratique (après action affine) :*

(i) ***Harmoniques.*** *Si chaque φ_i est harmonique ($\Delta\varphi_i = 0$), alors par (†) (applied to $H_i \in \mathcal{A}$ trivialement) et $G_{ij} \in \mathcal{A}$, on obtient $\Delta G_{ij} = 2\sum_{k,l}(\partial_k\partial_l\varphi_i)(\partial_k\partial_l\varphi_j) = 2\,\mathrm{tr}(\mathrm{Hess}(\varphi_i)\mathrm{Hess}(\varphi_j)) \in \Delta\mathcal{A}$. Or $\Delta\varphi_k = 0$ force $\Delta\mathcal{A} \subseteq \mathbb{R}$; par itération, $\mathrm{Hess}(\varphi_i) \equiv 0$, soit φ affine.*

(ii) ***Exponentielles*** *$\varphi_i(x) = e^{a_i^\top x}$. Alors $G_{ii}(x) = \|a_i\|^2 e^{2a_i^\top x}$, qui contient la fréquence $2a_i$. Par (†), $\mathcal{A}$ doit contenir $e^{2a_i^\top x}$; par minimalité, soit $2a_i = a_j$ pour un j, soit $a_i = 0$. L'itération génère une tour $a \to 2a \to 4a \to \cdots$ finie impossible sans $a = 0$.*

(iii) ***Trigonométriques*** *$\varphi_i(x) = \cos(a_i^\top x)$. G_{ii} produit une fréquence $2a_i$ par identité trigonométrique ; même argument de tour que (ii), forçant $a_i = 0$.*

(iv) ***Radial lisse sur $\mathbb{R}^d$.*** *Pour $\varphi(x) = h(\|x\|^2)$ avec $h \in C^2(\mathbb{R}_+)$, un calcul direct donne $G(x) = 4\|x\|^2 h'(\|x\|^2)^2$ et $H(x) = 4\|x\|^2 h''(\|x\|^2) + 2d\,h'(\|x\|^2)$. Par (†), $\mathcal{A} = \mathrm{Span}\{h(\|x\|^2), 1\}$ doit absorber G et H : cette contrainte équivaut à une EDO linéaire sur h dont les seules solutions bornées produisant une densité normalisable sont $h(r) = \alpha r$, soit la quadratique radiale.*

(v) ***Entière à croissance polynomiale.*** *Si chaque φ_i est entière sur $\mathbb{C}^d$ de croissance polynomiale $\leq K$, le Théorème de Liouville généralisé (Phragmén–Lindelöf) force φ_i polynomiale de degré $\leq K$. Puis le Théorème 4.3 donne $K \leq 2$.*

Hypothèses. $\varphi \in C^2$, minimalité, stabilité régulière, et la classe structurelle (harmonique, exponentielle, trigonométrique, radiale, entière) correspondante.

Démonstration (esquisse). Idée : Dans chaque classe, on applique (†) et on remarque que G_{ij} ou H_i produit un objet structurellement plus grand *que φ — degré plus élevé, fréquence doublée, nouvelle exponentielle — incompatible avec l'inclusion $\mathcal{A}$. L'itération force la dégénérescence vers l'affine ou la quadratique.* Les cas (i)–(iii) sont des calculs directs du carré du champ dans chaque classe. Le cas (iv) aboutit à une EDO du second ordre dont les solutions analytiques sont classifiées. Le cas (v) combine Liouville généralisé et la borne polynomiale du Théorème 4.3. □

Remarque 7.2 (Reformulation invariante (M, g, L))**.** La condition (†) est invariante par le groupe $\mathrm{Aff}(d) \rtimes O(d)$. Plus largement, dans le cadre invariant (M, g, L) où M est une variété riemannienne, g la métrique et L un générateur elliptique de diffusion satisfaisant la règle de chaîne diffusive [BÉ85 ; BGL14], la condition de stabilité du semigroupe (P_t) associé s'écrit $\Gamma(\varphi_i, \varphi_j) \in \mathcal{A}$ et $L\varphi_i \in \mathcal{A}$. Cela redonne (†) pour $(M, g, L) = (\mathbb{R}^d, I_d, \frac{1}{2}\Delta)$, et donne un flot de Riccati *modifié* pour l'Ornstein–Uhlenbeck $L = \frac{1}{2}\Delta - x \cdot \nabla$ [Nel73]. L'exemple OU illustre que la classe stable coïncide avec Q2 ; seule la partie dérive du flot change.

8 Conclusion

La stabilité des familles exponentielles par convolution gaussienne admet, dans toutes les classes naturelles (polynomiale, harmonique, exponentielle, trigonométrique, radiale, entière à croissance polynomiale), une seule réponse structurelle : la classe quadratique. Cette classe est caractérisée algébriquement (fermeture (†) de Cole–Hopf, Théorème 4.2), analytiquement (bord du domaine, contre-exemple x^4, Théorème 5.1) et dynamiquement (équivalence Hamilton–Jacobi sous stabilité régulière, Théorème 6.2). La reparamétrisation explicite $(M, b) \mapsto (M(I_d +$

$\sigma^2 M)^{-1}, (I_d + \sigma^2 M)^{-1} b)$ et son flot de Riccati $\dot{M} = -M^2$ réconcilient filtrage linéaire gaussien (Kalman), géométrie de l'information (Amari–Nagaoka, m-géodésie), flot de gradient Wasserstein de l'entropie relative (JKO) et canal AWGN (Shannon–Stam).

Questions ouvertes.

1. Existence d'un φ stable dans la classe C^∞ pur, non analytique, à croissance surpolynomiale — toutes les constructions explicites connues échouent mais un argument général de distribution manque.
2. Classification complète des (φ, M, g, L) stables au sens invariant (§7).
3. Extensions Lévy avec mesure de sauts non purement gaussienne.

Références

[AN00] Shun-ichi Amari et Hiroshi Nagaoka. *Methods of Information Geometry.* T. 191. Translations of Mathematical Monographs. American Mathematical Society et Oxford University Press, 2000. isbn : 978-0-8218-0531-2.

[BÉ85] Dominique Bakry et Michel Émery. “Diffusions hypercontractives”. In : *Séminaire de Probabilités XIX 1983/84*. Sous la dir. de J. Azéma et M. Yor. T. 1123. Lecture Notes in Mathematics. Springer, 1985, p. 177-206. doi : 10.1007/BFb0075847.

[BGL14] Dominique Bakry, Ivan Gentil et Michel Ledoux. *Analysis and Geometry of Markov Diffusion Operators.* T. 348. Grundlehren der mathematischen Wissenschaften. Springer, 2014. doi : 10.1007/978-3-319-00227-9.

[Bar78] Ole E. Barndorff-Nielsen. *Information and Exponential Families in Statistical Theory.* Wiley, 1978. doi : 10.1002/9781118857281.

[Bla65] N. M. Blachman. “The Convolution Inequality for Entropy Powers”. In : *IEEE Transactions on Information Theory* 11.2 (1965), p. 267-271. doi : 10.1109/TIT.1965.1053768.

[Bro86] Lawrence D. Brown. *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory.* T. 9. IMS Lecture Notes — Monograph Series. Institute of Mathematical Statistics, 1986. doi : 10.1214/lnms/1215466757.

[Col51] Julian D. Cole. “On a Quasi-Linear Parabolic Equation Occurring in Aerodynamics”. In : *Quarterly of Applied Mathematics* 9.3 (1951), p. 225-236. doi : 10.1090/qam/42889.

[DCT91] Amir Dembo, Thomas M. Cover et Joy A. Thomas. “Information-Theoretic Inequalities”. In : *IEEE Transactions on Information Theory* 37.6 (1991), p. 1501-1518. doi : 10.1109/18.104312.

[Gey] Charles J. Geyer. *Stat 8053 Notes : Exponential Families.* Course notes, School of Statistics, University of Minnesota. Open-access lecture notes on minimal representations, log-partition convexity, and identifiability of exponential families. Available at https://www.stat.umn.edu/geyer/8053/.

[Hop50] Eberhard Hopf. “The Partial Differential Equation $u_t + uu_x = \mu u_{xx}$”. In : *Communications on Pure and Applied Mathematics* 3.3 (1950), p. 201-230. doi : 10.1002/cpa.3160030302.

[Jay57] E. T. Jaynes. “Information Theory and Statistical Mechanics”. In : *Physical Review* 106.4 (1957), p. 620-630. doi : 10.1103/PhysRev.106.620.

[JKO98] Richard Jordan, David Kinderlehrer et Felix Otto. “The Variational Formulation of the Fokker–Planck Equation”. In : *SIAM Journal on Mathematical Analysis* 29.1 (1998), p. 1-17. doi : 10.1137/S0036141096303359.

[Kal60] Rudolf E. Kalman. “A New Approach to Linear Filtering and Prediction Problems”. In : *Journal of Basic Engineering* 82.1 (1960), p. 35-45. doi : 10.1115/1.3662552.

[Nel73] Edward Nelson. “The Free Markoff Field”. In : *Journal of Functional Analysis* 12.2 (1973), p. 211-227. doi : 10.1016/0022-1236(73)90025-6.

[Ott01] Felix Otto. “The Geometry of Dissipative Evolution Equations : The Porous Medium Equation”. In : *Communications in Partial Differential Equations* 26.1-2 (2001), p. 101-174. doi : 10.1081/PDE-100002243.

[Sta59] A. J. Stam. “Some Inequalities Satisfied by the Quantities of Information of Fisher and Shannon”. In : *Information and Control* 2.2 (1959), p. 101-112. doi : 10.1016/S0019-9958(59)90348-1.

[Tyc35] A. N. Tychonoff. “Théorèmes d’unicité pour l’équation de la chaleur”. In : *Matematicheskii Sbornik* 42.2 (1935), p. 199-216.

[WJ08] Martin J. Wainwright et Michael I. Jordan. “Graphical Models, Exponential Families, and Variational Inference”. In : *Foundations and Trends in Machine Learning* 1.1–2 (2008). See Proposition 3.1 for identifiability of minimal exponential families., p. 1-305. doi : 10.1561/2200000001.

[Wid44] David V. Widder. “Positive Temperatures on an Infinite Rod”. In : *Transactions of the American Mathematical Society* 55 (1944), p. 85-95. doi : 10.1090/S0002-9947-1944-0009795-2.

A Convention de temps

Trois conventions circulent dans la littérature : $\sigma^2 = 2t$ (chaleur $\partial_t = \frac{1}{2}\Delta$; notre choix), $\sigma^2 = t$ (chaleur $\partial_t = \Delta$), $s = \sigma^2$ (Riccati normalisé). Le tableau ci-dessous traduit les équations principales.

Convention	$\dot{M}$	$\tilde{\theta}(\sigma)$
$s := \sigma^2,\ \partial_t = \frac{1}{2}\Delta$	$-M^2$	$M(I_d + sM)^{-1}$
$t = \sigma^2/2,\ \partial_t = \frac{1}{2}\Delta$	$-2M^2$	$M(I_d + 2tM)^{-1}$
$t = \sigma^2,\ \partial_t = \Delta$	$-2M^2$	$M(I_d + 2tM)^{-1}$

Le choix $s = \sigma^2$ absorbe les facteurs 2 ; il est utilisé partout dans le corps du papier.

B Lemme de Woodbury commutatif

Lemme B.1 (Woodbury commutatif). *Soient $M \in \mathrm{SPD}_d$ et $\sigma^2 > 0$. Alors M commute avec $(I_d + \sigma^2 M)^{-1}$ et*

$$(M^{-1} + \sigma^2 I_d)^{-1} \;=\; M(I_d + \sigma^2 M)^{-1} \;=\; (I_d + \sigma^2 M)^{-1} M.$$

Démonstration. Idée : Un inverse neumannien de $I_d + \sigma^2 M$ est un polynôme en M, donc commute avec M ; la vérification Woodbury se réduit alors à des simplifications commutatives. $(I_d + \sigma^2 M)^{-1} = \sum_{k\geq 0} (-\sigma^2)^k M^k$ (série convergente pour $\sigma^2 M$ dont toutes les valeurs propres sont > -1, ici strictement > 0), polynôme en M, donc commute avec M. L’identité Woodbury suit par multiplication : $M(I_d+\sigma^2 M)^{-1}(M^{-1}+\sigma^2 I_d) = M(I_d+\sigma^2 M)^{-1} M^{-1}(I_d+\sigma^2 M) = MM^{-1} = I_d$ par commutativité. $\square$