VARIATIONAL MODULUS OF CONVEXITY UNDER MOREAU ENVELOPES: ATTENUATION, RECOVERY, AND ORACLE TUNING

SAMIR ADLY* AND R. TYRRELL ROCKAFELLAR†

Abstract. The variational modulus of convexity at a pair in the graph of the subgradient mapping for a generally nonconvex extended-real-valued function captures in a single number the degree of variational convexity, uniform quadratic growth, and subdifferential monotonicity seeded at that location. It has fundamental implications for understanding local optimality and variational stability, but its behavior with respect to taking Moreau envelopes is explored here for the first time, the motivation being the importance of those envelopes in smoothing and proximal-point methodology. This exploration results in an exact second-order transmission rule reflecting a rational attenuation together with recovery in the limit of vanishing smoothing. The attenuation law is applied to produce a diagnostic, computable from proximal-oracle schemes, for selecting a smoothing parameter while maintaining explicit guarantees of admissibility and inner accuracy.

Key words. Moreau envelope, prox-regularity, strict second subderivative, variational modulus of convexity, spectral attenuation law, second-order variational analysis, curvature estimation.

AMS subject classifications. 49J52; 49J53; 49J50; 90C26; 90C30.

1. Introduction.

Moreau envelopes are important for smoothing, or at least enhancing, functions that lack classically demanded properties of differentiability, or even continuity. They are employed for that in many areas such as signal and image processing, PDE regularization, and numerical optimization. But they are valuable also as tools in second-order variational analysis, the setting for this paper. They provide a route to global dualization beyond global convexity that will be seen to bring out, and make more available for use in applications, the infinitesimally local properties associated with *variational* convexity.

For an extended-real-valued and lower semicontinuous (lsc) function f on $\mathbb{R}^n$ that is proper (nowhere $-\infty$ and not the constant function with value ∞), and any $\lambda > 0$, the *Moreau envelope at level* λ is the function $e_\lambda f$ defined by

$$e_\lambda f(x) \;:=\; \inf_{y\in\mathbb{R}^n} \Big\{ f(y) + \tfrac{1}{2\lambda}\|y-x\|^2 \Big\}.$$

The set-valued mapping $P_\lambda f$ given by

$$P_\lambda f(x) \;:=\; \operatorname*{arg\,min}_{y\in\mathbb{R}^n} \Big\{ f(y) + \tfrac{1}{2\lambda}\|y-x\|^2 \Big\}$$

is the corresponding *proximal mapping*. Moreau himself looked only at $\lambda = 1$ and convex f, although in a Hilbert space instead of $\mathbb{R}^n$ [9, 10], but the concepts were extended to general λ and f in [12] in initializing the topic of prox-regularity. Deep ties to second-order theory can be seen in overview there and in [19], with more about smoothing in [2], and much can also be articulated in infinite dimensions . Here, though, we keep to $\mathbb{R}^n$ because infinite-dimensional counterparts are, so far, not available for a number of results about variational convexity that will be essential to our developments.

Global dualization as an interpretation comes from viewing the Moreau envelope formula in the framework of generalized conjugacy [19, Chap. 11] as $-e_\lambda f(x) = \sup_y\{\, [x,y]_\lambda - f(y)\,\}$ for the "pairing" $[x,y]_\lambda := -\frac{1}{2\lambda}\|x-y\|^2$ and $-e_\lambda f(x)$ as a function of (x,λ) on $\mathbb{R}^n \times (0,\infty)$. As long as that function on $\mathbb{R}^n \times (0,\infty)$ isn't $\equiv \infty$, the case of f being *prox-bounded* (admitting a quadratic minorant) [19, Ex. 1.24], the generalized biconjugate in this pairing framework is again f. (More simply, in fact, $f(x)$ is the limit of $e_\lambda f(x)$ as λ decreases to 0.) Therefore, every first-order or second-order aspect of f is transferred over and retained when passing to Moreau envelopes and, in principle, can be recovered from them. This is our central theme, with a particular focus that will be described next.

Prox-regularity, which marks the threshold for stable behaviors in variational analysis [12, 13, 15], furnishes a class of proper lsc functions f that exhibits quadratic support uniformly around a pair $(\bar x,\bar v)$, where $\bar v$ is a subgradient of f at $\bar x$ in the general (limiting, Mordukhovich) sense, $\bar v \in \partial f(\bar x)$. It makes $\bar v$ be, more specifically, a *proximal* subgradient. It can be studied in terms of the second-order difference functions associated with pairs (x,v) in $\operatorname{gph}\partial f$, the graph of the subdifferential mapping ∂f, that are given by

$$\Delta^2_\tau f(x\,|\,v)(\xi) \;=\; \frac{2}{\tau^2}\Big(f(x+\tau\xi) - f(x) - \tau\langle v,\xi\rangle \Big), \qquad \tau > 0.$$

*Laboratoire XLIM, Université de Limoges, 123 avenue Albert Thomas, 87060 Limoges CEDEX, France. Email: samir.adly@unilim.fr, https://www.unilim.fr/pages_perso/samir.adly/.

†University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195-4350. Email: rtr@uw.edu, https://sites.math.washington.edu/~rtr/mypage.html.

Taking the lower limit as $\tau \downarrow 0$ and $\xi' \to \xi$ defines the basic *second subderivative*

$$d^2 f(x \,|\, v)(\xi) := \liminf_{\substack{\tau \downarrow 0 \\ \xi' \to \xi}} \Delta_\tau^2 f(x \,|\, v)(\xi').$$

The *strict second subderivative* $d_*^2 f(\bar{x} \,|\, \bar{v})$ featured in [18] is obtained instead by also requiring that the pair $(x, v) \in \operatorname{gph} \partial f$ approach $(\bar{x}, \bar{v})$ attentively, i.e., with $f(x) \to f(\bar{x})$. Prox-regularity of f at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ fails [18] if and only if the function $d_*^2 f(\bar{x} \,|\, \bar{v})$ fails to be proper, which is detectable as

$$d_*^2 f(\bar{x} \,|\, \bar{v})(0) = -\infty.$$

Here, our interest in the strict second subderivative is directed at the formula in [18, Thm. 2.2] that it gives for the *variational modulus of convexity*, namely

$$\text{(1.1)} \qquad \operatorname{cnv} f(\bar{x} \,|\, \bar{v}) = \inf_{\|w\|=1} d_*^2 f(\bar{x} \,|\, \bar{v})(w).$$

The modulus $\operatorname{cnv} f(\bar{x} \,|\, \bar{v})$, taken up further in Section 2.3, simultaneously captures "curvature" properties of f in their presence relative to $(\bar{x}, \bar{v})$ from several different angles. At the level $\bar{s} \in \overline{\mathbb{R}}$, $\operatorname{cnv} f(\bar{x} \,|\, \bar{v})$ is the upper bound of the interval of s-values for which f is variationally s-convex, but also the upper bound of the interval of s-values for which f has uniform quadratic growth at level s and the upper bound of the interval of s-values for which ∂f is s-monotone in an f-attentive localization (see [18, (1.12), Def. 1.2]). In this it builds on an earlier modulus of Gfrerer [1] for the first of these intervals. Through the formula (1.1), which we deem to be a *spectral* characterization (and also present a new self-contained proof in Theorem A.2 in Appendix A), $\operatorname{cnv} f(\bar{x} \,|\, \bar{v})$, it acts moreover as a generalized minimal eigenvalue for f at $(\bar{x}, \bar{v})$. It thus encodes a range of information about the second-order structure of f at $\bar{x}$ for $\bar{v}$, along with prox-regularity as identified with having $\operatorname{cnv} f(\bar{x} \,|\, \bar{v}) > -\infty$.

A fundamental question then arises: how does this variational modulus of convexity behave when passing to Moreau envelopes, and what might it say about such regularization? In the convex case, the relationship between f and $e_\lambda f$ follows classical interactions when taking conjugates in convex analysis (see, for example, [19, Chap. 1-2]). Beyond the convex case, the duality supporting that is unavailable and the broader conjugacy outlined above has to be faced instead. The way in which information is preserved or transformed by the envelope requires then a kind of analysis that relies on prox-regularity, which has long been understood anyway as the appropriate setting for second-order structure in nonsmooth and nonconvex contexts, as in [19, Chap. 13]). Constructs like generalized double differentiability and quadratic bundles, have already been developed for prox-regular functions and linked to the differentiability properties of Moreau envelopes (see, for example, [2, 3, 4, 17]).

Main contribution. For a proper lsc function f that is prox-bounded and prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$, we prove that

$$d_*^2 \big[e_\lambda f\big](x_\lambda \,|\, \bar{v}) = e_{\lambda/2}\big[d_*^2 f(\bar{x} \,|\, \bar{v})\big], \qquad x_\lambda = \bar{x} + \lambda \bar{v},$$

for all $\lambda > 0$ sufficiently small. From this commutation formula, we derive the rational law

$$\operatorname{cnv}(e_\lambda f; x_\lambda) = \frac{\operatorname{cnv} f(\bar{x} \,|\, \bar{v})}{1 + \lambda \operatorname{cnv} f(\bar{x} \,|\, \bar{v})}, \quad \text{with } 1 + \lambda \operatorname{cnv} f(\bar{x} \,|\, \bar{v}) > 0,$$

together with the recovery formula

$$\text{(1.2)} \qquad \operatorname{cnv} f(\bar{x} \,|\, \bar{v}) = \frac{\operatorname{cnv}(e_\lambda f; x_\lambda)}{1 - \lambda \operatorname{cnv}(e_\lambda f; x_\lambda)}.$$

The variational modulus of convexity of f at $(\bar{x}, \bar{v})$ is thus modified by a rational rescaling of the variational modulus of convexity of $e_\lambda f$ for $(x_\lambda, \bar{v})$. In some situations, $e_\lambda f$ is of class $C^{1,1}$ locally around x_λ, and (1.2) can be used to estimate $\operatorname{cnv} f(\bar{x} \,|\, \bar{v})$ numerically.

When f is variationally strongly convex with modulus $\sigma > 0$ (the meaning of which will be recalled in Section 2.3), this result recovers the result of Khahn, Mordukhovich, and Phat in [3, Thm. 4.4] as a particular case (see Corollary 3.7 and Remark 3.8 for more details). Here, however, f may only be variationally hypoconvex, and the proof by way of the commutation formulas follows a different track, different from [3] by using the new tool, the strict second subderivative, introduced in [18].

Computational interpretation. Even though our approach is variational, the attenuation-recovery law (1.2) gives a concrete criterion for choosing the regularization parameter λ in proximal computations for nonsmooth, possibly nonconvex, models. When $(1 + \lambda \operatorname{cnv} f(\bar{x} \,|\, \bar{v})) \downarrow 0$, the recovery map is ill-conditioned, so λ must be kept within an interval buffered from such degeneracy. This leads to a curvature estimation based on a proximal oracle: the curvature is deduced from numerical evaluations of the envelope gradient, carried by (1.2), and used to select λ under an explicit admissibility condition. Proposition 5.1 quantifies this envelope-based curvature estimation and the recovery of the base-pair modulus. Proposition 5.4

justifies the secant accuracy assumption a posteriori in the model example, in both the exact and inexact settings. We discuss this novel development in Section 5.

Organization of the paper. Section 2 fleshes out details of the concepts in our second-order framework. Section 3 establishes the commutation formula and associated recovery formulas under Moreau regularization. In Section 4, we offer a worked-out set of examples involving smooth to nonsmooth and nonconvex functions which illustrate the interaction between the strict second subderivative, the variational modulus of convexity, and the Moreau envelopes. In Section 5, as an application from the perspective of proximal computations, we develop a curvature estimation procedure based on a proximal oracle and accompany it with an explicit admissibility rule for the regularization parameter. The paper ends in Section 6 with a summary of the main contributions and suggestions for further research.

2. Second-Order Tools and Variational Structure.

In what follows, $\mathbb{R}^n$ is equipped with its Euclidean structure and, unless otherwise stated, the functions $f : \mathbb{R}^n \to (-\infty, +\infty]$ are taken to be proper and lsc. All constructions are local in a neighborhood of a base pair $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n : v \in \partial f(x)\}$ with $\bar{v}$ being a proximal subgradient. For background on general and proximal subgradients, we refer to the books [8, 19].

2.1. Prox-regularity and resolvent representation.

The function f is *prox-regular* at $\bar{x}$ for $\bar{v}$ if f is finite and locally lsc at $\bar{x}$ with $\bar{v} \in \partial f(\bar{x})$, and there exists $\varepsilon > 0$ and $r \in \mathbb{R}$ such that

$$(2.1)\qquad f(x') \geq f(x) + \langle v,\, x' - x\rangle - \frac{r}{2}\,\|x' - x\|^2 \quad \text{for all } x' \in \mathbb{B}(\bar{x}, \varepsilon),$$

whenever $v \in \partial f(x)$, $\|v - \bar{v}\| < \varepsilon$, $\|x - \bar{x}\| < \varepsilon$, and $f(x) < f(\bar{x}) + \varepsilon$. When this holds for all $\bar{v} \in \partial f(\bar{x})$, f is said to be *prox-regular* at $\bar{x}$. (This ε-ball formulation is equivalent to the f-attentive neighborhood formulation in the paper [18, (1.3)-(1.4)] frequently cited below.)

For $\varepsilon > 0$, the *canonical f-attentive ε-localization of ∂f around* $(\bar{x}, \bar{v})$ is the set-valued mapping $T_\varepsilon : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by

$$(2.2)\qquad T_\varepsilon(x) := \begin{cases} \{\, v \in \partial f(x) \mid \|v - \bar{v}\| < \varepsilon \,\}, & \text{if } \|x - \bar{x}\| < \varepsilon \text{ and } f(x) < f(\bar{x}) + \varepsilon, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Equivalently,

$$(2.3)\qquad \operatorname{gph} T_\varepsilon = \Big\{ (x, v) \in \operatorname{gph} \partial f \;\Big|\; x \in \bar{x} + \varepsilon\mathbb{B},\ v \in \bar{v} + \varepsilon\mathbb{B},\ f(x) < f(\bar{x}) + \varepsilon \Big\}.$$

In taking $\lambda > 0$ sufficiently small and setting

$$x_\lambda := \bar{x} + \lambda\bar{v}, \qquad U_\lambda := \operatorname{rge}(\mathrm{I} + \lambda T_\varepsilon),$$

we get, according to Lemma 3.1 coming below, a neighborhood U_λ of x_λ and which the proximal mapping $P_\lambda f$ and Moreau envelope $e_\lambda f$ in Section 1 satisfy

$$P_\lambda f = (\mathrm{I} + \lambda T_\varepsilon)^{-1} \;\text{ and }\; \nabla e_\lambda f = \lambda^{-1}(\mathrm{I} - P_\lambda f) = (\lambda \mathrm{I} + T_\varepsilon^{-1})^{-1}.$$

Therefore, $\nabla e_\lambda f(x_\lambda) = \bar{v}$.

2.2. Strict second subderivative.

The precise expression for the subderivative $d^2_* f(\bar{x} \,|\, \bar{v})$ described in Section 1, as defined in [18, Def. 2.1], is

$$(2.4)\qquad d^2_* f(\bar{x} \,|\, \bar{v})(\xi) := \liminf_{\substack{\xi' \to \xi,\ \tau \downarrow 0 \\ (x,v) \to (\bar{x}, \bar{v}),\ f(x) \to f(\bar{x})}} \Delta^2_\tau f(x \,|\, v)(\xi').$$

When the function f is of class C^1, we have $\bar{v} = \nabla f(\bar{x})$, in this case we note simply $d^2_* f(\bar{x})(\xi)$.

Strict-admissible quadruples. For a given $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ and $\xi \in \mathbb{R}^n$, we say that a quadruple $(x_k, v_k, \tau_k, \xi_k)_{k \in \mathbb{N}}$ is *strict-admissible for* $d^2_* f(\bar{x} \,|\, \bar{v})(\xi)$ iff there exists $N \in \mathbb{N}$ such that, for all $k \geq N$,

$$v_k \in \partial f(x_k),\ (x_k, v_k) \to (\bar{x}, \bar{v}),\ f(x_k) \to f(\bar{x}),\ \tau_k \downarrow 0,\ \text{and } \xi_k \to \xi \text{ as } k \to +\infty.$$

When f is C^2 around $\bar{x}$ and $\bar{v} = \nabla f(\bar{x})$, according to [18], we have

$$(2.5)\qquad d^2 f(\bar{x} \,|\, \bar{v})(\xi) = d^2_* f(\bar{x} \,|\, \bar{v})(\xi) = \langle \nabla^2 f(\bar{x})\xi, \xi\rangle, \quad \text{for every } \xi \in \mathbb{R}^n.$$

The form of the limit in the definition ensures that the function $q(\xi) := d^2_* f(\bar{x} \,|\, \bar{v})(\xi)$ is lsc and positively 2-homogeneous, in the sense that $q(t\xi) = t^2 q(\xi)$ $(t > 0)$. Here q may take the value $-\infty$ and is therefore *not* necessarily proper. The potential lack of properness of q, and in particular the occurrence of the value $-\infty$, reflects the failure of prox-regularity at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$. The condition $q(0) > -\infty$ is equivalent, on the other hand, to such prox-regularity. This follows directly from [18, Cor. 1.3, Thm. 2.2], together with q being lsc and positively 2-homogeneous.

2.3. Variational modulus of convexity. For a given $s \in \mathbb{R}$, a function $f : \mathbb{R}^n \to (-\infty, +\infty]$ is said to be *s-convex* if $f - \frac{s}{2}\|\cdot\|^2$ is convex. Positive values of s signal strong convexity, while hypoconvexity allows s to be negative. In the terminology of [17], f is *variationally s-convex* at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if, for some convex neighborhood $\mathcal{U} \times \mathcal{V}$ of $(\bar{x}, \bar{v})$ and $\varepsilon > 0$, there is an lsc proper s-convex function $\hat{f}$ such that, in the notation $\mathcal{U}_\varepsilon := \{\, x \in \mathcal{U} \mid f(x) < f(\bar{x}) + \varepsilon \,\}$,

$$\hat{f} \le f \text{ on } \mathcal{U} \text{ with } [\mathcal{U} \times \mathcal{V}] \cap \operatorname{gph} \partial \hat{f} = [\mathcal{U}_\varepsilon \times \mathcal{V}] \cap \operatorname{gph} \partial f$$
$$\text{and } \hat{f}(x) = f(x) \text{ at all } x \text{ such that } (x, v) \text{ belongs to the joint intersection.}$$

This is *variational strong convexity* at level s when $s > 0$, plain *variational convexity* when $s = 0$, but only *variational hypoconvexity* when $s < 0$.

Variational s-convexity entails variational s'-convexity for all $s' < s$, so the s-values for this property at $\bar{x}$ for $\bar{v}$ form an interval in $\mathbb{R}$ that's unbounded from below (but nonempty under prox-regularity), although higher values may require smaller neighborhoods in the definition. The upper bound for that interval, possibly ∞, is the *modulus of convexity* $\operatorname{cnv} f(\bar{x} \,|\, \bar{v})$ in [18, Def. 1.2]. That upper bound had already been designated by Gfrerer in [1] as the modulus of variational convexity, but the contribution in [18, (1.12)-(1.15)] was recognizing that the upper bound of that interval coincides with the the upper bounds from the intervals associated with several other "curvature" properties of f localized to $(\bar{x}, \bar{v})$. Rather than have several different names for the same value, the idea was to have a single name that covers them all and indicates, quantitatively, the infinitesimally local manifestation there of degree of convexity. For more about this, see Appendix A.

3. Moreau-spectral recovery via a commutation identity.

In this section, we show that the Moreau envelope and the strict second subderivative of f commute at the transported pair, in a way that the second-order information of f is transmitted to its envelope. Throughout the paper, convergence of sequences is understood as $k \to +\infty$, we write simply $z_k \to z$.

3.1. Commutation of the Moreau envelope with the strict second subderivative. As before, let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be proper, lsc and prox-bounded. We fix a *base pair* $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$, and assume that f is prox-regular at $\bar{x}$ for $\bar{v}$. Fix $r \in \mathbb{R}$ and $\varepsilon > 0$ as in the prox-regularity inequality (2.1). In what follows, we choose $\lambda > 0$ sufficiently small so that $\frac{1}{\lambda} > r$. Then (2.1) gives the estimate

$$f(x') \ge f(x) + \langle v, x' - x \rangle - \frac{1}{2\lambda} \|x' - x\|^2 \quad \text{for all } x' \in \mathbb{B}(\bar{x}, \varepsilon), \tag{3.1}$$

whenever $v \in \partial f(x)$, $\|v - \bar{v}\| < \varepsilon$, $\|x - \bar{x}\| < \varepsilon$, and $f(x) < f(\bar{x}) + \varepsilon$.

The following lemma records local properties of $P_\lambda f$ and $e_\lambda f$ around the shifted point $x_\lambda = \bar{x} + \lambda \bar{v}$, for small $\lambda > 0$.

LEMMA 3.1 (Tilt lemma). *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc, and prox-bounded, as well as f is prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$. Then there exist $\varepsilon > 0$ and $\lambda_0 > 0$ such that, for every $\lambda \in (0, \lambda_0)$, there is a neighborhood V_λ of $x_\lambda := \bar{x} + \lambda \bar{v}$ satisfying:*

(a) *$P_\lambda f$ is single-valued and Lipschitz on V_λ, with $P_\lambda f(x_\lambda) = \bar{x}$.*

(b) *$e_\lambda f \in C^{1,1}(V_\lambda)$ and $\nabla e_\lambda f(u) = \lambda^{-1}\big(u - P_\lambda f(u)\big)$ for all $u \in V_\lambda$, hence $\nabla e_\lambda f(x_\lambda) = \bar{v}$.*

(c) *Let T_ε be the canonical f-attentive ε-localization of ∂f around $(\bar{x}, \bar{v})$ defined in (2.2), and set $U_\lambda := \operatorname{rge}(\mathrm{I} + \lambda T_\varepsilon)$. Then U_λ is a neighborhood of x_λ, and on U_λ, we have*

$$P_\lambda f = (\mathrm{I} + \lambda T_\varepsilon)^{-1} \quad \text{and} \quad \nabla e_\lambda f = (\lambda \mathrm{I} + T_\varepsilon^{-1})^{-1}.$$

Proof. It is a tilt reduction of [19, Prop 13.37]. Parts **(a)** and **(b)** follow from straightforward algebraic manipulations, while **(c)** is more technical. The subtle point is the meaning of "f-attentive localization," which is defined topologically in [19, pp. 611-612] while we work with the canonical mapping T_ε in (2.2). For this reason, in **(c)** we verify directly that $U_\lambda = \operatorname{rge}(\mathrm{I} + \lambda T_\varepsilon)$ is a neighborhood of x_λ, and that the resolvent identities hold on U_λ. For completeness, the full proof is given in Appendix B. ◻

Next we pin down how $\operatorname{gph} \partial f$ links with $\operatorname{gph} \nabla e_\lambda f$ in passing from (x, v) to $(x + \lambda v, v)$.

LEMMA 3.2 (Local chart at $(\bar{x}, \bar{v})$). *Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be proper, lsc, prox-bounded, and prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$. Then there exist $\varepsilon > 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{v}$ such that, for sufficiently small $\lambda > 0$, there is a neighborhood U_λ of $x_\lambda := \bar{x} + \lambda \bar{v}$ such that:*

(i) *The proximal mapping $P_\lambda f$ is single-valued and Lipschitz on the neighborhood U_λ, and $e_\lambda f$ is of class $C^{1,1}$ on U_λ with $\nabla e_\lambda f(y) = \lambda^{-1}\big(y - P_\lambda f(y)\big)$, $y \in U_\lambda$.*

(ii) *The mapping $\Phi_\lambda : (x, v) \mapsto (x + \lambda v,\, v)$ is a bijection between*

$$\operatorname{gph} \partial f \cap (U^\varepsilon \times V) \quad \text{and} \quad \operatorname{gph} \nabla e_\lambda f \cap (U_\lambda \times V), \quad \text{here } U^\varepsilon := \{x \in U \mid f(x) < f(\bar{x}) + \varepsilon\}.$$

If $(x,v) \in \operatorname{gph} \partial f \cap (U^\varepsilon \times V)$ and $y = x + \lambda v$, then $P_\lambda f(y) = x$ and $\nabla e_\lambda f(y) = v$. On the other hand, if $(y,v) \in \operatorname{gph} \nabla e_\lambda f \cap (U_\lambda \times V)$, then $x := P_\lambda f(y)$ satisfies $(x,v) \in \operatorname{gph} \partial f \cap (U^\varepsilon \times V)$ and $y = x + \lambda v$.

Proof. By Lemma 3.1-**(c)**, there exist $\varepsilon > 0$ and a canonical f-attentive ε-localization T_ε of ∂f around $(\bar{x}, \bar{v})$ such that, for every $\lambda > 0$ sufficiently small, we can take

$$U_\lambda = \operatorname{rge}(\mathrm{I} + \lambda T_\varepsilon), \qquad \nabla e_\lambda f = [\lambda \mathrm{I} + T_\varepsilon^{-1}]^{-1} \ \text{ on } U_\lambda.$$

Set $U := \bar{x} + \varepsilon \mathbb{B}$, $V := \bar{v} + \varepsilon \mathbb{B}$, $U^\varepsilon := \{x \in U \mid f(x) < f(\bar{x}) + \varepsilon\}$. From (2.3) we have $\operatorname{gph} T_\varepsilon = \operatorname{gph} \partial f \cap (U^\varepsilon \times V)$.

(i) Fix $\lambda > 0$ sufficiently small. By Lemma 3.1-**(c)**, for λ small enough the set $U_\lambda = \operatorname{rge}(\mathrm{I} + \lambda T_\varepsilon)$ is a neighborhood of x_λ, and on U_λ we have

$$P_\lambda f = (\mathrm{I} + \lambda T_\varepsilon)^{-1} \ \text{ and } \ \nabla e_\lambda f = (\lambda \mathrm{I} + T_\varepsilon^{-1})^{-1}.$$

By Lemma 3.1-**(a)**-**(b)**, the mapping $P_\lambda f$ is Lipschitz on U_λ and that $e_\lambda f \in C^{1,1}(U_\lambda)$, with

$$\nabla e_\lambda f(y) = \lambda^{-1}\big(y - P_\lambda f(y)\big) \qquad (y \in U_\lambda).$$

(ii) Define $\Phi_\lambda(x,v) := (x + \lambda v, v)$. We prove that Φ_λ is a bijection between $\operatorname{gph} T_\varepsilon$ and $\operatorname{gph} \nabla e_\lambda f \cap (U_\lambda \times V)$. Take $(x,v) \in \operatorname{gph} T_\varepsilon$ and set $y := x + \lambda v$. Then $y \in \operatorname{rge}(\mathrm{I} + \lambda T_\varepsilon) = U_\lambda$. Since $x \in T_\varepsilon^{-1}(v)$, we get $y \in (\lambda \mathrm{I} + T_\varepsilon^{-1})(v)$. This means that

$$v \in [\lambda \mathrm{I} + T_\varepsilon^{-1}]^{-1}(y) = \nabla e_\lambda f(y),$$

and hence $(y,v) \in \operatorname{gph} \nabla e_\lambda f \cap (U_\lambda \times V)$. By **(i)**,

$$P_\lambda f(y) = y - \lambda \nabla e_\lambda f(y) = y - \lambda v = x.$$

Conversely, take $(y,v) \in \operatorname{gph} \nabla e_\lambda f \cap (U_\lambda \times V)$ and set $x := P_\lambda f(y)$. By **(i)**,

$$v = \nabla e_\lambda f(y) = \lambda^{-1}(y - x), \ \text{ so } \ y = x + \lambda v.$$

Since $y \in U_\lambda$ and $\nabla e_\lambda f = [\lambda \mathrm{I} + T_\varepsilon^{-1}]^{-1}$ on U_λ, we have $y \in (\lambda \mathrm{I} + T_\varepsilon^{-1})(v)$, so there exists $x' \in T_\varepsilon^{-1}(v)$ such that $y = x' + \lambda v$. Comparing with $y = x + \lambda v$ gives $x' = x$, hence $x \in T_\varepsilon^{-1}(v)$, that is, $(x,v) \in \operatorname{gph} T_\varepsilon$. Thus $(x,v) \in \operatorname{gph} \partial f \cap (U^\varepsilon \times V)$ and $y = x + \lambda v$. This proves **(ii)**. ∎

Lemma 3.3 (Commutation identity). [1] *Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be proper, lsc, prox-bounded, and prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$. Fix $\lambda > 0$ sufficiently small and set $x_\lambda := \bar{x} + \lambda \bar{v}$. Then, for every $\xi \in \mathbb{R}^n$, we have*

$$d_*^2[e_\lambda f](x_\lambda \,|\, \bar{v})(\xi) = e_{\lambda/2}\big[d_*^2 f(\bar{x} \,|\, \bar{v})\big](\xi). \tag{3.2}$$

Proof. Fix $\lambda > 0$ sufficiently small so that Lemma 3.2 applies and $1/\lambda > r$. Let $\varepsilon > 0$, U, V, and U_λ be the corresponding elements. For $(x,v) \in \operatorname{gph} \partial f \cap (U^\varepsilon \times V)$ with $y := x + \lambda v \in U_\lambda$, Lemma 3.2 gives $P_\lambda f(y) = x$ and $\nabla e_\lambda f(y) = v$. We have also $e_\lambda f(y) = f(x) + \frac{1}{2\lambda}\|x - y\|^2 = f(x) + \frac{\lambda}{2}\|v\|^2$.

Take $(x,v) \in \operatorname{gph} \partial f \cap (U^\varepsilon \times V)$ and set $y := x + \lambda v \in U_\lambda$. For $\tau > 0$ and $\xi \in \mathbb{R}^n$, by definition of $e_\lambda f$, we get

$$e_\lambda f(y + \tau \xi) = \inf_{\eta \in \mathbb{R}^n} \Big\{ f(x + \tau \eta) + \frac{1}{2\lambda}\| - \lambda v + \tau(\eta - \xi)\|^2 \Big\}.$$

Expanding the square and using $e_\lambda f(y) = f(x) + \frac{\lambda}{2}\|v\|^2$ and $\nabla e_\lambda f(y) = v$ gives, for all $\tau > 0$,

$$\Delta_\tau^2[e_\lambda f](y \,|\, v)(\xi) = \inf_{\eta \in \mathbb{R}^n} \Big\{ \Delta_\tau^2 f(x \,|\, v)(\eta) + \frac{1}{\lambda}\|\xi - \eta\|^2 \Big\}. \tag{3.3}$$

We first prove "$\geq$" in (3.2). Fix $\xi \in \mathbb{R}^n$ and consider any strict-admissible quadruple $(y_k, w_k, \tau_k, \xi_k) \to (x_\lambda, \bar{v}, 0, \xi)$ in the definition of $d_*^2[e_\lambda f](x_\lambda \,|\, \bar{v})(\xi)$. For k large, $y_k \in U_\lambda$ and $y_k + \tau_k \xi_k \in U_\lambda$. Since $e_\lambda f$ is $C^{1,1}$ on U_λ, $w_k = \nabla e_\lambda f(y_k)$ for k large. Set $v_k := w_k$ and $x_k := P_\lambda f(y_k)$. Then $(x_k, v_k) \in \operatorname{gph} \partial f \cap (U^\varepsilon \times V)$, $y_k = x_k + \lambda v_k$, and $e_\lambda f(y_k) = f(x_k) + \frac{\lambda}{2}\|v_k\|^2$. Since $(y_k, v_k) \to (x_\lambda, \bar{v})$ and $e_\lambda f(y_k) \to e_\lambda f(x_\lambda) = f(\bar{x}) + \frac{\lambda}{2}\|\bar{v}\|^2$, it follows that $f(x_k) \to f(\bar{x})$, hence $(x_k, v_k) \to (\bar{x}, \bar{v})$ is f-attentive.

Define $u_k := P_\lambda f(y_k + \tau_k \xi_k)$ and $\eta_k := (u_k - x_k)/\tau_k$. Applying (3.3) at (x_k, v_k) with $(\tau, \xi) = (\tau_k, \xi_k)$ and using that the infimum is attained at $\eta = \eta_k$ gives

$$\Delta_{\tau_k}^2[e_\lambda f](y_k \,|\, v_k)(\xi_k) = \Delta_{\tau_k}^2 f(x_k \,|\, v_k)(\eta_k) + \frac{1}{\lambda}\|\xi_k - \eta_k\|^2. \tag{3.4}$$

Because $P_\lambda f$ is Lipschitz on U_λ, the sequence (η_k) is bounded. Passing to a subsequence (not relabeled) such that

$$\Delta_{\tau_k}^2[e_\lambda f](y_k \,|\, v_k)(\xi_k) \to \liminf_j \Delta_{\tau_j}^2[e_\lambda f](y_j \,|\, v_j)(\xi_j),$$

[1] Since $e_\lambda f$ is $C^{1,1}$ near x_λ, one could write $d_*^2(e_\lambda f)(x_\lambda)(\xi)$. We keep the pair $(x_\lambda, \bar{v})$ in the notation to track the base point.

and extracting again if needed, we may assume $\eta_k \to \eta$. Passing to the lim inf in (3.4), we obtain

$$\liminf_k \Delta^2_{\tau_k}[e_\lambda f](y_k \,|\, v_k)(\xi_k) \geq \liminf_k \Delta^2_{\tau_k} f(x_k \,|\, v_k)(\eta_k) + \frac{1}{\lambda}\|\xi - \eta\|^2.$$

Since $(x_k, v_k, \tau_k, \eta_k) \to (\bar{x}, \bar{v}, 0, \eta)$ is strict-admissible for $d^2_* f(\bar{x} \,|\, \bar{v})(\eta)$,

$$d^2_* f(\bar{x} \,|\, \bar{v})(\eta) \leq \liminf_k \Delta^2_{\tau_k} f(x_k \,|\, v_k)(\eta_k),$$

and therefore $\liminf_k \Delta^2_{\tau_k}[e_\lambda f](y_k \,|\, v_k)(\xi_k) \geq d^2_* f(\bar{x} \,|\, \bar{v})(\eta) + \frac{1}{\lambda}\|\xi - \eta\|^2$. Taking the infimum over $\eta \in \mathbb{R}^n$ and then over all admissible quadruples gives

$$(3.5) \qquad d^2_*[e_\lambda f](x_\lambda \,|\, \bar{v})(\xi) \geq \inf_{\eta \in \mathbb{R}^n} \left\{ d^2_* f(\bar{x} \,|\, \bar{v})(\eta) + \frac{1}{\lambda}\|\xi - \eta\|^2 \right\} = e_{\lambda/2}\big[d^2_* f(\bar{x} \,|\, \bar{v})\big](\xi).$$

We now prove "$\leq$" in (3.2). Set $q(\eta) := d^2_* f(\bar{x} \,|\, \bar{v})(\eta)$ and

$$\psi(\xi) := \inf_{\eta \in \mathbb{R}^n} \left\{ q(\eta) + \frac{1}{\lambda}\|\xi - \eta\|^2 \right\} = e_{\lambda/2}[q](\xi).$$

From the prox-regularity at $\bar{x}$ for $\bar{v}$, (2.1) gives the uniform estimate

$$q(\eta) \geq -r\|\eta\|^2, \ \eta \in \mathbb{R}^n,$$

hence, since $1/\lambda > r$, the function $\eta \mapsto q(\eta) + \frac{1}{\lambda}\|\xi - \eta\|^2$ is coercive and $\psi(\xi) \in \mathbb{R}$. Fix $\delta > 0$ and choose $\bar{\eta} \in \mathbb{R}^n$ such that

$$q(\bar{\eta}) + \frac{1}{\lambda}\|\xi - \bar{\eta}\|^2 \leq \psi(\xi) + \delta.$$

By the definition of $q(\bar{\eta})$, there exist an f-attentive sequence $(x_k, v_k) \to (\bar{x}, \bar{v})$ with $v_k \in \partial f(x_k)$ and $f(x_k) \to f(\bar{x})$, a sequence $\tau_k \downarrow 0$, and directions $\eta_k \to \bar{\eta}$ such that

$$\liminf_k \Delta^2_{\tau_k} f(x_k \,|\, v_k)(\eta_k) \leq q(\bar{\eta}) + \delta.$$

Set $y_k := x_k + \lambda v_k$. For k large, $y_k \in U_\lambda$ and $(y_k, v_k) \in \operatorname{gph} \nabla e_\lambda f \cap (U_\lambda \times V)$, and $e_\lambda f(y_k) = f(x_k) + \frac{\lambda}{2}\|v_k\|^2 \to e_\lambda f(x_\lambda)$. Define $\xi_k := \eta_k + (\xi - \bar{\eta})$, so $\xi_k \to \xi$ and $\xi_k - \eta_k = \xi - \bar{\eta}$. Applying (3.3) at (x_k, v_k) with $(\tau, \xi) = (\tau_k, \xi_k)$ and evaluating at $\eta = \eta_k$ gives

$$\Delta^2_{\tau_k}[e_\lambda f](y_k \,|\, v_k)(\xi_k) \leq \Delta^2_{\tau_k} f(x_k \,|\, v_k)(\eta_k) + \frac{1}{\lambda}\|\xi - \bar{\eta}\|^2.$$

Taking lim inf and using the choice of the sequences, we obtain

$$d^2_*[e_\lambda f](x_\lambda \,|\, \bar{v})(\xi) \leq q(\bar{\eta}) + \frac{1}{\lambda}\|\xi - \bar{\eta}\|^2 + \delta \leq \psi(\xi) + 2\delta.$$

Letting $\delta \downarrow 0$ gives

$$(3.6) \qquad d^2_*[e_\lambda f](x_\lambda \,|\, \bar{v})(\xi) \leq \inf_{\eta \in \mathbb{R}^n} \left\{ d^2_* f(\bar{x} \,|\, \bar{v})(\eta) + \frac{1}{\lambda}\|\xi - \eta\|^2 \right\}.$$

Combining the two inequalities (3.5)-(3.6), we obtain (3.2). □

Using Lemma 3.3, we have the following spectral recovery formula for cnv.

THEOREM 3.4 (Moreau-spectral recovery of the variational modulus of convexity). *Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be proper, lsc, prox-bounded, and prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$. For each sufficiently small $\lambda > 0$, set $x_\lambda := \bar{x} + \lambda \bar{v}$ and define*

$$m := \operatorname{cnv} f(\bar{x} \,|\, \bar{v}) = \inf_{\|w\|=1} d^2_* f(\bar{x} \,|\, \bar{v})(w).$$

Assume that $m \in \mathbb{R}$. Then, for every sufficiently small $\lambda > 0$ such that $1 + \lambda m > 0$, we have

$$(3.7) \qquad \operatorname{cnv}(e_\lambda f; x_\lambda) = \inf_{\|w\|=1} d^2_*[e_\lambda f](x_\lambda \,|\, \bar{v})(w) = \frac{m}{1 + \lambda m},$$

and therefore

$$\lim_{\lambda \downarrow 0} \operatorname{cnv}(e_\lambda f; x_\lambda) = \operatorname{cnv} f(\bar{x} \,|\, \bar{v}),$$

with monotone convergence as $\lambda \downarrow 0$.

Proof. Fix $\lambda > 0$ sufficiently small so that Lemma 3.3 and Lemma 3.2 apply, and $1/\lambda > r$. Set $q(\cdot) := d^2_* f(\bar{x} \,|\, \bar{v})(\cdot)$. By prox-regularity (2.1), there exist $\varepsilon > 0$ and $r \in \mathbb{R}$ such that, along any strict-admissible $(x_k, v_k, \tau_k, \xi_k) \to (\bar{x}, \bar{v}, 0, \xi)$,

$$\Delta^2_{\tau_k} f(x_k \,|\, v_k)(\xi_k) \geq -r\|\xi_k\|^2.$$

Taking lim inf gives

$$(3.8) \qquad q(\xi) \geq -r\|\xi\|^2 \qquad (\xi \in \mathbb{R}^n),$$

hence $m = \inf_{\|w\|=1} q(w) > -\infty$.
The mapping q is lsc and positively 2-homogeneous. By Lemma 3.3,

$$d^2_*[e_\lambda f](x_\lambda \,|\, \bar v)(\xi) = e_{\lambda/2}[q](\xi) \qquad (\xi \in \mathbb{R}^n),$$

and therefore

$$\inf_{\|w\|=1} d^2_*[e_\lambda f](x_\lambda \,|\, \bar v)(w) = \inf_{\|w\|=1} \inf_{\eta\in\mathbb{R}^n} \Big\{ q(\eta) + \frac{1}{\lambda}\|w-\eta\|^2 \Big\}. \tag{3.9}$$

We first prove "$\geq$" in (3.7). By 2-homogeneity and the definition of m, $q(\eta) \geq m\|\eta\|^2$ for all η (with $+\infty$ allowed). Hence, for $\|w\| = 1$,

$$e_{\lambda/2}[q](w) \geq \inf_{\eta\in\mathbb{R}^n} \Big\{ m\|\eta\|^2 + \frac{1}{\lambda}\|w-\eta\|^2 \Big\}.$$

If $1 + \lambda m > 0$, the right-hand side is a coercive quadratic and a direct minimization gives

$$\inf_{\eta\in\mathbb{R}^n} \Big\{ m\|\eta\|^2 + \frac{1}{\lambda}\|w-\eta\|^2 \Big\} = \frac{m}{1+\lambda m}.$$

Consequently,

$$\inf_{\|w\|=1} e_{\lambda/2}[q](w) \geq \frac{m}{1+\lambda m}. \tag{3.10}$$

We now prove "$\leq$" in (3.7). Fix $\delta > 0$ and choose w_δ with $\|w_\delta\| = 1$ and $q(w_\delta) \leq m + \delta$. Then

$$e_{\lambda/2}[q](w_\delta) \leq \inf_{\alpha\geq 0} \Big\{ \alpha^2 q(w_\delta) + \frac{1}{\lambda}(1-\alpha)^2 \Big\}.$$

For δ small, $1 + \lambda q(w_\delta) > 0$ and the minimizer $\alpha = (1 + \lambda q(w_\delta))^{-1}$ is feasible, hence

$$e_{\lambda/2}[q](w_\delta) \leq \frac{q(w_\delta)}{1+\lambda q(w_\delta)} \leq \frac{m+\delta}{1+\lambda(m+\delta)}.$$

Taking $\inf_{\|w\|=1}$ and letting $\delta \downarrow 0$ gives

$$\inf_{\|w\|=1} e_{\lambda/2}[q](w) \leq \frac{m}{1+\lambda m}. \tag{3.11}$$

Combining (3.10)-(3.11) with (3.9) gives

$$\inf_{\|w\|=1} d^2_*[e_\lambda f](x_\lambda \,|\, \bar v)(w) = \frac{m}{1+\lambda m}.$$

By Lemma 3.2**(i)**, $e_\lambda f$ is $C^{1,1}$ near x_λ, hence $\partial(e_\lambda f)(x_\lambda) = \{\nabla e_\lambda f(x_\lambda)\} = \{\bar v\}$. Applying Theorem A.2 to $g := e_\lambda f$ at $(x_\lambda, \bar v)$ gives

$$\operatorname{cnv}(e_\lambda f; x_\lambda) = \inf_{\|w\|=1} d^2_*[e_\lambda f](x_\lambda \,|\, \bar v)(w) = \frac{m}{1+\lambda m},$$

which is (3.7). The map $\lambda \mapsto m/(1+\lambda m)$ is nonincreasing on $\{\, \lambda > 0 :\ 1 + \lambda m > 0 \,\}$ and converges to m as $\lambda \downarrow 0$, so the convergence is monotone. ∎

Remark 3.5. The assumption $m := \operatorname{cnv} f(\bar x \mid \bar v) \in \mathbb{R}$ in Theorem 3.4 is used only to state the attenuation law in its standard real-valued form. The case $m = +\infty$ admits a natural extended-real interpretation. Indeed, under the assumptions of Theorem 3.4 and for every sufficiently small $\lambda > 0$ for which the commutation identity holds, we set $q := d^2_* f(\bar x \mid \bar v)$. If $\inf_{\|\xi\|=1} q(\xi) = +\infty$, then $q(\xi) = +\infty$ for every $\|\xi\| = 1$. By the positive 2-homogeneity of q, it follows that $q(\eta) = +\infty$ for every $\eta \neq 0$. Moreover, prox-regularity gives $q(0) \geq 0$, while the constant admissible sequence in the definition of the strict second subderivative gives $q(0) \leq 0$. Hence $q(0) = 0$. Therefore, the commutation identity (3.2) yields

$$d^2_*(e_\lambda f)(x_\lambda \mid \bar v)(\xi) = e_{\lambda/2} q(\xi) = \inf_{\eta\in\mathbb{R}^n} \left\{ q(\eta) + \frac{1}{\lambda}\|\xi-\eta\|^2 \right\} = \frac{1}{\lambda}\|\xi\|^2.$$

Consequently,

$$\operatorname{cnv}(e_\lambda f; x_\lambda) = \inf_{\|\xi\|=1} d^2_*(e_\lambda f)(x_\lambda \mid \bar v)(\xi) = \frac{1}{\lambda}.$$

This agrees with the formal limit

$$\lim_{m\to+\infty} \frac{m}{1+\lambda m} = \frac{1}{\lambda}.$$

The inverse recovery formula is then understood in the limiting sense

$$\lim_{c_\lambda \uparrow 1/\lambda} \frac{c_\lambda}{1-\lambda c_\lambda} = +\infty.$$

Remark 3.6 (Curvature estimation from proximal evaluations). Fix $\lambda > 0$ admissible at $x_\lambda = \bar{x} + \lambda\bar{v}$ and set $c_\lambda := \mathrm{cnv}(e_\lambda f; x_\lambda)$. The transmission rule (3.7) gives

$$m = \mathrm{cnv} f(\bar{x} \,|\, \bar{v}) = \frac{c_\lambda}{1 - \lambda c_\lambda}, \qquad 1 - \lambda c_\lambda > 0.$$

Thus estimating the variational modulus m reduces to estimating c_λ for the $C^{1,1}$ Moreau envelope at x_λ. Moreover, by Lemma 3.2,

$$\nabla e_\lambda f(x) = \lambda^{-1}\bigl(x - P_\lambda f(x)\bigr) \quad \text{near } x_\lambda,$$

so finite differences of $\nabla e_\lambda f$ are directly accessible from proximal evaluations and provide a practical diagnostic for choosing the smoothing parameter λ while keeping explicit admissibility and inner-accuracy control (we refer to Section 5 for more details).

3.2. Recovering the variational modulus of convexity via the Moreau envelope. The attenuation formula (3.7) describes curvature under Moreau regularization and recovers the base-pair modulus from $\mathrm{cnv}(e_\lambda f; x_\lambda)$. By the characterization [3, Thm. 3.2], f is variationally convex at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ iff it is prox-regular there and $e_\lambda f$ is locally convex around $x_\lambda = \bar{x} + \lambda\bar{v}$ for all sufficiently small $\lambda > 0$.

Corollary 3.7 (Inverse spectral recovery). *Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be proper, lsc, and prox-bounded. Assume that f is variationally convex at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ and that*

$$m := \mathrm{cnv} f(\bar{x} \,|\, \bar{v}) < +\infty.$$

For $\lambda > 0$ sufficiently small set $x_\lambda := \bar{x} + \lambda\bar{v}$ and define $c_\lambda := \mathrm{cnv}(e_\lambda f; x_\lambda)$. Then, for all sufficiently small $\lambda > 0$,

$$c_\lambda = \frac{m}{1 + \lambda m} \qquad \Longleftrightarrow \qquad m = \frac{c_\lambda}{1 - \lambda\, c_\lambda}.$$

In particular, $0 \le c_\lambda < 1/\lambda$ and the function $\lambda \mapsto c_\lambda$ is nonincreasing.

Proof. Using [3, Thm. 3.2], we know that variational convexity of f at $(\bar{x}, \bar{v})$ implies prox-regularity at $\bar{x}$ for $\bar{v}$ and local convexity of $e_\lambda f$ around x_λ for all $\lambda > 0$ small. Hence $c_\lambda = \mathrm{cnv}(e_\lambda f; x_\lambda) \ge 0$. Moreover, $s = 0$ is admissible at $(\bar{x}, \bar{v})$, so $m = \mathrm{cnv} f(\bar{x} \,|\, \bar{v}) \ge 0$, and $m < +\infty$ by assumption. In this case, Theorem 3.4 applies and gives, for $\lambda > 0$ sufficiently small, $c_\lambda = \frac{m}{1+\lambda m}$. In particular $0 \le c_\lambda \le m$ and $c_\lambda < 1/\lambda$. Solving for m gives $m = \frac{c_\lambda}{1-\lambda c_\lambda}$, and the function $\lambda \mapsto c_\lambda$ is nonincreasing. □

Remark 3.8 (Strong variational convexity). Assume that f is variationally strongly convex at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ with modulus $\sigma > 0$. By [3, Thm. 4.4], for every sufficiently small $\lambda > 0$ the envelope $e_\lambda f$ is locally strongly convex around $x_\lambda = \bar{x} + \lambda\bar{v}$ with modulus $\sigma/(1 + \sigma\lambda)$. In particular,

$$\mathrm{cnv}(e_\lambda f; x_\lambda) \ge \frac{\sigma}{1 + \sigma\lambda}.$$

On the other hand, Theorem 3.4 gives the exact identity

$$\mathrm{cnv}(e_\lambda f; x_\lambda) = \frac{m}{1 + \lambda m}, \quad \text{with} \;\; m = \mathrm{cnv} f(\bar{x} \,|\, \bar{v}),$$

and in the variationally strongly convex case one has $m \ge \sigma$, so the above bound follows (and is sharpened when $\sigma = m$).

3.3. Smooth perturbations and spectral transmission. We consider a C^2 perturbation $f = f_1 + f_2$. When $f_2 \in C^2$ near $\bar{x}$, the strict second subderivative obeys the addition rule of [18, Prop. 2.3]. Combined with Theorem 3.4, this gives a spectral response bound for envelope moduli; see Proposition 3.9.

Proposition 3.9 (Moreau-spectral response for a C^2 decomposition). *Let $f_1 : \mathbb{R}^n \to (-\infty, +\infty]$ be proper, lsc, prox-bounded, and prox-regular at $\bar{x}$ for $\bar{v}_1 \in \partial f_1(\bar{x})$, and let $f_2 \in C^2$ in a neighborhood of $\bar{x}$. Set*

$$f := f_1 + f_2, \qquad \bar{v} := \bar{v}_1 + \nabla f_2(\bar{x}),$$

so that $\bar{v} \in \partial f(\bar{x}) = \partial f_1(\bar{x}) + \nabla f_2(\bar{x})$. Define

$$m_1 := \mathrm{cnv} f_1(\bar{x} \,|\, \bar{v}_1), \qquad m := \mathrm{cnv} f(\bar{x} \,|\, \bar{v}), \qquad H := \nabla^2 f_2(\bar{x}).$$

Then:

(a) *For all $\xi \in \mathbb{R}^n$,*

$$d^2_* f(\bar{x} \,|\, \bar{v})(\xi) = d^2_* f_1(\bar{x} \,|\, \bar{v}_1)(\xi) + \langle H\xi, \xi\rangle,$$

and, in particular,

$$m \ge m_1 + \lambda_{\min}(H). \tag{3.12}$$

(b) *Assume, in addition, that Theorem 3.4 applies to f_1 at $(\bar x, \bar v_1)$ and to f at $(\bar x, \bar v)$. Set*

$$x^1_\lambda := \bar x + \lambda \bar v_1, \qquad x_\lambda := \bar x + \lambda \bar v.$$

For all sufficiently small $\lambda > 0$ such that $1 + \lambda m_1 > 0$, $1 + \lambda m > 0$, and

$$1 + \lambda\big(m_1 + \lambda_{\min}(H)\big) > 0, \tag{3.13}$$

one has

$$\operatorname{cnv}(e_\lambda f_1; x^1_\lambda) = \frac{m_1}{1+\lambda m_1}, \qquad \operatorname{cnv}(e_\lambda f; x_\lambda) = \frac{m}{1+\lambda m},$$

and, in particular,

$$\operatorname{cnv}(e_\lambda f; x_\lambda) \ge \frac{m_1 + \lambda_{\min}(H)}{1 + \lambda\big(m_1 + \lambda_{\min}(H)\big)}. \tag{3.14}$$

Proof. (a) By [18, Prop. 2.3], with $\bar v_1 := \bar v - \nabla f_2(\bar x) \in \partial f_1(\bar x)$ and $H := \nabla^2 f_2(\bar x)$,

$$d^2_* f(\bar x \,|\, \bar v)(\xi) = d^2_* f_1(\bar x \,|\, \bar v_1)(\xi) + \langle H\xi, \xi\rangle \qquad (\xi \in \mathbb{R}^n).$$

Taking $\inf_{\|w\|=1}$ gives

$$m = \inf_{\|w\|=1} d^2_* f(\bar x \,|\, \bar v)(w) \ge \inf_{\|w\|=1} d^2_* f_1(\bar x \,|\, \bar v_1)(w) + \inf_{\|w\|=1} \langle Hw, w\rangle = m_1 + \lambda_{\min}(H),$$

which is (3.12).

(b) By Theorem 3.4, for all sufficiently small $\lambda > 0$ with $1 + \lambda m_1 > 0$ and $1 + \lambda m > 0$,

$$\operatorname{cnv}(e_\lambda f_1; x^1_\lambda) = \frac{m_1}{1+\lambda m_1}, \qquad \operatorname{cnv}(e_\lambda f; x_\lambda) = \frac{m}{1+\lambda m}.$$

Under (3.13), the map $t \mapsto t/(1+\lambda t)$ is increasing on $\{t :\ 1 + \lambda t > 0\}$. Using (3.12),

$$\operatorname{cnv}(e_\lambda f; x_\lambda) = \frac{m}{1+\lambda m} \ge \frac{m_1 + \lambda_{\min}(H)}{1 + \lambda\big(m_1 + \lambda_{\min}(H)\big)},$$

which is (3.14). ∎

4. Worked-out examples: smooth, nonsmooth, and nonconvex.

This section collects some examples, covering smooth, nonsmooth, and prox-regular nonconvex models. When the strict second subderivative $d^2_* f(\bar x \,|\, \bar v)$ is available in closed form, we compute it in order to make the local second-order geometry concrete and to validate the spectral attenuation-recovery identity (3.7) of Theorem 3.4 on transparent test cases. We then use this theorem to describe the local behavior of the Moreau envelope $e_\lambda f$ near the transported point $x_\lambda = \bar x + \lambda \bar v$, without requiring an explicit formula.

The examples below illustrate complementary aspects of the theory. The one-dimensional ℓ^1-quadratic example shows that, at a kink, the choice of the base subgradient may determine whether the variational modulus is finite or equal to $+\infty$. The nonsmooth nonconvex example shows that negative curvature is also transmitted by the Moreau envelope, according to the same rational law. Finally, the spectral examples show that, when extreme eigenvalues are not simple, obtaining a closed-form expression for the strict second subderivative may become difficult, which motivates the proximal-oracle diagnostic developed in Section 5.

4.1. A one-dimensional ℓ^1-quadratic example.

Consider the function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x)$, defined by

$$f(x) := \alpha|x| + \tfrac{\mu}{2}(x - x_0)^2, \qquad \alpha > 0,\ \mu \in \mathbb{R},$$

where $x_0 \in \mathbb{R}$ is a fixed shift. Let $(\bar x, \bar v)$ be a base pair, i.e. $\bar v \in \partial f(\bar x)$, and set

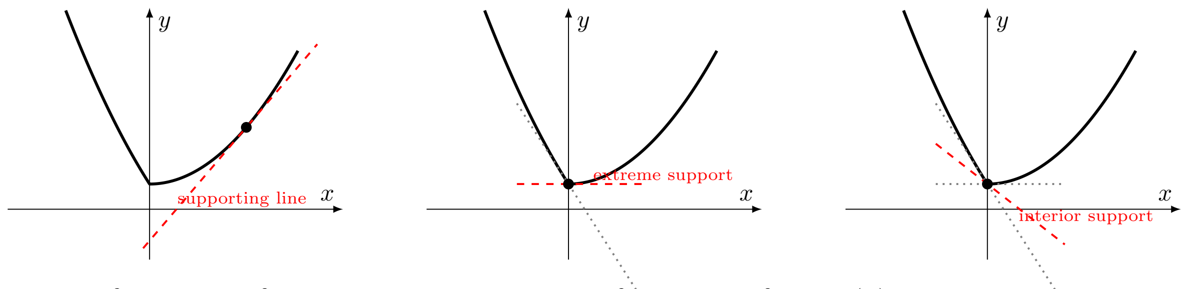
$$\bar s := \frac{\bar v - \mu(\bar x - x_0)}{\alpha}.$$

Then $\bar s \in \partial|\cdot|(\bar x)$ and, for every $\xi \in \mathbb{R}$,

$$d^2_* f(\bar x \,|\, \bar v)(\xi) = \alpha\, d^2_* |\cdot|(\bar x \,|\, \bar s)(\xi) + \mu \xi^2. \tag{4.1}$$

We note that f is prox-bounded since it admits a quadratic minorant; see [19, Ex. 1.24]. Moreover, f is prox-regular at $\bar x$ for $\bar v$ (in the sense of (2.1)): indeed $|\cdot|$ is convex (hence prox-regular with $r = 0$), and prox-regularity is stable under addition of a C^2 function (see [19, Ex. 13.35]).

In this example, three regimes occur at the kink: the smooth case ($\bar x \ne 0$) and the aligned extreme case ($\bar x = 0$, $\bar v + \mu x_0 = \pm\alpha$) give a finite modulus $\operatorname{cnv} f(\bar x \,|\, \bar v) = \mu$ and hence $\operatorname{cnv}(e_\lambda f; x_\lambda) = \mu/(1+\lambda\mu)$, whereas the interior case ($\bar x = 0$, $|\bar v + \mu x_0| < \alpha$) gives $d^2_* f(0 \,|\, \bar v)(\xi) = +\infty$ for $\xi \ne 0$, hence $\operatorname{cnv} f(0 \,|\, \bar v) = +\infty$. The extended-real interpretation in Remark 3.5 then gives $\operatorname{cnv}(e_\lambda f; \lambda \bar v) = 1/\lambda$ (cf. Figures 1-2).



(i) $\bar{x} \neq 0$, $d_*^2 f(\bar{x} \mid \bar{v})(\xi) = \mu\xi^2$ (ii) $\bar{x} = 0$, extreme, $d_*^2 f(0 \mid \bar{v})(\xi) = \mu\xi^2$ (iii) $\bar{x} = 0$, interior

Fig. 1: The three regimes in Proposition 4.1 for $f(x) = \alpha|x| + \frac{\mu}{2}(x - x_0)^2$: (i) $\bar{x} \neq 0$; (ii) $\bar{x} = 0$ and $\bar{v} + \mu x_0 = \pm\alpha$; (iii) $\bar{x} = 0$ and $|\bar{v} + \mu x_0| < \alpha$ (in which case $d_*^2 f(0 \,|\, \bar{v})(\xi) = +\infty$ for $\xi \neq 0$).

PROPOSITION 4.1. (i) *If $\bar{x} \neq 0$, then*

$$\bar{s} = \operatorname{sign}(\bar{x}), \ d_*^2 |\cdot|(\bar{x} \,|\, \bar{s})(\xi) = 0, \ \forall\, \xi \in \mathbb{R}, \quad \textit{and therefore } d_*^2 f(\bar{x} \,|\, \bar{v})(\xi) = \mu\xi^2.$$

(ii) *If $\bar{x} = 0$, then $\bar{s} = (\bar{v} + \mu x_0)/\alpha \in [-1, 1]$ and*

$$d_*^2 f(0 \,|\, \bar{v})(\xi) = \begin{cases} \mu\xi^2, & \text{if } \bar{v} + \mu x_0 = \pm\alpha, \\ 0, & \text{if } |\bar{v} + \mu x_0| < \alpha \text{ and } \xi = 0, \\ +\infty, & \text{if } |\bar{v} + \mu x_0| < \alpha \text{ and } \xi \neq 0. \end{cases} \tag{4.2}$$

Moreover, for all $\lambda > 0$ sufficiently small such that $1 + \lambda\mu > 0$,

$$\operatorname{cnv}\big(e_\lambda f; \, \bar{x} + \lambda\bar{v}\big) = \frac{\mu}{1 + \lambda\mu},$$

in the cases $\bar{x} \neq 0$ or $\bar{x} = 0$ with $\bar{v} + \mu x_0 = \pm\alpha$

Proof. Write $f = \alpha g + h$ with $g(x) = |x|$ and $h(x) = \frac{\mu}{2}(x - x_0)^2$. Then $h \in C^2$, $h'(x) = \mu(x - x_0)$, and $d_*^2 h(\bar{x} \,|\, h'(\bar{x}))(\xi) = \mu\xi^2$. Since $\bar{v} \in \partial f(\bar{x}) = \alpha\, \partial g(\bar{x}) + h'(\bar{x})$, the scalar $\bar{s} := (\bar{v} - h'(\bar{x}))/\alpha$ belongs to $\partial g(\bar{x})$ and $\bar{v} = \alpha\bar{s} + h'(\bar{x})$. By [18, Prop. 2.3], $d_*^2 f(\bar{x} \,|\, \bar{v})(\xi) = \alpha\, d_*^2 g(\bar{x} \,|\, \bar{s})(\xi) + \mu\xi^2$, which is (4.1).

For $g(x) = |x|$, we have

$$d_*^2 g(\bar{x} \,|\, \bar{s})(\xi) = \begin{cases} 0, & \bar{x} \neq 0 \text{ and } \bar{s} = \operatorname{sign}(\bar{x}), \\ 0, & \bar{x} = 0, \ \bar{s} \in \{\pm 1\}, \\ 0, & \bar{x} = 0, \ |\bar{s}| < 1, \ \xi = 0, \\ +\infty, & \bar{x} = 0, \ |\bar{s}| < 1, \ \xi \neq 0, \end{cases}$$

hence (i) and (4.2) follow from (4.1).

In the cases $\bar{x} \neq 0$ or $\bar{x} = 0$ with $\bar{v} + \mu x_0 = \pm\alpha$, one has $d_*^2 f(\bar{x} \,|\, \bar{v})(\xi) = \mu\xi^2$, hence $\operatorname{cnv} f(\bar{x} \,|\, \bar{v}) = \mu \in \mathbb{R}$ and Theorem 3.4 gives

$$\operatorname{cnv}\big(e_\lambda f; \bar{x} + \lambda\bar{v}\big) = \frac{\mu}{1 + \lambda\mu} \quad (\lambda > 0 \text{ sufficiently small}, \ 1 + \lambda\mu > 0).$$

Remark 4.2 (Degenerate interior case at the kink). Assume $\bar{x} = 0$ and

$$|\bar{v} + \mu x_0| < \alpha \qquad \text{equivalently} \qquad s := \frac{\bar{v} + \mu x_0}{\alpha} \in (-1, 1).$$

Then

$$d_*^2 |\cdot|(0 \,|\, s)(\xi) = +\infty \ \ (\xi \neq 0), \qquad d_*^2 |\cdot|(0 \,|\, s)(0) = 0,$$

and (4.1) gives

$$d_*^2 f(0 \,|\, \bar{v})(\xi) = +\infty \ \ (\xi \neq 0).$$

Hence

$$\inf_{\|\xi\|=1} d_*^2 f(0 \,|\, \bar{v})(\xi) = +\infty, \qquad \text{so} \qquad \operatorname{cnv} f(0 \,|\, \bar{v}) = +\infty.$$

In this interior regime, the condition $m \in \mathbb{R}$ in Theorem 3.4 fails, so its standard real-valued formulation does not apply. Nevertheless, Remark 3.5 gives $\operatorname{cnv}(e_\lambda f; \lambda\bar{v}) = \frac{1}{\lambda}$ for every sufficiently small $\lambda > 0$. Thus the modulus is transmitted in the extended-real sense.

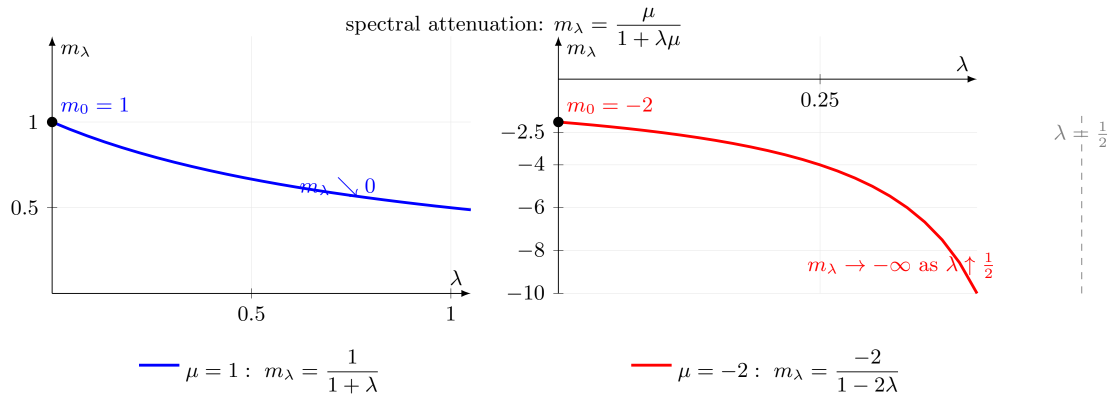


Fig. 2: Moreau transmission of the modulus for $f(x) = \alpha|x| + \frac{\mu}{2}(x-x_0)^2$: $m_\lambda = \mu/(1+\lambda\mu)$ on $\{\,\lambda > 0 :\ 1+\lambda\mu > 0\,\}$. Left: $\mu = 1$ (blue). Right: $\mu = -2$ (red), $\lambda \in (0, \frac{1}{2})$ and $m_\lambda \to -\infty$ as $\lambda \uparrow \frac{1}{2}$.

4.2. A nonsmooth, nonconvex example with negative curvature. Consider $f : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto f(x) := \|x\| - \sin\big(\|x\|^2\big)$, and a base pair $(\bar{x}, \bar{v}) = (0, \bar{v})$ with $\bar{v} \in \partial\|\cdot\|(0) = \overline{\mathbb{B}}(0,1)$ and $\|\bar{v}\| = 1$. Let's write $f = g + h$ with $g(x) := \|x\|$ and $h(x) := -\sin(\|x\|^2)$.

In this example, $d^2_* f(0 \,|\, \bar{v})$ is finite only on the active line $\mathbb{R}\,\bar{v}$: there the C^2 term gives a negative quadratic curvature, while transverse directions are excluded by the norm term. This gives a negative variational modulus of convexity, and Theorem 3.4 transmits it to $e_\lambda f$.

Lemma 4.3. *Let $g(x) = \|x\|$ on $\mathbb{R}^n$. Fix $(\bar{x}, \bar{v}) = (0, \bar{v})$ with $\|\bar{v}\| = 1$. Then, for every $w \in \mathbb{R}^n$,*

$$d^2_* g(0 \,|\, \bar{v})(w) = \begin{cases} 0, & \text{if } w \in \mathbb{R}\,\bar{v}, \\ +\infty, & \text{otherwise.} \end{cases}$$

Proof of Lemma 4.3. Recall

$$\Delta^2_\tau g(x \,|\, v)(\xi) = \frac{2}{\tau^2}\Big(\|x + \tau\xi\| - \|x\| - \tau\langle v, \xi\rangle\Big).$$

Since g is convex, $\|x + \tau\xi\| \geq \|x\| + \tau\langle v, \xi\rangle$ for $v \in \partial g(x)$, hence

$$\Delta^2_\tau g(x \,|\, v)(\xi) \geq 0 \qquad (\tau > 0).$$

- *Case $w \in \mathbb{R}\,\bar{v}$.* Write $w = a\bar{v}$. Take $r_k \downarrow 0$, set $x_k := r_k\bar{v} \neq 0$, $v_k := x_k/\|x_k\| = \bar{v}$, and

$$\tau_k := \frac{r_k}{2(|a|+1)}, \qquad \xi_k := w.$$

Then $x_k + \tau_k\xi_k = (r_k + \tau_k a)\bar{v}$ with $r_k + \tau_k a \geq r_k/2$, so

$$\|x_k + \tau_k\xi_k\| = \|x_k\| + \tau_k\langle v_k, \xi_k\rangle,$$

and $\Delta^2_{\tau_k} g(x_k \,|\, v_k)(\xi_k) = 0$. Thus $d^2_* g(0 \,|\, \bar{v})(w) \leq 0$, and the previous nonnegativity gives $d^2_* g(0 \,|\, \bar{v})(w) = 0$.

- *Case $w \notin \mathbb{R}\,\bar{v}$.* Let $c := \|P_{\bar{v}^\perp} w\| > 0$. Consider any strict-admissible choice of (ξ', τ, x, v) for $d^2_* g(0 \,|\, \bar{v})(w)$ such that $\xi' \to w$, $\tau \downarrow 0$, $(x, v) \to (0, \bar{v})$, with $v \in \partial g(x)$ and $g(x) \to 0$. If $x = 0$, then $v \in \partial g(0) = \overline{\mathbb{B}}(0,1)$ and $\Delta^2_\tau g(0 \,|\, v)(\xi') = \frac{2}{\tau}\big(\|\xi'\| - \langle v, \xi'\rangle\big)$. Since $\delta := \|w\| - \langle \bar{v}, w\rangle > 0$, for (v, ξ') close to $(\bar{v}, w)$ one has $\|\xi'\| - \langle v, \xi'\rangle \geq \delta/2$, hence $\Delta^2_\tau g(0 \,|\, v)(\xi') \to +\infty$.

If $x \neq 0$, then $v = x/\|x\|$ and

$$\|x + \tau\xi'\| - \|x\| - \tau\langle v, \xi'\rangle = \frac{\tau^2\|P_{v^\perp}\xi'\|^2}{\|x + \tau\xi'\| + \|x\| + \tau\langle v, \xi'\rangle}.$$

Therefore,

$$\Delta^2_\tau g(x \,|\, v)(\xi') = \frac{2\|P_{v^\perp}\xi'\|^2}{\|x + \tau\xi'\| + \|x\| + \tau\langle v, \xi'\rangle} \;\geq\; \frac{\|P_{v^\perp}\xi'\|^2}{\|x\| + \tau\|\xi'\|}.$$

Along the net, $v \to \bar{v}$ and $\xi' \to w$, so $\|P_{v^\perp}\xi'\| \to c > 0$, while $\|x\| + \tau\|\xi'\| \to 0$, hence $\Delta^2_\tau g(x \,|\, v)(\xi') \to +\infty$.

Thus every strict-admissible net gives $\Delta^2_\tau g(x \,|\, v)(\xi') \to +\infty$, so $d^2_* g(0 \,|\, \bar{v})(w) = +\infty$. ∎

Proposition 4.4. *Fix a base pair $(\bar{x}, \bar{v}) = (0, \bar{v})$ with $\|\bar{v}\| = 1$. Set $x_\lambda := \bar{x} + \lambda\bar{v} = \lambda\bar{v}$. Then*

$$d^2_* f(0 \,|\, \bar{v})(w) = \begin{cases} -2\,\|w\|^2, & \text{if } w \in \mathbb{R}\,\bar{v}, \\ +\infty, & \text{otherwise,} \end{cases}$$

and therefore $\operatorname{cnv} f(0 \,|\, \bar{v}) = -2$. Moreover, for every $\lambda > 0$ sufficiently small such that $1 - 2\lambda > 0$,

$$\operatorname{cnv}\big(e_\lambda f;\, x_\lambda\big) = \frac{-2}{1-2\lambda}, \qquad \textit{and} \qquad \operatorname{cnv}(e_\lambda f; x_\lambda) \nearrow -2 \ \textit{as} \ \ \lambda \downarrow 0.$$

Proof. Write $f = g + h$ with $g(x) = \|x\|$ and $h(x) = -\sin(\|x\|^2)$. One has $\nabla h(0) = 0$ and $\nabla^2 h(0) = -2I$, hence

$$d_*^2 h(0 \,|\, 0)(w) = \langle \nabla^2 h(0)w, w\rangle = -2\|w\|^2 \qquad (w \in \mathbb{R}^n).$$

Fix $(0, \bar v)$ with $\|\bar v\| = 1$. By Lemma 4.3,

$$d_*^2 g(0 \,|\, \bar v)(w) = 0 \text{ if } w \in \mathbb{R}\,\bar v, \qquad d_*^2 g(0 \,|\, \bar v)(w) = +\infty \text{ otherwise.}$$

Since $h \in C^2$ near 0, the C^2 addition rule [18, Prop. 2.3] gives

$$d_*^2 f(0 \,|\, \bar v)(w) = d_*^2 g(0 \,|\, \bar v)(w) + d_*^2 h(0 \,|\, 0)(w) = \begin{cases} -2\|w\|^2, & w \in \mathbb{R}\,\bar v, \\ +\infty, & \text{otherwise.} \end{cases}$$

Hence $\inf_{\|w\|=1} d_*^2 f(0 \,|\, \bar v)(w) = -2$, so $\operatorname{cnv} f(0 \,|\, \bar v) = -2$.

Moreover, $-\sin(\|x\|^2) \geq -1$ gives $f(x) \geq \|x\| - 1 \geq -\frac{1}{2}\|x\|^2 - \frac{3}{2}$, so f is prox-bounded. Prox-regularity of f at $(0, \bar v)$ follows from convexity of g and C^2 stability; see [19, Ex. 13.35]. Thus Theorem 3.4 applies with $m = -2$ and $x_\lambda = \lambda \bar v$, and for $\lambda > 0$ small with $1 - 2\lambda > 0$,

$$\operatorname{cnv}(e_\lambda f; x_\lambda) = \frac{m}{1 + \lambda m} = \frac{-2}{1 - 2\lambda}.$$

Therefore $\operatorname{cnv}(e_\lambda f; x_\lambda) \nearrow -2$ as $\lambda \downarrow 0$. ∎

Remark 4.5. This example shows that the variational modulus of convexity can be strictly negative at a base pair, and it is transmitted to the Moreau envelope by Theorem 3.4.

4.3. A double-well minimum of quadratics and its soft-min regularization. Consider the pointwise minimum of two quadratic wells $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) \;:=\; \min\left\{\tfrac{1}{2}\|x - u\|^2,\; \tfrac{1}{2}\|x + u\|^2\right\}, \qquad u \in \mathbb{R}^n \setminus \{0\}.$$

An equivalent closed form is given by

$$f(x) = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|u\|^2 - |\langle x, u\rangle|. \tag{4.3}$$

For any $x \in \mathbb{R}^n$, the limiting subdifferential of f at x is given by

$$\partial f(x) = \begin{cases} \{x + u\}, & \text{if } \langle x, u\rangle < 0, \\ \{x - u\}, & \text{if } \langle x, u\rangle > 0, \\ \{x + u,\; x - u\}, & \text{if } \langle x, u\rangle = 0. \end{cases} \tag{4.4}$$

Let's choose $\bar x = 0$, in this case by (4.4), we have $\partial f(0) = \{-u, u\}$. Let's choose $\bar v = u$ (the case $\bar v = -u$ is symmetric). The function f is proper, lsc, prox-bounded and locally Lipschitz, but it fails to be prox-regular at $\bar x = 0$ for $\bar v = u$.
In this example, the switching across the hyperplane $\{\langle x, u\rangle = 0\}$ creates unboundedly negative strict second-order behavior at the admissible base pairs $(0, \pm u)$: $d_*^2 f(0 \,|\, u)$ is quadratic on the half-space $\{\langle u, \xi\rangle < 0\}$ but drops to $-\infty$ as soon as $\langle u, \xi\rangle \geq 0$, hence $\operatorname{cnv} f(0 \,|\, u) = -\infty$ and no spectral transmission for $e_\lambda f$ is available (see Figure 3 for the regularized proxy below).

Proposition 4.6. *For every $\xi \in \mathbb{R}^n$, we have*

$$d_*^2 f(0 \,|\, u)(\xi) = \begin{cases} \|\xi\|^2, & \text{if } \langle u, \xi\rangle < 0, \\ -\infty, & \text{if } \langle u, \xi\rangle \geq 0, \end{cases} \tag{4.5}$$

Consequently,

$$\operatorname{cnv} f(0 \,|\, u) = \inf_{\|\xi\|=1} d_*^2 f(0 \,|\, u)(\xi) = -\infty,$$

so f is not prox-regular at $\bar x = 0$ for $\bar v = u \in \partial f(0)$ and the Moreau spectral attenuation law (3.7) does not apply at this pair.

Proof. For $\tau > 0$ and $\xi \in \mathbb{R}^n$, using (4.3) and $f(0) = \frac{1}{2}\|u\|^2$, we have

$$\Delta_\tau^2 f(0 \,|\, u)(\xi) = \|\xi\|^2 - \frac{2}{\tau}\Big(|\langle u, \xi\rangle| + \langle u, \xi\rangle\Big). \tag{4.6}$$

Set $b := \langle u, \xi\rangle$.

If $b > 0$, then (4.6) gives $\Delta_\tau^2 f(0 \,|\, u)(\xi) = \|\xi\|^2 - \frac{4b}{\tau} \to -\infty$ as $\tau \downarrow 0$, hence $d_*^2 f(0 \,|\, u)(\xi) = -\infty$.

If $b = 0$, let $\hat u := u/\|u\|$, take $\xi_k := \xi + \frac{1}{k}\hat u \to \xi$ and $\tau_k := k^{-2} \downarrow 0$. Then $\langle u, \xi_k\rangle = \|u\|/k > 0$ and (4.6) gives $\Delta_{\tau_k}^2 f(0 \,|\, u)(\xi_k) = \|\xi_k\|^2 - 4\|u\|\,k \to -\infty$, hence $d_*^2 f(0 \,|\, u)(\xi) = -\infty$.

If $b<0$. For the constant admissible choice $(x,v)=(0,u)$ one gets from (4.6) that $\Delta_\tau^2 f(0\,|\,u)(\xi)=\|\xi\|^2$ for all $\tau>0$, so $d_*^2 f(0\,|\,u)(\xi)\le\|\xi\|^2$. Let now $(x_k,v_k,\tau_k,\xi_k)\to(0,u,0,\xi)$ be strict-admissible, i.e. $v_k\in\partial f(x_k)$ and $f(x_k)\to f(0)$. Set $a_k:=\langle u,\xi_k\rangle$ and $b_k:=\langle x_k,u\rangle$. For k large, $a_k<0$. Moreover, (4.4) and $v_k\to u$ give $b_k\le 0$ and $v_k=x_k+u$. A direct expansion using (4.3) gives

$$\Delta_{\tau_k}^2 f(x_k\,|\,v_k)(\xi_k)=\|\xi_k\|^2+\frac{2}{\tau_k^2}\Big(-|b_k+\tau_k a_k|+|b_k|-\tau_k a_k\Big).$$

For k large, $b_k\le 0$ and $a_k<0$ imply $b_k+\tau_k a_k<0$, hence $|b_k+\tau_k a_k|=-(b_k+\tau_k a_k)$ and $|b_k|=-b_k$, so the parenthesis is 0. Thus $\Delta_{\tau_k}^2 f(x_k\,|\,v_k)(\xi_k)=\|\xi_k\|^2$ for k large, and $\liminf_{k\to\infty}\Delta_{\tau_k}^2 f(x_k\,|\,v_k)(\xi_k)=\|\xi\|^2$. Since the strict-admissible sequence is arbitrary, $d_*^2 f(0\,|\,u)(\xi)\ge\|\xi\|^2$, hence $d_*^2 f(0\,|\,u)(\xi)=\|\xi\|^2$ when $\langle u,\xi\rangle<0$.

This proves (4.5). Taking $w_0:=u/\|u\|$ gives $d_*^2 f(0\,|\,u)(w_0)=-\infty$, hence $\inf_{\|w\|=1}d_*^2 f(0\,|\,u)(w)=-\infty$, i.e. $\operatorname{cnv} f(0\,|\,u)=-\infty$. By [18, Cor. 1.3], f is not prox-regular at $(0,u)$, so Theorem 3.4 cannot be applied at this base pair. ∎

Remark 4.7 (Geometry of the kink (limiting subdifferential)). At $x=0$ the function $f=\min\{q_1,q_2\}$ switches across the hyperplane $\{\langle x,u\rangle=0\}$. On each open half-space it coincides with a smooth quadratic with Hessian I, but the switch creates the singular term $-\frac{2}{\tau}\big(|\langle u,\zeta\rangle|+\langle u,\zeta\rangle\big)$ in the second-order quotient at the base pair $(0,u)$, active exactly when $\langle u,\zeta\rangle\ge 0$.

In the limiting setting, $\partial f(0)=\{\pm u\}$, so $(0,0)$ is not an admissible base pair. For any admissible base pair $(0,\bar v)$ with $\bar v\in\{\pm u\}$, Proposition 4.6 gives

$$\operatorname{cnv} f(0\,|\,\bar v)=\inf_{\|w\|=1}d_*^2 f(0\,|\,\bar v)(w)=-\infty.$$

Thus the local second-order behavior is quadratic on each side, but becomes unboundedly negative as soon as directions cross the switching side: $\langle u,w\rangle\ge 0$ for $(0,u)$, and $\langle u,w\rangle\le 0$ for $(0,-u)$.

In this regularization, the double-well geometry is preserved while the curvature at the origin becomes finite and explicitly computable, so the spectral attenuation law applies at the admissible smooth base pair $(0,0)$ (see Figure 3).

Proposition 4.8 (Soft-min regularization). *For $\varepsilon>0$, define the smooth approximation*

$$f_\varepsilon(x):=-\varepsilon\log\big(e^{-\frac{1}{2\varepsilon}\|x-u\|^2}+e^{-\frac{1}{2\varepsilon}\|x+u\|^2}\big).$$

Then $f_\varepsilon\in C^\infty(\mathbb{R}^n)$, is even, and satisfies $f_\varepsilon(x)\uparrow f(x)$ for every $x\in\mathbb{R}^n$ as $\varepsilon\downarrow 0$.
At $(\bar x,\bar v)=(0,0)$ one has $\nabla f_\varepsilon(0)=0$ and $\nabla^2 f_\varepsilon(0)=I-\frac{1}{\varepsilon}(uu^\top)$, where $uu^\top$ denotes the rank-one operator $w\mapsto\langle u,w\rangle\,u$. Consequently,

$$d_*^2 f_\varepsilon(0\,|\,0)(\xi)=\langle\nabla^2 f_\varepsilon(0)\xi,\xi\rangle,\qquad \operatorname{cnv} f_\varepsilon(0\,|\,0)=\lambda_{\min}\big(\nabla^2 f_\varepsilon(0)\big)=1-\frac{1}{\varepsilon}\|u\|^2>-\infty.$$

In particular, the spectral attenuation law applies at $(0,0)$:

$$\operatorname{cnv}(e_\lambda f_\varepsilon;0)=\frac{\operatorname{cnv} f_\varepsilon(0\,|\,0)}{1+\lambda\operatorname{cnv} f_\varepsilon(0\,|\,0)},\qquad 1+\lambda\operatorname{cnv} f_\varepsilon(0\,|\,0)>0.$$

Proof. For $i=1,2$, set

$$q_1(x)=\tfrac12\|x-u\|^2,\qquad q_2(x)=\tfrac12\|x+u\|^2,\qquad Z(x)=e^{-q_1(x)/\varepsilon}+e^{-q_2(x)/\varepsilon},\qquad p_i(x)=\frac{e^{-q_i(x)/\varepsilon}}{Z(x)}.$$

Let $m(x):=\min\{q_1(x),q_2(x)\}$. Since $e^{-m(x)/\varepsilon}\ \le\ Z(x)\ \le\ 2e^{-m(x)/\varepsilon}$, we have

$$m(x)-\varepsilon\log 2\ \le\ f_\varepsilon(x)=-\varepsilon\log Z(x)\ \le\ m(x)=f(x)\qquad(x\in\mathbb{R}^n).$$

For fixed x, set

$$a(x):=\big(e^{-q_1(x)},e^{-q_2(x)}\big),\qquad p:=\frac{1}{\varepsilon}.$$

Then $f_\varepsilon(x)=-\log\|a(x)\|_p$. Since $p\mapsto\|a(x)\|_p$ is nonincreasing and $\|a(x)\|_p\to\|a(x)\|_\infty$ as $p\to+\infty$, it follows that $f_\varepsilon(x)\uparrow f(x)$ as $\varepsilon\downarrow 0$.
Since $f_\varepsilon=-\varepsilon\log Z$, one has

$$\nabla f_\varepsilon(x)=\sum_{i=1}^2 p_i(x)\,\nabla q_i(x)=:\bar g(x),$$

and differentiating gives

$$\nabla^2 f_\varepsilon(x)=\sum_{i=1}^2 p_i(x)\,\nabla^2 q_i(x)-\frac{1}{\varepsilon}\sum_{i=1}^2 p_i(x)\big(\nabla q_i(x)-\bar g(x)\big)\big(\nabla q_i(x)-\bar g(x)\big)^\top.$$

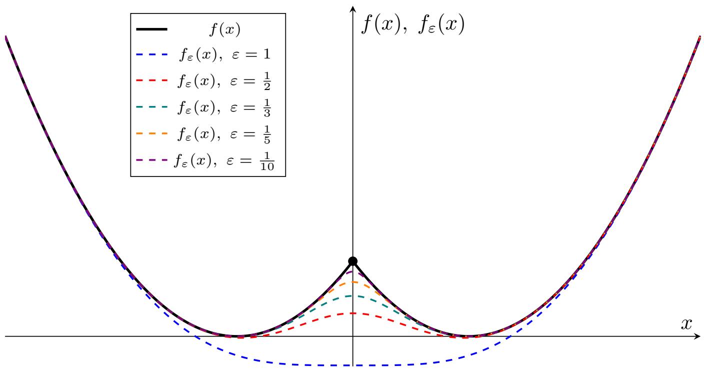


Fig. 3: Soft-min regularization ($u = 1$). The functions f_ε are smooth, even, and satisfy $f_\varepsilon \uparrow f$ as $\varepsilon \downarrow 0$. At $x = 0$, $f'_\varepsilon(0) = 0$ and $f''_\varepsilon(0) = 1 - \varepsilon^{-1}$.

At $x = 0$ we have $q_1(0) = q_2(0) = \frac{1}{2}\|u\|^2$, hence $p_1(0) = p_2(0) = \frac{1}{2}$, and

$$\nabla q_1(0) = -u, \qquad \nabla q_2(0) = u, \qquad \bar{g}(0) = 0, \qquad \nabla^2 q_1(0) = \nabla^2 q_2(0) = \mathrm{I},$$

so

$$\nabla^2 f_\varepsilon(0) = \sum_{i=1}^{2} p_i(0)\, \nabla^2 q_i(0) - \frac{1}{\varepsilon} \sum_{i=1}^{2} p_i(0)\, \nabla q_i(0) \nabla q_i(0)^\top = \mathrm{I} - \frac{1}{\varepsilon}\, uu^\top.$$

Therefore $\nabla^2 f_\varepsilon(0)$ has eigenvalue 1 on $u^\perp$ and eigenvalue $1 - \|u\|^2/\varepsilon$ on $\operatorname{span}\{u\}$, hence

$$\operatorname{cnv} f_\varepsilon(0 \,|\, 0) = \lambda_{\min}\big(\nabla^2 f_\varepsilon(0)\big) = 1 - \frac{1}{\varepsilon}\|u\|^2.$$

Since $f_\varepsilon \in C^\infty$, it is prox-regular at 0 for $\nabla f_\varepsilon(0) = 0$, and

$$d^2_* f_\varepsilon(0 \,|\, 0)(\xi) = \langle \nabla^2 f_\varepsilon(0)\xi, \xi \rangle \qquad (\xi \in \mathbb{R}^n).$$

Moreover, from $f_\varepsilon \geq f - \varepsilon \log 2$ and the fact that f admits a quadratic minorant, f_ε is prox-bounded. Applying Theorem 3.4 with $m = \operatorname{cnv} f_\varepsilon(0 \,|\, 0)$ gives, for all $\lambda > 0$ small with $1 + \lambda m > 0$,

$$\operatorname{cnv}(e_\lambda f_\varepsilon; 0) = \frac{m}{1 + \lambda m}.$$

Remark 4.9 (Limit as $\varepsilon \downarrow 0$). As $\varepsilon \downarrow 0$, the matrix $\nabla^2 f_\varepsilon(0) = \mathrm{I} - \varepsilon^{-1}(uu^\top)$ has eigenvalue $1 - \|u\|^2/\varepsilon$ on $\operatorname{span}\{u\}$ and eigenvalue 1 on $u^\perp$. Hence

$$\operatorname{cnv} f_\varepsilon(0 \,|\, 0) = \lambda_{\min}\big(\nabla^2 f_\varepsilon(0)\big) = 1 - \frac{\|u\|^2}{\varepsilon} \longrightarrow -\infty \qquad (\varepsilon \downarrow 0).$$

For the nonsmooth limit f, one has $\partial f(0) = \{\pm u\}$, so $(0, 0) \notin \operatorname{gph} \partial f$ and $(0, 0)$ is not an admissible base pair. For each admissible base pair $(0, \bar{v})$ with $\bar{v} \in \{\pm u\}$, Proposition 4.6 gives

$$\operatorname{cnv} f(0 \,|\, \bar{v}) = \inf_{\|w\|=1} d^2_* f(0 \,|\, \bar{v})(w) = -\infty.$$

Fix $\varepsilon > 0$. For every $\lambda > 0$ such that $1 + \lambda \operatorname{cnv} f_\varepsilon(0 \,|\, 0) > 0$, Theorem 3.4 gives

$$\operatorname{cnv}(e_\lambda f_\varepsilon; 0) = \frac{\operatorname{cnv} f_\varepsilon(0 \,|\, 0)}{1 + \lambda \operatorname{cnv} f_\varepsilon(0 \,|\, 0)}.$$

Thus, for each fixed $\varepsilon > 0$, the smooth function f_ε admits the base pair $(0, 0)$ with finite variational modulus of convexity, while the limit function f admits only the base pairs $(0, \pm u)$ at the origin and the corresponding modulus equals $-\infty$.

4.4. Spectral functions on $\mathcal{S}^n$: a nonconvex prox-regular example. Let $\mathcal{S}^n$ denote the space of real symmetric $n \times n$ matrices, endowed with the Frobenius inner product $\langle X, Y \rangle = \operatorname{tr}(XY)$ and the associated norm $\|\cdot\|_F$. When the extreme eigenvalues of the reference matrix $\bar{X}$ are simple, the local second-order behavior can be written in closed-form. We prove that $\operatorname{cnv} F_{\rho,\alpha}(\bar{X} \,|\, \bar{V}) = -\rho$ and hence, for λ small with $1 - \lambda\rho > 0$, the transmitted envelope modulus $\operatorname{cnv}(e_\lambda F_{\rho,\alpha}; X_\lambda) = -\rho/(1 - \lambda\rho)$.

Proposition 4.10. *Fix parameters $\rho > 0$ and $\alpha \geq 0$, and define*

$$F_{\rho,\alpha}(X) := \lambda_{\max}(X) + \alpha\, \lambda_{\max}(-X) - \frac{\rho}{2}\|X\|_F^2, \qquad X \in \mathcal{S}^n.$$

Let $\bar{X} \in \mathcal{S}^n$ and $\bar{V} \in \partial F_{\rho,\alpha}(\bar{X})$ such that $\lambda_{\max}(\bar{X})$ and $\lambda_{\min}(\bar{X})$ are simple eigenvalues. Then the following assertions hold.

(a) *$F_{\rho,\alpha}$ is prox-bounded with prox-threshold $1/\rho$, and it is prox-regular at $\bar{X}$ for $\bar{V}$.*
(b) *The strict second subderivative satisfies*

$$d_*^2 F_{\rho,\alpha}(\bar{X} \,|\, \bar{V})(H) \geq -\rho \|H\|_F^2 \text{ for all } H \in \mathcal{S}^n,$$

and equality holds for directions H commuting with $\bar{X}$.
(c) *The variational modulus of convexity is given by $\operatorname{cnv} F_{\rho,\alpha}(\bar{X} \,|\, \bar{V}) = -\rho$.*
(d) *For every $\lambda \in (0, 1/\rho)$ small enough, setting $X_\lambda = \bar{X} + \lambda \bar{V}$, the Moreau envelope satisfies*

$$\operatorname{cnv}(e_\lambda F_{\rho,\alpha}; X_\lambda) = \frac{-\rho}{1 - \lambda\rho}.$$

Proof. (a) Write

$$F_{\rho,\alpha} = G_\alpha - \frac{\rho}{2}\|\cdot\|_F^2, \qquad G_\alpha(X) := \lambda_{\max}(X) + \alpha\,\lambda_{\max}(-X) = \lambda_{\max}(X) - \alpha\,\lambda_{\min}(X).$$

For $\alpha \geq 0$, the scalar function $f_\alpha(\lambda) = \max_i \lambda_i + \alpha \max_i(-\lambda_i)$ is symmetric, convex, and lsc on $\mathbb{R}^n$. By Corollary 2.4 in [5], the induced spectral function

$$G_\alpha(X) = f_\alpha(\lambda(X)) = \lambda_{\max}(X) + \alpha\,\lambda_{\max}(-X)$$

is a proper lsc convex function on $\mathcal{S}^n$. Moreover, for every $X \in \mathcal{S}^n$,

$$|\lambda_{\max}(X)| \leq \|X\|_2 \leq \|X\|_F \quad \text{and} \quad |\lambda_{\max}(-X)| \leq \|X\|_2 \leq \|X\|_F,$$

hence

$$|G_\alpha(X)| \leq (1+\alpha)\|X\|_F. \tag{4.7}$$

∎

Consider, for $\lambda > 0$,

$$F_{\rho,\alpha}(X) + \frac{1}{2\lambda}\|X\|_F^2 = G_\alpha(X) + \frac{1}{2}\Big(\frac{1}{\lambda} - \rho\Big)\|X\|_F^2.$$

If $0 < \lambda < 1/\rho$, then $a := \frac{1}{2}(\frac{1}{\lambda} - \rho) > 0$ and by (4.7),

$$F_{\rho,\alpha}(X) + \frac{1}{2\lambda}\|X\|_F^2 \geq -(1+\alpha)\|X\|_F + a\|X\|_F^2 \geq -\frac{(1+\alpha)^2}{4a},$$

where the last inequality is the elementary bound $at^2 - bt \geq -\frac{b^2}{4a}$ with $t = \|X\|_F$ and $b = 1 + \alpha$. Thus $\inf_X \big(F_{\rho,\alpha}(X) + \frac{1}{2\lambda}\|X\|_F^2\big) > -\infty$ for every $\lambda \in (0, 1/\rho)$.

If $\lambda > 1/\rho$, then $\frac{1}{2}(\frac{1}{\lambda} - \rho) < 0$. Taking $X = tI$ with $t \to +\infty$ gives $\|X\|_F^2 = nt^2$ and $G_\alpha(tI) = \lambda_{\max}(tI) + \alpha\,\lambda_{\max}(-tI) = t - \alpha t = (1-\alpha)t$, so

$$F_{\rho,\alpha}(tI) + \frac{1}{2\lambda}\|tI\|_F^2 = (1-\alpha)t + \frac{n}{2}\Big(\frac{1}{\lambda} - \rho\Big)t^2 \to -\infty.$$

Hence $\inf_X \big(F_{\rho,\alpha}(X) + \frac{1}{2\lambda}\|X\|_F^2\big) = -\infty$ for every $\lambda > 1/\rho$. Therefore, $F_{\rho,\alpha}$ is prox-bounded and its prox-threshold is exactly $1/\rho$.
Since G_α is proper lsc and convex, it is prox-regular at $\bar{X}$ for every $V_G \in \partial G_\alpha(\bar{X})$ (with parameter $r = 0$). Moreover, $\phi(X) := -\frac{\rho}{2}\|X\|_F^2$ is C^∞, hence C^2 around $\bar{X}$. Because ϕ is C^1, we have the subdifferential sum rule

$$\partial(G_\alpha + \phi)(\bar{X}) = \partial G_\alpha(\bar{X}) + \nabla\phi(\bar{X}),$$

so for $\bar{V} \in \partial F_{\rho,\alpha}(\bar{X})$ there exists $V_G \in \partial G_\alpha(\bar{X})$ such that $\bar{V} = V_G + \nabla\phi(\bar{X})$. Applying the stability of prox-regularity under C^2 perturbations [19, Ex. 13.35] gives that $F_{\rho,\alpha} = G_\alpha + \phi$ is prox-regular at $\bar{X}$ for $\bar{V}$.
(b) Second-order structure. Let $u_1, \ldots, u_n$ be an orthonormal eigenbasis of $\bar{X}$ with eigenvalues

$$\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_{n-1} > \lambda_n,$$

so that $\lambda_{\max}(\bar{X}) = \lambda_1$ and $\lambda_{\min}(\bar{X}) = \lambda_n$ are simple. Then the subdifferential reduces to singletons,

$$\partial\lambda_{\max}(\bar{X}) = \{u_1 u_1^\top\}, \qquad \partial\lambda_{\min}(\bar{X}) = \{u_n u_n^\top\},$$

and therefore $\bar{V} = u_1 u_1^\top - \alpha u_n u_n^\top - \rho\bar{X}$, see [6, Cor. 10].
Under the simplicity assumptions, $\lambda_{\max}$ and $\lambda_{\min}$ are twice continuously differentiable in a neighborhood of $\bar{X}$, and hence so is G_α; see [7, Thm. 4.2]. Consequently, $F_{\rho,\alpha}$ is C^2 near $\bar{X}$. In this case, the strict second subderivative coincides with the classical second directional derivative at $(\bar{X}, \bar{V})$. Moreover, the second-order perturbation formula for simple eigenvalues gives

$$\frac{d^2}{dt^2}\lambda_{\max}(\bar{X} + tH)\bigg|_{t=0} = 2\sum_{j>1}\frac{\langle u_j, Hu_1\rangle^2}{\lambda_1 - \lambda_j},$$

and an analogous expression holds for $\lambda_{\min}$; see [7, Thm. 3.3 and Section 5, p. 385].

Combining these expansions with the exact second derivative of the quadratic term, we obtain

$$d_*^2 F_{\rho,\alpha}(\bar X \mid \bar V)(H) = 2\sum_{j>1} \frac{\langle u_j, Hu_1\rangle^2}{\lambda_1-\lambda_j} + 2\alpha \sum_{j<n} \frac{\langle u_j, Hu_n\rangle^2}{\lambda_j-\lambda_n} - \rho\,\|H\|_F^2.$$

Each spectral sum is nonnegative, hence

$$d_*^2 F_{\rho,\alpha}(\bar X \mid \bar V)(H) \geq -\rho\|H\|_F^2 \qquad \text{for all } H \in \mathcal{S}^n.$$

If H commutes with $\bar X$, then each eigenspace of $\bar X$ is H-invariant. Since $\lambda_{\max}(\bar X) = \lambda_1$ and $\lambda_{\min}(\bar X) = \lambda_n$ are simple, we have $Hu_1 \in \operatorname{span}\{u_1\}$ and $Hu_n \in \operatorname{span}\{u_n\}$, hence $\langle u_j, Hu_1\rangle = 0$ for $j \neq 1$ and $\langle u_j, Hu_n\rangle = 0$ for $j \neq n$. Therefore both spectral sums vanish and

$$d_*^2 F_{\rho,\alpha}(\bar X \mid \bar V)(H) = -\rho\|H\|_F^2.$$

(c) variational modulus of convexity. Taking the Rayleigh infimum over $\|H\|_F = 1$ and using (b) gives

$$\operatorname{cnv} F_{\rho,\alpha}(\bar X \mid \bar V) \geq -\rho.$$

Choosing any commuting $H \neq 0$ and normalizing to $\|H\|_F = 1$ gives $\operatorname{cnv} F_{\rho,\alpha}(\bar X \mid \bar V) \leq -\rho$, hence $\operatorname{cnv} F_{\rho,\alpha}(\bar X \mid \bar V) = -\rho$.

(d) Spectral attenuation. We view $\mathcal{S}^n$ as a Euclidean space $\mathbb{R}^N$, $N = n(n+1)/2$, endowed with the Frobenius inner product. By part (a), $F_{\rho,\alpha}$ is proper, lsc, prox-bounded, and prox-regular at $\bar X$ for $\bar V \in \partial F_{\rho,\alpha}(\bar X)$. By part (c), $m := \operatorname{cnv} F_{\rho,\alpha}(\bar X \mid \bar V) = -\rho \in \mathbb{R}$. Therefore, Theorem 3.4 applies: for all $\lambda > 0$ sufficiently small such that $1 + \lambda m > 0$ (equivalently $1 - \lambda\rho > 0$), with $X_\lambda = \bar X + \lambda \bar V$,

$$\operatorname{cnv}(e_\lambda F_{\rho,\alpha}; X_\lambda) = \frac{m}{1+\lambda m} = \frac{-\rho}{1-\lambda\rho}.$$

Remark 4.11 (When the extreme eigenvalues are not simple). The simplicity of $\lambda_{\max}(\bar X)$ and $\lambda_{\min}(\bar X)$ are assumed only to keep Proposition 4.10 fully explicit. It ensures that $G_\alpha = \lambda_{\max} - \alpha\lambda_{\min}$ is C^2 near $\bar X$ and allows the classical second-order eigenvalue expansions in part (b); see [7]. Without simplicity, the active eigenspaces have dimension > 1 and the spectral subdifferentials are no longer singletons: $\partial\lambda_{\max}(\bar X)$ and $\partial\lambda_{\min}(\bar X)$ are convex hulls of rank-one projectors on the corresponding active eigenspaces [6]. In this regime, G_α is typically not differentiable at $\bar X$, hence there is no classical Hessian to read off the curvature. More importantly for our purposes, the strict second-order object $d_*^2 G_\alpha(\bar X \mid \bar W)$ depends on the choice of $\bar W \in \partial G_\alpha(\bar X)$ and on the geometry of the active eigenspaces, so the value of $\operatorname{cnv} G_\alpha(\bar X \mid \bar W)$ is no longer given by a simple eigenvalue formula. This is precisely a situation where the attenuation-recovery identity (3.7) in Theorem 3.4 becomes useful: instead of attempting a closed-form computation of $\operatorname{cnv} F_{\rho,\alpha}(\bar X \mid \bar V)$ at a nonsmooth base pair, we can estimate it from proximal oracle. Indeed, for $\lambda > 0$ small enough (and in particular $\lambda < 1/\rho$ so that $F_{\rho,\alpha}$ is prox-bounded), we can form a base pair $(\bar X, \bar V) \in \operatorname{gph} \partial F_{\rho,\alpha}$ and its transported point $X_\lambda = \bar X + \lambda \bar V$, compute the envelope curvature proxy $c_\lambda \approx \operatorname{cnv}(e_\lambda F_{\rho,\alpha}; X_\lambda)$, and then recover a proxy for the base-pair modulus by the inverse map $m \approx c_\lambda/(1-\lambda c_\lambda)$, keeping a fixed margin away from the barrier $1 - \lambda c_\lambda = 0$.

How $e_\lambda F_{\rho,\alpha}$ enters the estimation. In the nonsimple-eigenvalue regime we do not attempt a closed-form evaluation of $\operatorname{cnv} F_{\rho,\alpha}(\bar X \mid \bar V)$. Instead, we follow Section 5: assuming access to proximal evaluations (equivalently, to $\nabla e_\lambda F_{\rho,\alpha}$ on the local chart), we estimate the envelope modulus $\operatorname{cnv}(e_\lambda F_{\rho,\alpha}; X_\lambda)$ from directional secants of the envelope gradient, and then recover a proxy of the base-pair modulus through the inverse attenuation map. The goal here is not an exact spectral formula, but a practical local curvature diagnostic when λ is small and kept away from the barrier (see Section 5 for more details about the curvature estimation).

5. Curvature estimation from Moreau envelopes and admissible smoothing.

This section gives an operational use of the attenuation-recovery identity in Theorem 3.4. Starting from a base pair $(\bar x, \bar v) \in \operatorname{gph} \partial f$, our goal is to extract local curvature information from the Moreau envelopes $e_\lambda f$ by using proximal evaluations. Rather than computing $\operatorname{cnv} f(\bar x \mid \bar v)$ directly, we first estimate the variational modulus of convexity of the envelope at the transported point $x_\lambda = \bar x + \lambda \bar v$, and then recover an approximation of the modulus at the base pair through the inverse recovery formula. Subsection 5.1 develops the directional proxy and its safeguards, while Subsection 5.2 tests them on a nonconvex composite model and quantifies exact and inexact secant errors.

5.1. A general framework based on proximal evaluations.

For an admissible $\lambda > 0$, set $x_\lambda := \bar x + \lambda \bar v$. Theorem 3.4 also requires the feasibility condition $1 + \lambda m > 0$, which cannot be certified since m is unknown. We therefore use two safeguards: a proxy feasibility check based on $\widehat m_\lambda$ (see (5.6)), and an independent conditioning check based on $\widehat{c}_\lambda$ (see (5.7)). By Lemma 3.2, for each admissible λ, the envelope $e_\lambda f$ is $C^{1,1}$ near x_λ and the proximal map $P_\lambda f$ is single-valued and Lipschitz there. Theorem 3.4 gives

$$c_\lambda := \operatorname{cnv}(e_\lambda f; x_\lambda) = \frac{m}{1+\lambda m}, \qquad m = \operatorname{cnv} f(\bar{x}\,|\,\bar{v}), \ \text{ with } \ 1+\lambda m > 0. \tag{5.1}$$

We estimate c_λ from proximal evaluations, invert the relation to get a proxy for m (the base-pair variational modulus of convexity), and then choose λ under (5.7).

Prox-oracle and access to $\nabla e_\lambda f$. Given (x, λ), an exact prox-oracle returns

$$y_\lambda(x) := P_\lambda f(x) \in \arg\min_y \Big\{ f(y) + \tfrac{1}{2\lambda}\|y-x\|^2 \Big\}.$$

In many nonsmooth models, $y_\lambda(x)$ is obtained by an inner iterative scheme, hence an *inexact* oracle returns $y^\varepsilon_\lambda(x)$, possibly together with an inner iteration count. Once such a proximal point is available, the envelope gradients satisfy

$$\nabla e_\lambda f(x) = \frac{1}{\lambda}\big(x - y_\lambda(x)\big), \ \text{ and } \ \nabla e^\varepsilon_\lambda f(x) := \frac{1}{\lambda}\big(x - y^\varepsilon_\lambda(x)\big),$$

so the construction relies only on proximal evaluations; see [20] for related envelope-based approaches in nonconvex settings.

Curvature estimation by directional secants. Fix a stepsize $t > 0$ and a unit direction w. At the transported point x_λ, define the centered secant slope of the Lipschitz gradient by

$$\widehat{c}_\lambda(w) := \frac{1}{2t}\,\big\langle w, \nabla e_\lambda f(x_\lambda + tw) - \nabla e_\lambda f(x_\lambda - tw)\big\rangle. \tag{5.2}$$

If $e_\lambda f$ is twice differentiable at x_λ, then $\widehat{c}_\lambda(w)$ is the standard centered finite-difference approximation of $\langle w, \nabla^2 e_\lambda f(x_\lambda) w\rangle$ from gradient values; see, e.g., [11, Ch. 7, Eq. (7.10)]. In general (with the assumption of Lemma 3.2), $e_\lambda f$ is only $C^{1,1}$, we view $\widehat{c}_\lambda(w)$ as a directional curvature proxy based on secants of $\nabla e_\lambda f$; see [14] for envelope regularity and second-order structure under prox-regularity. Since $\nabla e_\lambda f$ is locally Lipschitz near x_λ, $\widehat{c}_\lambda(w)$ is well-defined and locally bounded. Given a finite set of unit directions $\mathcal{W} \subset \mathbb{S}^{n-1}$ (the unit sphere in $\mathbb{R}^n$), define the aggregated proxy

$$\widehat{c}_\lambda := \min_{w \in \mathcal{W}} \widehat{c}_\lambda(w). \tag{5.3}$$

We choose $\mathcal{W}$ as a finite random set and take the minimum as a conservative proxy for worst-case curvature. The aim is not to approximate the exact smallest eigenvalue of $\nabla^2 e_\lambda f(x_\lambda)$, but to detect when the inverse map in (5.6) becomes ill-conditioned. It reflects the most negative directional curvature of $e_\lambda f$ near x_λ revealed by these secants.

Scope of the directional oracle diagnostic. The estimate above is deterministic and therefore requires the finite set $\mathcal{W} \subset \mathbb{S}^{n-1}$ to approximate the unit sphere through its covering radius. Consequently, the number of directions needed for a uniform guarantee may grow rapidly with the dimension. For this reason, the present oracle-based procedure should be viewed primarily as a local diagnostic in low or moderate dimension, or as a diagnostic restricted to a prescribed family of meaningful directions.

In order to link quantitatively the directional proxy $\hat{c}_\lambda$ to $\operatorname{cnv}(e_\lambda f; x_\lambda)$, we have the following proposition. For a given $\varepsilon > 0$, a set $\mathcal{W} \subset \mathbb{S}^{n-1}$ is an ε-cover of the unit sphere if $\forall w \in \mathbb{S}^{n-1}$, $\exists \tilde{w} \in \mathcal{W}$ such that $\|w - \tilde{w}\| \le \varepsilon$. We note that by compactness of $\mathbb{S}^{n-1}$, we can always choose $\mathcal{W}$ finite.

Proposition 5.1 (Accuracy of the directional secant proxy). *Adopt the setting of Lemma 3.2. Fix $\lambda > 0$ small and let U_λ be the neighborhood of $x_\lambda = \bar{x} + \lambda\bar{v}$ given by Lemma 3.2. Choose $t > 0$ such that $x_\lambda \pm tw \in U_\lambda$ for all $w \in \mathbb{S}^{n-1}$. Define: $q(w) := d^2_*[e_\lambda f](x_\lambda\,|\,\bar{v})(w)$, $w \in \mathbb{S}^{n-1}$. Let $\varepsilon > 0$ and $\mathcal{W} \subset \mathbb{S}^{n-1}$ be a finite ε-cover of the unit sphere. Assume that, for the chosen stepsize $t > 0$,*

$$\max_{w \in \mathcal{W}} \big|\widehat{c}_\lambda(w) - q(w)\big| \le \omega(t), \qquad \omega(t) \ge 0, \tag{5.4}$$

where $\widehat{c}_\lambda(w)$ is defined in (5.2) and $\widehat{c}_\lambda := \inf_{w \in \mathcal{W}} \widehat{c}_\lambda(w)$ as in (5.3). Then there exists $L_\lambda > 0$ such that

$$\big|\widehat{c}_\lambda - \operatorname{cnv}(e_\lambda f; x_\lambda)\big| \le \omega(t) + 2L_\lambda\,\varepsilon. \tag{5.5}$$

Proof. By Lemma 3.2, $e_\lambda f$ is of class $C^{1,1}$ on U_λ and $\bar{v} = \nabla e_\lambda f(x_\lambda)$. Hence $\nabla e_\lambda f$ is Lipschitz on U_λ: there exists $L_\lambda > 0$ such that

$$\|\nabla e_\lambda f(u) - \nabla e_\lambda f(v)\| \le L_\lambda \|u - v\| \qquad \forall u, v \in U_\lambda.$$

Moreover, fix any $x \in U_\lambda$ and any $\tau > 0$. Assume that $x + \tau w_i \in U_\lambda$ $(i = 1, 2)$. Assume in addition that the whole segment $\{\, x + \tau((1-s)w_1 + sw_2) : \ s \in [0,1] \,\}$ is contained in U_λ (which holds for $\tau > 0$ small enough since U_λ contains a convex ball around x). Then, for any $w_1, w_2 \in \mathbb{S}^{n-1}$, an integral argument shows that the maps $w \mapsto \Delta^2_\tau[e_\lambda f](x\,|\,\nabla e_\lambda f(x))(w)$ are $2L_\lambda$-Lipschitz on $\mathbb{S}^{n-1}$, uniformly in (x, τ). Hence, by stability of Lipschitz constants under $\liminf$, q is $2L_\lambda$-Lipschitz on $\mathbb{S}^{n-1}$. Using (5.4) and taking the minimum over $w \in \mathcal{W}$, it is easy to check that $-\omega(t) \le \widehat{c}_\lambda - c_\lambda$, with $c_\lambda := \operatorname{cnv}(e_\lambda f; x_\lambda) = \inf_{\|w\|=1} q(w)$ as defined in (5.1). On the other hand, fix any $\eta > 0$. By the definition of the infimum, there exists $\bar{w} \in \mathbb{S}^{n-1}$ such that: $q(\bar{w}) \le c_\lambda + \eta$. Since, $\mathcal{W} \subset \mathbb{S}^{n-1}$ be an ε-cover of the unit sphere, there exists $\tilde{w} \in \mathcal{W}$ with $\|\tilde{w} - \bar{w}\| \le \varepsilon$. Since q is $2L_\lambda$-Lipschitz on $\mathbb{S}^{n-1}$, we get $q(\tilde{w}) \le q(\bar{w}) + 2L_\lambda\|\tilde{w} - \bar{w}\| \le c_\lambda + \eta + 2L_\lambda\varepsilon$. Using (5.4) at $\tilde{w} \in \mathcal{W}$ and passing to the minimum over $w \in \mathcal{W}$, we get: $\widehat{c}_\lambda \le c_\lambda + \eta + 2L_\lambda\varepsilon + \omega(t)$. Letting $\eta \downarrow 0$ gives $\widehat{c}_\lambda - c_\lambda \le 2L_\lambda\varepsilon + \omega(t)$. Therefore, (5.5) is satisfied. ∎

Remark 5.2. In practice, (5.4) is only required on the sampled directions $w \in \mathcal{W}$. It should be understood as a discretization/inexactness model for the directional secant proxy, summarized by the error bound $\omega(t)$ (typically small for small t). In Proposition 5.4, we give a concrete example (double-well-ℓ^1) where (5.4) is satisfied with an explicit $\omega(t)$ in both exact and inexact prox (see (5.11) and (5.12)).

Inverse recovery and a margin condition. We apply the inverse recovery map only when $1 - \lambda\widehat{c}_\lambda > 0$. We recover a proxy of the base-pair variational modulus of convexity via

$$\widehat{m}_\lambda := \frac{\widehat{c}_\lambda}{1 - \lambda\widehat{c}_\lambda}, \tag{5.6}$$

which is the inverse of $c_\lambda = m/(1 + \lambda m)$. The barrier $1 + \lambda m > 0$ corresponds to $1 - \lambda c_\lambda > 0$. We have

$$1 + \lambda\widehat{m}_\lambda = \frac{1}{1 - \lambda\widehat{c}_\lambda},$$

and enforce the margin condition

$$1 + \lambda\widehat{m}_\lambda \le \delta^{-1}, \qquad \delta \in (0, 1), \tag{5.7}$$

Under this positivity condition, this is equivalent to $1 - \lambda\widehat{c}_\lambda \ge \delta$. To keep a fixed margin away from the theorem feasibility threshold when the recovered proxy is negative, we also impose

$$1 + \lambda\,\widehat{m}_\lambda \ge \delta, \qquad \text{whenever} \quad \widehat{m}_\lambda < 0, \tag{5.8}$$

equivalently $\lambda \le (1 - \delta)/|\widehat{m}_\lambda|$ when $\widehat{m}_\lambda < 0$. This safeguard is different from (5.7): (5.8) keeps λ away from the boundary $1 + \lambda m = 0$, while (5.7) keeps the inverse map (5.6) well conditioned. In computations, we replace c_λ by $\widehat{c}_\lambda$ from (5.3) and impose (5.7). We also restrict $0 < \lambda \le \lambda_{\max}$ to avoid poorly conditioned proximal subproblems. Since the quadratic term $\frac{1}{2\lambda}\|y - x\|^2$ adds curvature of size $1/\lambda$, this effect becomes weaker when λ increases. As a result, the proximal subproblems are usually harder to solve and may need more inner iterations. The choice of $\lambda_{\max}$ depends on the application. Starting from a current value of λ, we update $\lambda \leftarrow \min\{\tau\lambda, \lambda_{\max}\}$ and then enforce the safeguards. If $\widehat{m}_\lambda < 0$, we first apply the feasibility safeguard (5.8), i.e., $\lambda \leftarrow \min\{\lambda, (1 - \delta)/|\widehat{m}_\lambda|\}$. We then enforce the conditioning safeguard (5.7), equivalently $1 - \lambda\widehat{c}_\lambda \ge \delta$. In the numerical illustration, we fix $t > 0$ and restrict $\lambda \in (0, \lambda_{\max}]$ so that $x_\lambda \pm tw$ stays in the local chart of Lemma 3.2 for every $w \in \mathcal{W}$.

Remark 5.3. (i) Assume that, at the points used in the finite differences, $\|y_\lambda^\varepsilon(x) - y_\lambda(x)\| \le \varepsilon_{\text{prox}}$. Then

$$\|\nabla e_\lambda^\varepsilon f(x) - \nabla e_\lambda f(x)\| = \frac{1}{\lambda}\|y_\lambda^\varepsilon(x) - y_\lambda(x)\| \le \frac{\varepsilon_{\text{prox}}}{\lambda},$$

and for $\|w\| = 1$, $\left|\widehat{c}_\lambda^\varepsilon(w) - \widehat{c}_\lambda(w)\right| \le \frac{\varepsilon_{\text{prox}}}{\lambda t}$. So a small error in the proximal point gives a controlled error in the gradient secants.

(ii) Let $R_\lambda(\mu) := \mu/(1 - \lambda\mu)$. On the set $\{\mu : 1 - \lambda\mu \ge \delta\}$ we have $|R_\lambda'(\mu)| = (1 - \lambda\mu)^{-2} \le \delta^{-2}$. Hence, $|R_\lambda(\mu_1) - R_\lambda(\mu_2)| \le \delta^{-2}|\mu_1 - \mu_2|$. In particular, on $(-\infty, 0]$ the map R_λ is 1-Lipschitz. This means: if we keep the margin $1 - \lambda\mu \ge \delta$, the inverse recovery step cannot amplify errors too much.

5.2. Numerical illustration for a composite model. We illustrate Section 5.1 on the composite model

$$f(x) = \phi(Bx) + \alpha\|x\|_1, \qquad x \in \mathbb{R}^n,\ \alpha > 0, \tag{5.9}$$

where ϕ is smooth and nonconvex (double-well type). In this setting the base-pair modulus $m = \operatorname{cnv} f(\bar{x} \,|\, \bar{v})$ may be negative. We work at a base pair $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ and at the transported point $x_\lambda = \bar{x} + \lambda\bar{v}$. We report the quantities defined in (5.2)-(5.7), computed from proximal evaluations.

Numerical proximal solver and restriction on λ. Given x and $\lambda > 0$, we compute $y_\lambda(x) = P_\lambda f(x)$ by approximately solving

$$\min_{y \in \mathbb{R}^n} \ \phi(By) + \alpha\|y\|_1 + \frac{1}{2\lambda}\|y - x\|^2$$

with an inner proximal-gradient method (backtracking and soft-thresholding), initialized at $y = x$. Each call returns the computed proximal point and the associated inner iteration count, used as a cost indicator.

We restrict $0 < \lambda \le \lambda_{\max}$, where $\lambda_{\max}$ is chosen from a local smoothness estimate of $\nabla(\phi \circ B)$ on the region explored by the inner solver.

Directional curvature estimate and recovered modulus. Directional curvature is evaluated at x_λ via (5.2) with stepsize $t = 10^{-4}$ and aggregated by (5.3) using N_w random unit directions. The base-pair modulus is then recovered through (5.6). We repeat the construction for two inner stopping tolerances (for instance 10^{-6} and 10^{-8}) and report $\widehat{c}_\lambda$, $\widehat{m}_\lambda$, the margin indicator in (5.7), and the average inner iteration count. In practice we take t small, but not extremely small. In our tests, $\widehat{c}_\lambda$ changes very little when we replace t by $10t$ or by $t/10$.

Estimated envelope curvature and recovered modulus. Figure 4 shows $\widehat{c}_\lambda$ together with the margin indicator in (5.7) as λ varies in $(0, \lambda_{\max}]$, and the recovered proxy $\widehat{m}_\lambda$ in (5.6). In this experiment, $\widehat{c}_\lambda$ is negative and $|\widehat{c}_\lambda|$ increases with λ. The curves obtained with the two stopping tolerances are essentially identical, suggesting that the proximal accuracy is sufficient at this scale.

Selection of λ under the margin constraint. Starting from $\lambda_1 > 0$, we generate a smoothing sequence by a multiplicative increase and a projection step enforcing $\lambda \le \lambda_{\max}$ together with (5.7). When $\widehat{m}_\lambda < 0$, (5.8) gives $\lambda \le \frac{1-\delta}{|\widehat{m}_\lambda|}$.

Numerical example and parameters. We illustrate numerically the local attenuation-recovery and its characteristics. For the figures, we take $n = 300$ and $m = 200$, and we draw $B \in \mathbb{R}^{m\times n}$ with i.i.d. standard Gaussian entries rescaled as $B_{ij} \sim \mathcal{N}(0, 1/n)$, with a fixed random seed. We set $\alpha = 0.6$ and, for the illustrative nonconvex potential, we use the smooth double-well

$$\phi(z) = \tfrac{1}{4}\|z\|^4 - \tfrac{\mu}{2}\|z\|^2, \qquad \mu = 2.0, \tag{5.10}$$

together with its explicit gradient in the inner solver. The admissible range of λ is restricted by a cap

$$\lambda_{\max} = 0.90\big(\mu\|B\|^2\big)^{-1},$$

where $\|B\|$ is estimated by 80 steps of power iteration, and we sample λ on a grid of $N_\lambda = 40$ values in

$$[\lambda_{\min}, \lambda_{\max}], \qquad \lambda_{\min} = 0.02\,\lambda_{\max}.$$

Directional secants use stepsize $t = 10^{-4}$ and an aggregated minimum over $N_w = 25$ random unit directions. The margin parameter is $\delta = 0.20$. The proximal subproblems are solved by a proximal-gradient method with backtracking, with a maximum of 4000 inner iterations. We report two inner stopping tolerances, 10^{-6} and 10^{-8}, to assess the sensitivity of the recovered quantities to proximal accuracy. The inner tolerance refers to the stopping rule of the proximal-gradient solver (relative iterate change and prox-gradient mapping), and it is used only to assess the sensitivity of $\widehat{c}_\lambda$ and $\widehat{m}_\lambda$ to inexact proximal points. For the sequence of smoothing parameters, we start from $\lambda_1 = 0.20\,\lambda_{\max}$, apply a multiplicative growth factor 1.25, and project back onto the constraints $\lambda \le \lambda_{\max}$ and (5.7) over $K = 20$ steps.

We can justify a posteriori the hypothesis (5.4) on the secant error in this composite model. The following proposition is in this sense and establishes an explicit bound (in exact and inexact prox) under a strict complementarity condition. Adopt the model in (5.9) with (5.10) (with parameter $\mu > 0$), and denote its smooth part by $g(x) := \phi(Bx)$, so that $f(x) = g(x) + \alpha\|x\|_1$.

Proposition 5.4 (Directional secant accuracy for the double-well-ℓ^1 model). *Assume $0 < \lambda < (\mu\|B\|^2)^{-1}$ and fix $(\bar{x}, \bar{v}) \in \operatorname{gph}\partial f$, set $x_\lambda := \bar{x} + \lambda\bar{v}$, and denote $y_\lambda(x) := P_\lambda f(x)$, $y_\lambda := y_\lambda(x_\lambda)$,*

$$r_\lambda := \nabla g(y_\lambda) + \tfrac{1}{\lambda}(y_\lambda - x_\lambda), \qquad I_0 := \{i : (y_\lambda)_i = 0\}, \qquad S := \{i : (y_\lambda)_i \neq 0\}.$$

Assume the following strict complementarity condition: $\exists\, \kappa, \eta > 0$ such that

$$|(r_\lambda)_i| \le \alpha - \kappa \quad \forall i \in I_0, \qquad |(y_\lambda)_i| \ge \eta \quad \forall i \in S.$$

Then $\exists$ a neighborhood U of x_λ such that $e_\lambda f \in C^{2,1}(U)$ and, for $\|w\| = 1$,

$$q(w) := d^2_*[e_\lambda f]\big(x_\lambda \,|\, \bar{v}\big)(w) = \langle \nabla^2 e_\lambda f(x_\lambda) w, w\rangle.$$

Let $\mathcal{W} \subset \mathbb{S}^{n-1}$ be finite and choose $t > 0$ such that $x_\lambda + sw \in U$ for all $|s| \le t$ and $w \in \mathcal{W}$. With $\widehat{c}_\lambda(w)$ defined by (5.2), set

$$R := \max_{w\in\mathcal{W}} \max_{|s|\le t} \|B\, y_\lambda(x_\lambda + sw)\|, \qquad L_e := \frac{6R\,\|B\|^3}{\big(1 - \lambda\mu\|B\|^2\big)^3}.$$

Then

$$\max_{w\in\mathcal{W}} \big|\widehat{c}_\lambda(w) - q(w)\big| \le \frac{L_e}{2}\, t. \tag{5.11}$$

If, at $x_\lambda \pm tw$, an inexact oracle returns $y^\varepsilon_\lambda(x)$ with $\|y^\varepsilon_\lambda(x) - y_\lambda(x)\| \le \varepsilon_{\text{prox}}$, and $\widehat{c}^\varepsilon_\lambda(w)$ is the associated secant proxy, then

$$\max_{w\in\mathcal{W}} \big|\widehat{c}^\varepsilon_\lambda(w) - q(w)\big| \le \frac{L_e}{2}\, t + \frac{\varepsilon_{\text{prox}}}{\lambda t}. \tag{5.12}$$

Proof. Since $\bar v \in \partial f(\bar x)$ and $x_\lambda = \bar x + \lambda \bar v$, we have $0 \in \partial f(\bar x) + \frac{1}{\lambda}(\bar x - x_\lambda)$. Equivalently, $\bar x$ satisfies the first-order optimality condition for the proximal subproblem $\min_{y\in\mathbb{R}^n}\Big\{f(y) + \frac{1}{2\lambda}\|y - x_\lambda\|^2\Big\}$. We have for all $x \in \mathbb{R}^n$,

$$\nabla^2 g(x) = B^\top \nabla^2\phi(Bx)\, B \succeq -\mu B^\top B \succeq -\mu\|B\|^2 I. \tag{5.13}$$

Fix $x \in \mathbb{R}^n$ and consider: $y \mapsto F_x(y) := g(y) + \frac{1}{2\lambda}\|y - x\|^2 + \alpha\|y\|_1$. The Hessian of the smooth part in F_x is given by: $\nabla^2 g(y) + \lambda^{-1} I \succeq (\lambda^{-1} - \mu\|B\|^2)I =: \sigma I$ with $\sigma > 0$. Hence, F_x is σ-strongly convex (as the sum of a strongly convex and a convex function). Therefore, $y_\lambda(x) = P_\lambda f(x)$ is uniquely defined. Applying this with $x = x_\lambda$, and using the optimality condition established above at $\bar x$, we deduce by uniqueness that $P_\lambda f(x_\lambda) = \bar x$. Choose x, x' and set $y = y_\lambda(x)$, $y' = y_\lambda(x')$. Select $s \in \partial\|\cdot\|_1(y)$ and $s' \in \partial\|\cdot\|_1(y')$ such that

$$\nabla g(y) + \tfrac{1}{\lambda}(y - x) + \alpha s = 0 \ \text{ and } \ \nabla g(y') + \tfrac{1}{\lambda}(y' - x') + \alpha s' = 0.$$

Subtracting the last equalities and taking the inner product with $y - y'$, we have

$$\langle \nabla g(y) - \nabla g(y'), y - y'\rangle + \tfrac{1}{\lambda}\|y - y'\|^2 + \alpha\langle s - s', y - y'\rangle = \tfrac{1}{\lambda}\langle x - x', y - y'\rangle.$$

Using the monotonicity of $\partial\|\cdot\|_1$ and (5.13), we obtain

$$\langle \nabla g(y) - \nabla g(y'), y - y'\rangle \geq -\mu\|B\|^2\|y - y'\|^2.$$

Therefore,

$$\Big(\tfrac{1}{\lambda} - \mu\|B\|^2\Big)\|y - y'\|^2 \leq \tfrac{1}{\lambda}\|x - x'\|\,\|y - y'\| \qquad \Rightarrow \qquad \|y_\lambda(x) - y_\lambda(x')\| \leq \frac{1}{1 - \lambda\mu\|B\|^2}\,\|x - x'\|.$$

Consequently, $y_\lambda(\cdot)$ is k-Lipschitz with $k = (1 - \lambda\mu\|B\|^2)^{-1}$. Now set $y_\lambda := y_\lambda(x_\lambda)$ and $r_\lambda := \nabla g(y_\lambda) + \lambda^{-1}(y_\lambda - x_\lambda)$. Since $y_\lambda = P_\lambda f(x_\lambda)$, by the definition of the proximal mapping, this is equivalent to write that: $-r_\lambda \in \alpha\,\partial\|\cdot\|_1(y_\lambda)$. This means that for each i, $(y_\lambda)_i \neq 0 \Rightarrow |(r_\lambda)_i| = \alpha$, while $(y_\lambda)_i = 0 \Rightarrow |(r_\lambda)_i| \leq \alpha$. For x in a neighborhood of x_λ, let us define

$$r(x) := \nabla g\big(y_\lambda(x)\big) + \tfrac{1}{\lambda}\big(y_\lambda(x) - x\big).$$

Since $y_\lambda(\cdot)$ and ∇g are continuous, $r(\cdot)$ is also continuous. By the strict complementarity assumption at y_λ and the continuity of $y_\lambda(\cdot)$ and $r(\cdot)$ at x_λ, there exists a neighborhood U of x_λ such that for all $x \in U$,

$$|(r(x))_i| \leq \alpha - \tfrac{\kappa}{2} \ \ \forall i \in I_0, \qquad |(y_\lambda(x))_i - (y_\lambda)_i| \leq \tfrac{\eta}{2} \ \ \forall i \in S.$$

If $i \in I_0$ and $(y_\lambda(x))_i \neq 0$, then optimality at x forces $|(r(x))_i| = \alpha$, contradiction; hence $(y_\lambda(x))_i = 0$ for all $i \in I_0$. If $i \in S$, the second inequality implies $(y_\lambda(x))_i \neq 0$ and the sign is fixed on U. In fact, for $i \in S$ we have

$$|(y_\lambda(x))_i| \geq |(y_\lambda)_i| - |(y_\lambda(x))_i - (y_\lambda)_i| \geq \eta - \frac{\eta}{2} = \frac{\eta}{2} > 0.$$

Hence the ℓ^1-subdifferential is single-valued on the S-coordinates and constant on U. Setting $\sigma_S := \operatorname{sign}\big((y_\lambda)_S\big) \in \{-1, 1\}^{|S|}$, we have $\partial\|\cdot\|_1\big(y_\lambda(x)\big)_S = \{\sigma_S\}$, $\forall x \in U$. Thus, on U, the ℓ^1 subdifferential is locally single-valued on S (a constant sign vector), and the optimality system reduces on S. Let $P_S : \mathbb{R}^n \to \mathbb{R}^{|S|}$ denote the canonical coordinate projection, $P_S x = x_S$. Since $(y_\lambda(x))_{I_0} = 0$ on U, we define the reduced smooth function

$$g_S : \mathbb{R}^{|S|} \to \mathbb{R},\ \xi \mapsto g_S(\xi) := g(\xi, 0_{I_0}).$$

Then the reduced optimality equation on S reads

$$\nabla g_S(y_{\lambda,S}(x)) + \tfrac{1}{\lambda}\big(y_{\lambda,S}(x) - x_S\big) + \alpha\,\sigma_S = 0.$$

Its Jacobian with respect to the variable $y_{\lambda,S}$ is

$$\nabla^2 g_S(y_{\lambda,S}(x)) + \lambda^{-1}\mathrm{I}_{|S|} = \nabla^2_{SS} g(y_\lambda(x)) + \lambda^{-1}\mathrm{I}_{|S|} \succeq \sigma\mathrm{I}_{|S|},$$

hence invertible. By the implicit function theorem, the map $x \mapsto y_{\lambda,S}(x)$ is C^1 on U. Since $(y_\lambda(x))_i = 0$ for all $i \in I_0$ and all $x \in U$, it follows that $x \mapsto y_\lambda(x)$ is C^1 on U. On U, the standard envelope identities give

$$\nabla e_\lambda f(x) = \tfrac{1}{\lambda}\big(x - y_\lambda(x)\big) \ \text{ and } \ \nabla^2 e_\lambda f(x) = \tfrac{1}{\lambda}\big(I - Dy_\lambda(x)\big),$$

so $e_\lambda f \in C^2(U)$ and

$$q(w) = d^2_*[e_\lambda f]\big(x_\lambda \,|\, \bar v\big)(w) = \langle \nabla^2 e_\lambda f(x_\lambda) w, w\rangle \qquad (\|w\| = 1).$$

Moreover, since $\nabla^2 g$ is locally Lipschitz on bounded sets and $y_\lambda(\cdot)$ is Lipschitz on U, the map $x \mapsto \nabla^2 g_S(y_{\lambda,S}(x))$ is locally Lipschitz on U,and differentiating the reduced equation on S gives

$$\big(\mathrm{I}_{|S|} + \lambda\nabla^2 g_S(y_{\lambda,S}(x))\big) Dy_{\lambda,S}(x) = P_S.$$

Hence,
$$Dy_{\lambda,S}(x) = \big(\mathrm{I}_{|S|} + \lambda\nabla^2 g_S(y_{\lambda,S}(x))\big)^{-1} P_S.$$
Since $\nabla^2 g_S(y_{\lambda,S}(x)) = \nabla^2_{SS} g(y_\lambda(x))$ on U, we equivalently have
$$Dy_{\lambda,S}(x) = \big(\mathrm{I}_{|S|} + \lambda\nabla^2_{SS} g(y_\lambda(x))\big)^{-1} P_S.$$
Therefore, $Dy_\lambda(\cdot)$ is locally Lipschitz on U and $e_\lambda f \in C^{2,1}(U)$. Since U is open and $\mathcal{W}$ is finite, we may choose $t > 0$ so that $x_\lambda + sw \in U$ for all $w \in \mathcal{W}$ and all $|s| \le t$. Set $R := \max_{w\in\mathcal{W}} \max_{|s|\le t} \|B\, y_\lambda(x_\lambda + sw)\|$. If $\|z\|, \|z'\| \le R$, then $\|\nabla^2\phi(z) - \nabla^2\phi(z')\| \le 6R\|z - z'\|$. We can then deduce that
$$\|\nabla^2 g(u) - \nabla^2 g(u')\| \le 6R\,\|B\|^3\,\|u - u'\| \qquad \text{whenever} \quad \|Bu\|, \|Bu'\| \le R.$$
On U, differentiating the reduced optimality equation on S gives
$$Dy_{\lambda,S}(x) = \big(\mathrm{I}_{|S|} + \lambda H_S(x)\big)^{-1} P_S$$
with $H_S(x) = \nabla^2_{SS} g(y_\lambda(x))$ and
$$\big\|\big(\mathrm{I}_{|S|} + \lambda H_S(x)\big)^{-1}\big\| \le (1 - \lambda\mu\|B\|^2)^{-1}.$$
Using $A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}$, the above Lipschitz bound on $\nabla^2 g$, and the Lipschitz bound on $y_\lambda(\cdot)$, one obtains on each sampled segment $\{x_\lambda + sw : \ |s| \le t\}$,
$$\|\nabla^2 e_\lambda f(x) - \nabla^2 e_\lambda f(x')\| \le L_e\,\|x - x'\|, \qquad L_e := \frac{6R\,\|B\|^3}{(1 - \lambda\mu\|B\|^2)^3}.$$
Fix $w \in \mathcal{W}$ and define $\psi(s) := \langle w, \nabla e_\lambda f(x_\lambda + sw)\rangle$ for $|s| \le t$. Then $\psi'(s) = \langle w, \nabla^2 e_\lambda f(x_\lambda + sw)w\rangle$ and $|\psi'(s) - \psi'(0)| \le L_e|s|$. Since $\widehat{c}_\lambda(w) = \frac{1}{2t}\big(\psi(t) - \psi(-t)\big)$, we get
$$|\widehat{c}_\lambda(w) - \psi'(0)| = \frac{1}{2t}\left|\int_{-t}^{t} (\psi'(s) - \psi'(0))\,ds\right| \le \frac{1}{2t}\int_{-t}^{t} L_e|s|\,ds = \frac{L_e}{2}\,t.$$
As $\psi'(0) = \langle w, \nabla^2 e_\lambda f(x_\lambda)w\rangle = q(w)$, this gives exactly (5.11) after taking $\max_{w\in\mathcal{W}}$.

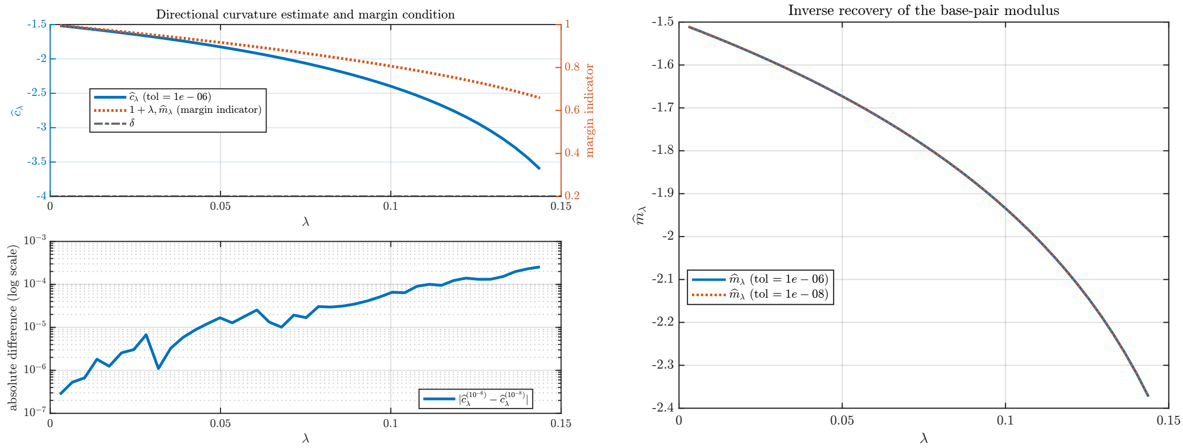
For the inexact oracle, if $\|y^\varepsilon_\lambda(x) - y_\lambda(x)\| \le \varepsilon_{\mathrm{prox}}$ at $x = x_\lambda \pm tw$, then define $g^\varepsilon_\lambda(x) := \frac{1}{\lambda}\big(x - y^\varepsilon_\lambda(x)\big)$. Then
$$\|g^\varepsilon_\lambda(x) - \nabla e_\lambda f(x)\| = \tfrac{1}{\lambda}\|y^\varepsilon_\lambda(x) - y_\lambda(x)\| \le \tfrac{\varepsilon_{\mathrm{prox}}}{\lambda},$$
and, writing the exact and inexact secant proxies at the two points $x_\lambda \pm tw$, we get
$$\begin{aligned}|\widehat{c}^\varepsilon_\lambda(w) - \widehat{c}_\lambda(w)| &\le \frac{1}{2t}\Big(\|g^\varepsilon_\lambda(x_\lambda + tw) - \nabla e_\lambda f(x_\lambda + tw)\| + \|g^\varepsilon_\lambda(x_\lambda - tw) - \nabla e_\lambda f(x_\lambda - tw)\|\Big)\\ &\le \frac{1}{2t}\left(\frac{\varepsilon_{\mathrm{prox}}}{\lambda} + \frac{\varepsilon_{\mathrm{prox}}}{\lambda}\right) = \frac{\varepsilon_{\mathrm{prox}}}{\lambda t}.\end{aligned}$$
Combining with (5.11) gives (5.12). ∎

Remark 5.5. The bound (5.12) highlights a classical trade-off. The truncation term $L_e t/2$ decreases, while the inexactness term $\varepsilon_{\mathrm{prox}}/(\lambda t)$ increases when t is small. By minimizing the right hand side in (5.12) for $t > 0$, we obtain the optimal choice
$$t^\star = \sqrt{\frac{2\,\varepsilon_{\mathrm{prox}}}{\lambda\, L_e}}, \text{ and the corresponding error is of order } \sqrt{\frac{2L_e\,\varepsilon_{\mathrm{prox}}}{\lambda}}.$$
Decreasing t is meaningful only if we also reduce $\varepsilon_{\mathrm{prox}}$ at the same time. The numerical experiment on the example (5.9)-(5.10), shows that the strict complementarity condition in Proposition 5.4 is satisfied uniformly on the sampled λ-grid with $|I_0| = 90$ et $|S| = 210$, and the margin $\widehat{\kappa}_\lambda$ and $\widehat{\eta}_\lambda$ are positive. This strict complementarity assumption could be replaced by other assumptions (polyhedral error bounds of Hoffman type, or metric regularity), in order to obtain stability of the active set. For the double-well-ℓ^1 example, it would be interesting to treat the general $C^{1,1}$ regime by replacing the classical Hessian with appropriate nonsmooth second-order objects.

Figure 5(a) displays the resulting sequence $(\lambda_k)_k$ and the margin indicator, while Figure 5(b) reports the average inner iteration count as a function of λ (for the two inner tolerances). This illustrates the balance between conditioning, controlled by (5.7), and inner cost as λ increases.

6. Conclusion.

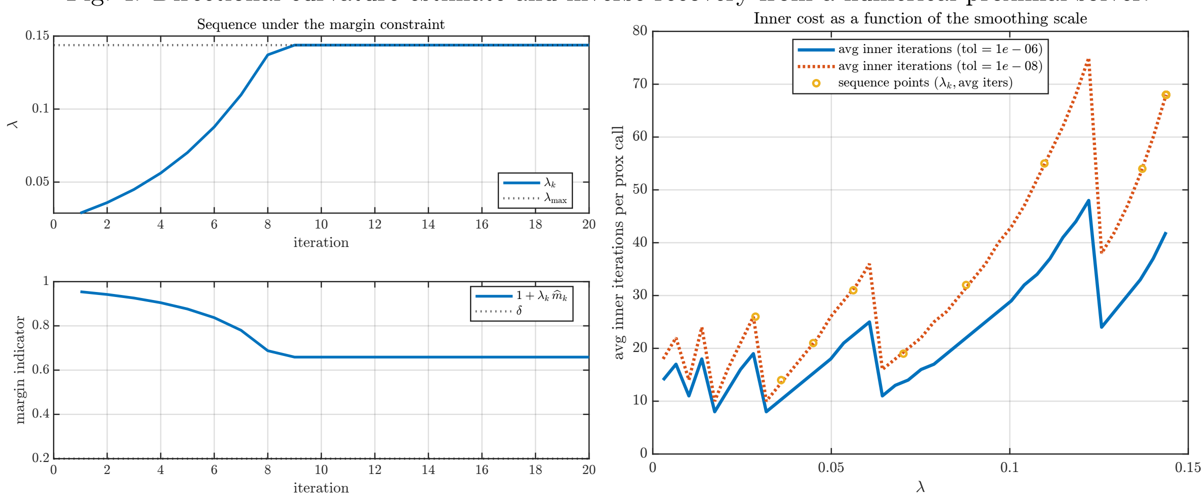
The results obtained in this paper clarify the second-order role of Moreau envelopes in the variational geometry of epigraphs. Beyond smoothing, the passage to an envelope carries out a precisely quantified transformation of local curvature. At the level of strict second subderivatives this is captured by a commutation formula, while at the scalar level it is manifested in an explicit attenuation-recovery law for the variational modulus of convexity. A direct means of transporting and



(a) $\widehat{c}_\lambda$ and margin indicator

(b) Recovered $\widehat{m}_\lambda$ (stability)

Fig. 4: Directional curvature estimate and inverse recovery from a numerical proximal solver.



(a) Margin-constrained sequence $(\lambda_k)_k$ and margin indicator

(b) Inner proximal cost vs. λ (two tolerances)

Fig. 5: Margin-constrained selection of the smoothing scale and computational cost.

recovering curvature data under Moreau regularization has thereby been provided, including under smooth C^2 perturbations.

From a practical standpoint, the attenuation-recovery law also suggests a quantitative rule for selecting the smoothing scale in prox-based nonconvex contexts. Section 5 illustrates this idea in an oracle-driven setting, where curvature information is inferred from proximal evaluations through the envelope gradient and used to enforce an admissible choice of λ.

Several analytical questions remain open. In one directions, there is the potential for extensions of the spectral framework beyond the Euclidean Moreau envelope, for instance to Bregman-type regularizations or other proximal constructions adapted to nonquadratic geometries. In another direction there is perhaps the incorporation of these curvature identities into the analysis of proximal algorithms, where the attenuated modulus naturally enters stability and rate estimates. A still further possibility is viewing the envelope as a numerical probe of local curvature in nonsmooth and nonconvex problems. In that context, the numerical evaluation of the Moreau envelope and of its gradient in prox-regular, nonconvex settings needs a better understanding beyond the local chart, so as to encompass the effect of inexact proximal solvers and finite-difference approximations on curvature estimates. All these are promising topics for follow-up research.

Appendix A. A self-contained proof of the modulus formula.

Throughout this appendix, let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be proper and lsc, and fix $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ with $f(\bar{x}) \in \mathbb{R}$. The objects $\Delta_\tau^2 f(x \,|\, v)$ and $d_*^2 f(\bar{x} \,|\, \bar{v})$ are defined in Subsection 2.2. We give here an alternative proof of the spectral identity of [18, Thm 2.2]. We use a direct approach combined with the compactness of the unit sphere in $\mathbb{R}^n$.

In [18, Def 1.2], the variational modulus $\operatorname{cnv} f(\bar{x} \,|\, \bar{v})$ is defined as a common value associated with several equivalent formulations (1.12)-(1.13)-(1.14) and (1.15) (with opposite sign). The equivalences are guaranteed by [18, Thm 1.1]. In order to lighten the presentation, we adopt here the formulation (1.14) based on uniform quadratic growth. The next definition is in this sense.

Definition A.1 (variational modulus of convexity at $(\bar{x},\bar{v})$). *A number $s \in \mathbb{R}$ is said to be* admissible *at $(\bar{x},\bar{v}) \in \operatorname{gph}\partial f$ if there exist $\alpha > f(\bar{x})$, neighborhoods X of $\bar{x}$ and V of $\bar{v}$ such that*

$$f(x') \ \geq \ f(x) + \langle v, x'-x\rangle + \frac{s}{2}\|x'-x\|^2 \tag{A.1}$$

for all $(x,v) \in (X_f^\alpha \times V) \cap \operatorname{gph}\partial f$ and all $x' \in X$, where $X_f^\alpha := \{x \in X : \ f(x) < \alpha\}$. The variational modulus of convexity *of f at $(\bar{x},\bar{v})$ is*

$$\operatorname{cnv} f(\bar{x}\,|\,\bar{v}) := \sup\{\, s \in \mathbb{R} : \ s \text{ is admissible at } (\bar{x},\bar{v})\,\},$$

with the convention $\sup\emptyset = -\infty$.

We note that the s-admissibility does not affect (A.1) if we add the condition $\|x'-x\| \leq \beta$ for some $\beta > 0$. It is implied by the condition $\forall x' \in X$, and conversely, for $\|x'-x\| \leq \beta$, we shrink X (e.g. $\tilde{X} \subset X \cap \mathbb{B}(\bar{x},\beta/2)$ such that $\|x'-x\| \leq \beta$ is satisfied for every $x, x' \in \tilde{X}$, which gives $\forall x' \in \tilde{X}$). This reparametrization of the localization by a radius $\beta \downarrow 0$ was used in (2.6) and (2.7) in [18].

Theorem A.2 (Modulus formula from strict second subderivatives). *One has*

$$\operatorname{cnv} f(\bar{x}\,|\,\bar{v}) = \inf_{\|w\|=1} d_*^2 f(\bar{x}\,|\,\bar{v})(w).$$

Proof. Let us set

$$m := \inf_{\|w\|=1} d_*^2 f(\bar{x} \mid \bar{v})(w) \in [-\infty, +\infty].$$

We first show that $m \geq \operatorname{cnv} f(\bar{x} \mid \bar{v})$. Consider an arbitrary admissible value $s \in \mathbb{R}$ at $(\bar{x},\bar{v})$, so that (A.1) holds on some $(X_f^\alpha \times V) \cap \operatorname{gph}\partial f$, for neighborhoods X and V that we may assume open without loss of generality. Fix $w \in \mathbb{R}^n$ such that $\|w\| = 1$ and consider arbitrary sequences

$$\xi_k \to w, \quad \tau_k \downarrow 0, \quad (x_k, v_k) \to (\bar{x},\bar{v}), \quad v_k \in \partial f(x_k), \text{ and } f(x_k) \to f(\bar{x}).$$

For k sufficiently large, we have $(x_k, v_k) \in (X_f^\alpha \times V) \cap \operatorname{gph}\partial f$ and $x_k + \tau_k \xi_k \in X$. Applying (A.1) with $x' = x_k + \tau_k\xi_k$, we get

$$\Delta_{\tau_k}^2 f(x_k \mid v_k)(\xi_k) \geq s\,\|\xi_k\|^2.$$

Taking the lim inf as $k \to +\infty$ in the last inequality, we obtain $d_*^2 f(\bar{x} \mid \bar{v})(w) \geq s$. Which implies that $m \geq s$. Since s was an arbitrary admissible value, we conclude

$$m \geq \operatorname{cnv} f(\bar{x} \mid \bar{v}). \tag{A.2}$$

Let us show now that $m \leq \operatorname{cnv} f(\bar{x} \mid \bar{v})$. The case $m = -\infty$ being obvious, we assume $m \in \mathbb{R} \cup \{+\infty\}$ and fix $s < m$. We prove that s is admissible at $(\bar{x},\bar{v})$. Suppose by contradiction that this last claim is not true. Then, for every α, X, V as in Definition A.1, there exist $(x,v) \in (X_f^\alpha \times V) \cap \operatorname{gph}\partial f$ and $x' \in X$ violating (A.1). Since (A.1) is an equality when $x' = x$, any violating point must satisfy $x' \neq x$. Hence $\tau = \|x'-x\| > 0$, and let us define $w = \frac{1}{\tau}(x'-x)$. It is clear that $x' = x + \tau w$, $\|w\| = 1$ and

$$\Delta_\tau^2 f(x \mid v)(w) < s.$$

Let us choose $X_k = \mathbb{B}(\bar{x}, 1/k)$, $V_k = \mathbb{B}(\bar{v}, 1/k)$, and $\alpha_k = f(\bar{x}) + 1/k$. This gives a sequence $(x_k, v_k) \in ((X_k)_f^{\alpha_k} \times V_k) \cap \operatorname{gph}\partial f$ and $x_k' \in X_k$ such that, setting $\tau_k := \|x_k' - x_k\|$ and $w_k := (x_k' - x_k)/\tau_k$,

$$\Delta_{\tau_k}^2 f(x_k \mid v_k)(w_k) < s, \qquad \|w_k\| = 1.$$

By construction, we have $\tau_k \to 0$, $(x_k, v_k) \to (\bar{x},\bar{v})$, and $f(x_k) \to f(\bar{x})$ (since f is lsc and $x_k \in (X_k)_f^{\alpha_k}$). Passing to a further subsequence if necessary, we may assume that $\tau_k \downarrow 0$. By compactness of the unit sphere in $\mathbb{R}^n$, along a subsequence still denoted by w_k, with $\|w_k\| = 1$, we have $w_k \to w$. Hence, the quadriplet (x_k, v_k, τ_k, w_k) is strict-admissible for $d_*^2 f(\bar{x}\,|\,\bar{v})(w)$. Consequently,

$$d_*^2 f(\bar{x} \mid \bar{v})(w) \leq \liminf_{k\to\infty} \Delta_{\tau_k}^2 f(x_k \mid v_k)(w_k) \leq s,$$

which gives $m \leq s$, contradicting $s < m$. Hence every $s < m$ is admissible at $(\bar{x},\bar{v})$, and so $\operatorname{cnv} f(\bar{x}\,|\,\bar{v}) \geq m$. Combining this last inequality with (A.2), we conclude that $\operatorname{cnv} f(\bar{x} \mid \bar{v}) = m$. ∎

Appendix B. Proof of Lemma 3.1.

Proof. We consider the tilt function $x \mapsto g(x) := f(x) - \langle \bar{v}, x - \bar{x}\rangle$. Then g is proper, lsc, and prox-bounded. In this case, we have $\partial g(x) = \partial f(x) - \bar{v}$, which means that $(\bar{x}, 0) \in \operatorname{gph}\partial g$. It is clear that g is prox-regular at $\bar{x}$ for 0. By [19, Prop. 13.37], there exists $\lambda_{\mathrm{RW}} > 0$ such that, for every $\lambda \in (0, \lambda_{\mathrm{RW}})$, there exist an g-attentive localization (in the sense of [19, page 611]) S of ∂g around $(\bar{x}, 0)$ and a neighborhood $U_\lambda^g = \operatorname{rge}(\mathrm{I} + \lambda S)$ of $\bar{x}$ on which $P_\lambda g$ is single-valued and Lipschitz, $e_\lambda g$ is of class $C^{1,1}$, and

$$\nabla e_\lambda g = \lambda^{-1}\big[\mathrm{I} - P_\lambda g\big] = [\lambda \mathrm{I} + S^{-1}]^{-1} \text{ on } U_\lambda^g.$$

Fix $\lambda \in (0, \lambda_{\mathrm{RW}})$, and consider any $z, u \in \mathbb{R}^n$,

$$f(u) + \frac{1}{2\lambda}\|u - z\|^2 = g(u) + \langle \bar{v}, u - \bar{x}\rangle + \frac{1}{2\lambda}\|u - z\|^2$$

$$= g(u) + \frac{1}{2\lambda}\|u - (z - \lambda \bar{v})\|^2 + \langle \bar{v}, z - \bar{x}\rangle - \frac{\lambda}{2}\|\bar{v}\|^2. \tag{B.1}$$

By taking $\arg\min$ and $\inf$ in u in (B.1), we obtain for every z,

$$P_\lambda f(z) = P_\lambda g(z - \lambda \bar{v}) \text{ and } e_\lambda f(z) = e_\lambda g(z - \lambda \bar{v}) + \langle \bar{v}, z - \bar{x}\rangle - \frac{\lambda}{2}\|\bar{v}\|^2. \tag{B.2}$$

By setting $V_\lambda := U_\lambda^g + \lambda \bar{v}$, it is clear that V_λ is a neighborhood of $x_\lambda := \bar{x} + \lambda \bar{v}$ and $z \in V_\lambda$ implies $z - \lambda \bar{v} \in U_\lambda^g$.

(a) Let $z \in V_\lambda$. We have $z - \lambda \bar{v} \in U_\lambda^g$, so $P_\lambda g$ is single-valued and Lipschitz at $z - \lambda \bar{v}$. By (B.2), $P_\lambda f(z) = P_\lambda g(z - \lambda \bar{v})$. Which implies that, $P_\lambda f$ is single-valued and Lipschitz on V_λ, and $P_\lambda f(x_\lambda) = P_\lambda g(\bar{x}) = \bar{x}$.

(b) Let $z \in V_\lambda$. By the second part of (B.2), $e_\lambda f$ is of class $C^{1,1}$ on V_λ and

$$\nabla e_\lambda f(z) = \nabla e_\lambda g(z - \lambda \bar{v}) + \bar{v}.$$

Using the fact that $\nabla e_\lambda g = \lambda^{-1}(\mathrm{I} - P_\lambda g)$ on U_λ^g, we obtain

$$\nabla e_\lambda f(z) = \lambda^{-1}\big[(z - \lambda \bar{v}) - P_\lambda f(z)\big] + \bar{v} = \lambda^{-1}\big[z - P_\lambda f(z)\big] = \lambda^{-1}\big[\mathrm{I} - P_\lambda f\big](z).$$

For $z = x_\lambda$, we get $\nabla e_\lambda f(x_\lambda) = \lambda^{-1}\big[x_\lambda - P_\lambda f(x_\lambda)\big] = \bar{v}$.

(c) By [19, Prop. 13.37] applied to g at $(\bar{x}, 0)$ and [19, (13.30)], there exist $\lambda_g > 0$ and a g-attentive localization S of ∂g around $(\bar{x}, 0)$ such that

$$g(x') > g(x) + \langle w, x' - x\rangle - \frac{1}{2\lambda_g}\|x' - x\|^2, \quad \text{for all } x' \neq x \tag{B.3}$$

holds for every $(x, w) \in \operatorname{gph} S$. Since g is lsc at $\bar{x}$, the sets

$$\mathcal{N}_\eta := \{(x, w) : \ \|x - \bar{x}\| < \eta, \ \|w\| < \eta, \ g(x) < g(\bar{x}) + \eta\}, \ (\eta > 0)$$

form a neighborhood basis of $(\bar{x}, 0)$ for the product topology where w carries the usual topology and x carries the g-attentive topology (see [19, p. 611]). As S is a g-attentive localization, there exists a neighborhood $\mathcal{N}$ of $(\bar{x}, 0)$ in this topology such that $\operatorname{gph} S = \operatorname{gph} \partial g \cap \mathcal{N}$. Hence, there exists $\varepsilon_g > 0$ with $\mathcal{N}_{\varepsilon_g} \subset \mathcal{N}$, and then (B.3) holds for every $(x, w) \in \operatorname{gph} \partial g$ satisfying

$$\|x - \bar{x}\| < \varepsilon_g, \qquad \|w\| < \varepsilon_g, \qquad g(x) < g(\bar{x}) + \varepsilon_g.$$

Fix $\varepsilon > 0$ such that

$$\varepsilon \leq \varepsilon_g, \qquad (1 + \|\bar{v}\|)\varepsilon \leq \varepsilon_g, \tag{B.4}$$

and let T_ε be the canonical f-attentive ε-localization around $(\bar{x}, \bar{v})$ as defined in (2.2). Let $\lambda_{\mathrm{pb}} > 0$ be a prox-boundedness threshold for f as in [19, Ex. 10.32], and let $\lambda_{\mathrm{RW}} > 0$ be the constant already obtained in parts **(a)-(b)** (from [19, Prop. 13.37]). We define $\lambda_0 > 0$ such that

$$\lambda_0 := \min\{\lambda_{\mathrm{RW}}, \lambda_g, \lambda_{\mathrm{pb}}\}, \ \lambda \in (0, \lambda_0) \text{ and } U_\lambda := \operatorname{rge}(\mathrm{I} + \lambda T_\varepsilon).$$

Let $u \in U_\lambda$, which implies that $u = x + \lambda v$ for some $(x, v) \in \operatorname{gph} T_\varepsilon$. Let's set $w := v - \bar{v}$. Since $\partial g = \partial f - \bar{v}$, $(x, w) \in \operatorname{gph} \partial g$. On the other hand, $\|x - \bar{x}\| < \varepsilon \leq \varepsilon_g$ and $\|w\| < \varepsilon \leq \varepsilon_g$, and by (B.4),

$$g(x) - g(\bar{x}) = f(x) - f(\bar{x}) - \langle \bar{v}, x - \bar{x}\rangle < f(x) - f(\bar{x}) + \|\bar{v}\| \, \|x - \bar{x}\| < \varepsilon + \|\bar{v}\|\varepsilon \leq \varepsilon_g.$$

Consequently, (B.3) applies at (x, w) and gives, after substituting $g = f - \langle \bar{v}, \cdot - \bar{x}\rangle$,

$$f(x') > f(x) + \langle v, x' - x\rangle - \frac{1}{2\lambda_g}\|x' - x\|^2, \ \forall x' \neq x.$$

Since $\lambda < \lambda_g$, we have

$$f(x') > f(x) + \langle v, x' - x\rangle - \frac{1}{2\lambda}\|x' - x\|^2, \ \forall x' \neq x.$$

Since $x' - u = (x' - x) - \lambda v$, we have for every $x' \neq x$,

$$f(x') + \frac{1}{2\lambda}\|x' - u\|^2 > f(x) + \frac{1}{2\lambda}\|x - u\|^2.$$

We deduce that x is the unique minimizer of $y \mapsto f(y) + \frac{1}{2\lambda}\|y-u\|^2$, so

(B.5) $$P_\lambda f(u) = \{x\}, \; P_\lambda f(u) = (\mathrm{I} + \lambda T_\varepsilon)^{-1}(u) \text{ and } (u \in U_\lambda).$$

Since $(\bar{x}, \bar{v}) \in \operatorname{gph} T_\varepsilon$, we have $x_\lambda = \bar{x} + \lambda\bar{v} \in U_\lambda$ and $P_\lambda f(x_\lambda) = \bar{x}$.

We check now that that U_λ is a neighborhood of x_λ. Because $\lambda < \lambda_{\mathrm{RW}}$, parts **(a)**-**(b)** provide a neighborhood V_λ of x_λ on which $P_\lambda f$ is single-valued and Lipschitz and $e_\lambda f \in C^{1,1}(V_\lambda)$, with

$$\nabla e_\lambda f(u) = \lambda^{-1}\big(u - P_\lambda f(u)\big), \; u \in V_\lambda.$$

Let us define $x(u) := P_\lambda f(u)$ and $v(u) := \lambda^{-1}(u - x(u))$ on V_λ. Then $u \mapsto x(u)$ and $u \mapsto v(u)$ are continuous on V_λ, and

(B.6) $$e_\lambda f(u) = f(x(u)) + \frac{1}{2\lambda}\|x(u) - u\|^2, \; u \in V_\lambda.$$

From (B.6), we deduce that $u \mapsto f(x(u))$ is continuous on V_λ. At $u = x_\lambda$, we have $x(x_\lambda) = \bar{x}$ and $v(x_\lambda) = \bar{v}$. There exists $\rho_\lambda > 0$ such that $\mathbb{B}(x_\lambda, \rho_\lambda) \subset V_\lambda$ and, for every $u \in \mathbb{B}(x_\lambda, \rho_\lambda)$,

$$\|x(u) - \bar{x}\| < \varepsilon, \; \|v(u) - \bar{v}\| < \varepsilon, \; f(x(u)) < f(\bar{x}) + \varepsilon.$$

By the definitions of $x(u)$ and $v(u)$ above, we have $v(u) \in \partial f(x(u))$. Hence, $(x(u), v(u)) \in \operatorname{gph} T_\varepsilon$ and $u = x(u) + \lambda v(u) \in U_\lambda$. Thus $\mathbb{B}(x_\lambda, \rho_\lambda) \subset U_\lambda$, and U_λ is a neighborhood of x_λ.
Since $\lambda < \lambda_{\mathrm{pb}}$, applying [19, Ex. 10.32], we deduce that $-e_\lambda f$ is strictly continuous and

$$\partial[-e_\lambda f](u) = \lambda^{-1}\big(\operatorname{con} P_\lambda f(u) - u\big).$$

On U_λ, $P_\lambda f(u)$ is a singleton by (B.5), hence $\partial[-e_\lambda f](u)$ (and so $\partial[e_\lambda f](u)$) is a singleton. By [19, Thm. 9.18], $e_\lambda f$ is differentiable on U_λ with

$$\nabla e_\lambda f(u) = \lambda^{-1}\big(u - P_\lambda f(u)\big), \; u \in U_\lambda.$$

For $u \in U_\lambda$, set $x := P_\lambda f(u)$ and $v := \nabla e_\lambda f(u) = \lambda^{-1}(u - x)$. Then $(x, v) \in \operatorname{gph} T_\varepsilon$ and $u = x + \lambda v$, so $u \in (\lambda\mathrm{I} + T_\varepsilon^{-1})(v)$ and hence $v \in (\lambda\mathrm{I} + T_\varepsilon^{-1})^{-1}(u)$. Therefore,

$$\nabla e_\lambda f(u) = \big(\lambda\mathrm{I} + T_\varepsilon^{-1}\big)^{-1}(u), \qquad u \in U_\lambda,$$

together with (B.5).

In this last part, we prove the Lipschitz and $C^{1,1}$ properties of e_λ on U_λ. For $i = 1, 2$, let $u_i \in U_\lambda$ and set $x_i := P_\lambda f(u_i)$ and $v_i := \lambda^{-1}(u_i - x_i)$, so that $(x_i, v_i) \in \operatorname{gph} T_\varepsilon$ and $u_i = x_i + \lambda v_i$. By applying the transported form of (B.3) at (x_1, v_1) with $x' = x_2$ and at (x_2, v_2) with $x' = x_1$, and adding the inequalities, we obtain $\langle v_1 - v_2, x_1 - x_2\rangle \geq -\frac{1}{\lambda_g}\|x_1 - x_2\|^2$. Therefore,

$$\langle u_1 - u_2, x_1 - x_2\rangle = \|x_1 - x_2\|^2 + \lambda\langle v_1 - v_2, x_1 - x_2\rangle \geq \Big(1 - \frac{\lambda}{\lambda_g}\Big)\|x_1 - x_2\|^2,$$

and Cauchy-Schwarz implies

$$\|x_1 - x_2\| \leq \frac{1}{1 - \lambda/\lambda_g}\,\|u_1 - u_2\|.$$

Consequently, $P_\lambda f$ is Lipschitz on U_λ, and since $\nabla e_\lambda f = \lambda^{-1}(\mathrm{I} - P_\lambda f)$ on U_λ, $\nabla e_\lambda f$ is also Lipschitz on U_λ, so $e_\lambda f \in C^{1,1}(U_\lambda)$, which what we needed to complete the proof of Lemma 3.1. ∎

Remark B.1. An alternative way to justify the continuity step in the proof of Lemma 3.1**(c)** is to use the localized subdifferential continuity property of prox-regular functions. In fact, Proposition 2.3 in [12] ensures that, on the f-attentively truncated graph of ∂f around $(\bar{x}, \bar{v})$, convergence $(x_k, v_k) \to (x, v)$ with $v_k \in \partial f(x_k)$ and $f(x_k)$ uniformly bounded above near $f(\bar{x})$ forces $f(x_k) \to f(x)$. In our setting this can be applied to the pairs $(x(u), v(u))$ arising from the proximal relation $u = x(u) + \lambda v(u)$, with $v(u) = \lambda^{-1}(u - x(u))$, to obtain $f(x(u)) \to f(\bar{x})$ as $u \to x_\lambda$. We preferred the argument based on the identity $e_\lambda f(u) = f(x(u)) + \frac{1}{2\lambda}\|x(u) - u\|^2$, since it gives the same conclusion directly from the local single-valuedness of $P_\lambda f$ and the $C^{1,1}$ regularity of $e_\lambda f$ near x_λ.

Acknowledgements. The authors warmly thank the two anonymous referees for their careful reading of the manuscript, and their constructive comments and suggestions, which helped us improve the presentation and clarity of the paper.
The first author gratefully acknowledges support from the Math AmSud project N°51756TF (VIPS), ECOS Project C24E06 and the FMJH Gaspard Monge Program for Optimization and Data Science.

REFERENCES

[1] H. Gfrerer. *On second-order variational analysis of variational convexity of prox-regular functions. Set-Valued and Variational Analysis*, forthcoming. Available at arXiv:2408.13795.

[2] A. Jourani, L. Thibault, and D. Zagrodny. *Differential properties of the Moreau envelope.* J. Funct. Anal. **266** (2014), 1185-1237.

[3] P. D. Khanh, B. S. Mordukhovich, V. T. Phat. *Variational convexity of functions and variational sufficiency in optimization.* SIAM J. Optim. **33** (2023), 1121-1158.

[4] P. D. Khanh, B. S. Mordukhovich, V. T. Phat, L. D. Viet. *Generalized twice differentiability and quadratic bundles in second-order variational analysis.* Preprint (2024), arXiv:2501.02067.

[5] A. S. Lewis. *Convex analysis on the Hermitian matrices.* SIAM J. Optim. **6** (1996), no. 1, 164-177.

[6] A. S. Lewis. *Nonsmooth analysis of eigenvalues.* Math. Program. **84** (1999), 1-24.

[7] A. S. Lewis, H. S. Sendov. *Twice differentiable spectral functions.* SIAM J. Matrix Anal. Appl. **23** (2001), no. 2, 368-386.

[8] B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation, I: Basic Theory.* Springer, Berlin, 2006.

[9] J. J. Moreau. *Fonctions convexes duales et points proximaux dans un espace hilbertien.* C. R. Acad. Sci. Paris **255** (1962), 2897-2899.

[10] J. J. Moreau. *Proximité et dualité dans un espace hilbertien.* Bull. Soc. Math. France **93** (1965), 273-299.

[11] J. Nocedal and S. J. Wright. *Numerical Optimization.* 2nd ed., Springer, 2006.

[12] R. A. Poliquin, R. T. Rockafellar. *Prox-regular functions in variational analysis.* Trans. Amer. Math. Soc. **348** (1996), 1805-1838.

[13] R. A. Poliquin, R. T. Rockafellar. *Tilt stability of a local minimum.* SIAM J. Optim. **8** (1998), 287-299.

[14] R. A. Poliquin and R. T. Rockafellar. *Generalized Hessian properties of regularized nonsmooth functions.* SIAM J. Optim. **6**, No. 4, (1996), 1121-1137.

[15] R. A. Poliquin, R. T. Rockafellar, L. Thibault. *Local differentiability of distance functions.* Trans. Amer. Math. Soc. **352** (2000), 5231-5249.

[16] R. T. Rockafellar. *Variational convexity and the local monotonicity of subgradient mappings.* Vietnam J. Math. **47** (2019), 547-561.

[17] R. T. Rockafellar. *Variational convexity and prox-regularity.* J. Convex Anal., **33** (2026), 1049–1060.

[18] R. T. Rockafellar. *Derivative Tests for Prox-Regularity and the Modulus of Convexity.* J. Convex Anal., **33** (2026), 243–266.

[19] R. T. Rockafellar, R. J.-B. Wets. *Variational Analysis.* Springer, Berlin, 1998.

[20] A. Themelis, L. Stella, and P. Patrinos. *Forward-backward envelope for the sum of two nonconvex functions: Further properties and nonmonotone linesearch algorithms.* SIAM J. Optim. **28** (2018), 2274-2303.