The Discrete Legendre-Fenchel Transform and its application to phase separation in electrolytes

Lorenzo Contento

The Discrete Legendre–Fenchel Transform and its application to phase separation in electrolytes

Lorenzo Contento

11 December 2012

Dipartimento di Matematica e Informatica, Università degli Studi di Udine

Contents

Introduction

The study of equilibrium electrolytes in a porous and electrically charged medium is of interest in the design of semiconductor devices or the study of clay rocks at the nanometre scale. The bulk free-energy f of these systems, which is a function of the ionic concentrations, contains a non-convex contribution due to the electrostatic interactions between the ions; it has been shown that for small ion diameters this contribution is strong enough to make the function f non-convex. In such cases the region $\mathcal{T}$ of the state space where f differs from its convex hull is thermodynamically unstable and $\mathcal{T}$ divides the state space in different phases. For a one-dimensional state space (only one ionic species), the convex hull of the bulk free-energy f can be determined using Maxwell equal area rule. For more than one ionic species, we need to determine the convex hull of a multidimensional function f. In the thesis we study the Legendre-Fenchel transform and we propose an algorithm to approximate the convex hull, which is based on its discrete version. This approach is known in the literature, but here we develop and test an improved version of it. We apply this algorithm to the bulk free energy in the symmetric and asymmetric case, finding the shape of the region $\mathcal{T}$. Finally we solve the model equations by a finite elements discretization, obtaining solutions in which phase separation clearly arises.

The thesis is organized as follows.

In Chapter 1 we study a non-confined electrolyte. We describe its free-energy, called *bulk free-energy*; in the non-ideal case, i.e. when interactions between ions are not negligible, such free-energy contains terms due to coulombian interaction and steric effects. Finally, we observe that if the free-energy is not convex, phase separation may occur.

In Chapter 2 we introduce and prove some basic results on the Legendre-Fenchel transform and its discrete counterpart with particular attention to the construction of the convex hull; more details can be found in Appendices A and B. Convergence properties are assessed and an exact formula is given for the one-dimensional case.

In Chapter 3 we study the properties of a double application of the discrete Legendre-Fenchel transform, which furnishes an approximation of the convex hull. Attention is payed to the choice of the discretization for the second transform, as it has great importance in the convergence results.

In Chapter 4 we present an algorithm for the computation of the convex hull based on the dimensional factorization of the Legendre-Fenchel transform. Some improvements to the standard approach studied in the literature are proposed and the numerical results of the algorithms are analyzed on some test problems.

In Chapter 5 we unveil the shape of the non-convexity region $\mathcal{T}$ of the bulk free-energy density for electrolytes of varying ion diameter; both symmetric and asymmetric mixtures of ions are considered.

In Chapter 6 the full model for a confined electrolyte is introduced; the total free-energy now includes a contribution of the electrostatic potential. A numerical method for the solution of the model equations is briefly presented and actual solutions computed for some test cases, showing that phase separation indeed arises in the presence of charged walls.

To conclude, the proposed algorithm allows to study any confined electrolyte, no matter what the valences and the ionic diameter are. However further improvements are needed to approximate the convex-full of the energy in case of asymmetric mixture with small ion diameters, because the iterative method used thus not converge.

1. The bulk free-energy of electrolytes

The phenomenon we are going to study is the behaviour of electrolytes in a porous and electrically charged medium; an electrolyte is a solvent, usually water, in which one or more ionic species are dissolved. This behaviour is of interest in many fields of application, such as the design of semiconductor devices or the study of clay rocks at the nanometre scale. The model used in this work is the one considered by Ern, Joubaud and Lelièvre [4].

1.1. Bulk electrolytes

Consider an electrolyte composed by a cation and an anion contained in a region $\Omega \subseteq \mathbb{R}^n$, where $n = 2, 3$, whose boundary $\partial\Omega$ is electrically charged; we will denote respectively by Z_+ and Z_- their valences, i.e. the number of extra/missing electrons, and with $c = (c_+, c_-)$ their concentrations. We will initially suppose that Ω is sufficiently large to make the influence on the electrolyte of the boundary $\partial\Omega$ negligible; this is equivalent to consider the limit case $\Omega = \mathbb{R}^n$ and an electrolyte in this condition is called a bulk electrolyte. At equilibrium, we have by symmetry that the concentration $c : \Omega \to \mathbb{R}^2_{\geq 0}$ and the electrostatic potential $\psi : \Omega \to \mathbb{R}$ are constant; for simplicity we can take $\psi = 0$, being the potential determined up to an additive constant.

At the equilibrium ψ and c must be such that $\mu_+(\psi, c)$ and $\mu_-(\psi, c)$ are both constant on Ω, where $\mu_\pm(\psi, c)$ are the electrochemical potentials of the two ions; these potentials are given by

$$\mu_i(\psi, c) = k_\mathrm{B}T \log(\sigma^3 c_i) + k_\mathrm{B}T \log(\gamma_i(c)) + Z_i e\psi, \quad i = \pm,$$

where k_B is the Boltzmann constant, e is the elementary charge, T is the temperature, σ is the mean ion diameter, which we suppose the same for both species, and $\gamma_i(c)$ is the activity coefficient of species i which is used to account for non-ideal behaviour (in the ideal case we have $\gamma_\pm(c) \equiv 1$). In the case $\Omega = \mathbb{R}^n$ we have seen that $\psi = 0$ and c is constant and thus the condition on the electrochemical potentials is satisfied for every c constant.

1.2. Non-dimensional formulation

It is convenient to have a non-dimensional formulation, for example in order to normalize numerical values and simplify the equations through the omission of physical constants. This means substituting each quantity with the dimensionless ratio between it and a

reference value and then finding the equations that this ratios satisfy; these reference values are

length	L_*
concentration c	$c_* := L_*^{-3}$
electrostatic potential ψ	$\psi_* := \frac{k_{\mathrm{B}}T}{e}$
electrochemical potential μ	$\mu_* := \mathrm{k}_B T$

where the reference length L_* can be chosen arbitrarily. We will also define the so called Bjerrum length

$$L_{\mathrm{B}} := \frac{e^2}{4\pi\varepsilon k_{\mathrm{B}}T},$$

where $\varepsilon = \varepsilon_0\varepsilon_r$ is the solvent permittivity (ε_0 is the vacuum permittivity and ε_r the solvent relative permittivity), and the dimensionless ratio

$$\lambda := \frac{L_*}{4\pi L_{\mathrm{B}}}.$$

Denoting for simplicity the new dimensionless quantities with the same symbols used before, the expression for the electrochemical potentials becomes

$$\mu_i(\psi, c) = \log(\sigma^3 c_i) + \log(\gamma_i(c)) + Z_i\psi, \quad i = \pm.$$

1.3. Non-ideality

In this section we specify the activity coefficients $\gamma_\pm(c)$. Each of them can be split into two parts as

$$\log(\gamma_i(c)) = \log(\gamma_i^{\mathrm{Coul}}(c)) + \log(\gamma^{\mathrm{Steric}}(c_+ + c_-)), \quad i = \pm;$$

the first contribution is due to Coloumb interactions, while the second accounts for steric effects, i.e. effects which appear when the ions are very close to each another. The Coulomb term is given by

$$\log(\gamma_i^{\mathrm{Coul}}(c)) = Z_i^2 \log(\gamma_0(I(c))), \quad i = \pm,$$

where $I(c)$ is the ionic strength defined as

$$\begin{aligned} \eta_i &:= \frac{1}{2}Z_i^2, \\ I(c) &:= \sum_{i=\pm} \eta_i c_i, \end{aligned}$$

while the function $\gamma_0 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is defined as

$$\begin{aligned} \Upsilon_{\mathrm{MSA}}(\theta) &:= \frac{1}{2\sigma}\left(\sqrt{2\sigma\sqrt{\frac{2\theta}{\lambda}} + 1} - 1\right), \\ \log(\gamma_0(\theta)) &:= -\frac{1}{4\pi\lambda}\frac{\Upsilon_{\mathrm{MSA}}(\theta)}{1 + \sigma\Upsilon_{\mathrm{MSA}}(\theta)}, \end{aligned}$$

where Υ_{MSA} is called the screening parameter and the acronym MSA stands for "mean spherical approximation". We present two version of the steric term. The first is called hard-sphere approximation and is given by

$$\log(\gamma^{\mathrm{HS}}(u)) := \frac{4\pi\sigma^3}{3}u.$$

We thus have that $\log(\gamma^{\mathrm{HS}}(u))$ is proportional to the fraction of the volume occupied by the ions, but it is finite for every concentration $u = c_+ + c_-$, no matter how high; this is not what we expect because it should not be possible to fit any number of ions in the same space and thus there should be a limit concentration. This problem is resolved by using the Carnahan-Starling approximation, defined as

$$\begin{aligned} r(u) &:= \frac{\pi\sigma^3}{6}u \\ \log(\gamma^{\mathrm{CS}}(u)) &:= \begin{cases} \frac{3r(u)^3-9r(u)^2+8r(u)}{(1-r(u))^3} & \text{if } 0 \le r(u) < 1 \\ +\infty & \text{if } r(u) \ge 1 \end{cases}, \end{aligned}$$

where r is the so called packing fraction. With this approximation we have that $\log(\gamma^{\mathrm{CS}}(u))$ is $+\infty$ when beyond a certain total concentration, while for small concentrations, i.e. for $r(u)$ near 0, we have that $\log(\gamma^{\mathrm{CS}}(u)) \approx 8r(u) = \log(\gamma^{\mathrm{HS}}(u))$.

1.4. Bulk free-energy density

We indicate with $f(c)$ the bulk free-energy density; f is a function $\mathbb{R}^2_{\ge 0} \to \mathbb{R}$ such that

$$\frac{\partial f}{\partial c_i}(c) = \mu_i(0, c) \text{ on } \mathbb{R}^2_{>0}, \quad i = \pm. \tag{1.1}$$

Note the f is thus defined up to an additive constant. We can decompose f as

$$f(c) = f_{\mathrm{id}}(c_+) + f_{\mathrm{id}}(c_-) + f_{\mathrm{ex}}(c),$$

where $f_{\mathrm{id}} : \mathbb{R}_{\ge 0} \to \mathbb{R}$ is the contribution due to the ideal term in $\mu_\pm$ and $f_{\mathrm{ex}} : \mathbb{R}^2_{\ge 0} \to \mathbb{R}$ (ex stands for "excess") the one resulting from non-ideality; in particular we ask that

$$\frac{\partial f_{\mathrm{id}}}{\partial u}(u) = \log(\sigma^3 u)$$

and

$$\frac{\partial f_{\mathrm{ex}}}{\partial c_i}(c) = \log(\gamma_i(c)), \quad i = \pm.$$

Then we have that also f_{id} and f_{ex} are defined up to an additive constant.

We can take f_{id} as

$$f_{\mathrm{id}}(u) = \begin{cases} u(\log(\sigma^3 u) - 1) & \text{if } u > 0 \\ 0 & \text{if } u = 0 \end{cases},$$

while we further decompose $f_{\rm ex}$ as

$$f_{\rm ex}(c) = 2\Gamma^{\rm Coul}(I(c)) + \Gamma^{\rm Steric}(c_+ + c_-),$$

where we require that

$$\frac{\partial \Gamma^{\rm Coul}}{\partial \theta}(\theta) = \log(\gamma_0(\theta))$$

and

$$\frac{\partial \Gamma^{\rm Steric}}{\partial u}(u) = \log(\gamma^{\rm Steric}(u)).$$

We then have

$$\begin{aligned}
\frac{\partial f_{\rm ex}}{\partial c_i}(c) &= 2\frac{\partial \Gamma^{\rm Coul}}{\partial \theta}(I(c))\frac{\partial I}{\partial c_i}(c) + \frac{\partial \Gamma^{\rm Steric}}{\partial u}(c_+ + c_-) \\
&= Z_i^2 \log(\gamma_0(I(c))) + \log(\gamma^{\rm Steric}(c_+ + c_-)) \\
&= \log(\gamma_i(c)),
\end{aligned}$$

for $i = \pm$, as wanted. We can then take

$$\begin{aligned}
\Gamma_0(\theta) &= -\frac{1}{4\pi\sigma\lambda}\left(\theta - \frac{8\lambda\sigma}{3}(\Upsilon_{\rm MSA}(\theta))^3 - 2\lambda(\Upsilon_{\rm MSA}(\theta))^2\right), \\
\Gamma_{\rm HS}(u) &= \frac{2\pi\sigma^3}{3}u^2, \\
\Gamma_{\rm CS}(u) &= \begin{cases} -\frac{6r(u)^2}{\pi\sigma^3}\frac{3r(u)-4}{(1-r(u))^2} & \text{if } 0 \le r(u) < 1 \\ +\infty & \text{if } r(u) \ge 1 \end{cases},
\end{aligned}$$

and the requested relations hold; in the following we will always use the more accurate Carnahan-Starling approximation. Observe finally that the ideal free-energy density is continuous in $\mathbb{R}_{\ge 0}$ and continuously differentiable in $\mathbb{R}_{>0}$, while the excess free-energy is continuously differentiable in $\mathbb{R}^2_{\ge 0}$; the bulk free-energy density is thus continuous in $\mathbb{R}_{\ge 0}$ and continuously differentiable in $\mathbb{R}_{>0}$ and

$$\frac{\partial f}{\partial c_i}(c) \to -\infty \text{ when } c_i \to 0, \quad i = \pm.$$

1.5. Convexity of f

If f is convex, then in [4] it is proved that there exists one and only one equilibrium solution when the effect of the charged walls is no more negligible; moreover they showed that f is convex for values of σ which are above a threshold σ_0 depending on the valences $Z_\pm$, on the temperature T and on the relative permittivity ε_r (the values for σ_0 for certain values of the parameters are reported in Table 1.1).

When f is not convex, the region $\mathcal{T}$ of the state space $c = (c_+, c_-)$ where f differs from its convex hull is thermodynamically instable; we then have that $\mathcal{T}$ divides the state space in different phases, which can instead coexist inside $\mathcal{T}$. When one species of ion is present,

$Z_+ : Z_-$	$T = 300$ K and $\varepsilon_r = 78.3$	$T = 350$ K and $\varepsilon_r = 62.0$
$+1 : -1$	0.0560 nm	0.0606 nm
$+2 : -1$	0.2329 nm	0.2521 nm
$+2 : -2$	0.2239 nm	0.2424 nm
$+3 : -3$	0.5039 nm	0.5454 nm
$+3 : -1$	0.7352 nm	0.7958 nm

Table 1.1.: Minimum values of σ for which f is convex.

this construction is equivalent to applying Maxwell equal area rule to the derivative of f, i.e. the electrochemical potential μ, in order to make the function μ monotone; since μ and c are conjugate variables, as well as P and V in gas thermodynamics, this situation is identical to the Maxwell's original construction, which has been applied to the isotherm $P = P(V)$ of the van der Waals equation (for Maxwell's original article see [11]). In order to study the behaviour of the electrolyte when f is not convex, it is then essential to find the region $\mathcal{T}$. Since f is difficult to treat analytically, we proceed numerically and derive an algorithm to compute the convex hull of f.

2. The Discrete Legendre-Fenchel Transform

As highlighted in Section 1.5, in order to study the behaviour of electrolytes with an arbitrary ionic diameter σ, it is necessary to devise an algorithm to calculate the convex hull of the bulk free-energy density. Our approach to build such algorithm is based on the properties of the Legendre-Fenchel transformation (LFT), an important concept in convex analysis, and of its discrete counterpart, the Discrete Legendre-Fenchel Transform (DLFT). In the first part of this chapter we present the main results concerning the LFT and DLFT, the second part deals with the exact form of the DLFT in one-dimension, while the last part is devoted to the study of the convergence of the DLFT to the LFT.

2.1. Legendre-Fenchel transformation: basics

We present a brief introduction on convex functions and on Legendre-Fenchel transformation; for more details see respectively Appendices A and B, or any introductory book on convex analysis, such as [13, 2].

Definition 2.1. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. The set

$$\operatorname{epi} f := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid y \geq f(x)\}$$

is called *epigraph* of f.

Definition 2.2. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be *convex* if $\operatorname{epi} f$ is convex as a subset of $\mathbb{R}^{n+1}$.

Definition 2.3. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function. The set

$$\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$$

is called the *effective domain* of f.

Definition 2.4. A convex function f is said to be *proper* if

- $f(x) < +\infty$ for at least one $x \in \mathbb{R}^n$
- $f(x) > -\infty$ for all $x \in \mathbb{R}^n$

Definition 2.5. The *convex hull* of a function f, denoted by $\operatorname{conv} f$, is the greatest convex function majorized by f.

Definition 2.6. The *lower semicontinuous hull* of a function f, denoted by $\underline{f}$, is the greatest lower semicontinuous function majorized by f.

Definition 2.7. The *closure* of a convex function f, denoted by $\operatorname{cl} f$, is

- the lower semicontinuous hull of f if $f(x) > -\infty$ for all $x \in \mathbb{R}^n$
- the constant function $-\infty$ otherwise

Definition 2.8. Given $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convex and $x \in \mathbb{R}^n$, a vector $\xi \in \mathbb{R}^n$ is called a *subgradient* of f at x if the affine function $z \mapsto f(x) + \langle \xi, z - x \rangle$ is majorized by f.

Definition 2.9. The collection of all subgradients of f at x is called the *subdifferential* of f at x and is denoted by $\partial f(x)$.

Definition 2.10. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. The function

$$\begin{array}{rcl} f^* : \mathbb{R}^n & \to & \overline{\mathbb{R}} \\ \xi & \mapsto & \sup\limits_{x \in \mathbb{R}^n} \left[\langle x, \xi \rangle - f(x) \right] = - \inf\limits_{x \in \mathbb{R}^n} \left[f(x) - \langle x, \xi \rangle \right], \end{array}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product of $\mathbb{R}^n$, is called the *Legendre-Fenchel Transform (LFT)*, or *conjugate*, of f.

Proposition 2.11. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. The function f^* is then a closed convex function.*

The two main results involving the LFT are the following, respectively linking the double transformation with the convex hull operation and showing how the transformation can be factorized along each dimension.

Theorem 2.12. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. We then have that $f^{**} = \operatorname{cl}(\operatorname{conv} f)$.*

Theorem 2.13. *Given $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and an index $i = 1, \dots n$, let f^{*i} be the LFT along the ith-dimension, i.e.*

$$\begin{array}{rcl} f^{*i} : \mathbb{R}^n & \to & \overline{\mathbb{R}} \\ (x_1, \dots, x_{i-1}, \xi_i, x_{i+1}, \dots, x_n) & \mapsto & \sup\limits_{x_i \in \mathbb{R}} \left[x_i \xi_i - f(x_1, \dots, x_i, \dots x_n) \right] \\ & = & \left[f(x_1, \dots, x_{i-1}, \cdot, x_i, \dots x_n) \right]^* (\xi_i). \end{array}$$

We then have that

$$f^* = \left(- \left(\cdots \left(- \left(-f^{*1} \right)^{*2} \right)^{*3} \cdots \right)^{*(n-1)} \right)^{*n}.$$

2.2. Discrete Legendre-Fenchel Transform

To compute the convex hull of a given function f, the number of evaluation points must be finite. Therefore, it is useful to introduce the discrete version of the LFT.

Definition 2.14. Given a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\Omega \subseteq \mathbb{R}^n$ we define the function $f_\Omega : \mathbb{R}^n \to \overline{\mathbb{R}}$ as

$$f_\Omega(x) = \begin{cases} f(x) & \text{if } x \in \Omega \\ +\infty & \text{elsewhere} \end{cases}.$$

Proposition 2.15. *Given a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\Omega \subseteq \mathbb{R}^n$ we have that*

$$f^*_\Omega(\xi) = \sup_{x \in \Omega} \left[\langle x, \xi \rangle - f(x)\right].$$

Proof. Fix $\xi \in \mathbb{R}^n$; by Definition B.1 we have that

$$f^*_\Omega(\xi) = \sup_{x \in \mathbb{R}^n} \left[\langle x, \xi \rangle - f_\Omega(x)\right].$$

If $\Omega = \emptyset$, then $f_\Omega \equiv +\infty$ and the thesis follows trivially from the fact that the supremum of an empty set is $-\infty$. If there is $x_0 \in \Omega$, we then have for every $x \in \mathbb{R}^n \setminus \Omega$ that $f_\Omega(x) = +\infty$ and thus

$$\langle x, \xi \rangle - f_\Omega(x) \leq \langle x_0, \xi \rangle - f(x_0);$$

this implies that the two suprema have the same value. □

Proposition 2.16. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\Omega \subseteq \Omega' \subseteq \mathbb{R}^n$. Then*

$$f^*_{\Omega'} \geq f^*_\Omega.$$

Proof. We have that $f_{\Omega'} \leq f_\Omega$ by construction; we can then apply Proposition B.5 and obtain the thesis. □

Proposition 2.17. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega, \Omega' \subseteq \mathbb{R}^n$. Then*

$$f^*_{\Omega \cup \Omega'} = \max\left\{f^*_\Omega, f^*_{\Omega'}\right\}.$$

Proof. By Proposition 2.15 we have that

$$\begin{aligned} f^*_{\Omega \cup \Omega'}(\xi) &= \sup_{x \in \Omega \cup \Omega'} \left[\langle x, \xi \rangle - f(x)\right] \\ &= \max\left\{\sup_{x \in \Omega} \left[\langle x, \xi \rangle - f(x)\right], \sup_{x \in \Omega'} \left[\langle x, \xi \rangle - f(x)\right]\right\} \\ &= \max\left\{f^*_\Omega(\xi), f^*_{\Omega'}(\xi)\right\}. \end{aligned}$$

□

Definition 2.18. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega_N \subset \mathbb{R}^n$ be finite; the subscript N is used to highlight the finiteness of Ω_N and in general does not mean that $|\Omega_N| = N$. The function $f^*_{\Omega_N}$ is then called the *Discrete Legendre-Fenchel Transform* (DLFT) of f with respect to the set Ω_N; in particular we have

$$\begin{aligned} f^*_{\Omega_N} : \mathbb{R}^n &\to \overline{\mathbb{R}} \\ \xi &\mapsto \max_{x \in \Omega_N} \left[\langle x, \xi \rangle - f(x) \right]. \end{aligned}$$

Remark 2.19. We could in principle have that $f(x) = +\infty$ for some $x \in \Omega_N$, but in this case we would have $f^*_{\Omega_N} = f^*_{\Omega_N \setminus \{x\}}$; we could also have $f(x) = -\infty$ for some $x \in \Omega_N$, but then we would have $f^*_{\Omega_N} = +\infty$. For simplicity we will omit in all results involving the DLFT the assumption "$f(x)$ is finite for all $x \in \Omega_N$" which excludes these trivial cases.

2.3. Conjugation of piecewise linear functions

In the one-dimensional case the DLFT is equivalent to a LFT of a piecewise linear function; in order to highlight the fact that we are considering functions defined on $\mathbb{R}$, we will denote finite sets in $\mathbb{R}$ by the symbol X_N.

Definition 2.20. Given $f : \mathbb{R} \to \overline{\mathbb{R}}$ and $X_N = \{x_1, \dots, x_N\} \subset \mathbb{R}$ with $x_1 < \cdots < x_N$, we denote by $I_{X_N}(f)$ the function which coincides on the interval $[x_1, x_N]$ with the piecewise linear interpolant of f on the nodes X_N and is $+\infty$ elsewhere.

Proposition 2.21. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ and let $X_N \subset \mathbb{R}$ be finite. Then* $\operatorname{conv} f_{X_N} = \operatorname{conv} I_{X_N}(f)$.

Proof. Because $I_{X_N}(f) \le f_{X_N}$, we have by definition of convex hull that also $\operatorname{conv} I_{X_N}(f) \le \operatorname{conv} f_{X_N}$. It can be easily seen that $\operatorname{epi} I_{X_N}(f) \subseteq \operatorname{conv} \operatorname{epi} f_{X_N}$; thus $\operatorname{conv} f_{X_N} \le I_{X_N}(f)$ and, by applying the convex hull operator, we obtain $\operatorname{conv} f_{X_N} \le \operatorname{conv} I_{X_N}(f)$. □

Corollary 2.22. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ and let $X_N \subset \mathbb{R}$ be finite. Then we have* $[\operatorname{conv} I_{X_N}(f)]^* = [I_{X_N}(f)]^* = f^*_{X_N}$.

Proof. The thesis follows from Proposition 2.21 and Lemma B.15. □

This characterization is very important because the LFT of a convex piecewise linear function can be easily computed; the following theorem is the basis for an efficient one-dimensional algorithm for the DLFT ([10]).

Theorem 2.23. *Given a finite collection of points $P_i = (x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, N$, such that $x_1 < \cdots < x_N$ and given $c_0, c_N \in \overline{\mathbb{R}}$, let g be the piecewise linear continuous function which interpolates the points P_i and is extended to the left and the right with slopes respectively c_0 and c_N (possibly infinite), i.e.*

$$g(x) = \begin{cases} y_1 + c_0(x - x_1) & \text{if } x < x_1 \\ y_i + c_i(x - x_i) & \text{if } x_i \le x \le x_{i+1},\ i = 1, \dots, N-1 \\ y_N + c_N(x - x_N) & \text{if } x > x_N \end{cases}, \tag{2.1}$$

where the values c_i, the slopes of the linear segments of g, are given by

$$c_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \quad i = 1, \dots, N-1.$$

Suppose that g is convex, i.e. $c_0 \leq \cdots \leq c_N$; then

$$g^*(\xi) = \begin{cases} +\infty & \text{if } \xi < c_0 \text{ or } \xi > c_N \\ x_i\xi - y_i & \text{if } c_{i-1} \leq \xi \leq c_i,\ i = 1, \dots, N \end{cases}. \tag{2.2}$$

Proof. By Proposition A.35 we have that

$$\partial g(x_i) = [c_{i-1}, c_i], \quad i = 1, \dots, N;$$

then the thesis follows for each $c_0 \leq \xi \leq c_N$ by applying Theorem B.26.

If there exists $\xi < c_0$, by (2.1) we obtain that

$$\begin{aligned} g^*(\xi) &= \sup_{x \in \mathbb{R}} [x\xi - g(x)] \\ &\geq \sup_{x < x_1} [x\xi - y_1 - c_0(x - x_1)] \\ &= \sup_{x < x_1} [(\xi - c_0)x - y_1 + c_0 x_1], \end{aligned}$$

which is $+\infty$ because $\xi - c_0 < 0$. The case $\xi > c_N$ is similar. □

By Theorem 2.23 the conjugate of a piecewise-linear convex function g with nodes $\{x_i\}$ and slopes $\{c_i\}$ is a piecewise-linear convex function with nodes $\{c_i\}$ and slopes $\{x_i\}$. The conjugate is just a different "coding" of the some information: the values defining the function, i.e. its slopes and its nodes, do not vary after the conjugation, but just swap roles. In Figure 2.1 we plot g and g^* defined as in Theorem 2.23 where $x_i = -3 + i$ for $i = 1, \dots 5$, $y_1 = 1.75$, $y_2 = 0.6$, $y_3 = 0.1$, $y_4 = 0$, $y_5 = 0.33$, $c_0 = -\infty$, $c_5 = 1$; observe in particular that on the side on which g is $+\infty$ its LFT is extended linearly, and conversely on the side on which g is extended linearly its FLT is $+\infty$.

2.4. Convergence to the Legendre-Fenchel transform

The convergence properties of the DLFT, including its more complex versions considered in the next chapter, are based on Theorem (2.27). This result is due to Corrias 3; our proof has been adapted to our notation and extended explicitly to the case Ω is not closed.

Lemma 2.24. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be such that $f(x_0) < +\infty$ for at least one point $x_0 \in \mathbb{R}^n$. Then $f^* > -\infty$.*

Proof. Let $\xi \in \mathbb{R}^n$; by definition of conjugation we then have

$$\begin{aligned} f^*(\xi) &= \sup_{x \in \mathbb{R}^n} [\langle x, \xi \rangle - f(x)] \\ &\geq \langle x_0, \xi \rangle - f(x_0) > -\infty. \end{aligned}$$

□

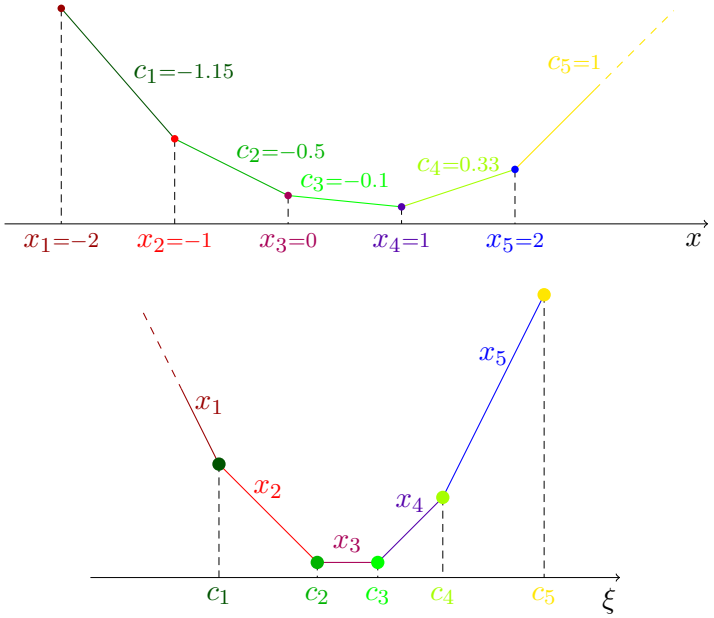

Figure 2.1.: The plots of the piecewise linear convex function g (top) and its LFT g^* (bottom); the correspondence between nodes and slopes is highlighted.

Lemma 2.25. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega \subseteq \mathbb{R}^n$. Then we have*

$$\left(\overline{f_\Omega}\right)^*(\xi) = \sup_{x\in\Omega} \left[\langle x, \xi\rangle - \overline{f_\Omega}(x)\right].$$

Proof. Since the right hand side is equal to $\left(\overline{f_\Omega}\right)^*_\Omega$, it suffices to prove that $\overline{f_\Omega} = \left(\overline{f_\Omega}\right)_\Omega$. Since $\overline{f_\Omega} \geq f_\Omega$, we have that $\overline{f_\Omega}(x) = +\infty$ for every $x \notin \Omega$; being $\overline{f_\Omega} = \left(\overline{f_\Omega}\right)_\Omega$ on Ω, the thesis follows. □

Definition 2.26. Let $(\Omega_N)_N$ be an increasing sequence of finite subsets of $\Omega \subseteq \mathbb{R}^n$. We say that the sequence converges to Ω if $\min_{x'\in\Omega_N} \|x' - x\| \to 0$ as $N \to \infty$ for all $x \in \Omega$; we denote this fact by the notation $\Omega_N \to \Omega$. If we also have that $\sup_{x\in\Omega} \min_{x'\in X_N} \|x' - x\| \to 0$ as $N \to \infty$, we say that the convergence is uniform.

Theorem 2.27. *Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ and let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be such that $\left(\overline{f_\Omega}\right) = \underline{f_\Omega}$. Let $(\Omega_N)_N$ a sequence of finite subsets of $\mathbb{R}^n$ such that $\Omega_N \to \Omega$. Then, $f^*_{\Omega_N}$ converges pointwise to f^*_Ω. Moreover, if we also have that $f|_\Omega$ is uniformly continuous and that $\Omega_N \to \Omega$ uniformly, then the convergence is uniform on every bounded subset S of* $\operatorname{dom} f^*_\Omega$.

Proof. If $f = +\infty$ on Ω we have that both f_Ω and f_{Ω_N} are $+\infty$ on $\mathbb{R}^n$; thus $f^*_{\Omega_N} = f^*_\Omega = -\infty$ on $\mathbb{R}^n$ and the thesis follows. Assume then that there exists $x_\mathrm{f} \in \Omega$ such that $f(x_\mathrm{f}) < +\infty$.

Fix $\xi \in \mathbb{R}^n$. Firstly we prove that $f^*_{\Omega_N}(\xi)$ has a limit. Since $\Omega_N \subset \Omega_{N+1} \subset \Omega$ for every N, we have by Proposition 2.16 that

$$f^*_\Omega(\xi) \geq f^*_{\Omega_{N+1}}(\xi) \geq f^*_{\Omega_N}(\xi).$$

The sequence $f^*_{\Omega_N}(\xi)$ is increasing and the limit of $f^*_{\Omega_N}(\xi)$ as $N \to \infty$ exists.

Now we prove that this limit is $f^*_\Omega(\xi)$. By Lemma B.9 we have $f^*_\Omega = (\underline{f_\Omega})^*$ and $\left(\overline{f_\Omega}\right)^* = \left[\underline{\left(\overline{f_\Omega}\right)}\right]^*$. By the hypothesis we get $f^*_\Omega = \left(\overline{f_\Omega}\right)^*$ and applying Lemma 2.25 we obtain that

$$f^*_\Omega(\xi) = \sup_{x\in\Omega}\left[\langle x,\xi\rangle - \overline{f_\Omega}(x)\right] \qquad \forall \xi \in \mathbb{R}^n. \tag{2.3}$$

Fix $\varepsilon > 0$. By Lemma (2.24) we have that $f^*_\Omega(\xi) > -\infty$; we then have to consider only two cases: $f^*_\Omega(\xi) = +\infty$ and $f^*_\Omega(\xi)$ finite. First suppose that $f^*_\Omega(\xi)$ is finite; by the properties of the supremum in (2.3), we can find $z \in \Omega$ such that

$$f^*_\Omega(\xi) \le \langle z,\xi\rangle - \overline{f_\Omega}(z) + \frac{\varepsilon}{3} \le f^*_\Omega(\xi) + \frac{\varepsilon}{3},$$

which implies that $\overline{f_\Omega}(z)$ is finite. Then, by the upper semicontinuity of $\overline{f_\Omega}$, there exists $r > 0$ such that

$$\overline{f_\Omega}(x) - \overline{f_\Omega}(z) \le \frac{\varepsilon}{3} \qquad \forall x \in B(z,r),$$

where $B(z,r)$ is the open ball with center z and radius r.

Let $\tilde\varepsilon \le \min\{\frac{\varepsilon}{3\|\xi\|}, r\}$. We denote by $\hat z_N$ the point of Ω_N nearest to z, i.e. $\hat z_N := \arg\min_{x\in\Omega_N}\|x - z\|$. By hypothesis we have that $\|z - \hat z_N\| \le \tilde\varepsilon$ for sufficiently large values of N and thus $\hat z_N \in B(z,r)$ and $\|\hat z_N - z\| \le \frac{\varepsilon}{3\|\xi\|}$. Since $f_\Omega \le \overline{f_\Omega}$ and $f = f_\Omega$ on $\Omega_N \subset \Omega$, we have that

$$\begin{aligned}
0 &\le f^*_\Omega(\xi) - f^*_{\Omega_N}(\xi)\\
&= \sup_{x\in\Omega}\left[\langle x,\xi\rangle - \overline{f_\Omega}(x)\right] - \max_{x\in\Omega_N}\left[\langle x,\xi\rangle - f_\Omega(x)\right]\\
&\le \langle z,\xi\rangle - \overline{f_\Omega}(z) + \frac{\varepsilon}{3} - \langle \hat z_N,\xi\rangle + f_\Omega(\hat z_N)\\
&\le \langle z - \hat z_N,\xi\rangle + \overline{f_\Omega}(\hat z_N) - \overline{f_\Omega}(z) + \frac{\varepsilon}{3}\\
&\le \|z - \hat z_N\|\,\|\xi\| + \frac{2}{3}\varepsilon \le \varepsilon,
\end{aligned} \tag{2.4}$$

and thus $f^*_{\Omega_N}(\xi)$ converges to $f^*_\Omega(\xi)$.

Now consider $f^*_\Omega(\xi) = +\infty$ and fix $\gamma \in \mathbb{R}$. By the properties of the supremum in (2.3), we can find $z \in \Omega$ such that $\langle z,\xi\rangle - \overline{f_\Omega}(z) > \gamma$. Being $\langle\cdot,\xi\rangle - \overline{f_\Omega}$ lower semicontinuous, we can find, for N sufficiently large, $\hat z_N \in \Omega_N$ sufficiently close to z to make $\langle \hat z_N,\xi\rangle - \overline{f_\Omega}(\hat z_N) > \gamma$. We have

$$\begin{aligned}
f^*_{\Omega_N}(\xi) &= \max_{x\in\Omega_N}\left[\langle x,\xi\rangle - f(x)\right]\\
&\ge \langle \hat z_N,\xi\rangle - f_\Omega(\hat z_N) \ge\\
&= \langle \hat z_N,\xi\rangle - \overline{f_\Omega}(\hat z_N) > \gamma,
\end{aligned}$$

and thus $f^*_{\Omega_N}(\xi)$ converges to $+\infty$.

Assume that $f|_\Omega$ is uniformly continuous and consider a bounded subset S of $\operatorname{dom} f^*_\Omega$. By uniform continuity we can find $\delta > 0$ such that for any $x, z \in \Omega$ with $\|x - z\| < \delta$ we

have that $|f(x) - f(z)| \leq \varepsilon/3$. For any $z \in \Omega$ we have that $\|\hat{z}_N - z\| < \min\{\delta, \frac{\varepsilon}{3K}\}$ for sufficiently large N, where $K = \sup_{\xi \in S} \|\xi\| < +\infty$. The inequality (2.4) holds for any $\xi \in S$, thus proving uniform convergence on S. □

2.5. Sufficient conditions for convergence

It is interesting to note we cannot substitute the hypothesis $\underline{(\overline{f_\Omega})} = \underline{f_\Omega}$ of Theorem (2.27) with the mere lower semicontinuity of f or f_Ω, as is shown in the following example from 3.

Example 2.28 (Corrias, 1996)**.** Consider $f(x) = |x - k|$ with $k \in]0, 1[$; it is continuous, thus also upper semicontinuous and moreover satisfies the hypothesis $\underline{(\overline{f_\Omega})} = \underline{f_\Omega}$. By Theorem 2.23 we have

$$f^*(\xi) = \begin{cases} k\xi & \text{if } \xi \in [-1, 1] \\ +\infty & \text{elsewhere} \end{cases},$$

$$f^*_{[0,1]}(\xi) = \begin{cases} k\xi & \text{if } \xi \in [-1, 1] \\ -k & \text{if } \xi \leq -1 \\ \xi + k - 1 & \text{if } \xi \geq 1 \end{cases}.$$

Let $(X_N)_N$ be a sequence of finite grids such that $X_N \to [0, 1]$. We can choose $k \in]0, 1[$ such that $k \notin X_N$ for every N and consider the lower semicontinuous function

$$g(x) = \begin{cases} |x - k| & \text{if } x \neq k \\ -1 & \text{if } x = k \end{cases};$$

observe that $\underline{(\overline{g_\Omega})} \neq \underline{g_\Omega}$. By geometric construction, we can easily find

$$g^*(\xi) = \begin{cases} k\xi + 1 & \text{if } \xi \in [-1, 1] \\ +\infty & \text{elsewhere} \end{cases},$$

$$g^*_{[0,1]}(\xi) = \begin{cases} k\xi + 1 & \text{if } \xi \in [-\frac{k+1}{k}, \frac{2-k}{1-k}] \\ -k & \text{if } \xi < -\frac{k+1}{k} \\ \xi + k - 1 & \text{if } \xi > \frac{2-k}{1-k} \end{cases}.$$

By Theorem 2.27 we have $f^*_{X_N} \to f^*_{[0,1]}$ pointwise as $N \to \infty$. Being $k \notin X_N$ for every N, we have $f_{X_N} = g_{X_N}$ for every N and thus also $f^*_{X_N} = g^*_{X_N}$ for every N; this means that $g^*_{X_N} \to f^*_{[0,1]} \neq g^*_{[0,1]}$ and thus Theorem 2.27 cannot be applied to functions which are just lower semicontinuous.

Now we give a necessary condition on the set Ω in order to have $\underline{(\overline{f_\Omega})} = \underline{f_\Omega}$.

Theorem 2.29. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\Omega \subseteq \mathbb{R}^n$ be such that $\underline{(\overline{f_\Omega})} = \underline{f_\Omega}$. Then $\Omega' \subseteq \operatorname{cl} \operatorname{int} \Omega'$, where $\Omega' = \{x \in \Omega \mid f(x) < +\infty\}$.*

Proof. Since $f_\Omega = f_{\Omega'}$ we can suppose without loss of generality that $f(x) < +\infty$ for all $x \in \Omega$, i.e., $\Omega = \Omega'$.

Suppose that $\Omega \nsubseteq \operatorname{cl}\operatorname{int}\Omega$ and let $x_0 \in \Omega$ be such that $x_0 \notin \operatorname{cl}\operatorname{int}\Omega$.

Since we have that $f_\Omega(x_0) < +\infty$, also $\underline{f_\Omega}(x_0) < +\infty$.

Since $x_0 \notin \operatorname{int}\Omega$, in every neighbourhood of x_0 there is a point x such that $x \notin \Omega$ and thus $f_\Omega(x) = +\infty$. This implies that $\overline{f_\Omega}(x_0) = +\infty$.

There exists an open neighbourhood U of x_0 such that $x \notin \operatorname{cl}\operatorname{int}\Omega$ for every $x \in U$. If this were not true, then in every open neighbourhood V of x_0 there would be a point of $\operatorname{cl}\operatorname{int}\Omega$ and thus, by openness of V, also a point of $\operatorname{int}\Omega$. This would mean that $x_0 \in \operatorname{cl}\operatorname{int}\Omega$ and we have reached a contradiction.

As we have seen, we have that $\overline{f_\Omega}(x) = +\infty$ for every $x \in U \cap \Omega$ and, being $\overline{f_\Omega} \geq f_\Omega$, this holds also for every $x \in U \setminus \Omega$. This means that $\overline{f_\Omega}$ is $+\infty$ on a neighbourhood of x_0 and thus $\underline{\left(\overline{f_\Omega}\right)}(x_0) = +\infty$. We then get that $\underline{\left(\overline{f_\Omega}\right)}(x_0) \neq \underline{f_\Omega}(x_0)$, which is in contradiction with our hypothesis. □

We also give a sufficient condition on f in order for $\underline{\left(\overline{f_\Omega}\right)} = \underline{f_\Omega}$ to hold.

Lemma 2.30. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega \subseteq \mathbb{R}^n$ be an open set. Suppose that $f|_\Omega$ is upper semicontinuous; then f_Ω is upper semicontinuous.*

Proof. Since Ω is open, any point $x \in \Omega$ is an internal point of Ω and thus there is a neighbourhood U of x such that $U \subset \Omega$ and $f_\Omega = f$ on U; being semicontinuity a local property, we then have that f_Ω is upper semicontinuous at x. If instead $x \in \mathbb{R}^n \setminus \Omega$, we have $f_\Omega(x) = +\infty$, which gives us that f_Ω is upper semicontinuous at x. □

Lemma 2.31. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega \subseteq \mathbb{R}^n$. Suppose that $f|_{\operatorname{int}\Omega}$ is upper semicontinuous; then $\overline{f_\Omega} = f_{\operatorname{int}\Omega}$.*

Proof. By Lemma 2.30 we have that $f_{\operatorname{int}\Omega}$ is upper semicontinuous. Then we have $f_{\operatorname{int}\Omega} \geq \overline{f_\Omega} \geq f_\Omega$; since $f_\Omega = f_{\operatorname{int}\Omega}$ on $\operatorname{int}\Omega \cup (\mathbb{R}^n \setminus \Omega)$, we also have $f_{\operatorname{int}\Omega} = \overline{f_\Omega}$ on the same set.

Let $x_0 \in \Omega \setminus \operatorname{int}\Omega$. This means that for every neighbourhood of x_0 there is a point $x \notin \Omega$ for which we have $f_\Omega(x) = +\infty$; then, we have that $\overline{f_\Omega}(x_0) = +\infty$. Since we trivially have that $f_{\operatorname{int}\Omega}(x_0) = +\infty$, we get the thesis. □

Theorem 2.32. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega \subseteq \mathbb{R}^n$ be such that $\Omega \subseteq \operatorname{cl}\operatorname{int}\Omega$. Suppose that $f|_\Omega$ is upper semicontinuous; then $\underline{\left(\overline{f_\Omega}\right)} = \underline{f_\Omega}$.*

Proof. By Lemma 2.31 we have that $\overline{f_\Omega} = f_{\operatorname{int}\Omega}$; thus we have to prove that $\underline{f_{\operatorname{int}\Omega}} = \underline{f_\Omega}$.

Being $f_{\operatorname{int}\Omega} \geq f_\Omega$, we have that $\underline{f_{\operatorname{int}\Omega}} \geq \underline{f_\Omega}$. Thus, if we prove that $f_\Omega \geq \underline{f_{\operatorname{int}\Omega}}$, then we have that $\underline{f_{\operatorname{int}\Omega}} = \underline{f_\Omega}$.

Since $f_\Omega = f_{\operatorname{int}\Omega}$ on $\operatorname{int}\Omega \cup (\mathbb{R}^n \setminus \Omega)$, we have that $f_\Omega \geq \underline{f_{\operatorname{int}\Omega}}$ on the same set.

Let $x_0 \in \Omega \setminus \operatorname{int}\Omega$ and suppose that $f_\Omega(x_0) < \underline{f_{\operatorname{int}\Omega}}(x_0)$. This implies that there exists $c \in \mathbb{R}$ such that $f_\Omega(x_0) < c < \underline{f_{\operatorname{int}\Omega}}(x_0)$.

By lower semicontinuity of $\underline{f_{\operatorname{int}\Omega}}$, there exists a neighbourhood U of x_0 such that $\underline{f_{\operatorname{int}\Omega}}(x) > c$ for every $x \in U$. Let U_N a sequence of open balls centered in x_0, having

decreasing radius and contained in U. Since by hypothesis $x_0 \in \operatorname{cl}\operatorname{int}\Omega$, there is a point $x_N \in \operatorname{int}\Omega \cap U_N$ for all N; such points satisfy $f_\Omega(x_N) \geq \underline{f_{\operatorname{int}\Omega}}(x_N) > c$, where the left inequality follows from the fact that $x_N \in \operatorname{int}\Omega$.

By upper semicontinuity of $f|_\Omega$, there exists a neighbourhood V of x_0 such that $f_\Omega(x) = f(x) < c$ for every $x \in V \cap \Omega$. But for N sufficiently large we have $U_N \subset V$ and thus $x_N \in V \cap \Omega$, which leads to a contradiction. □

Theorem 2.32 shows that upper semicontinuity is a sufficient condition for $\underline{\left(\overline{f_\Omega}\right)} = \underline{f_\Omega}$; however, it is not a necessary one, as shown in the next example.

Example 2.33. Let

$$f(x) = \begin{cases} \sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

and $\Omega = [-1, 1]$. The function f is continuous everywhere but in 0, where it is neither lower or upper semicontinuous. Nonetheless, we have that $\underline{\left(\overline{f_\Omega}\right)} = \underline{f_\Omega}$.

3. The double Discrete Legendre-Fenchel Transform

According to the results in the previous chapter, the application of two successive LFT to a function gives its convex hull. Similarly the algorithm to approximate the convex hull can be constructed by applying two successive DLFT; we call it *double DLFT*. In the literature the most of the papers focuses on the computation of the conjugate, which has many pratical applications in physics; on the contrary there are few theoretical results on the double DLFT, mainly dealing with convex functions. For the same reasons, one of the main problems regarding a correct discretization of the double LFT, i.e. the choice of the grid on which the second DLFT is computed, called the *dual grid*, is rarely treated. In this chapter we present an attempt to generalize the results of the previous chapter to the double DLFT, with particular attention payed to the choice of the dual grid.

3.1. Properties of the double DLFT

Lemma 3.1. *Let $\Omega \subset \mathbb{R}^n$ be bounded and let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Suppose that there exists $L \in \mathbb{R}$ such that $f(x) > L$ for every $x \in \Omega$. Then* $\operatorname{dom} f^*_\Omega = \mathbb{R}^n$.

Proof. Fix $\xi \in \mathbb{R}^n$. We have that

$$\begin{aligned} \xi \in \operatorname{dom} f^*_\Omega \quad &\iff \quad f^*_\Omega(\xi) < +\infty \\ &\iff \quad \sup_{x\in\Omega} [\langle \xi, x\rangle - f(x)] < +\infty. \end{aligned}$$

Being Ω bounded, there is a constant $K > 0$ such that for every $x \in \Omega$ we have $\|x\| \leq K$; we then have that for any $x \in \Omega$

$$\langle \xi, x\rangle - f(x) < |\langle \xi, x\rangle| - L \leq \|\xi\| \, \|x\| - L \leq K \, \|\xi\| - L,$$

and thus $\sup_{x\in\Omega} [\langle \xi, x\rangle - f(x)] \leq K \, \|\xi\| - L < +\infty$. □

Corollary 3.2. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be such that $f > -\infty$. Then* $\operatorname{dom} f^*_{\Omega_N} = \mathbb{R}^n$ *for any $\Omega_N \subset \mathbb{R}^n$ finite.*

Corollary 3.3. *Let $\Omega_N \subset \mathbb{R}^n$ be finite and let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be such that $f(x_0) < +\infty$ for at least one point $x_0 \in \Omega_N$. Then* $\operatorname{dom} \left(f^*_{\Omega_N}\right)^*_{S_N} = \mathbb{R}^n$ *for any $S_N \subseteq \mathbb{R}^n$ finite.*

Proof. Follows from Lemma 2.24 and Corollary 3.2. □

Lemma 3.4. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega, S, S' \subseteq \mathbb{R}^n$ be such that $S \subseteq S'$. Then $(f^*_\Omega)^*_S \leq (f^*_\Omega)^*_{S'} \leq \operatorname{cl}\operatorname{conv} f_\Omega$ on $\mathbb{R}^n$.*

Proof. By observing that $\operatorname{cl}\operatorname{conv} f_\Omega = (f^*_\Omega)^* = (f^*_\Omega)^*_{\mathbb{R}^n}$, the thesis follows immediately from Proposition 2.16. □

Corollary 3.5. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega_N, S, S' \subseteq \mathbb{R}^n$ be such that Ω_N is finite and $S \subseteq S'$. Then $\left(f^*_{\Omega_N}\right)^*_S \leq \left(f^*_{\Omega_N}\right)^*_{S'} \leq \operatorname{conv} f_{\Omega_N}$ on $\mathbb{R}^n$.*

Proof. By Theorem A.56 $\operatorname{conv} f_{\Omega_N}$ is closed; the thesis follows from Lemma 3.4. □

Remark 3.6. The last two corollaries show that if we enlarge the dual grid then the resulting double DLFT/LFT always improves as an approximation of the convex hull.

Corollary 3.7. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega, S \subseteq \mathbb{R}^n$. Then $(f^*_\Omega)^*_S \leq f$ on Ω.*

Proof. We have that $\operatorname{cl}\operatorname{conv} f_\Omega \leq f_\Omega$ and by Lemma 3.4 that $(f^*_\Omega)^*_S \leq \operatorname{cl}\operatorname{conv} f_\Omega$; since $f_\Omega = f$ on Ω, we get the thesis. □

In general, the inequality of Corollary 3.7 does not hold everywhere; this is reasonable, because by restricting f to the set Ω every information on the behaviour of f outside Ω is discarded.

Example 3.8. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be the piecewise linear function with nodes $\{-1, -1/2, 0, 1/2, 1\}$ and values $\{2, 0, 1, 0, 2\}$. By choosing $X_N = \{-1, 0, 1\}$ and $C_N = \{-1, 1\}$, we have that $\left(f^*_{X_N}\right)^*_{C_N}$ is the piecewise linear function with nodes $\{-1, 0, 1\}$ and values $\{2, 1, 2\}$, which is strictly greater than f outside of X_N.

If we try to compare $(f^*_\Omega)^*_S$ with the exact convex hull $\operatorname{cl}\operatorname{conv} f = f^{**}$, no inequality can be established, not even by considering only the points of Ω.

Example 3.9. Let f and X_N be defined as in previous example, whereas let $C_N = \{0, 1\}$; then we have that

$$\left(f^*_{X_N}\right)^*_{C_N}(x) = \begin{cases} 1 & \text{if } x \leq \frac{1}{2} \\ 1 + 2(x - \frac{1}{2}) & \text{if } x \geq \frac{1}{2} \end{cases},$$

which is somewhere greater than f^{**} and somewhere smaller.

3.2. Optimal dual grids

In the applications we usually need to compute the convex hull of functions of the type $g : \Omega \to \overline{\mathbb{R}} \setminus \{-\infty\}$, where $\Omega \subset \mathbb{R}^n$ is bounded. We can easily translate this problem on Ω into a problem on the whole $\mathbb{R}^n$ by taking

$$f(x) := \begin{cases} g(x) & \text{if } x \in \Omega \\ +\infty & \text{elsewhere} \end{cases}$$

and observing that $\operatorname{conv} f = \operatorname{conv} g$ on Ω; then we discretize this problem by considering a finite set $\Omega_N \subset \Omega$ and supposing that the values of f are known only on the points of Ω_N, i.e. by taking the function f_{Ω_N}. Finally we take the double DLFT $\left(f^*_{\Omega_N}\right)^*_{S_N}$ as an approximation for $\operatorname{conv} f$, where $S_N \subset \mathbb{R}^n$ is a finite set.

A correct choice of the grid S_N is essential to obtain a good approximation. We begin by observing that, being the behaviour of f unknown outside Ω_N, the function we are actually trying to compute is $\operatorname{conv} f_{\Omega_N}$, which we then take as an approximation of $\operatorname{conv} f$. Since by Corollary 3.2 the domain $\operatorname{dom} f^*_{\Omega_N}$ is the whole $\mathbb{R}^n$, the functions $f^*_{\Omega_N}$ and $\operatorname{cl}\operatorname{conv}\left(f^*_{\Omega_N}\right)_{S_N}$ are necessarily different by finiteness of S_N; by applying Corollary B.16 and Corollary B.13 we obtain $\left(f^*_{\Omega_N}\right)^*_{S_N} \neq \left(f^*_{\Omega_N}\right)^* = \operatorname{conv} f_{\Omega_N}$ as functions defined on $\mathbb{R}^n$. Nonetheless in this section we will show that there exists a set S_N such that $\left(f^*_{\Omega_N}\right)^*_{S_N}(x) = \operatorname{conv} f_{\Omega_N}(x)$ for every $x \in \Omega_N$.

Lemma 3.10. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function, let $\Omega_N \subset \mathbb{R}^n$ be a finite set and let $\xi_0 \in \mathbb{R}^n$ such that $f^*_{\Omega_N}(\xi_0) = f^*(\xi_0)$. Then $\partial f^*(\xi_0) \cap \Omega_N \neq \emptyset$.*

Proof. By hypothesis we have

$$\max_{x\in\Omega_N} \left[\langle x, \xi_0\rangle - f(x)\right] = \sup_{x\in\mathbb{R}^n} \left[\langle x, \xi_0\rangle - f(x)\right].$$

Let $x_0 := \arg\max_{x\in\Omega_N} \left[\langle x, \xi_0\rangle - f(x)\right]$; then for every $x \in \mathbb{R}^n$ we have that

$$\begin{aligned} \langle x, \xi_0\rangle - f(x) &\leq \langle x_0, \xi_0\rangle - f(x_0) \\ f(x_0) + \langle x - x_0, \xi_0\rangle &\leq f(x), \end{aligned}$$

i.e. $\xi_0 \in \partial f(x_0)$. From Corollary B.20 we get $x_0 \in \partial f^*(\xi_0)$ and thus $\partial f^*(\xi_0) \cap \Omega_N \neq \emptyset$. □

In Lemma 3.10 the converse implication does not hold in general, as shown in the next example.

Example 3.11. Let $\Omega_N = \{0, 1\}$ and let $f : \mathbb{R} \to \mathbb{R}$ be such that

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } 0 < x \leq 1 \\ +\infty & \text{elsewhere} \end{cases};$$

observe that f is convex, but not closed. By Corollary B.10 we have that $f^* = (\operatorname{cl} f)^*$ and by applying Theorem 2.23 we obtain that

$$f^*(\xi) = \begin{cases} 0 & \text{if } \xi \leq 0 \\ \xi & \text{if } \xi > 0 \end{cases},$$

whereas, again by Theorem 2.23, we have that

$$f^*_{\Omega_N}(\xi) = \begin{cases} -1 & \text{if } \xi \leq -1 \\ \xi & \text{if } \xi > -1 \end{cases}.$$

If we take $\xi = -1$ we have that $0 \in \partial f^*(-1) \cap \Omega_N$ but $f^*_{\Omega_N}(-1) \neq f^*(-1)$.

Theorem 3.12. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a closed convex function, let $\Omega_N \subset \mathbb{R}^n$ be a finite set, and let $\xi_0 \in \mathbb{R}^n$. Then, $\partial f^*(\xi_0) \cap \Omega_N \neq \emptyset$ if and only if $f^*_{\Omega_N}(\xi_0) = f^*(\xi_0)$.*

Proof. Necessity follows from Lemma 3.10. In order to prove sufficiency suppose $\partial f^*(\xi_0) \cap \Omega_N \neq \emptyset$ and let $x_0 \in \partial f^*(\xi_0) \cap \Omega_N$. By Corollary B.23 we have $\xi_0 \in \partial f(x_0)$, i.e. $h \leq f$, where h is the affine function whose gradient is ξ_0 and whose value at x_0 is $f(x_0)$. Since $h \leq f \leq f_{\Omega_N}$ and $f(x_0) = f_{\Omega_N}(x_0)$, by Proposition A.20 we obtain that $\xi_0 \in \partial \operatorname{conv} f_{\Omega_N}(x_0)$. Applying Lemma B.15 and Corollary B.27 we get

$$f^*_{\Omega_N}(\xi_0) = \left(\operatorname{conv} f_{\Omega_N}\right)^*(\xi_0) = f^*(\xi_0).$$

□

Remark 3.13. Theorem 3.12 is inspired by Proposition 2.3 of [7].

Corollary 3.14. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, let $\Omega_N, S_N \subset \mathbb{R}^n$ be finite sets, and let $x_0 \in \Omega_N$. Then $\partial \operatorname{conv} f_{\Omega_N}(x_0) \cap S_N \neq \emptyset$ if and only if $\left(f^*_{\Omega_N}\right)^*_{S_N}(x_0) = \operatorname{conv} f_{\Omega_N}(x_0)$.*

Proof. We have that $f^*_{\Omega_N}$ is a closed convex function and that $\left(f^*_{\Omega_N}\right)^* = \operatorname{cl}\operatorname{conv} f_{\Omega_N} = \operatorname{conv} f_{\Omega_N}$ by applications of Theorem B.17 and Theorem A.56. By Theorem 3.12 we get the thesis. □

Definition 3.15. Given $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\Omega_N \subset \mathbb{R}^n$ finite, we call an *optimal dual grid* any set S_N such that $\partial \operatorname{conv} f_{\Omega_N}(x) \cap S_N \neq \emptyset$ for every $x \in \Omega_N$

By Corollary 3.14 requiring that S_N is optimal is equivalent to asking that $\left(f^*_{\Omega_N}\right)^*_{S_N}(x) = \operatorname{conv} f_{\Omega_N}(x)$ for every $x \in \Omega_N$. It can be shown that such an optimal set always exists.

Corollary 3.16. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega_N \subset \mathbb{R}^n$ be finite. Then, there exists a dual grid $S_N \subset \mathbb{R}^n$ finite and optimal.*

Proof. Fix $x \in \Omega_N$. By Theorem A.56 we have that $\partial \operatorname{conv} f_{\Omega_N}(x) \neq \emptyset$; then we can find $\xi_x \in \partial \operatorname{conv} f_{\Omega_N}(x)$. Thus we can take $S_N = \{\xi_x \mid x \in \Omega_N\}$, which is finite by finiteness of Ω_N. □

Corollary 3.17. *Given $f : \mathbb{R} \to \overline{\mathbb{R}}$ not necessarily convex and $X_N \subset \mathbb{R}$ finite, we have that $C := \{c_1, \ldots, c_{n-1}\}$ is an optimal dual grid, where $\{c_i\}$ are the slopes of* $\operatorname{conv} I_{X_N}(f)$ *as defined in Theorem 2.23.*

Proof. By Proposition 2.21 we have $\operatorname{conv} f_{X_N} = \operatorname{conv} I_{X_N}(f)$ and thus for every $x \in X_N$ we have that $\partial \operatorname{conv} f_{X_N}(x) \cap C \neq \emptyset$. □

3.3. Optimal dual grids in the non-finite case

Now we extend Theorem 3.12 to the case in which Ω and S are not necessarily finite.

Theorem 3.18. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a closed convex function, let $\Omega \subseteq \mathbb{R}^n$, and let $\xi_0 \in \mathbb{R}^n$ be such that $\partial f^*(\xi_0) \cap \Omega \neq \emptyset$. Then, $f^*_\Omega(\xi_0) = f^*(\xi_0)$.*

Proof. Let $x_0 \in \partial f^*(\xi_0) \cap \Omega$ and consider the finite set $\Omega_N = \{x_0\}$. Being $\Omega_N \subseteq \Omega \subseteq \mathbb{R}^n$, by applying Proposition 2.16 we obtain

$$f^*_{\Omega_N}(\xi_0) \leq f^*_\Omega(\xi_0) \leq f^*(\xi_0);$$

since by Theorem 3.12 we have $f^*_{\Omega_N}(\xi_0) = f^*(\xi_0)$, we get the thesis. □

Example 3.19. In Theorem 3.12 we had a necessary and sufficient condition, whereas for Theorem 3.18 only one direction is true. Consider for example

$$f(x) = \begin{cases} 0 & \text{if } x \in [-1, 1] \\ +\infty & \text{elsewhere} \end{cases};$$

by Theorem 2.23 we have that

$$f^*(\xi) = |\xi| .$$

Fix $\xi_0 > 0$; we have $\partial f^*(\xi_0) = \{1\}$. Take $\Omega = \mathbb{R} \setminus \{1\}$ open; we have $f = \operatorname{cl} f_\Omega$ and thus by Corollary B.10 we obtain that $f^*_\Omega = f^*$. Consequently we have $f^*_\Omega(\xi_0) = f^*(\xi_0)$, but $\partial f^*(\xi_0) \cap \Omega = \emptyset$.

Corollary 3.20. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $\Omega, S \subseteq \mathbb{R}^n$. Let $x_0 \in \Omega$ such that $\partial \operatorname{cl}\operatorname{conv} f_\Omega(x_0) \cap S \neq \emptyset$; then $(f^*_\Omega)^*_S(x_0) = \operatorname{cl}\operatorname{conv} f_\Omega(x_0)$.*

Proof. Being a LFT, the function f^*_Ω is a closed convex function. By Theorem B.17 we have that $(f^*_\Omega)^* = \operatorname{cl}\operatorname{conv} f_\Omega$. Then, by applying Theorem 3.18 to the function f^*_Ω and to the set S, we get the thesis. □

Corollary 3.21. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, let $\Omega_N \subset \mathbb{R}^n$ be finite and let $S \subseteq \mathbb{R}^n$. Let $x_0 \in \Omega_N$ such that $\partial \operatorname{conv} f_{\Omega_N}(x_0) \cap S \neq \emptyset$; then $\left(f^*_{\Omega_N}\right)^*_S(x_0) = \operatorname{conv} f_{\Omega_N}(x_0)$.*

Proof. It follows from Corollary 3.20 after observing that $\operatorname{cl}\operatorname{conv} f_{\Omega_N} = \operatorname{conv} f_{\Omega_N}$ by Theorem A.56. □

3.4. Convergence to the discretized convex hull

There are two main difficulties in the use of optimal dual grids as defined in the previous section. The first and more obvious one is that we need to know something about $\operatorname{conv} f_{\Omega_N}$ before being able to find an optimal grid. The second is that for efficiency reasons the algorithm for the DLFT in two dimensions works on grids of the form $\Omega_N = X_N \times Y_N$ and $S_N = C_N \times D_N \subset \mathbb{R}^2$, whereas the optimal grid is not necessarily of this

form; we could build opportune C_N and D_N such that S_N contains the optimal grid, but in the worst case we would have $|S_N| = |\Omega_N|^2$. In the general case when S_N is not optimal, $\left(f^*_{\Omega_N}\right)^*_{S_N}(x)$ is only an approximation of $\operatorname{conv} f_{\Omega_N}(x)$, but by refining S_N this approximation becomes better, as we will show in Theorem 3.22.

Theorem 3.22. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, let Ω_M be a finite subset of $\mathbb{R}^n$ and let $S \subseteq \mathbb{R}^n$ be such that $S \subseteq \operatorname{cl}\operatorname{int} S$ and $\partial \operatorname{conv} f_{\Omega_M}(x) \cap S \neq \emptyset$ for every $x \in \Omega_M$. Let $(S_N)_N$ be a sequence of finite subsets of $\mathbb{R}^n$ such that $S_N \to S$. Then $\left(f^*_{\Omega_M}\right)^*_{S_N}(x) \to \operatorname{conv} f_{\Omega_M}(x)$ for every $x \in \Omega_M$.*

Proof. By Lemma 3.1 $\operatorname{dom} f^*_{\Omega_M} = \mathbb{R}^n$; then, by Theorem A.25, $f^*_{\Omega_M}$ is continuous on all $\mathbb{R}^n$. By applying Theorem 2.32 we obtain that $\underline{\left(\overline{\left(f^*_{\Omega_M}\right)_S}\right)} = \underline{\left(f^*_{\Omega_M}\right)_S}$. By applying Theorem 2.27 to the function $f^*_{\Omega_M}$ and the set S, we obtain that $\left(f^*_{\Omega_M}\right)^*_{S_N}(x) \to \left(f^*_{\Omega_M}\right)^*_S(x)$ for every $x \in \Omega_M$. Finally, from Corollary 3.21 we get that $\left(f^*_{\Omega_M}\right)^*_S(x) = \operatorname{conv} f_{\Omega_M}(x)$ for every $x \in \Omega_M$. □

By Corollary 3.16 a set S with the required properties always exists. Since the points which are required to be in S are finite, S can be chosen bounded; for instance we could choose S as a union of neighbourhoods of these points or, should we be interested in the value of $\operatorname{conv} f_{\Omega_M}$ in only one $x_0 \in \Omega_M$, as a neighbourhood of a point of $\partial \operatorname{conv} f_{\Omega_M}(x_0)$. All these points are unknown, but if we are able to choose S sufficiently large to include them all, we can achieve convergence by simply refining the grid on which the second DLFT is computed.

Corollary 3.23. *In the hypothesis of Theorem 3.22, let $x_0, \xi_0 \in \mathbb{R}^n$ be such that $\xi_0 \in \partial \operatorname{conv} f_{\Omega_M}(x_0) \cap S$. We then have*

$$\begin{aligned}\left|\operatorname{conv} f_{\Omega_M}(x_0) - \left(f^*_{\Omega_M}\right)^*_{S_N}(x_0)\right| \;\leq\; & \left\|\xi_0 - \hat{\xi}_N\right\| \|x_0\| + \\ & \left|f^*_{\Omega_M}(\hat{\xi}_N) - f^*_{\Omega_M}(\xi_0)\right|,\end{aligned}$$

where $\hat{\xi}_N = \arg\min_{\xi \in S_N} \|\xi_0 - \xi\|$.

Proof. Since $\xi_0 \in \partial \operatorname{conv} f_{\Omega_M}(x_0) = \partial \left(f^*_{\Omega_M}\right)^*(x_0)$, we have by Corollary B.23 that $x_0 \in \partial f^*_{\Omega_M}(\xi_0)$; thus by Lemma B.26 follows that $\left(f^*_{\Omega_M}\right)^*(x_0) = \langle x_0, \xi_0\rangle - f^*_{\Omega_M}(\xi_0)$ and, being $\xi_0 \in S$, by Theorem 3.12 we have $\left(f^*_{\Omega_M}\right)^*_S(x_0) = \left(f^*_{\Omega_M}\right)^*(x_0)$. Following the

proof of Theorem 2.27, we can then write

$$\begin{aligned} 0 &\leq \left(f^*_{\Omega_M}\right)^*_S(x_0) - \left(f^*_{\Omega_M}\right)^*_{S_N}(x_0) \\ &= \langle x_0, \xi_0\rangle - f^*_{\Omega_M}(\xi_0) - \max_{\xi\in S_N}\left[\langle x_0,\xi\rangle - f^*_{\Omega_M}(\xi)\right] \\ &\leq \langle x_0, \xi_0\rangle - f^*_{\Omega_M}(\xi_0) - \left\langle x_0, \hat{\xi}_N\right\rangle + f^*_{\Omega_M}(\hat{\xi}_N) \\ &\leq \left\langle \xi_0 - \hat{\xi}_N, x_0\right\rangle + f^*_{\Omega_M}(\hat{\xi}_N) - f^*_{\Omega_M}(\xi_0) \\ &\leq \left\|\xi_0 - \hat{\xi}_N\right\| \|x_0\| + \left|f^*_{\Omega_M}(\hat{\xi}_N) - f^*_{\Omega_M}(\xi_0)\right|, \end{aligned}$$

and the thesis is proved. □

Corollary 3.23 shows that the error in the approximation of $\operatorname{conv} f_{\Omega_M}$ at the point x_0 is determined only by the properties of the grid S_N in its vicinities: a finer grid around ξ_0 gives us a point $\hat{\xi}_N$ nearer to ξ_0 and by continuity of $f^*_{\Omega_M}$ we can also expect $\left|f^*_{\Omega_M}(\hat{\xi}_N) - f^*_{\Omega_M}(\xi_0)\right|$ to be smaller; the points of S_N far away from ξ_0 do not modify our estimate. Therefore, we can use this result in the case we are interested in the value of $\operatorname{conv} f_{\Omega_M}$ on all points of Ω_M but want a better approximation for a particular point: this objective can be achieved by using a fine dual grid around a subgradient of such point and a coarse grid elsewhere.

4. An algorithm for the convex hull of two-dimensional functions

The multi-dimensional DLFT can be reduced by factorization to several one-dimensional DLFTs; since we are interested in the convex hull of the bulk free-energy density (see Section 1.5), which is a two-dimensional function, we study this case in detail. Then we consider a factorized double DLFT and we show that a better approximation of the convex hull can be obtained by swapping dimensions in the second transform; the symmetry properties of this modified transformation are also studied. The factorization is used as a basis of an algorithm for computing the two-dimensional convex hull, using as the fundamental building block the fast, i.e. $O(n)$, DLFT algorithm by Lucet (10); the algorithms based respectively on the standard and modified factorization are then compared. Finally, we propose an adaptive dual grid construction and explain how dual grids of arbitrary length can be used without incurring in an increased memory usage.

4.1. Factorization of the DLFT

The DLFT, being a special case of LFT, can be factorized in a similar way. For simplicity, we will restrict to the case $n = 2$, which is the case of interest in the physical model presented in Chapter 1. Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ and $\Omega \subset \mathbb{R}^2$, not necessarily finite; again for the sake of simplicity, we will suppose that $\Omega = X \times Y$. By Theorem B.18 we have

$$f_\Omega^* = \left[- \left(f_\Omega \right)^{*1} \right]^{*2}.$$

Since for every $y \in \mathbb{R}$ we can have $f_\Omega(x, y) < +\infty$ only if $x \in X$, we obtain

$$\begin{aligned} g(\xi, y) := (f_\Omega)^{*1}(\xi, y) &= [f_\Omega(\cdot, y)]^* (\xi) \\ &= [f(\cdot, y)]_X^* (\xi); \end{aligned}$$

in particular, when $y \notin Y$, we have that $f_\Omega(\cdot, y) \equiv +\infty$ and thus $g(\cdot, y) \equiv -\infty$. Finally we obtain that

$$\begin{aligned} f_\Omega^*(\xi, \eta) &= (-g)^{*2} (\xi, \eta) \\ &= [-g(\xi, \cdot)]^* (\eta) \\ &= [-g(\xi, \cdot)]_Y^* (\eta), \end{aligned}$$

where the last equality follows from the fact that, fixed ξ, the function $-g(\xi, \cdot)$ is $+\infty$ outside of Y.

In the case of the DLFT, i.e. when $\Omega = \Omega_N = X_N \times Y_N$, where $X_N = \{x_1, \ldots, x_n\}$ and $Y_N = \{y_1, \ldots, y_m\}$, we then have

$$f^*_{\Omega_N}(\xi, \eta) = [-g(\xi, \cdot)]^*_{Y_N}(\eta),$$

where

$$g(\xi, y) = [f(\cdot, y)]^*_{X_N}(\xi). \tag{4.1}$$

For any $y \in Y_N$ the function $g(\cdot, y)$ is a closed convex piecewise linear function whose domain is the entire real line and whose nodes are the slopes of conv $f_{\Omega_N}(\cdot, y)$; for every $\xi \in \mathbb{R}$ the function $f^*_{\Omega_N}(\xi, \cdot)$ is a piecewise linear function whose domain is the whole real line and whose slopes are a subset of Y_N, in particular they are the nodes of conv $[-g(\xi, \cdot)]$. Moreover, we could swap the ordering of the directions in the factorization and obtain the same resulting $f^*_{\Omega_N}$; this means that for every $\eta \in \mathbb{R}$ the function $f^*_{\Omega_N}(\cdot, \eta)$ is a piecewise linear function whose domain is the whole real line and whose slopes are a subset of X_N.

Now we have a factorization of the two-dimensional DLFT in one-dimensional DLFTs. In particular, in order to compute $f^*_{\Omega_N}(\xi, \eta)$ for a given $(\xi, \eta) \in \mathbb{R}^2$ we need to compute $g(\xi, y_j)$ for every $j = 1, \ldots, m$ using m one-dimensional DLFTs along the x-direction (each on data whose length is n) and then applying one one-dimensional DLFT along the y-direction (on data long m); if we want to compute the transform on a grid of size $n \times m$, we then have in total m DLFTs on data of size n and n DLFTs on data of size m. Having the one-dimensional DLFT algorithm linear complexity, the complexity of a two-dimensional algorithm based on this factorization is $O(nm)$, again linear in time.

4.2. Factorization of the double DLFT with alternating dimensions

Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ and let $\Omega_N = X_N \times Y_N \subset \mathbb{R}^2$ and $S_N = C_N \times D_N \subset \mathbb{R}^2$ both finite; by proceeding as in the previous section, we can compute $\left(f^*_{\Omega_N}\right)^*_{S_N}(x, y)$. However, now we decompose the second DLFT as

$$\left(f^*_{\Omega_N}\right)^*_{S_N} = \left[-\left(f^*_{\Omega_N}\right)^{*2}\right]^{*1},$$

obtaining that

$$\left(f^*_{\Omega_N}\right)^*_{S_N}(x, y) = [-h(\cdot, y)]^*_{C_N}(x), \tag{4.2}$$

where

$$\begin{aligned} h(\xi, y) &= \left[f^*_{\Omega_N}(\xi, \cdot)\right]^*_{D_N}(y) \\ &= \left[(-g(\xi, \cdot))^*_{Y_N}\right]^*_{D_N}(y). \end{aligned} \tag{4.3}$$

This means that $h(\xi, \cdot)$ is an approximation of the convex hull of $(-g(\xi, \cdot))_{Y_N}$; we want to know if substituting h with the true convex hull (which is easily calculable in one-dimension) gives us a better final result. Since by Theorem B.17 and Theorem A.56

$$\text{conv}\left[(-g(\xi, \cdot))_{Y_N}\right] = \left[(-g(\xi, \cdot))^*_{Y_N}\right]^*,$$

we have that this new approximation for conv f_{Ω_N} is the function $\left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}}$; we now show that it is a better approximation than $\left(f^*_{\Omega_N}\right)^*_{S_N}$.

Proposition 4.1. *Let $f:\mathbb{R}^2\to\overline{\mathbb{R}}$ and let $\Omega_N = X_N\times Y_N\subset\mathbb{R}^2$ and $S_N = C_N\times D_N\subset\mathbb{R}^2$ be both finite. Then*

$$\left(f^*_{\Omega_N}\right)^*_{S_N} \le \left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}} \le \operatorname{conv} f_{\Omega_N}.$$

Proof. It is a straightforward consequence of Corollary 3.5. □

One possible drawback of $\left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}}$ is that the transformation no longer preserves symmetry, due to the fact that the computation of $\left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}}$ is not invariant under swap of dimensions. Let $g^T(x,y) := g(y,x)$ for every $x,y\in\mathbb{R}^2$ and any function $g:\mathbb{R}\to\overline{\mathbb{R}}$. Being the order in which the dimensions are processed irrelevant to the final result, we have that

$$\left(f^*_{X_N\times Y_N}\right)^*_{C_N\times D_N} = \left(\left(\left(f^T\right)^*_{Y_N\times X_N}\right)^*_{D_N\times C_N}\right)^T;$$

in particular, when $f = f^T$ (i.e. the graph of f is symmetric with respect to the plane $x=y$), $X_N = Y_N$ and $C_N = D_N$ we also have that

$$\left(f^*_{X_N^2}\right)^*_{C_N^2} = \left(\left(f^*_{X_N^2}\right)^*_{C_N^2}\right)^T,$$

which is a desiderable property since the exact convex hull itself satisfies it (as we will prove in Corollary 4.6). In the general case, i.e. when the subsets of $\mathbb{R}^2$ on which the transform is computed are not factorizable, the same results hold.

Definition 4.2. Let $A\subseteq\mathbb{R}^2$. We will denote by A^T the set $\{(x,y)\in\mathbb{R}^2 \mid (y,x)\in A\}$.

Proposition 4.3. *Let $f:\mathbb{R}^2\to\overline{\mathbb{R}}$ and let $\Omega\subseteq\mathbb{R}^2$. Then we have*

$$f^*_\Omega = \left(\left(f^T\right)^*_{\Omega^T}\right)^T.$$

Proof. We have that

$$\begin{aligned}\left(f^T\right)^*_{\Omega^T}(\xi,\eta) &= \sup_{(x,y)\in\Omega^T}\left[x\xi+y\eta-f^T(x,y)\right]\\ &= \sup_{(y,x)\in\Omega}\left[y\eta+x\xi-f(y,x)\right]\\ &= f^*_\Omega(\eta,\xi).\end{aligned}$$

□

Corollary 4.4. *Let $f:\mathbb{R}^2\to\overline{\mathbb{R}}$ and let $\Omega, S\subseteq\mathbb{R}^2$. Then we have that*

$$\left(f^*_\Omega\right)^*_S = \left(\left(\left(f^T\right)^*_{\Omega^T}\right)^*_{S^T}\right)^T.$$

Proof. It follows from applying Proposition 4.3 twice. □

Corollary 4.5. *Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ such that $f^T = f$ and let $\Omega, S \subseteq \mathbb{R}^2$ such that $\Omega^T = \Omega$ and $S^T = S$. Then we have that*

$$(f^*_\Omega)^*_S = \left((f^*_\Omega)^*_S\right)^T .$$

Corollary 4.6. *Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ such that $f^T = f$. Then*

$$\operatorname{cl}\operatorname{conv} f = (\operatorname{cl}\operatorname{conv} f)^T .$$

In the case $D = \mathbb{R} \neq C_N$, this reasoning is no longer applicable and $\left(f^*_{X^2_N}\right)^*_{C_N\times\mathbb{R}}$ is not symmetric, even when f is such; nonetheless it can be applied to the transformation $\left(f^*_{X^2_N}\right)^*_{(C_N\times\mathbb{R})\cup(\mathbb{R}\times C_N)}$. This last transformation is easily obtained from transformations of the type $\left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}}$ thanks to the Proposition 2.17, which gives

$$\left(f^*_{\Omega_N}\right)^*_{(C_N\times\mathbb{R})\cup(\mathbb{R}\times D_N)} = \max\left\{\left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}}, \left(f^*_{\Omega_N}\right)^*_{\mathbb{R}\times D_N}\right\};$$

in addition the result is more accurate than the previous versions of the transformation, again by Proposition 2.16.

4.3. Estimation of the dual set S

In this section we provide an estimate for a set $S = C \times D$ containing an optimal dual grid for the two-dimensional problem. We begin with some results linking the subdifferentiability properties of one-dimensional convex functions (see Section A.5) with the general case.

Definition 4.7. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, let $x_0 \in \mathbb{R}^n$ and let $u \in \mathbb{R}^n$ such that $\|u\| = 1$. Then

$$f'_u(x_0) := \lim_{\lambda\to 0^+} \frac{f(x_0 + \lambda u) - f(x_0)}{\lambda}$$

is called the *one-sided directional derivative* of f at x_0 with respect to u.

Proposition 4.8. *If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex and $x_0 \in \operatorname{dom} f$ then $f'_u(x_0)$ exists for every $u \in \mathbb{R}^n$ such that $\|u\| = 1$. Moreover, if we also have that $x_0 \in \operatorname{int}\operatorname{dom} f$, then $f'_u(x_0)$ is finite.*

Proof. The restriction of f to the direction of u, i.e. the one-dimensional function $g(\lambda) := f(x_0 + \lambda u)$, is a convex function by convexity of f; in particular we have that $g'_+(0) = f'_u(x_0)$. Then the thesis follows from the differentiability properties of one-dimensional convex functions (see Theorem A.30). □

Theorem 4.9. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ a convex function, let $x_0 \in \operatorname{dom} f$ and let $\xi \in \mathbb{R}^n$. Then we have $f'_u(x_0) \geq \langle \xi, u\rangle$ for all $u \in \mathbb{R}^n$ such that $\|u\| = 1$ if and only if $\xi \in \partial f(x_0)$.*

Proof. Suppose $\xi \in \partial f(x_0)$. This is equivalent to requiring that for all $\lambda > 0$ and $u \in \mathbb{R}^n$ such that $\|u\| = 1$ the following inequality holds

$$\begin{aligned} f(x_0) + \langle \xi, x_0 + \lambda u - x_0 \rangle &\leq f(x_0 + \lambda u) \\ \langle \xi, u \rangle &\leq \frac{f(x_0 + \lambda u) - f(x_0)}{\lambda}; \end{aligned} \tag{4.4}$$

by taking the limit for $\lambda \to 0^+$ we obtain that $f'_u(x_0) \geq \langle \xi, u \rangle$ for every $u \in \mathbb{R}^n$ such that $\|u\| = 1$.

Now suppose that $f'_u(x_0) \geq \langle \xi, u \rangle$ for every $u \in \mathbb{R}^n$ such that $\|u\| = 1$. Consider the function $\lambda \mapsto f(x_0 + \lambda u)$, which is convex by convexity of f; then the function $\lambda \mapsto \frac{f(x_0+\lambda u)-f(x_0)}{\lambda}$ is increasing (see Theorem A.27). This implies that $f'_u(x_0) \leq \frac{f(x_0+\lambda u)-f(x_0)}{\lambda}$ for every $\lambda > 0$ and every $u \in \mathbb{R}^n$ such that $\|u\| = 1$; thus we have obtained inequality (4.4) which is equivalent to $\xi \in \partial f(x)$. □

Corollary 4.10. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function, let $x_0 \in \operatorname{dom} f$ and let $u \in \mathbb{R}^n$ be such that $\|u\| = 1$. Let $\alpha, \beta > 0$ be such that $x_\alpha := x_0 - \alpha u, x_\beta := x_0 + \beta u \in \operatorname{dom} f$. Then for every $\xi_0 \in \partial f(x_0)$, $\xi_\alpha \in \partial f(x_\alpha)$ and $\xi_\beta \in \partial f(x_\beta)$ we have that*

$$\langle \xi_\alpha, u \rangle \leq \langle \xi_0, u \rangle \leq \langle \xi_\beta, u \rangle .$$

Proof. Consider again the convex function $g(\lambda) := f(x_0 + \lambda u)$; the following equalities hold: $g'_+(0) = f'_u(x_0)$, $g'_-(0) = -f'_{-u}(x_0)$, $g'_+(\alpha) = f'_u(x_\alpha)$ and $g'_-(\beta) = -f'_{-u}(x_\beta)$. Since for a one-dimensional convex function the derivative is increasing (see Corollary A.34), we have that

$$g'_+(\alpha) \leq g'_-(0) \leq g'_+(0) \leq g'_-(\beta).$$

From Theorem 4.9 we obtain that $g'_+(\alpha) \geq \langle \xi_\alpha, u \rangle$, $g'_-(0) \leq \langle \xi_0, u \rangle$, $g'_+(0) \geq \langle \xi_0, u \rangle$ and $g'_-(\beta) \leq \langle \xi_\beta, u \rangle$. By combining all the inequalities the thesis follows. □

The following propositions give a method for the estimation of the subgradient inside the domain based on the subdifferential on the border of the domain.

Proposition 4.11. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function, let $u \in \mathbb{R}^n$ be such that $\|u\| = 1$ and let $x_0 \in \operatorname{dom} f$. Suppose for simplicity that $\operatorname{dom} f$ is closed and bounded. We then have that for every $\xi \in \partial f(x_0)$*

$$\inf_{x \in \partial \operatorname{dom} f} \sup_{\xi' \in \partial f(x)} \langle \xi', u \rangle \leq \langle \xi, u \rangle \leq \sup_{x \in \partial \operatorname{dom} f} \inf_{\xi' \in \partial f(x)} \langle \xi', u \rangle .$$

Proof. It is a straightforward application of Corollary 4.10, where we take α and β such that x_α and x_β belong to the border of $\operatorname{dom} f$. □

Proposition 4.12. *Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ and let $\Omega_N = X_N \times Y_N$ be a finite subset of $\mathbb{R}^2$. We will suppose for simplicity that f is finite on Ω_N. Then for every $(x, y) \in \Omega_N$ and every*

$(\xi, \eta) \in \partial \operatorname{conv} f_{\Omega_N}(x, y)$ *we have that*

$$\inf_{y' \in Y_N} \sup_{\substack{(\xi', \eta') \\ \in \\ \partial \operatorname{conv} f_{\Omega_N}(\min X_N, y')}} \xi' \leq \xi \leq \sup_{y' \in Y_N} \inf_{\substack{(\xi'', \eta'') \\ \in \\ \partial \operatorname{conv} f_{\Omega_N}(\max X_N, y')}} \xi'',$$

$$\inf_{x' \in X_N} \sup_{\substack{(\xi', \eta') \\ \in \\ \partial \operatorname{conv} f_{\Omega_N}(x', \min Y_N)}} \eta' \leq \eta \leq \sup_{x' \in X_N} \inf_{\substack{(\xi'', \eta'') \\ \in \\ \partial \operatorname{conv} f_{\Omega_N}(x', \max Y_N)}} \eta''.$$

Proof. Let $(x, y) \in \Omega_N$ and let $(\xi, \eta) \in \partial \operatorname{conv} f_{\Omega_N}(x, y)$. By applying Corollary 4.10, where we take $u = (1, 0)$, we obtain that

$$\xi' \leq \xi \leq \xi''$$

holds for every $(\xi', \eta') \in \partial \operatorname{conv} f_{\Omega_N}(\min X_N, y)$ and every $(\xi'', \eta'') \in \partial \operatorname{conv} f_{\Omega_N}(\max X_N, y)$. Then we have that

$$\sup_{(\xi', \eta') \in \partial \operatorname{conv} f_{\Omega_N}(\min X_N, y)} \xi' \leq \xi \leq \inf_{(\xi'', \eta'') \in \partial \operatorname{conv} f_{\Omega_N}(\max X_N, y)} \xi''$$

and by taking on each side respectively the infimum and the supremum over $y \in Y_N$ we obtain the first part of the thesis; the second part can be proved similarly. □

Now we can apply Proposition 4.11 to the problem of finding a set S which contains at least one optimal grid in the case of the factorized two-dimensional transformations; it is firstly necessary to provide estimates of $\partial \operatorname{conv} f$ on $\Omega_N \cap \partial \operatorname{conv} \Omega_N$.

Lemma 4.13. *Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ and let $\Omega_N = X_N \times Y_N$ be a finite subset of $\mathbb{R}^2$. We suppose for simplicity that f is finite on Ω_N. Given $y \in \mathbb{R}$, let $g_y := f_{\Omega_N}(\cdot, y)$, $\xi_y^- := \max \partial \operatorname{conv} g_y(\min X_N)$ and $\xi_y^+ := \min \partial \operatorname{conv} g_y(\max X_N)$. Then we have that for every $y \in Y_N$ there exist $\eta_y^-, \eta_y^+ \in \mathbb{R}$ such that*

$$\begin{aligned} (\xi^-, \eta_y^-) &\in \partial \operatorname{conv} f_{\Omega_N}(\min X_N, y), \\ (\xi^+, \eta_y^+) &\in \partial \operatorname{conv} f_{\Omega_N}(\max X_N, y), \end{aligned}$$

where $\xi^- = \min_{y \in Y_N} \xi_y^-$ and $\xi^+ = \max_{y \in Y_N} \xi_y^+$.

Proof. We consider only the boundary $\min X_N \times Y_N$; the proof for the other boundary is the same. Firstly we prove that ξ_y^- is well defined for every $y \in Y_N$. Since f is finite on Ω_N, for each $y \in Y_N$ the function g_y is finite on X_N and is $+\infty$ elsewhere; its convex hull is then a piecewise linear function finite on the interval $[\min X_N, \max X_N]$ and thus ξ_y^- exists and is equal to the right slope in $\min X_N$ (see Proposition A.35). We observe that being $\xi_y^- \in \partial \operatorname{conv} g_y(\min X_N)$ we have that for every $x \in \mathbb{R}$

$$\operatorname{conv} g_y(\min X) + \xi_y^-(x - \min X_N) \leq \operatorname{conv} g_y(x). \tag{4.5}$$

By convexity of $\operatorname{conv} f_{\Omega_N}$ we have that $\operatorname{conv} f_{\Omega_N}(x,y) \le \operatorname{conv} g_y(x)$ for every $(x,y) \in \mathbb{R}^2$ and thus $\operatorname{conv} f_{\Omega_N} = \operatorname{conv} g$, where $g(x,y) := \operatorname{conv} g_y(x)$ (observe that $g(x,y) = +\infty$ if $y \notin Y_N$); then inequality (4.5) becomes

$$g(\min X, y) + \xi_y^-(x - \min X) \le g(x,y) \quad \forall (x,y) \in \mathbb{R} \times Y_N. \tag{4.6}$$

Moreover, by convexity of $\operatorname{conv} g$ we have that $\operatorname{conv} g(\min X_N, \cdot) \le \operatorname{conv} h^-$, where $h^- = g(\min X_N, \cdot)$; observe that, being $h^- < +\infty$ only on Y_N, we have that $\operatorname{conv} h^- = I_{Y_N'}(h^-)$ for a certain $Y_N' \subseteq Y_N$. We have that $\operatorname{conv} g = \operatorname{conv} \tilde{g}$, where

$$\tilde{g}(x,y) := \begin{cases} \operatorname{conv} h^-(y) & \text{if } x = \min X_N \\ g(x,y) & \text{elsewhere} \end{cases}.$$

Since $g = \tilde{g}$ on $(\mathbb{R} \setminus \{\min X_N\}) \times \mathbb{R}$ and $\tilde{g} \le g$ on $\{\min X_N\} \times \mathbb{R}$, inequality (4.6) implies that

$$\tilde{g}(\min X_N, y) + \xi_y^-(x - \min X) \le \tilde{g}(x,y) \quad \forall (x,y) \in \mathbb{R} \times Y_N. \tag{4.7}$$

Now fix $y_0 \in Y_N$. Since $\tilde{g}(\min X_N, \cdot) = \operatorname{conv} h^-$ is convex and piecewise linear, there exists $\eta_{y_0} \in \partial \operatorname{conv} h^-(y_0)$; for such η_{y_0} we have that

$$\tilde{g}(\min X_N, y_0) + \eta_{y_0}(y - y_0) \le \tilde{g}(\min X_N, y) \quad \forall y \in \mathbb{R}. \tag{4.8}$$

By combining inequalities (4.7) and (4.8) we obtain that for every $(x,y) \in X_N \times Y_N$

$$\tilde{g}(\min X_N, y_0) + \xi_y^-(x - \min X_N) + \eta_{y_0}(y - y_0) \le \tilde{g}(x,y);$$

being $\xi^- \le \xi_y^-$ for every $y \in Y_N$ and $x - \min X \ge 0$ for every $x \in X_N$, we have that

$$\tilde{g}(\min X_N, y_0) + \xi^-(x - \min X_N) + \eta_{y_0}(y - y_0) \le \tilde{g}(x,y).$$

By Proposition A.20 we conclude that $(\xi^-, \eta_{y_0}) \in \partial \operatorname{conv} \tilde{g}(\min X_N, y_0)$ and thus $(\xi^-, \eta_{y_0}) \in \partial \operatorname{conv} f_{\Omega_N}(\min X_N, y_0)$ by equality of the two hulls. □

Lemma 4.14. *Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ and let $\Omega_N = X_N \times Y_N$ be a finite subset of $\mathbb{R}^2$. We suppose for simplicity that f is finite on Ω_N. Given $x \in \mathbb{R}$, let $h_x := f_{\Omega_N}(x, \cdot)$, $\eta_x^- := \max \partial \operatorname{conv} h_x(\min Y_N)$ and $\eta_x^+ := \min \partial \operatorname{conv} h_x(\max Y_N)$. Then we have that for every $x \in X_N$ there exist $\xi_x^-, \xi_x^+ \in \mathbb{R}$ such that*

$$\begin{aligned} (\xi_x^-, \eta^-) &\in \partial \operatorname{conv} f_{\Omega_N}(x, \min Y_N), \\ (\xi_x^+, \eta^+) &\in \partial \operatorname{conv} f_{\Omega_N}(x, \max Y_N), \end{aligned}$$

where $\eta^- = \min_{x \in X_N} \eta_x^-$ and $\eta^+ = \max_{x \in X_N} \eta_x^+$.

By combining Lemmas 4.13 and 4.14 with Proposition 4.12 we have an estimate for the set S.

Theorem 4.15. *With the hypothesis and notation of Lemmas 4.13 and 4.14, we have that the set $S = C \times D := [\xi^-, \xi^+] \times [\eta^-, \eta^+]$ contains an optimal dual grid relative to the primal grid Ω_N. In particular this means that for all $(x, y) \in \Omega_N$*

$$\operatorname{conv} f_{\Omega_N}(x, y) = \left(f^*_{\Omega_N}\right)^*_S(x, y).$$

Proof. Fix $(x, y) \in \Omega_N$; by Lemma A.56 there exists a subgradient $(\xi, \eta) \in \partial \operatorname{conv} f_{\Omega_N}(x, y)$. By Lemmas 4.13 and 4.14, we have that for all $x' \in X_N$ and all $y' \in Y_N$

$$\begin{aligned}
\sup_{(\xi',\eta') \in \partial \operatorname{conv} f_{\Omega_N}(\min X_N, y')} &\geq \xi^-, \\
\inf_{(\xi'',\eta'') \in \partial \operatorname{conv} f_{\Omega_N}(\max X_N, y')} &\leq \xi^+, \\
\sup_{(\xi',\eta') \in \partial \operatorname{conv} f_{\Omega_N}(x', \min Y_N)} &\geq \eta^-, \\
\inf_{(\xi'',\eta'') \in \partial \operatorname{conv} f_{\Omega_N}(x', \max Y_N)} &\leq \eta^+.
\end{aligned}$$

We can then apply Proposition 4.12 and obtain our thesis. □

4.4. Implementation details

As we have seen in Section 4.1, the basic building block for the multidimensional DLFT is the one-dimensional DLFT. We will use the fast algorithm for the one-dimensional DLFT developed by Lucet in 10, which has complexity $O(n)$; older algorithms inspired by the FFT have complexity $O(n \log n)$ ([9]). In 10 the extension to the multidimensional DLFT is also treated. An explicit convex hull algorithm based on Lucet's multidimensional DLFT is found in 6; in the following, we will call it the standard algorithm for the convex hull. We restrict to the two-dimensional case, mainly because the function we consider in the physical problem is two-dimensional (see Section 1.4) and because generalization to any number of dimensions is straightforward. Moreover, we present a modified version of the standard algorithm which takes into account the results of Section 4.2.

Firstly we describe the fast one-dimensional DLFT algorithm by Lucet (10). Let us consider a function $f : \mathbb{R} \to \overline{\mathbb{R}}$, a grid $X_N = \{x_1, \ldots, x_N\}$ and the set of values $\{f(x_1), \ldots, f(x_N)\}$. By Corollary 2.22 we have that $f^*_{X_N} = [\operatorname{conv} I_{X_N}(f)]^*$. We assume that the points of X_N are already ordered. Thus $\operatorname{conv} I_{X_N}(f)$ is easily calculable in linear time through a Graham's scan ([5, 1]) or a divide et impera approach ([12]); moreover, $\operatorname{conv} I_{X_N}(f)$ is still a piecewise linear function. Then, by Theorem 2.23, where we take $c_0 = -\infty$ and $c_N = +\infty$, the DLFT $f^*_{X_N}$ is a piecewise linear function and its nodes, values and external slopes can be easily computed by (2.2). A pseudocode description of this algorithm follows (`fast_dlft`). We denote by `convexhull_1d(grid, values)` the subroutine which computes the convex hull of the piecewise linear interpolant on the nodes `grid` and the values `values`. Moreover, we denote by `pcw_dlft(grid, values)` the subroutine which computes the DLFT of a convex interpolant. Both subroutines

return a piecewise linear function, which is represented by an object `pcw` whose components are `pcw.grid`, `pcw.values`, `pcw.left_slope` and `pcw.right_slope`. The last two values, which represents the external slopes, are infinite in the convex hull case and finite in the DLFT case.

```
function fast_dlft(grid, values)
  pcw = convexhull_1d(grid, values)
  pcw = pcw_dlft(pcw.grid, pcw.values)
  return pcw
end function
```

In what follows we denote by `pcw(grid)` the evaluation of the piecewise linear function on the points of `pcw(grid)`.

By the factorization of the DLFT presented in Section 4.1, the fast one-dimensional DLFT algorithm can be used as the basis for the computation of the multidimensional DLFT (10). Then we can apply the resulting algorithm twice, obtaining the convex hull algorithm presented in 6, which we describe in the two-dimensional case. Let us consider a function $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$, a primal grid $\Omega_N = X_N \times Y_N$, a dual grid $S_N = C_N \times D_N$ and suppose that the values of f on the grid Ω_N are given. By factorizing the DLFT and applying the fast one-dimensional algorithm, we can evaluate $f^*_{\Omega_N}$ on the grid S_N and then evaluate $(f^*_{\Omega_N})^*_{S_N}$ on the grid Ω_N. If the grids Ω_N and S_N have the same size, the matrix which initially contains the values of f on the grid Ω_N can be used for all the successive computations: we substitute each row with its DLFT and then do the same for the columns of the matrix so obtained. If the sizes of the grids differ, it is still possible to use one matrix by choosing it sufficiently large to accomodate both grids.

The grid S_N can be chosen as a uniform discretization of the grid $S = C \times D$ given in Theorem 4.15. We notice that C and D can be easily found easily by computing the interval containing all the natural grids of the one-dimensional DLFTs respectively along the rows and along the columns, since in one dimension those grids correspond to the slopes of the convex hull. We observe that, since we are computing the double DLFT $(f^*_{\Omega_N})^*_{S_N}$ exactly (except for numerical errors), the output data of the algorithm, i.e. the approximation of conv f_Ω, is always convex (being a Legendre-Fenchel transform) and less than or equal to the input data (by Corollary 3.7); the same is true also for the modified transforms.

Now we present the pseudocode for the standard two-dimensional convex hull algorithm (`convexhull_std`). We use an array notation where indexes start from 1 and where the index -1 is a shortcut for the last index; the function `size(array)` gives the size of an array. Moreover, we denote by `linspace(start,stop,N)` the grid which discretizes the interval [`start`, `end`] with `N` points evenly spaced. Finally, we denote by the keyword `parallel` the loops whose iterations are independent and which we compute in parallel.

```
function convexhull_std(xgrid, ygrid, values)
  ! compute the DLFT along each column
  lwb = +Inf
  upb = -Inf
  parallel do j = 1,size(ygrid)
```

```
    pcws(j) = fast_dlft(xgrid, values(:,j))
    lwb = min(lwb, pcws(j).grid(1))
    upb = max(upb, pcws(j).grid(-1))
  end do
  ! build the dual grid C_N
  ! and evaluate each of the DLFTs on it
  cgrid = linspace(lwb, upb, size(xgrid))
  pallel do j = 1,size(ygrid)
    values(:,j) = -pcws(j)(cgrid)
  end do
  ! now compute the DLFT
  ! along each row of the updated matrix
  lwb = +Inf
  upb = -Inf
  parallel do i = 1,size(cgrid)
    pcws(i) = fast_dlft(ygrid, values(i,:))
    lwb = min(lwb, pcws(i).grid(1))
    upb = max(upb, pcws(i).grid(-1))
  end do
  ! build the dual grid D_N
  ! and evaluate each of the DLFTs on it
  dgrid = linspace(lwb, upb, size(ygrid))
  parallel do i = 1,size(cgrid)
    values(i,:) = pcws(i)(dgrid)
  end do
  ! we have thus built the two-dimensional DLFT
  ! now compute the second 2d DLFT on Omega_N
  parallel do j = 1,size(dgrid)
    pcw = fast_dlft(cgrid, values(:,j))
    values(:,j) = -pcw(xgrid)
  end do
  parallel do i = 1,size(xgrid)
    pcw = fast_dlft(dgrid, values(i,:))
    values(i,:) = pcw(ygrid)
  end do
  return values
end function
```

In order to build an algorithm which computes the modified double DLFT $\left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}}$, it is sufficient to adapt the standard algorithm by condensing two one-dimensional DLFTs in a single application of the convex hull operation; thus the computation of $\left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}}$ is not only slightly better (by Theorem 4.1), but also slightly faster (see, for instance, the results in Table 4.3).

The same idea can be applied even if the number of dimensions is higher than two, but the improvements in speed and precision could be the less relevant the higher the number of dimensions, since the number of loops is always reduced by one independently from the number of dimensions.

As we have seen $\left(f^*_{\Omega_N}\right)^*_{(C_N\times\mathbb{R})\cup(\mathbb{R}\times D_N)}$ can be computed from $\left(f^*_{\Omega_N}\right)^*_{C_N\times\mathbb{R}}$ and $\left(f^*_{\Omega_N}\right)^*_{\mathbb{R}\times D_N}$; in order to compute $\left(f^*_{\Omega_N}\right)^*_{\mathbb{R}\times D_N}$ it is not necessary to implement a new function, because by Corollary 4.4 we can change the order of the dimension (i.e. transpose the data matrices) and compute it as $\left(\left(\left(f^T\right)^*_{Y_N\times X_N}\right)^*_{D_N\times\mathbb{R}}\right)^T$. Finally we present the pseudocodes of the two variants of the algorithm (respectively `convexhull_mod` and `convexhull_sym`).

```
function convexhull_mod(xgrid, ygrid, values)
  ! compute the DLFT along each column
  lwb = +Inf
  upb = -Inf
  parallel do j = 1,size(ygrid)
    pcws(j) = fast_dlft(xgrid, values(:,j))
    lwb = min(lwb, pcws(j).grid(1))
    upb = max(upb, pcws(j).grid(-1))
  end do
  ! build the dual grid C_N
  ! and evaluate each of the DLFTs on it
  cgrid = linspace(lwb, upb, size(xgrid))
  pallel do j = 1,size(ygrid)
    values(:,j) = -pcws(j)(cgrid)
  end do
  ! up to now it is identical to
  ! the standard algorithm;
  ! now we compute the 2nd and 3rd steps
  ! of the std algorithm in only one loop
  parallel do i = 1,size(cgrid)
    pcw = convexhull_1d(ygrid, values(i,:))
    values(i,:) = -pcw(ygrid)
  end do
  ! and finally the last pass of DLFTs
  parallel do j = 1,size(ygrid)
    pcw = fast_dlft(cgrid, values(:,j))
    values(:,j) = pcw(xgrid)
  end do
  return values
end function
function convexhull_sym(xgrid, ygrid, values)
  newValues = convexhull_mod(xgrid, ygrid, values)
```

```
  ! now we swap dimensions
  values = transpose(values)
  symValues = convexhull_mod(ygrid, xygrid, values)
  symValues = transpose(symValues)
  ! finally we merge the results with a maximum
  return max(newValues, symValues)
end function
```

4.5. Numerical comparisons

In this section we present some numerical tests to compare the three variants of the algorithm presented in the previous section, which we denote by dDLFT ("double DLFT"), mdDLFT ("modified" dDLFT) and smdDLFT ("symmetric" mdDLFT) respectively. For simplicity we test the algorithms on functions of the form $f(x, y) = f(r)$, where r is the distance from the origin of the point (x, y), i.e. $r = \sqrt{x^2 + y^2}$. The first test function (see Figure 4.1) we use is

$$f^1_{\text{test}}(x, y) = (r^2 - 1)^2,$$

whose convex hull is

$$\operatorname{conv} f^1_{\text{test}}(x, y) = \begin{cases} (r^2 - 1)^2 & \text{if } r > 1 \\ 0 & \text{if } 0 \leq r \leq 1 \end{cases};$$

the grid on which the function is evaluated has 1000×1000 points uniformly distributed on the square $[-1.5, 1.5]^2$. A simple way to assess the quality of the obtained results is by comparing the regions the different algorithms have classified as non-convex, i.e. where the computed convex hull differs from the value of the function; in the case of the exact hull this region is the unit circle. The regions detected by the dDLFT, mdDLFT and smdDLFT are shown in white in Figure 4.2; the loss of symmetry due to the mdDLFT is visible in the different shape of the region around the two axes. In Figure 4.3 the error between the computed hull and the exact hull is shown (here the asymmetric behaviour of mdDLFT is even more evident), whereas in Table 4.1 some quantitative information about the same error is reported. Finally we analyze the restriction of the transforms to the axes (see Figure 4.4). As an experimental proof for Corollary 3.14, we have also studied the case in which the value 0 is inserted in the dual grid manually; being 0 the value of the gradient in the non-convex region, we obtain an almost exact result inside the region of non-convexity as predicted by Corollary 3.14.

The second test function (see Figure 4.5) is

$$f^2_{\text{test}}(x, y) = \exp r + 25 \cdot \sin(2.5 - r) \cdot \exp\left[-(2.5 - r)^2\right],$$

and, by observing that

$$\operatorname{conv} f^2_{\text{test}}(x, y) = \operatorname{conv} f^2_{\text{test}}(r),$$

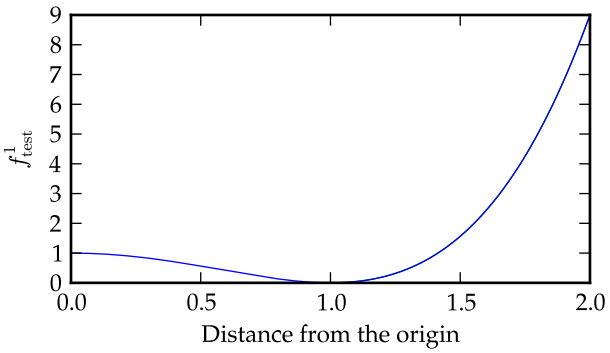

Figure 4.1.: The graph of the test function $f^1_{\text{test}}(x, y) = (r^2 - 1)^2$.

on $[-1.5, 1.5]^2$	dDLFT	mdDLFT	smdDLFT
Maximum value	0.0297	0.021	0.021
Medium value	0.0041	0.0031	0.0026
Standard deviation	0.0068	0.0054	0.0046
on the unit circle	dDLFT	mdDLFT	smdDLFT
Maximum value	0.0297	0.021	0.021
Medium value	0.0117	0.009	0.0074
Standard deviation	0.0068	0.0055	0.0051
outside the unit circle	dDLFT	mdDLFT	smdDLFT
Maximum value	0.0026	0.0025	$4.834 \cdot 10^{-5}$
Medium value	$2.883 \cdot 10^{-5}$	$1.035 \cdot 10^{-5}$	$8.763 \cdot 10^{-7}$
Standard deviation	$8.565 \cdot 10^{-5}$	$5.311 \cdot 10^{-5}$	$3.535 \cdot 10^{-6}$

Table 4.1.: Distribution of the error between the exact convex hull of f^1_{test} and its computed value.

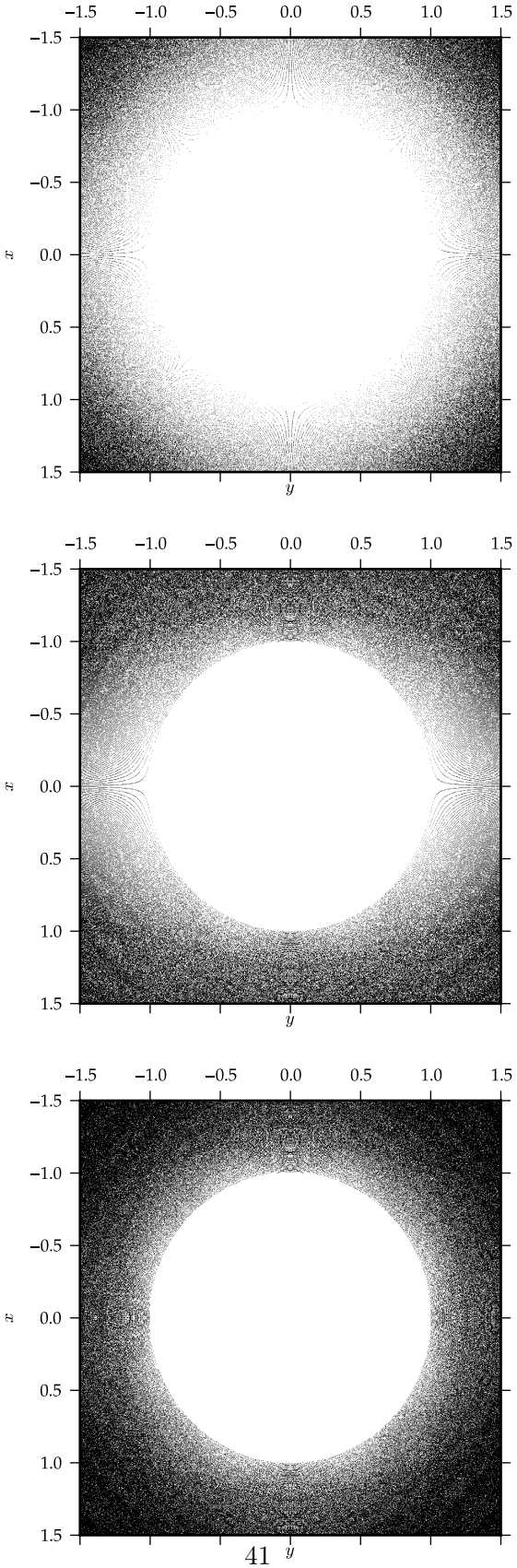

Figure 4.2.: Non-convexity region of the function f^1_{test} on the domain $[-1.5, 1.5]^2$, computed respectively through the dDLFT, mdDLFT and smdDLFT; the points where the computed convex hull differs from the function f^1_{test} are shown in white.

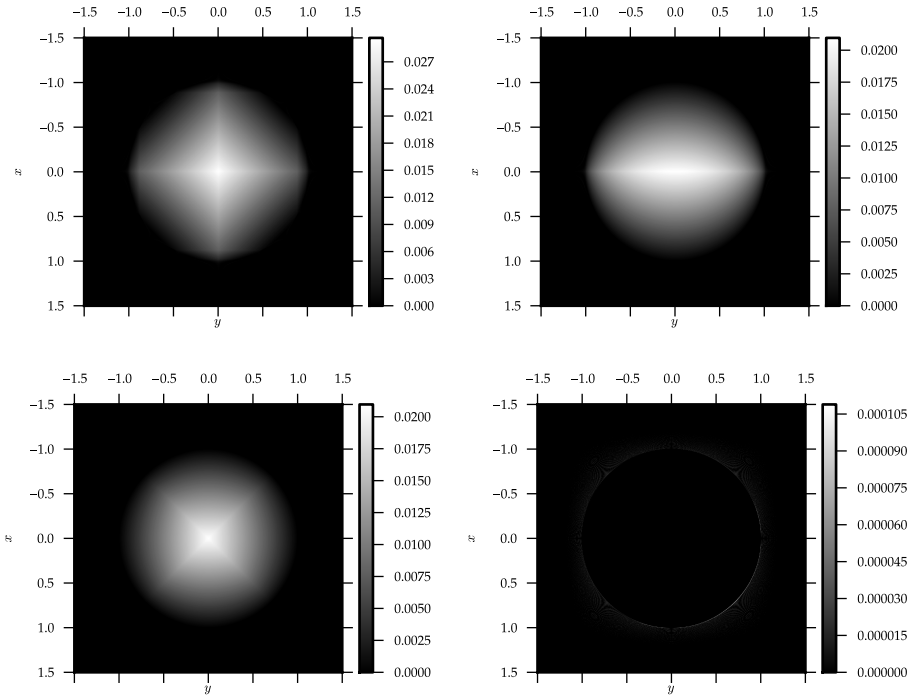

Figure 4.3.: Error between the exact convex hull of f^1_{test} and its value computed respectively from left to right and from top to bottom by the dDLFT, mdDLFT, smdDLFT and smdDLFT using a dual grid containing 0.

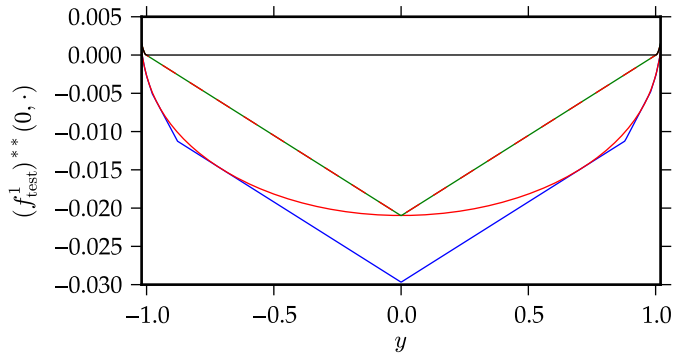

Figure 4.4.: Restriction to the axes of the convex hull of f^1_{test} computed respectively by dDLFT (blue line), mdDLFT (the solid red line is the restriction to the y-axis, while the dashed red line is the restriction to the x-axis), smdDLFT (green line) and smdDLFT using a dual grid containing 0 (black line); the true value of the convex hull in the interval $[-1, 1]$ is 0.

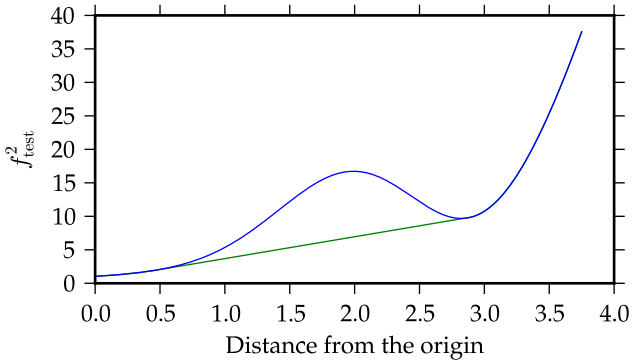

Figure 4.5.: The graph of the test function f^2_{test} (blue line) and of its convex hull (green line).

we can build a reasonable approximation of the convex hull by a computation in only one dimension (we denote by s the value of the one-dimensional derivative inside the non-convex region); we treat this hull as the "exact" one in order to test the two-dimensional codes. The domain is the square $[-3.75, 3.75]^2$ and is discretized by a 1000×1000 uniform grid. The exact non-convex region is an annulus with radii 0.47 and 2.87, whereas the regions computed numerically are presented in Figure 4.6. We take the opportunity to stress that the convex hull, and thus the shape of the non-convex region, depends heavily on the domain considered; this phenomenon is evident in Figure 4.7 where a different domain is used for computing the convex hull of f^2_{test}. Finally we present the same comparisons we did for $f^1_{\text{test}}(x, y)$. Since in this case the gradient in the non-convex region is not constant (only its norm is constantly s), we have chosen the s (positive) as the value to insert in the grid manually; thus the gain in accuracy is limited to the points of the non-convex region where one of the components of the gradient is s, as seen both in Figure 4.8 and Figure 4.9.

As a last comparison between the algorithms, in Table 4.3 the computation times in seconds for the two test functions are presented; they are averaged on several runs of the algorithms with the test functions f^1_{test} and f^2_{test} on the same domains considered above, but discretized with a $10^4 \times 10^4$ uniform grid. The convex hull algorithm has been implemented in Fortran and OpenMP has been used to provide easy parallelization. Because this compiled code is invoked from Python, there may be present some additional overhead; anyway, this overhead should be small and identical for all three versions of the algorithm. Then the code was run under Mac OS X 10.6 on a 2.93 GHz Intel Core i7 (4 cores, HT).

In order to properly compare the algorithms, we have to account not only for the computational time, but also for the quality of the results. A rough way to do so is the following. We fix the grid size at 1000^2 for the smdDLFT; the computation time (excluding grid generation and function evaluation) is 0.785 s for f^1_{test} and 0.788 s for f^2_{test}. Then gradually we increase the size of the grid used for the mdDLFT until a result of quality comparable to the smdDLFT is found. For the test function f^1_{test} this happens for a grid of 1212^2 points, when the mean error and the standard deviation become both

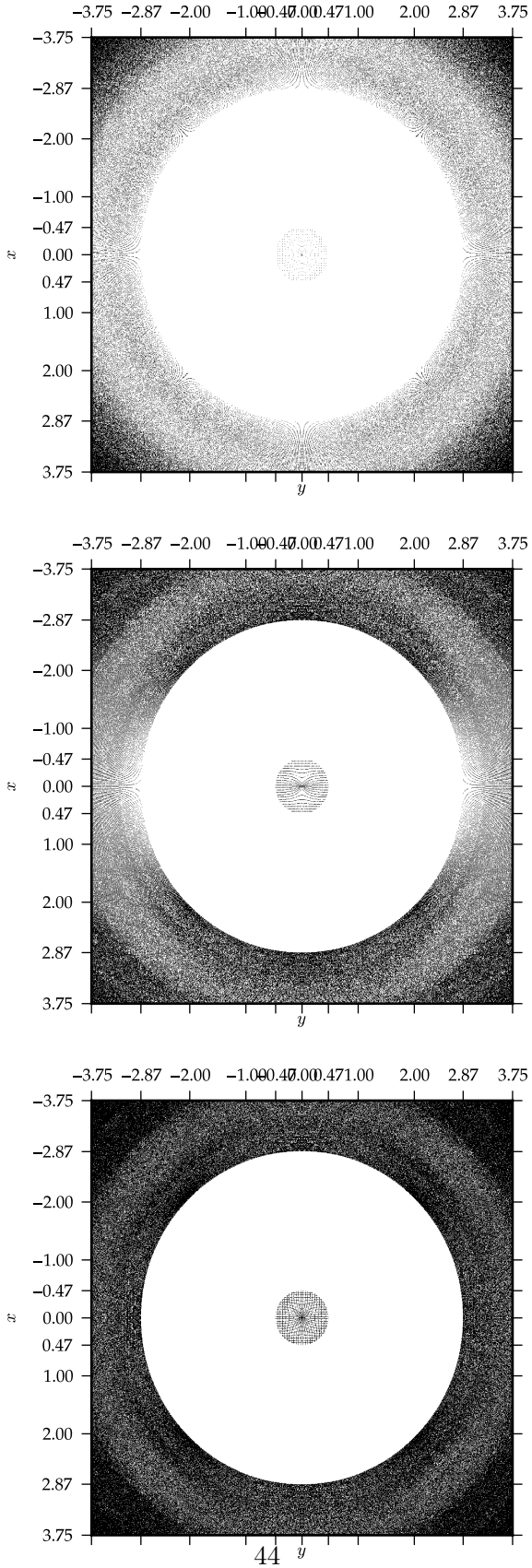

Figure 4.6.: Non-convexity region of the function f^2_{test} on the domain $[-3.75, 3.75]^2$, computed respectively through the dDLFT, mdDLFT and smdDLFT; the points where the computed convex hull differs from the function f^2_{test} are shown in white.

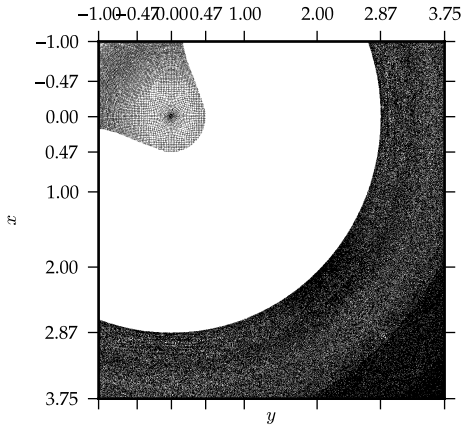

Figure 4.7.: Non-convexity region of the function f^2_{test} on the domain $[-1, 3.75]^2$ computed through the smdDLFT; the points where the computed convex hull differs from the function f^2_{test} are shown in white.

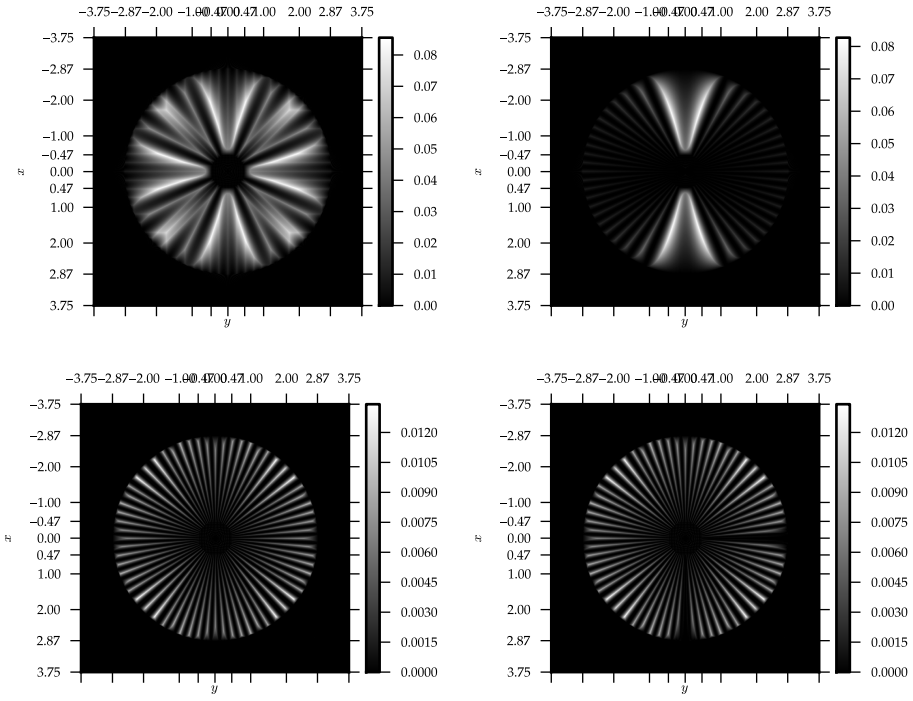

Figure 4.8.: Error between the exact convex hull of f^2_{test} and its value computed respectively from left to right and from top to bottom by the dDLFT, mdDLFT, smdDLFT and smdDLFT using a dual grid containing s.

on $[-3.75, 3.75]^2$	dDLFT	mdDLFT	smdDLFT
Maximum value	0.0855	0.0827	0.0134
Medium value	0.0125	0.0043	$9.542 \cdot 10^{-4}$
Standard deviation	0.0182	0.0111	0.0019
on the annulus	dDLFT	mdDLFT	smdDLFT
Maximum value	0.0855	0.0827	0.0134
Medium value	0.0276	0.0096	0.0021
Standard deviation	0.0181	0.015	0.0023
outside the annulus	dDLFT	mdDLFT	smdDLFT
Maximum value	0.0081	0.0081	0.0015
Medium value	$2.219 \cdot 10^{-4}$	$8.747 \cdot 10^{-5}$	$1.009 \cdot 10^{-5}$
Standard deviation	$3.725 \cdot 10^{-4}$	$2.466 \cdot 10^{-4}$	$5.898 \cdot 10^{-5}$

Table 4.2.: Distribution of the error between the exact convex hull of f^2_{test} and its computed value.

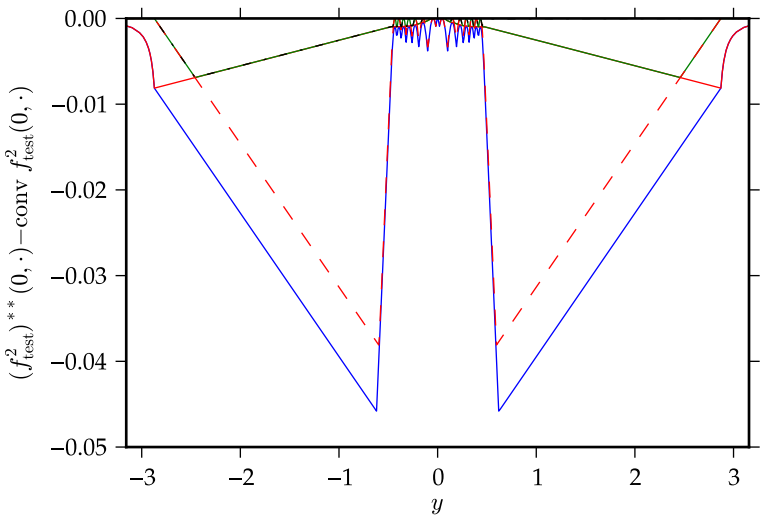

Figure 4.9.: Restriction to the axes of the error between the exact convex hull of f^2_{test} and the one computed respectively by dDLFT (blue line), mdDLFT (the solid red line is the restriction to the y-axis, while the dashed red line is the restriction to the x-axis), smdDLFT (green line) and smdDLFT using a dual grid containing s (black dashed line); observe that the black line is near 0 only on the right part of the plot, due to the positivity of s.

	f^1_{test}	f^2_{test}
dDLFT	73 s	73.4 s
mdDLFT	49.6 s −32% of dDLFT	50.4 s −31% of dDLFT
smdDLFT	108.6 s +119% of mdDLFT +49% of dDLFT	109.2 s +117% of mdDLFT +49% of dDLFT

Table 4.3.: Computational times in seconds of the three different versions of the convex hull algorithm.

better than their counterparts for the smdDLFT; the time expended is 0.575 s, better than the smdDLFT. In the case of f^2_{test} the mean value becomes better for the first time at 2498^2 points, but the maximum error at this stage is still about two times the one for the smdDLT; the time in this case is 2.628 s, much larger than for the smdDLFT, and again does not include grid generation and function evaluation. Thus we conclude tentatively that using the smdDLFT in the general case is worth the extra computational time; more rigorous and extensive tests should be done to confirm this assessment.

4.6. The choice of the dual grid

In the following we always use the more accurate smdDLFT transform, but similar problems and solutions arise also in the other two cases. The main problem in the computation of the convex hull is the choice of the dual grids C_N and D_N; as seen in Theorem 3.22 it is sufficient to cover a certain bounded set S, which can be estimated thanks to Theorem 4.15. Until now we have used a uniform discretization of this set; this can be a poor choice when the derivatives of the function grow rapidly, causing S to be large and the points of C_N and D_N to be too sparse. This is the case for both the test functions f^1_{test} and f^2_{test} (defined in Section 4.5) if we enlarge the domain on which the convex hull is computed (see respectively Figures 4.11 and 4.10); note that even increasing the size of the grid, while certainly improving the result, is not a definitive answer (see the right plot in Figure 4.11).

It is interesting to observe that the plots are composed of black lines whose spacing is greater the nearer they are to the origin: each of these lines corresponds to a point in the dual grid. Consider for example $\xi \in C_N$; being $C_N \times \mathbb{R}$ the actual set on which the second transform is computed, by Corollary 3.21 for every point $(x, y) \in \Omega_N$, such that $(\xi, \eta) \in \partial \operatorname{conv} f_{\Omega_N}(x, y)$ for a certain $\eta \in \mathbb{R}$, we have that $\operatorname{conv} f_{\Omega_N}(x, y) = \left(f^*_{\Omega_N}\right)^*_{C_N \times \mathbb{R}}$. Since in our case the test functions and their hulls are differentiable, the locus of the points where the first component of the gradient is ξ is one of the black lines, i.e. is composed of points where the algorithm gives the exact result. By observing that for our test functions the gradient is always directed radially and has module dependent only on the distance from the origin and increasing with it, the shape and origin of the black lines which are asymptotically tangent to the y-axis is explained easily; a similar reasoning

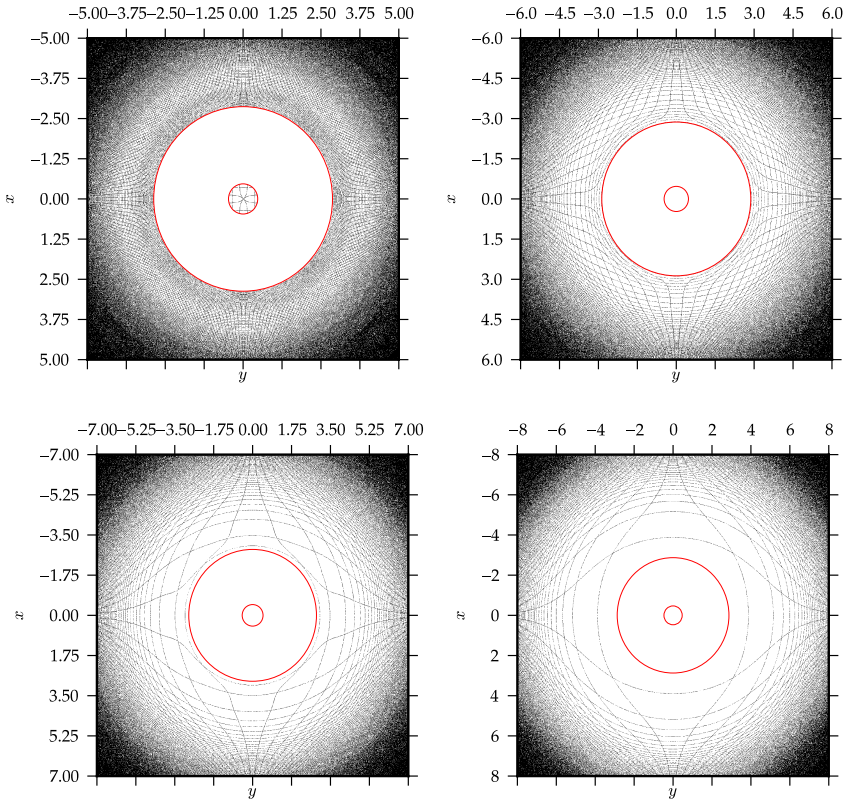

Figure 4.10.: Non-convexity region of the function f^2_{test} respectively from left to right and from top to bottom on the domains $[-5,5]^2$, $[-6,6]^2$, $[-7,7]^2$ and $[-8,8]^2$; the points where the computed convex hull differs from the function f^2_{test} are shown in white, while the boundary of the exact non-convexity region is shown in red.

can be done for the lines asymptotically tangent to the x-axis, which are related to the points $\eta \in D_N$. Since for our test functions the second derivative increases with the distance from the origin, our interpretation of the black lines explains also why they are more spaced close to the origin: the distance between successive points of the dual grids is always the same, but the distance in primal space needed for the same increase in the gradient is smaller as we move away from the origin. This explains why the lines are denser, and the result of the algorithm better, far away from the origin.

Since we are more interested in detecting the shape of the region of non-convexity and in computing the values of the convex hull inside it, we would like our dual grid to be denser around the slopes of the convex hull in that region. The problem is, as usual, that we cannot know the values of these slopes without knowing the convex hull. A possible solution to this problem consists in using a less precise approximation to the convex hull in order to detect the region of non-convexity and computing its slopes. A very simple approximation for the convex hull can be obtained by computing one-dimensional convex hulls on all the rows of the data matrix and then on all its columns. Since the convex

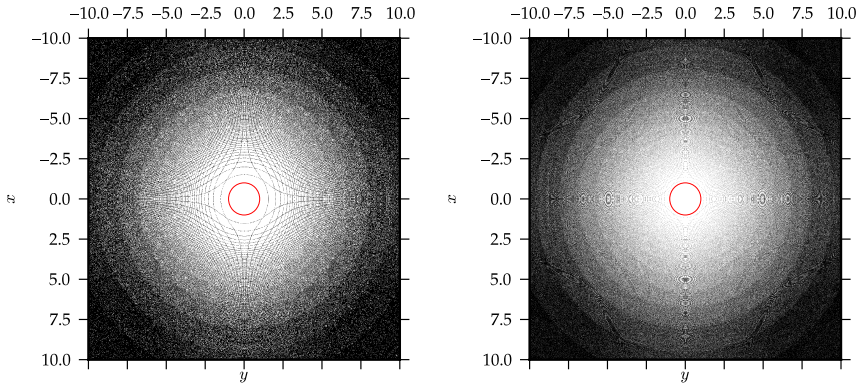

Figure 4.11.: Non-convexity region of the function f^1_{test} on the domain $[-10, 10]^2$ discretized with a uniform grid of size respectively 1000×1000 (left) and 5000×5000 (right); the points where the computed convex hull differs from the function f^1_{test} are shown in white, while the boundary of the exact non-convexity region is shown in red.

f	$y = 1$	$y = 2$	$y = 3$
$x = 1$	0	0	0
$x = 2$	3	1	0
$x = 3$	6	3	0

Table 4.4.: A function f and a grid $\{1, 2, 3\}^2$ for which the heuristic method does not give the exact result; f is convex along all the rows and all the columns, but not along the main diagonal.

hull, differently from the DLFT, is not factorizable, this method is purely heuristic: the result can be greater than the exact convex hull (see Table 4.4), which implies that the detected region of non-convexity is smaller than the true one. The reason for the failure of this simple algorithm is that it makes the function convex along all the rows and all the columns, but not along all the other directions; for a differentiable funtion, this is equivalent to forgetting the role of mixed derivatives in determining convexity at a given point.

The heuristic detection can be implemented inside the convex hull step of the first pass of DLFTs, by saving for each row which slopes are inserted in order to make it convex. These values are transformed in intervals by enlarging each of them by a fixed window size; then the resulting intervals are joined when they are close, for example nearer than a certain multiple of the windows size. In this way we have partitioned the set C defined in Theorem 4.15 in intervals relative to the region of non-convexity and intervals relative to the region where the function is already convex. Then we fix a fraction of the available points to be assigned to intervals of the first type: these points are distributed evenly among all such intervals independently from their length (we assign at least three points to each interval). Finally, the remaining points are distributed evenly among the intervals of the second type, again independently of their length. The results obtained using dual grids of this kind are shown in Figures 4.12 and 4.13 (observe that the shape

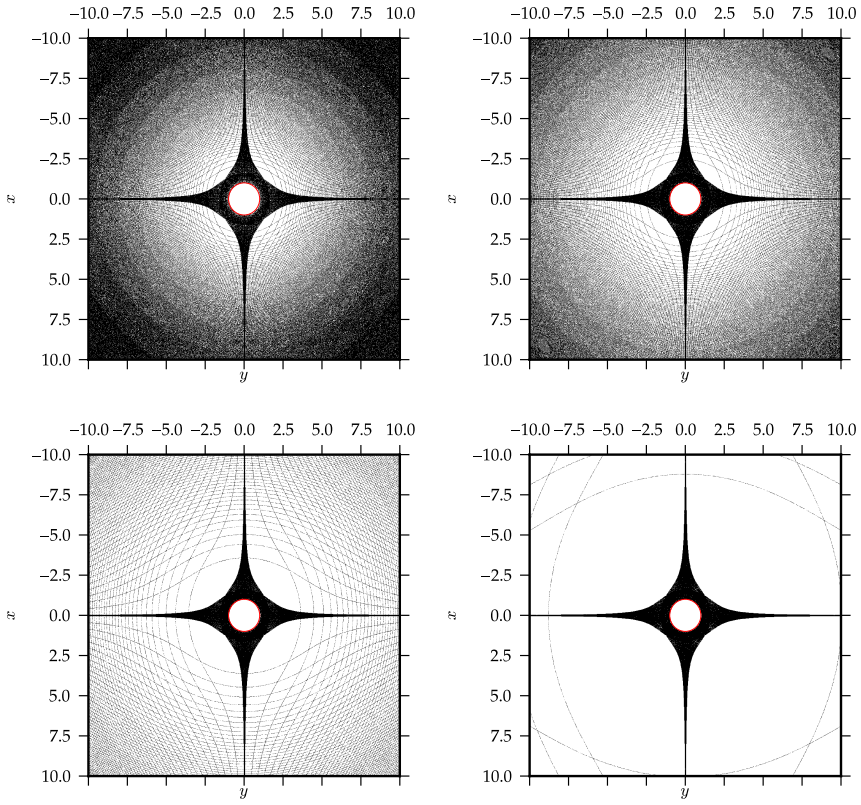

Figure 4.12.: Plots of the non-convex regions of the test function f^1_{test} on the domain $[-10, 10]^2$ covered by a uniform 1000×1000 grid. An adaptive dual grid is used; the fraction of points inside the intervals flagged as non-convex is respectively from left to right and from top to bottom 0.2, 0.6, 0.9 and 1.

of the adapted region is in accordance with the interpretation of the black lines as loci of points corresponding to each dual grid point); the same method works well also for larger domains than those considered so far, as shown in Figure 4.14.

We will now consider a more complex function f^3_{test} defined as

$$f^3_{\text{test}}(x, y) := \begin{cases} H_i & \text{if } (x, y) \in B(10i, 10i, 1), i \in \mathbb{N} \setminus \{0\} \\ f^1_{\text{test}}(x, y) & \text{elsewhere} \end{cases},$$

where $B(x, y, r)$ is the open ball with radius r centered in (x, y) and $H_i \in \mathbb{R}$ is such that $H_i \geq \max_{(x,y)\in B(10i,10i,1)} f^1_{\text{test}}(x, y)$; the region of non-convexity of f^3_{test} on $\mathbb{R}^2$ is $\bigcup_{i\in\mathbb{N}} B(10i, 10i, 1)$ and thus is not connected. As expected, a uniform discretization of the domain does not allow to clearly detect all the connected components of the region of non-convexity (see the first plot in Figure 4.15). Now consider an adaptive grid built with a window size of 10 and by reserving just the 10% of points for the intervals flagged as non convex; the corresponding result is shown in the second plot of Figure 4.15. Even

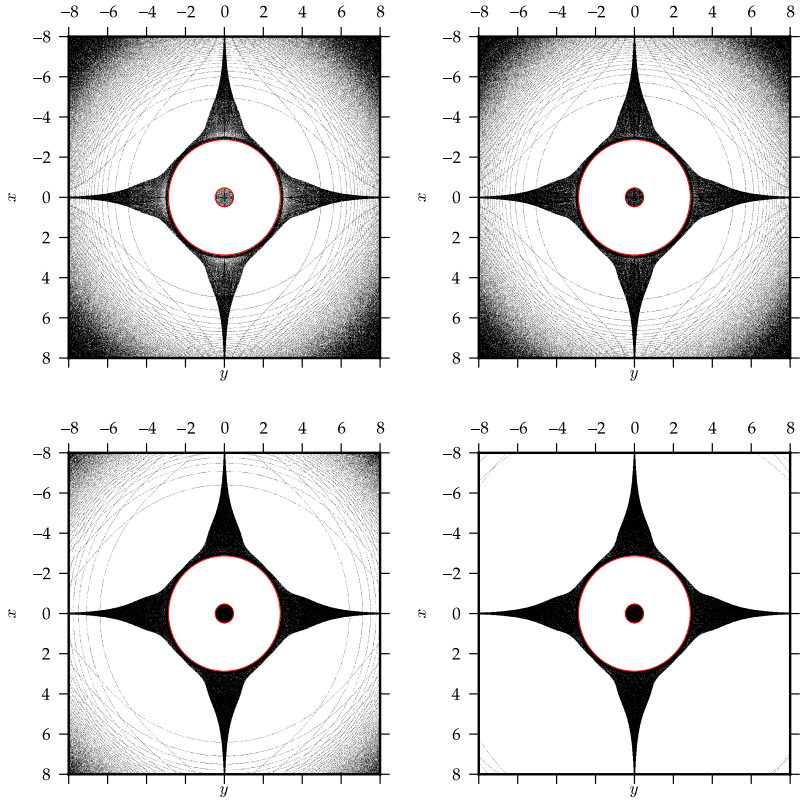

Figure 4.13.: Plots of the non-convex regions of the test function f^2_{test} on the domain $[-8, 8]^2$ covered by a uniform 1000×1000 grid. An adaptive dual grid is used; the fraction of points inside the intervals flagged as non-convex is respectively from left to right and from top to bottom 0.1, 0.2, 0.8 and 1.

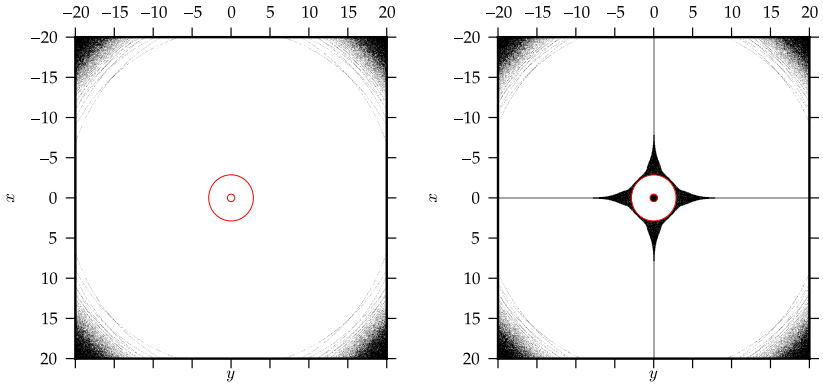

Figure 4.14.: Plots of the non-convex regions of the test function f^2_{test} on the domain $[-20, 20]^2$ covered by a uniform 1000×1000 grid. In the right figure an adaptive dual grid is used; the fraction of points inside the intervals flagged as non-convex is 0.2.

though only a small fraction of the points is assigned to the non-convex region, the density of the black points away from that region is very low, against our expectations. This is due to the fact the window size is very low compared to the distance between the flagged slopes, distance that increases rapidly away from the origin, and thus there is no merging of intervals; then we have a large number of intervals flagged as non-convex and, since we assign to each of them at least 3 points, there are few points left for the other intervals, explaining the poor result away from the non-convexity region. The situation is particularly bad in the upper left quadrant, which does not contain any part of the non-convexity region and thus corresponds to a single interval in dual space; being the number of points assigned to each interval the same independently of its length, the resulting density of black points in this quadrant is less than elsewhere. Moreover, by examining the region of non-convexity detected by the algorithm, we observe that their borders are not entirely captured; in order to better capture them we should increase the number of points outside the flagged regions, by increasing the window size. Doing so by modifying the adaptation parameters (see the third and fourth plot of Figure 4.15) improves the result, but is against the two main objectives of our adaptive method: being automatic, i.e. being independent of the function given, and using only a subset of the entire estimated dual set S for the computation.

It is interesting to understand why our method fails to wholly capture the border of the non-convexity region; in order to visualize better this situation, consider the upper left region of non-convexity in Figure 4.16, which is obtained by modifying f^1_{test} in a way similar to the one used for f^3_{test}. The red circle is the exact boundary of the region of non-convexity, while the colored area is the region detected as non-convex by the heuristic method; in this case the two coincide. Observe that the boundary lines of the black stripes, on which the convex hull is computed exactly, touch the non-convexity region respectively on its topmost, bottom, leftmost and rightmost point; this is due to the fact that they correspond to the values of the components of the gradient in those points. Since the change we made to f^1_{test} is discontinuous, the derivatives of the convex hull are discontinuos too; this means that the black stripes cannot cover the zone of non-convexity wholly, as was the case for f^1_{test}. A possible modification to the method in order to account for non-continuous derivatives of the hull could be implemented by flagging as non-convex also the slopes immediately preceding and following an actual one-dimensional non-convex region. Nonetheless, this modification is not resolutive, because the same problem can arise even if the change to the function is done preserving continuity and differentiability; consider for example the lower right region of non-convexity in Figure 4.16, which is obtained by modifying f^1_{test} by adding a gaussian function centered in $(20, 20)$. The problem in this case is due to the fact that the region detected by the heuristic is different from the exact region of non-convexity; as an effect of the adaptive construction of the dual grid, the black stripes would cover completely the colored region, but they are hidden by the larger true non-convexity region, which they are not able to cover. For certain functions, as we will see in Section 5.2, the difference between the two regions is even greater.

Until now we have searched a method to efficiently use a fixed number of points available for the dual grid. Another approach to deal with large dual sets S, which at first

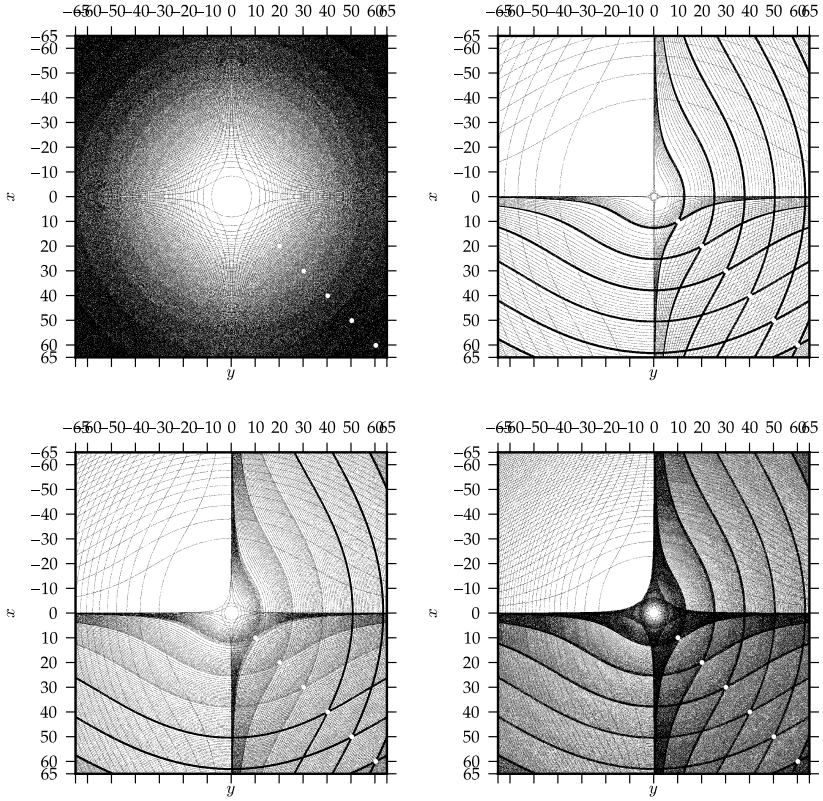

Figure 4.15.: Non-convexity region of the function f^3_{test} on the domain $[-65, 65]^2$ covered by a uniform 1000×1000 grid; the points where the computed convex hull differs from the function f^3_{test} are shown in white. In the first (from left to right and top to bottom) plot a uniform dual grid is used, whereas in the others an adaptive dual grid is used; the fraction of points inside the interval flagged as non-convex is respectively 0.1, 0.1 and 0.5 and the window size is respectively 10, 250 and 1000.

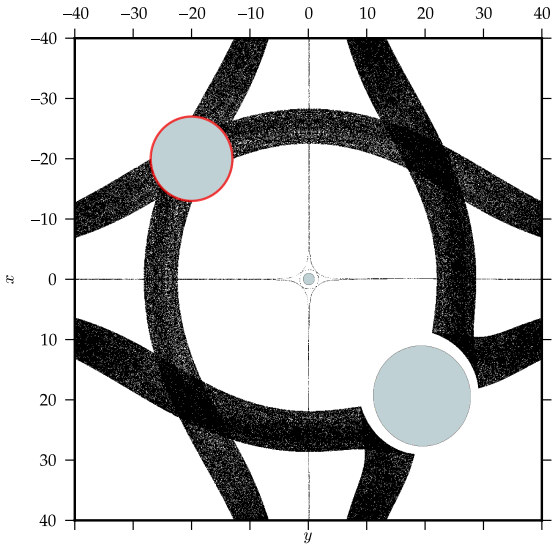

Figure 4.16.: A more visible example of the same behaviour exhibited in the right plot of Figure 4.15; the upper left region of non-convexity is due to a discontinuous change to the function f^1_{test} (as is the case in f^3_{test}), while the lower right one is caused by a continuous change. The red circle indicates the exact boundary of the first region, while the colored areas are the one flagged as non-convex by the heuristic method based on factorization of the convex hull.

sight may appear naïve, is using a dual grid with a higher number of points; in this case, a potential problem could be memory usage since a larger matrix has to be stored in memory. This problem can be easily resolved thanks to Proposition 2.17: the dual grid is partitioned in smaller grids, each of these grids is used for a computation of the convex hull and finally the partial results are merged by a maximum operation; it is even possible to perform the first pass of DLFTs only once and sample them each time on a different grid. The problem is how to partition the large grid; since equal intervals in dual space can correspond to varying intervals in primal space, this is again a problem which cannot easily resolved independently of the given function. A possible heuristic solution is using as the dual grid all the slopes of the functions obtained in the first pass of one-dimensional convex hulls; this time we take every slope, instead of only the non-convex ones. This large grid in dual space, which has about the same size as the whole primal grid, can be partitioned naturally in correspondence to each line on which the one-dimensional hull has been computed; not all the lines have to be taken and results are usually good even when only a few of them are taken (see Figure 4.17). It is worthwhile observing in Figure 4.17 that, when the number of steps is small, the quality of the results is better for odd numbers; in this case in fact the central line is also taken and thus we obtain a better result in the central region.

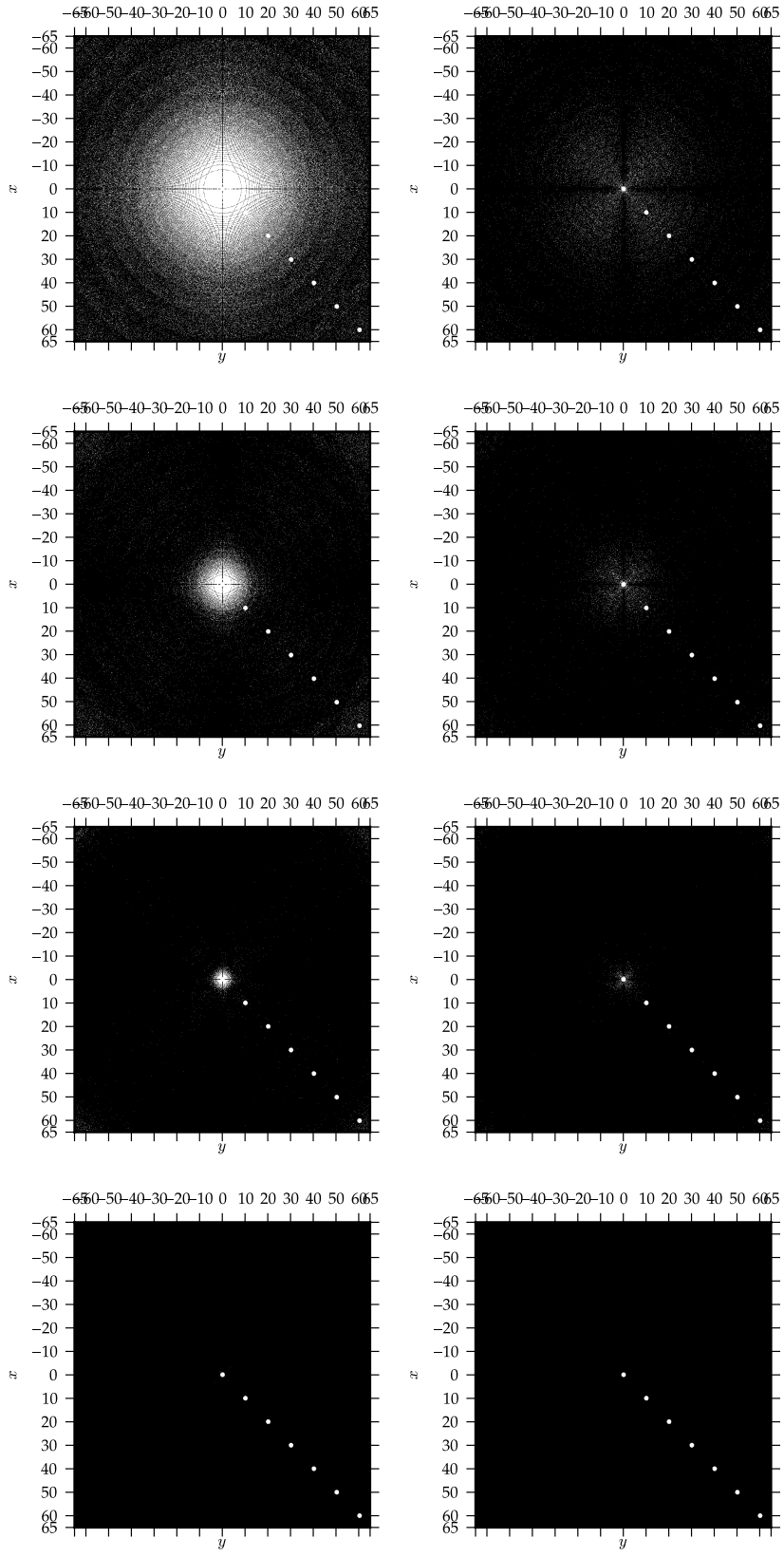

Figure 4.17.: Plots of the non-convex regions of the test function f^3_{test} on the domain $[-65, 65]^2$ covered by a uniform 1000×1000 grid. A progressive construction of the result is made, respectively from left to right and from top to bottom in 2, 3, 4, 5, 10, 11, 75 and 150 steps.

5. Non-convex free-energies

In this chapter we apply the convex hull algorithm to the bulk free-energy density f defined in Section 1.4. Our aim is finding the region of the state space $c = (c_+, c_-)$ where f differs from its convex hull, i.e. the region thermodynamically unstable, which we denote by $\mathcal{T}$.

5.1. Symmetric case

First we study the case $|Z_+| = |Z_-|$, which ensures that f is invariant with respect to the exchange of c_- and c_+, i.e. its graph (and also its convex hull) is symmetric with respect to the plane of equation $c_+ = c_-$. Consider for example the case $Z_\pm = \pm 2$: the qualitative behaviour of f on a large scale is the same whatever the value of the ion diameter σ is (see Figure 5.1). The main difference is the domain of f, which is the larger the smaller is σ; this is due to the fact that the smaller the ions, the higher the concentrations at which steric effects become relevant (smaller atoms can be more tightly "packed" together). All the three functions in Figure 5.1 seem to be convex, but this is actually true only in the case $\sigma = 0.3$ nm; we may see this by computing the difference between f and its convex hull.

Now we examine the case $\sigma = 0.2$ nm in detail. The region of non-convexity for such a diameter is shown in white in the left plot of Figure5.2: observe that the shape of $\mathcal{T}$ is symmetric as expected. What is not evident is whether the region $\mathcal{T}$ extends all the way to the origin. By observing the enlargement presented in the right plot of Figure 5.2, it appears that the region around the origin is indeed convex; the boundary of $\mathcal{T}$ closest to the origin appears to be contained in $[0, 0.25]^2$ and to have the same shape of the other boundary; in spite of that, by computing the convex hull of f on the domain $[0, 0.25]^2$, it seems that the function is convex there (see the top left image in Figure 5.3). If we gradually enlarge the domain (see the rest of Figure 5.3), a region of non-convexity appears and progressively increases, while its boundary tends to the line we observed in the right plot of Figure 5.2. This behaviour is not surprising: the convex hull of a function depends on the set on which that function is considered; if we enlarge this set, there is the possibility that the region of non-convexity increases (obviously it cannot become smaller). We are interested in the convex hull of f on its whole domain, which is finite because for large concentration f becomes $+\infty$; fortunately we can restrict the computation to a smaller set because the non-convex terms of f are negligible at high concentration and thus the growth of the region of non-convexity ceases for c sufficiently large (this enable us to use this method even if we had used the hard-sphere approximation which gives a free-energy defined on the whole $\mathbb{R}^2$).

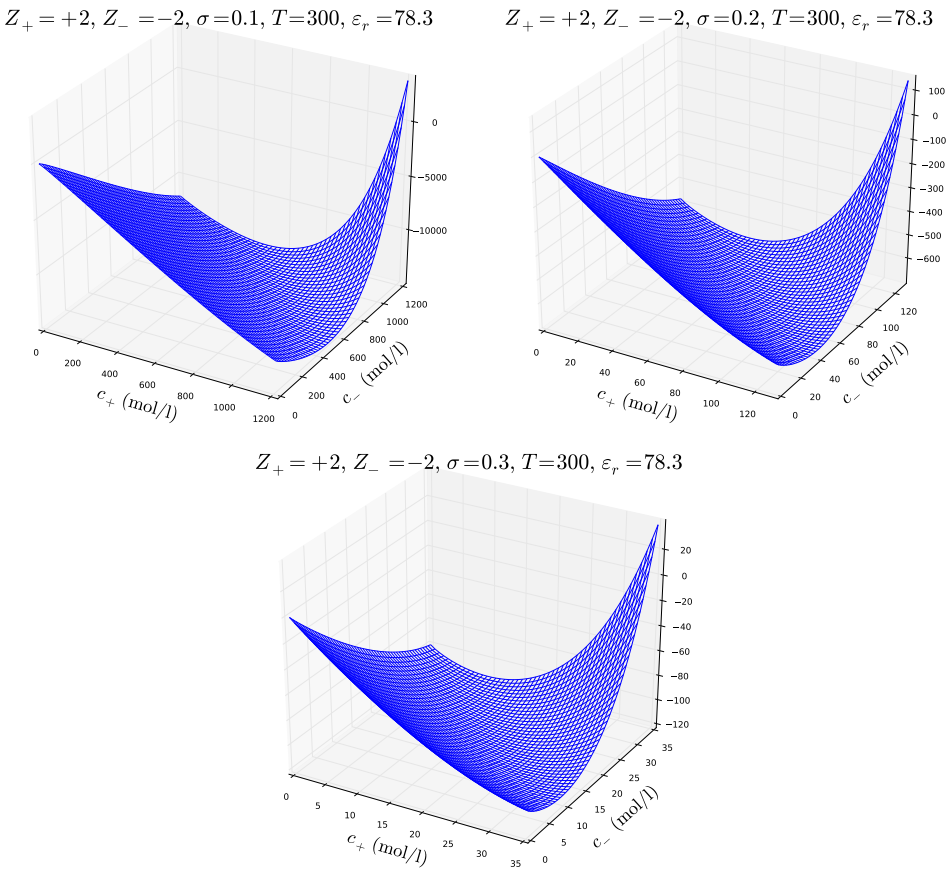

Figure 5.1.: The behavior of f at large scale when $Z_\pm = \pm 2$ and the ion diameter σ is equal respectively to 0.1 nm (left), 0.2 nm (right) and 0.3 nm (bottom); be aware that f is convex only in the last case.

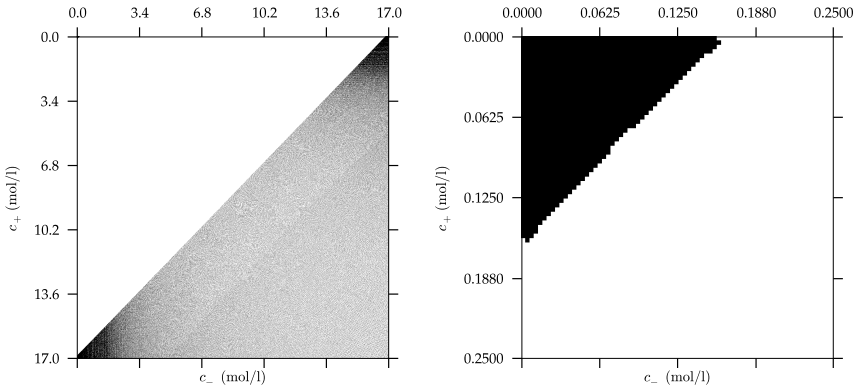

Figure 5.2.: The region of non convexity in the case $Z_\pm = \pm 2$ and $\sigma = 0.2$ nm; the black points are where $f = \operatorname{conv} f$. The convex hull is computed on a uniform 5000×5000 grid on $[0, 17]^2$; the right plot is just an enlargement of the left one.

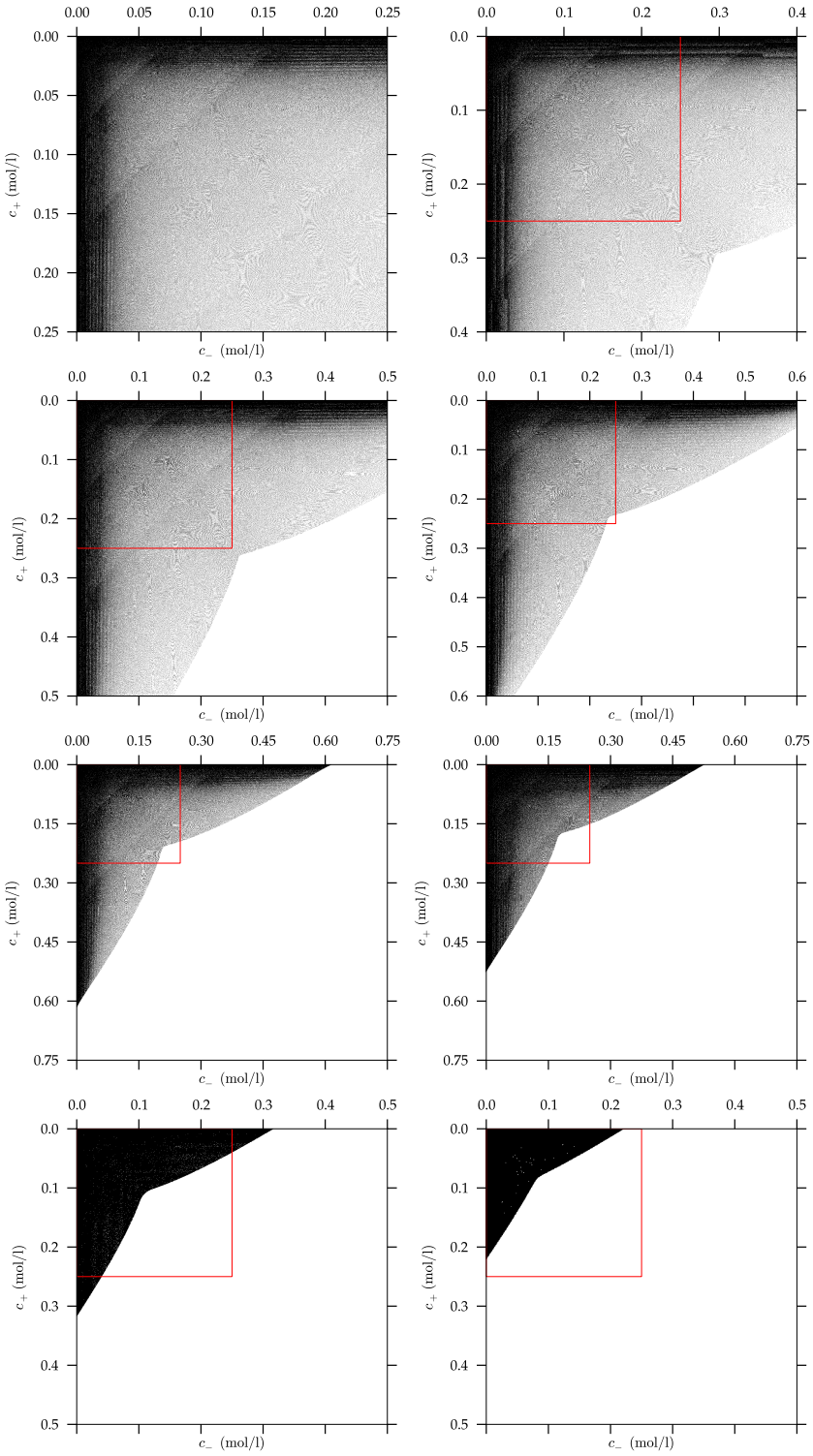

Figure 5.3.: These plots evidence the dependence of the region $\mathcal{T}$ on the set on which the convex hull is computed; model parameters are $Z_\pm = \pm 2$ and $\sigma = 0.2$ nm. Each figure (from left to right and from top to bottom) is computed on a increasingly larger set (always discretized by a 5000×5000 uniform grid); the first 5 plots are computed on the displayed region, while the last three figures are actually computed on a domain larger than the one shown, respectively $[0, 1]^2$, $[0, 2.5]^2$ and $[0, 5]^2$. The red square is the starting region $[0, 0.25]^2$.

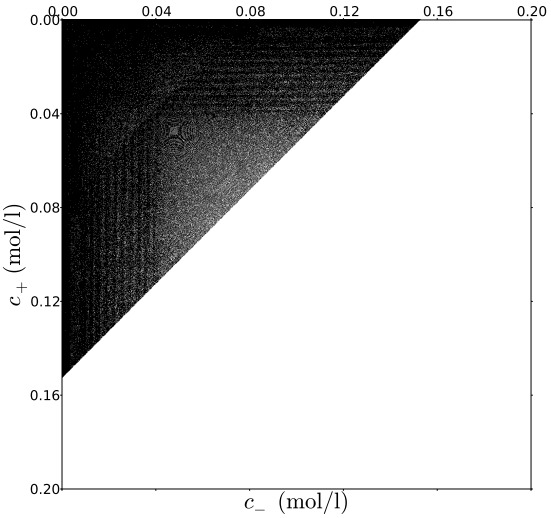

Figure 5.4.: The boundary of $\mathcal{T}$ close to the origin for $Z_\pm = \pm 2$ and $\sigma = 0.2$ nm. The grid used is a non-uniform grid composed of 3000 points between 0 and 0.4, 2000 points between 0.4 and 18 and 500 points between 18 and 50.

As we have observed, the correct shape of $\mathcal{T}$ is captured only when the set on which the convex hull is computed is sufficiently large. If the grid used is uniform, the larger the computational domain the fewer the points of the grid on which the boundary of $\mathcal{T}$ close to the origin lies. It is then useful to adopt a non-uniform grid in order to achieve higher resolution in the area near the origin, without reducing the computational domain and thus changing the shape of $\mathcal{T}$ (see Figure 5.4). We observe that if we restrict f to one of the axes, we obtain the free-energy density arising in the case only one ion species is present. Moreover, if we fix Z_+ and restrict f to the c_+ axis, we obtain the same function for every value of Z_-; its non-convexity region as a one-dimensional function coincides with the intersection of $\mathcal{T}$ with the axis considered (see Theorem A.57).

We may be interested in obtaining a characteristic function for the set $\mathcal{T}$, i.e. a function $\mathbf{1}_\mathcal{T}$ such that $\mathbf{1}_\mathcal{T}(x) = 1$ when $x \in \mathcal{T}$ and is equal to 0 elsewhere. This is useful for a memory-efficient representation of the convex hull of f, because we need to save only the values inside the region $\mathcal{T}$, whereas the values outside of $\mathcal{T}$ can be computed from the analytical expression of f. A simple approach to the computation of $\mathbf{1}_\mathcal{T}$ is the application of a threshold to the difference $f - \operatorname{conv} f$, i.e. taking $c \in \mathcal{T}$ if and only if $f(c) - \operatorname{conv} f(c) > \mathrm{TOL}$, where TOL is a given tolerance. In general, the value of TOL should be chosen on a case by case basis, because if it is too large the shape of $\mathcal{T}$ could be enlarged too much (see Figure 5.5). If the region $\mathcal{T}$ can be already recognized well from the plot of the binary matrix $f \neq \operatorname{conv} f$ (as it has been the case up to now), a better solution is using the closing operation from mathematical morphology ([14]), which eliminates the problem of the spurious enlargement of $\mathcal{T}$ typical of the threshold operation (see Figure 5.6).

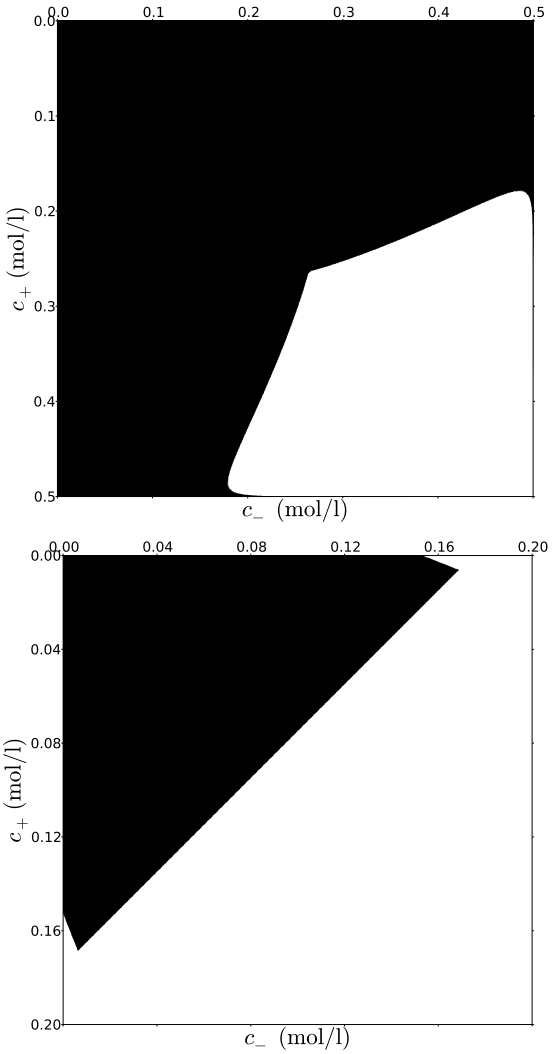

Figure 5.5.: These are the same computations as the third plot of Figure 5.3 and as Figure 5.4, where TOL is 10^{-3} times the maximum difference between f and $\operatorname{conv} f$ (computed separatedly for each figure).

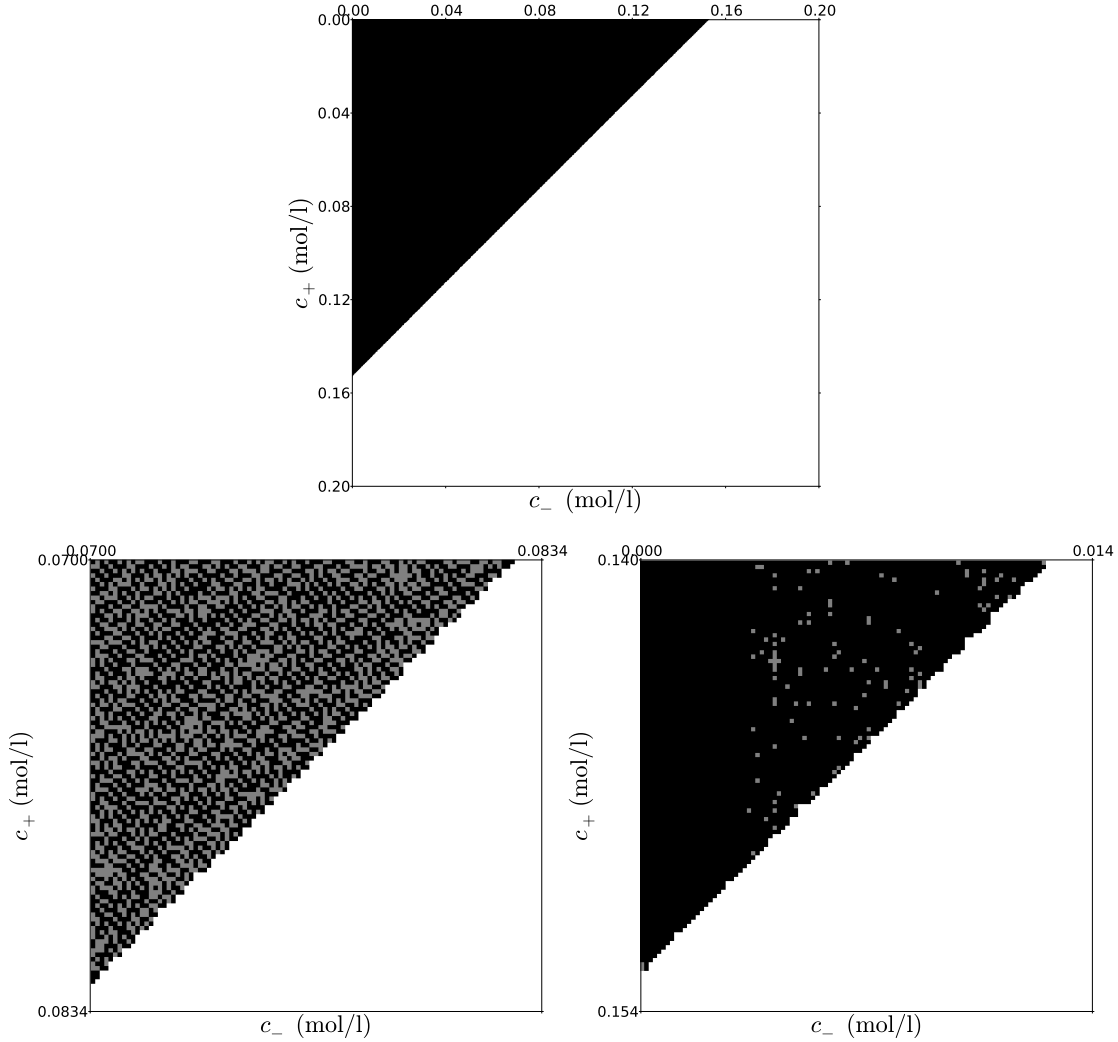

Figure 5.6.: The top plot is the closing of the binary image of Figure 5.4; the bottom plots are enlargement of the top one, where the closing is represented in grey and the original image superimposed in black.

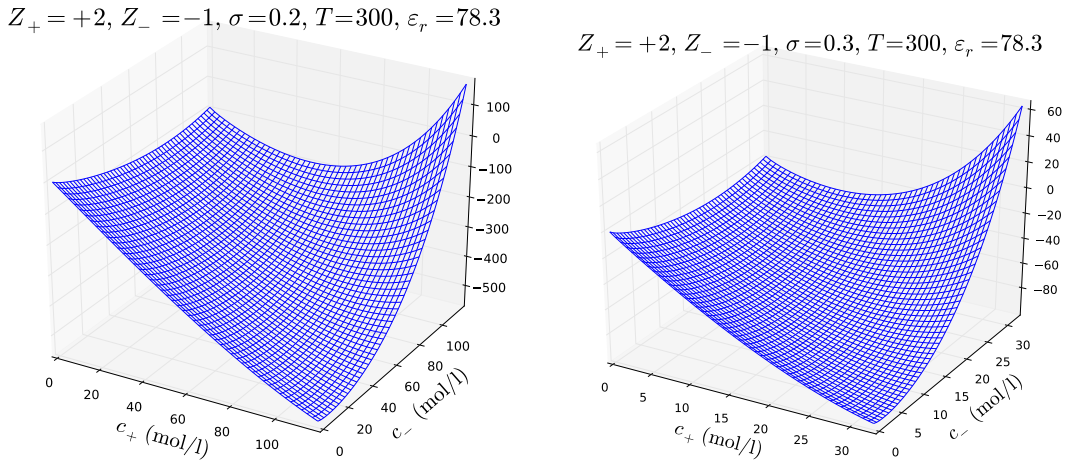

Figure 5.7.: The behavior of f at large scale for $Z_+ = +2$, $Z_- = -1$ and σ equal respectively to 0.2 nm (left) and 0.3 nm (right); be aware that f is convex only in the second case.

5.2. Non-symmetric case

Now we consider the case where $|Z_+| \neq |Z_-|$, which implies that f is no more symmetric; in particular, as an example we take $Z_+ = +2$ and $Z_- = -1$. In Figure 5.7 we present the plots of f for respectively $\sigma = 0.2$ nm (left) and $\sigma = 0.3$ nm (right): on a large scale the qualitative behavior of the convex ($\sigma = 0.3$ nm) and non-convex cases is still very similar. In Figure 5.8 is plotted the non-convexity region in the non-convex case: as expected, the region $\mathcal{T}$ becomes asymmetric. Observe that, differently from the symmetric case, the region $\mathcal{T}$ does not divide the plane (c_+, c_-) into two parts; the reason for this is easily understood by considering the restriction of f on the axes. The restriction of f on the c_+ axis is the free-energy density when we have only one ion species with $|Z| = 2$ and $\sigma = 0.2$ nm. Thus, by Theorem A.57, the restriction of conv f on the c_+ axis is the convex hull of such one-dimensional density; thus, if we fix Z_+, the intersection of $\mathcal{T}$ with the c_+ axis is the same for any value of Z_- (observe for example Figures 5.8 and 5.2). On the other hand, the restriction of conv f on the c_- axis is the convex hull of the free-energy density for $|Z| = 1$ and $\sigma = 0.2$ nm; as seen in Table 1.1, for such a combination the one-dimensional density is already convex and thus there is no intersection between $\mathcal{T}$ and the c_- axis. If σ decreases below the value of σ_0 corresponding to $|Z| = 1$, then the intersection between $\mathcal{T}$ and the c_- axis becomes non empty and $\mathcal{T}$ divides the state space into two parts. In this case $\sigma_0 = 0.0560$ nm is very small for an ion diameter and thus is not physically acceptable; however, if we consider the case $Z_+ = +3$ and $Z_- = -2$, the threshold is $\sigma_0 = 0.2239$ nm, which is acceptable (see Figure 5.9 for an evolution of the non-convexity region as σ varies).

Now consider the case $Z_+ = +3$, $Z_- = -1$ and $\sigma = 0.17$ nm. For these values of the parameters, the region of non-convexity covers an area so large that it is close to the boundary of the domain of the free energy density, where its values and derivatives go rapidly to infinite. Moreover, the area detected as non-convex by a heuristic factorization

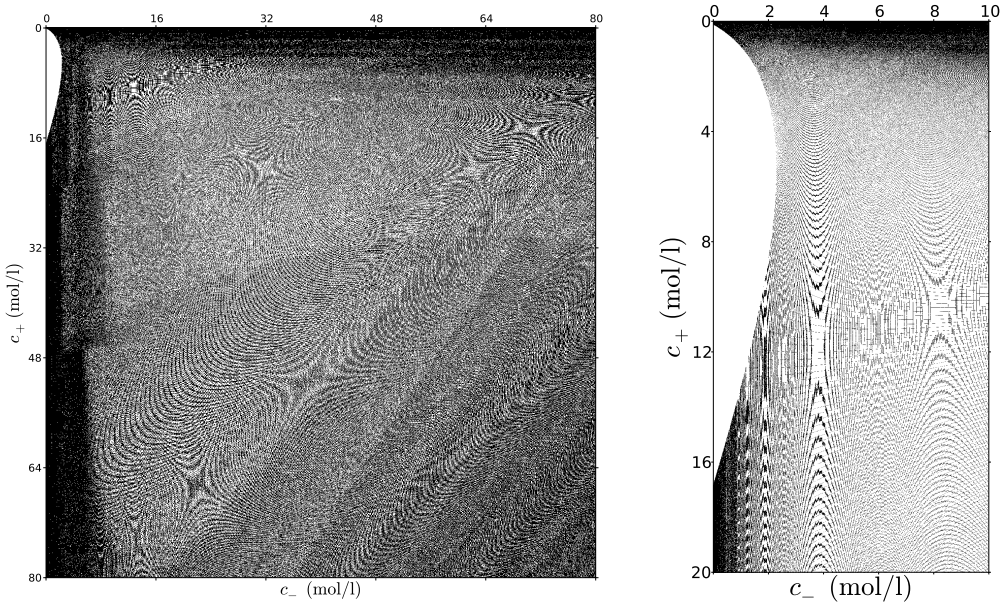

Figure 5.8.: The non-convexity region for model parameters $Z_+ = +2$, $Z_- = -1$ and $\sigma = 0.2$ nm.

of the convex hull is extremely different from the real one. For these reasons approaches based on uniform dual grids or adapted dual grids fail (see Figure 5.10) and approaches based on larger dual grids built progressively are the only way to capture the geometry of the region of non-convexity (see Figure 5.11).

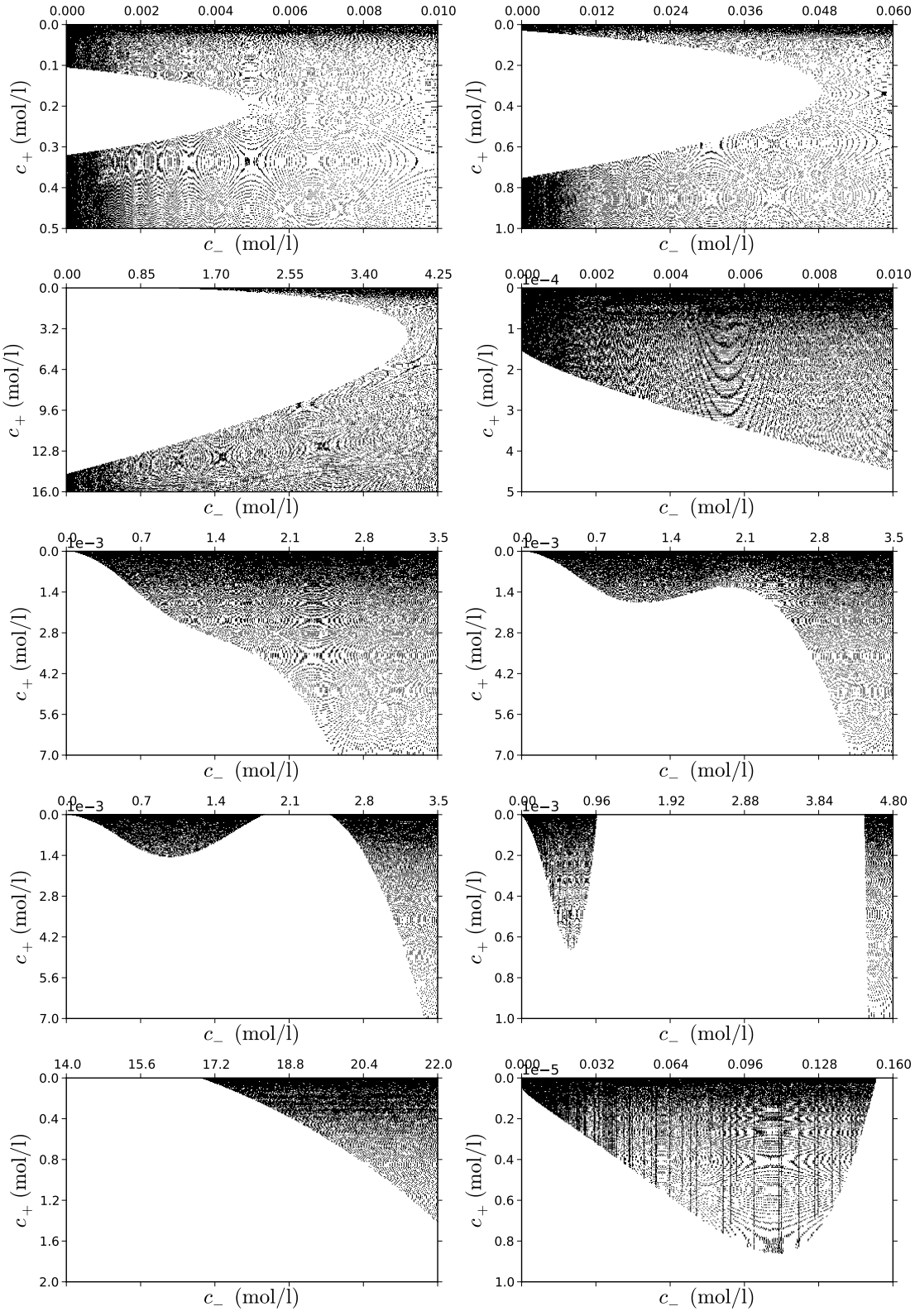

Figure 5.9.: The non-convexity region for model parameters $Z_+ = +3$ and $Z_- = -2$; from top to bottom and from left to right, the value of σ is respectively 0.5 nm, 0.48 nm, 0.3 nm, 0.3 nm (enlargement of the region close to the origin), 0.225 nm, 0.224 nm, 0.2235 nm, 0.221 nm, 0.2 nm and 0.2 nm (enlargement of the region close to the origin).

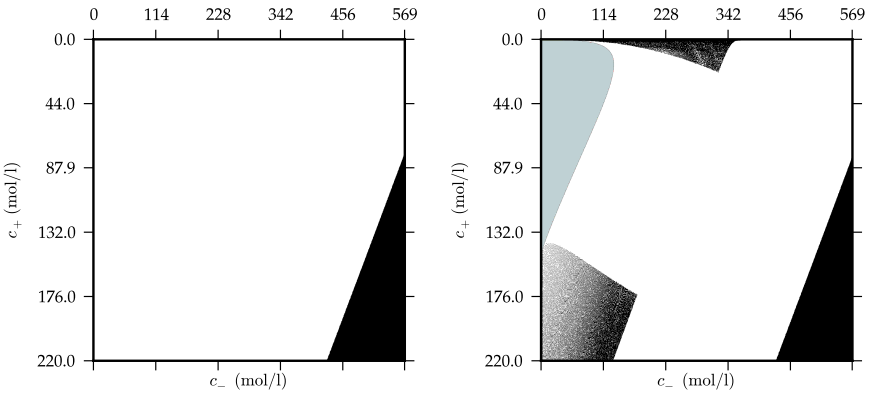

Figure 5.10.: The non-convexity convexity region for model parameters $Z_+ = +3$, $Z_- = -1$ and $\sigma = 0.17$. In the left figure a uniform dual grid is used, whereas in the right figure an adaptive dual grid is used; in color the region detected by the euristic flagging.

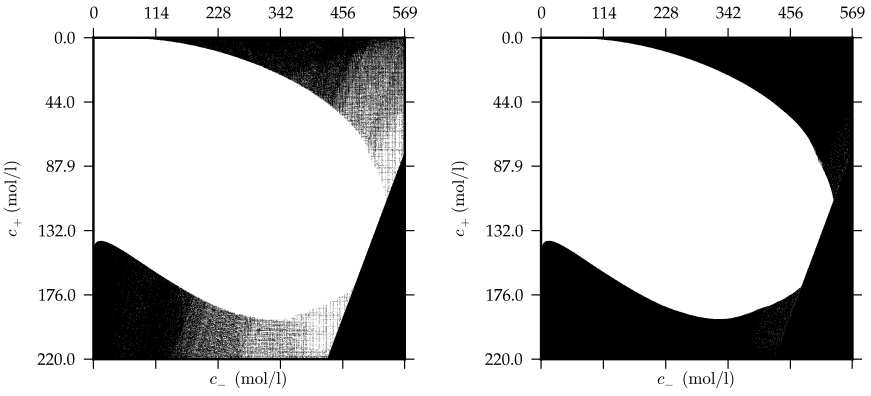

Figure 5.11.: The non-convexity convexity region for model parameters $Z_+ = +3$, $Z_- = -1$ and $\sigma = 0.17$. A progressive construction of the result is made, respectively in 5 (left) and 500 steps (right).

6. Phase separation in electrolytes

In this chapter the model presented in Chapter 1 is extended to a two-species electrolyte confined in a negatively charged medium ([4]); the equations for such a model involve the convex hull of the bulk free-energy density, which we have computed in Chapter 5. The equations are then solved with a numerical procedure based on a finite elements discretization and on the Newton-Raphson method (8).

6.1. The general model

We consider a two-species electrolyte for which the interaction with the charged walls is no more neglegible. For simplicity we assume a periodic setting consisting in the repetition of an elementary cell $[0, L_*]^n$, with $n = 2, 3$. Inside the elementary cell there are inclusions Ω_S whose boundary $\partial\Omega_S$ is negatively charged with surface density Σ_S; the electrolyte is contained in the remaining region $\Omega := [0, L_*]^n \setminus \Omega_S$. The resulting differential equations are

$$-\triangle\psi = \frac{e}{\varepsilon}\sum_{i=\pm} Z_i c_i \text{ in } \Omega, \tag{6.1}$$
$$\mu_+(\psi, c) \text{ and } \mu_-(\psi, c) \text{ are constant in } \Omega.$$

With respect to the bulk electrolyte case studied in Chapter 1, we have that ψ is not constant and thus the Poisson equation (6.1) is no longer satisfied trivially.

We have the following boundary conditions

$$\begin{aligned} &\psi \text{ is periodic on } \partial\Omega \setminus \partial\Omega_S, \\ &\nabla\psi \cdot \mathbf{n} = -\frac{1}{\varepsilon}\Sigma_S \text{ on } \partial\Omega_S, \end{aligned} \tag{6.2}$$

where $\mathbf{n}$ is the outward-pointing normal versor to $\partial\Omega_S$; we also fix the mean ionic concentration, i.e. require that

$$\langle c_i \rangle_\Omega = c_i^0, \quad i = \pm,$$

where c_+^0 and c_-^0 are given positive real numbers and where $\langle f \rangle_\Omega := \frac{1}{|\Omega|}\int_\Omega f$ for every $f \in L^1(\Omega)$. We must take $c_\pm^0$ such that they satisfy the global electro-neutrality condition

$$\sum_{i=\pm} Z_i c_i^0 = \frac{1}{|\Omega|}\int_{\partial\Omega_S} \frac{1}{e}\Sigma_S; \tag{6.3}$$

this is a necessary and sufficient condition for the solvability of the Poisson equation (6.1) together with the boundary conditions. Finally we require that

$$\langle \psi \rangle_\Omega = 0,$$

which is always possible because ψ is determined up to an addictive constant.

In the non-dimensional formulation we take the quantity $\Sigma_{S*} := \frac{k_B T \varepsilon}{e L_*}$ as the reference value for the surface charge density; we then have that equations (6.1), (6.2) and (6.3) become

$$\begin{aligned} -\lambda \triangle \psi &= \sum_{i=\pm} Z_i c_i \text{ in } \Omega, \\ \nabla \psi \cdot \mathbf{n} &= -\Sigma_S \text{ on } \partial\Omega_S, \\ \sum_{i=\pm} Z_i c_i^0 &= \frac{\lambda}{|\Omega|} \int_{\partial\Omega_S} \Sigma_S. \end{aligned}$$

6.2. The free-energy functional

We define the free-energy functional $\mathcal{F}$ as

$$\mathcal{F}(\psi, c) := \mathcal{S}(c) + \mathcal{B}(c, \psi) - \mathcal{U}(\psi), \tag{6.4}$$

where

$$\begin{aligned} \mathcal{S}(c) &:= \int_\Omega f(c) \\ \mathcal{B}(\psi, c) &:= \sum_{i=\pm} \int_\Omega Z_i c_i \psi \\ \mathcal{U}(\psi) &:= \frac{\lambda}{2} \int_\Omega |\nabla \psi|^2 + \lambda \int_{\partial\Omega_S} \Sigma_S \psi. \end{aligned}$$

We also define the correct functional spaces to which the electrostatic potential ψ and the concentration c must belong. We define

$$\begin{aligned} \mathfrak{H} &:= \{\psi \in H^1_{\text{per}}(\Omega) \mid \langle \psi \rangle_\Omega = 0\} \\ \mathfrak{K} &:= \{c \in \left[L^2_{\text{per}}(\Omega)\right]^2 \mid c \geq 0 \text{ a.e. in } \Omega \text{ and } \langle c \rangle_\Omega = c^0\}, \end{aligned}$$

where $H^1_{\text{per}}(\Omega)$ and $L^2_{\text{per}}(\Omega)$ are, respectively, the closure of $C^\infty_{\text{per}}(\overline{\Omega})$, the space of periodic and infinitely differentiable functions in $\overline{\Omega}$, for the norms $\|\cdot\|_{H^1(\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$. In 4 it is shown that if $\mathcal{F}$ admits a minimum point on the set $\mathfrak{H} \times \mathfrak{K}$, then that point is a solution of the model equations; it is also proved that, under some hypotheses which assure the convexity of f, there is one and only one minimum point of $\mathcal{F}$. As we mentioned in Section 1.5, these sufficient conditions for the convexity of f consist in a lower limit on the ion diameter σ; this limit value depends on the valences $Z_\pm$, on the temperature T

and on the relative permittivity ε_r (the values for some of this lower limits were given in Table 1.1).

When f is no longer convex, phase separation appears in the electrolyte. In order to solve the problem numerically, since we do not know where the sharp interface between the phases lies, we add the following regularization term to the free energy

$$\frac{\varepsilon_{\text{reg}}}{2}\int_\Omega \left(|\nabla c_+|^2 + |\nabla c_-|^2\right);$$

this term penalizes high values of the gradient of c, making thus the phase transition the smoother the greater the parameter ε_{reg} is. Thanks to the regularization, we can solve numerically our problem for each $\varepsilon_{\text{reg}} > 0$ without knowing a priori where the sharp interface lies and then consider the limit behaviour for $\varepsilon_{\text{reg}} \to 0$ as the correct solution for the phase separation problem; this procedure is common in phase-field theory and is known as the sharp interface limit.

6.3. Finite elements discretization

In order to resolve the minimization problem we follow the same approach of 8, which is based on a finite elements discretization and on a Newton-Raphson method. We will briefly highlight the main points of the procedure; for a more exhaustive presentation we refer to 8.

In order to account for the constraints contained in the definitions of $\mathfrak{H}$ and $\mathfrak{K}$, we will employ Lagrange multipliers; the actual minimization will involve the functional

$$\begin{aligned}\tilde{\mathcal{F}}(\psi, c, \lambda, \lambda_+, \lambda_-): \;&=\; \mathcal{F}(\psi) + \varepsilon_{\text{reg}}\int_\Omega \left(|\nabla c_+|^2 + |\nabla c_-|^2\right) + \\ &\quad \lambda\langle\psi\rangle_\Omega + \lambda_+\left(\langle c_+\rangle_\Omega - c_+^0\right) + \lambda_-\left(\langle c_-\rangle_\Omega - c_-^0\right)\end{aligned}$$

on the set $H^1_{\text{per}}(\Omega)\times\{c \in \left[L^2_{\text{per}}(\Omega)\right]^2 \mid c \geq 0 \text{ a.e. in } \Omega\}\times\mathbb{R}^3$. Now we write the derivatives of $\tilde{\mathcal{F}}$ with respect to each variable

$$\begin{aligned}\frac{\partial\tilde{\mathcal{F}}}{\partial\psi} &: \phi \mapsto \sum_{i=\pm}\int_\Omega Z_i c_i \phi - \lambda\int_\Omega \nabla\psi\cdot\nabla\phi - \lambda\int_{\partial\Omega_S}\Sigma_S\psi + \lambda\langle\phi\rangle_\Omega,\\ \frac{\partial\tilde{\mathcal{F}}}{\partial c_+} &: \phi_+ \mapsto \int_\Omega \frac{\partial f}{\partial c_+}(c)\phi_+ + \int_\Omega Z_+\psi\phi_+ + \varepsilon_{\text{reg}}\int_\Omega \nabla c_+\cdot\nabla\phi_+ + \lambda_+\langle\phi_+\rangle_\Omega,\\ \frac{\partial\tilde{\mathcal{F}}}{\partial c_-} &: \phi_- \mapsto \int_\Omega \frac{\partial f}{\partial c_-}(c)\phi_- + \int_\Omega Z_-\psi\phi_- + \varepsilon_{\text{reg}}\int_\Omega \nabla c_-\cdot\nabla\phi_- + \lambda_-\langle\phi_-\rangle_\Omega,\\ \frac{\partial\tilde{\mathcal{F}}}{\partial\lambda} &= \langle\psi\rangle_\Omega, \quad \frac{\partial\tilde{\mathcal{F}}}{\partial\lambda_+} = \langle c_+\rangle_\Omega - c_+^0, \quad \frac{\partial\tilde{\mathcal{F}}}{\partial\lambda_-} = \langle c_-\rangle_\Omega - c_-^0.\end{aligned}$$

We must require that all the derivatives are zero, obtaining

$$\begin{cases} \sum_{i=\pm} \int_\Omega Z_i c_i \phi - \lambda \int_\Omega \nabla\psi \cdot \nabla\phi - \lambda \int_{\partial\Omega_S} \Sigma_S \psi + \lambda \langle\phi\rangle_\Omega = 0 \\ \int_\Omega \frac{\partial f}{\partial c_+}(c)\phi_+ + \int_\Omega Z_+ \psi\phi_+ + \varepsilon_{\text{reg}} \int_\Omega \nabla c_+ \cdot \nabla\phi_+ + \lambda_+ \langle\phi_+\rangle_\Omega = 0 \\ \int_\Omega \frac{\partial f}{\partial c_-}(c)\phi_- + \int_\Omega Z_- \psi\phi_- + \varepsilon_{\text{reg}} \int_\Omega \nabla c_- \cdot \nabla\phi_- + \lambda_- \langle\phi_-\rangle_\Omega = 0 \\ \langle\psi\rangle_\Omega = 0 \\ \langle c_+\rangle_\Omega = c_+^0 \\ \langle c_-\rangle_\Omega = c_-^0 \end{cases}$$

for all $\phi \in H^1_{\text{per}}$ and all $\phi_+, \phi_- \in L^2_{\text{per}}$. In order to discretize these equations, we restrict them to a periodic finite elements space V_h of dimension N, for example the space of periodic conforming piecewise $\mathbb{P}_1$ functions; we can thus reduce the number of equations to $3N+3$, by taking ϕ only among the basis vectors of V_h, and change the domain of the problem from a functional one to $\mathbb{R}^{3N+3}$, by decomposing ψ, c_+ and c_- with respect to the basis of V_h. Unfortunately, the derivatives of f are not linear and thus the resulting system of equations is not linear; we have to use an iterative method for its solution, such as the Newton-Raphson method. Since the second derivatives of f are discontinuous, it may happen that after an iteration the values of the concentrations become negative or large enough to make $f(c) = +\infty$; in order to avoid this problem, we clip the function c at each step, forcing its values to lie in the correct domain of f. The sharp interface limit process is implemented by gradually decreasing the value of ε_{reg}, each time starting the iterative method from the last computed solution; moreover, the grid is refined in the area where the last solution for c had values inside the region of non-convexity of f, in order to be able to capture the phase transition, which is steeper for small values of ε_{reg}. When the regularization parameter is sufficiently small and the interface sufficiently sharp, it is also possible to set $\varepsilon_{\text{reg}} = 0$ and solve the problem without regularization.

6.4. Test cases

As a test case we will consider an electrolyte in a flat nanochannel, i.e. an electrolyte confined between two planes negatively charged whose distance is $L_* = 1$ nm; we will suppose that the solvent is water at a temperature of $T = 300$ K (at this temperature we have $\varepsilon_r = 78.3$). Because of the symmetry of model, ψ and c depend only on the distance from the nearest wall; thus we can consider the one-dimensional case in which Ω is a segment whose length is half the distance between the planes and only one extreme of Ω is charged. In 8 solutions are computed for the one-species and symmetric two-species cases: in the former the exact one-dimensional hull for f is used, while in the latter an approximation for the two-dimensional hull of f is obtained by reducing the problem to a one-dimensional one by making use of the symmetry of f. First we will check our code in the one-species case, then we will try to validate the approximate result in the symmetric two-species case and finally we will study an asymmetric two-species case not considered in 8.

We begin with the case in which only one ionic species is present; let $Z_+ = +3$, $\sigma = 0.45$ nm and $\Sigma_S = 0.1$ C/m^2. The results for decreasing values of ε_{reg} down to 0

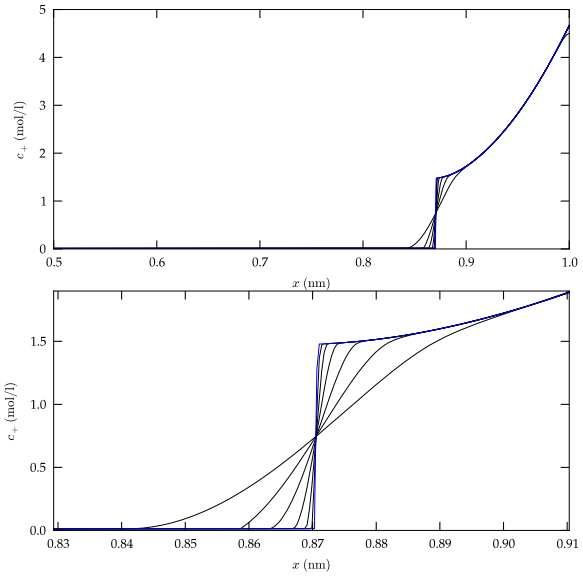

Figure 6.1.: Values of the concentration c_+ for a trivalent counterion of diameter 0.45 nm; the right plot is an enlargement of the interface between the two phases. The values of ε_{reg} used are 10^{-i} with $i = -5, \ldots, -10$ and 0 (blue line).

are shown in Figure 6.1 ($x = 0.5$ nm corresponds to the center of the nanochannel, while $x = 1$ nm corresponds to any one of the walls); as expected, they agree with the results found in 8.

Unlike the one-species case, where the mean concentration c^0 is completely determined by the electro-neutrality condition, in the two-species case we have to specify in what proportion the two ions are present; we will codify this information in a parameter $c_{\text{salt}} \geq 0$, which represents the concentration of added salt with respect to the situation in which there is only one ion species. Since there were no negatively charged ions before the addition of salt and the salt itself is electrically neutral, we must have that

$$Z_+ c_{\text{salt}} + Z_- c_-^0 = 0,$$

and thus

$$c_-^0 = -\frac{Z_+}{Z_-} c_{\text{salt}};$$

by substituting in the electro-neutrality relation, we obtain the other mean concentration as

$$c_+^0 = c_{\text{salt}} + \frac{\lambda}{Z_+ |\Omega|} \int_{\partial\Omega_S} \Sigma_S.$$

Now we consider a symmetric mixture with $|Z_+| = |Z_-| = 2$, $\sigma = 0.22$ nm, $\Sigma_S = 0.2$ C/m^2 and $c_{\text{salt}} = 0.1$ mol/l. The results for $\varepsilon_{\text{reg}} = 10^{-17}$ are shown in Figure 6.2; they are in excellent accordance with the approximate solution obtained in 8.

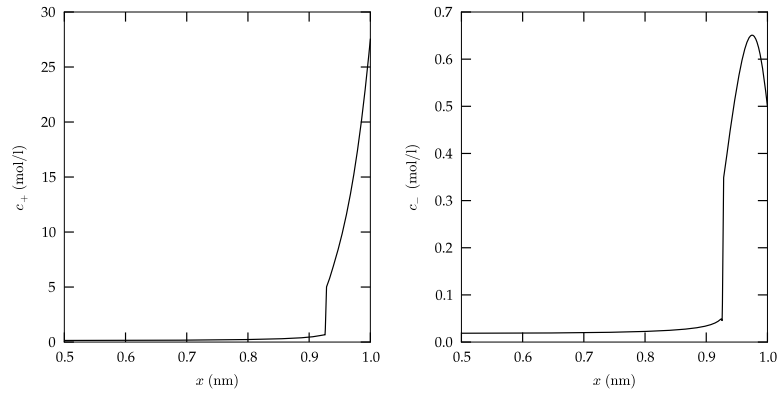

Figure 6.2.: Values of the concentrations c_+ and c_- for a symmetric mixture with $Z_\pm = \pm 2$ and an ion diameter of 0.22 nm; the regularization parameter has value $\varepsilon_{\text{reg}} = 10^{-17}$.

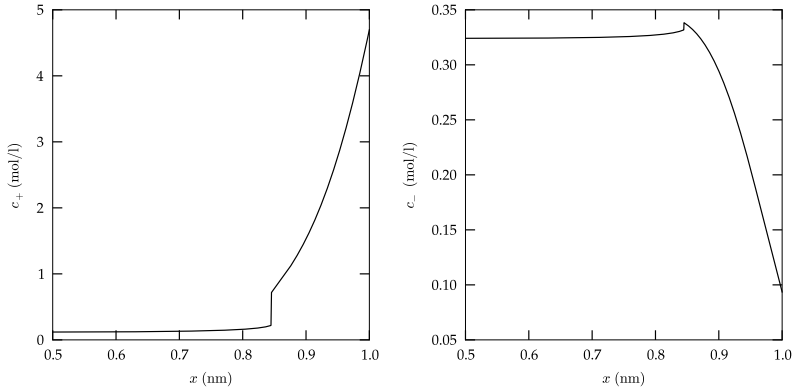

Figure 6.3.: Values of the concentrations c_+ and c_- for an asymmetric mixture with $Z_+ = +3$, $Z_- = -1$ and an ion diameter of 0.45 nm; the regularization parameter has value $\varepsilon_{\text{reg}} = 0$.

As a final example, we take an asymmetric electrolyte; this case was not considered in 8, because it requires the computation of the two-dimensional convex hull. The situation considered is described by $Z_+ = +3$, $Z_- = -1$, $\sigma = 0.45$ nm, $\Sigma_S = 0.1$ C/m^2 and $c_{\text{salt}} = 0.1$ mol/l; the solution, shown in Figure 6.3, is obtained for $\varepsilon_{\text{reg}} = 0$. The asymmetric case is increasingly challenging from the numerical point of view as we decrease the value of σ and the height of the phase transition grows; obtaining a solution for $\sigma \leq 0.4$ nm could be the focus of further studies.

Conclusions

In the thesis we have studied an algorithm for the approximation of convex hull of a multi-dimensional function which is based on the discrete Legendre-Fenchel transform. Our interest was motivated by the applications to the model of a two-species electrolyte

in a porous and electrically charged medium. After recalling the basic results for the Legendre-Fenchel transform and its discrete version, we have analyzed the properties of two successive applications of such transforms. We have also treated some problems, which are often neglected in the literature, such as the relevance of the dual grid, and proved some results. After the description of the standard algorithm for the convex hull approximation found in the literature, we have introduced some improvements and we have tested the new versions against the standard one. By applying it to the electrolyte bulk free-energy density in two dimensions, we have obtained the shape of the region of non-convexity; this region is thermodynamically unstable and divides the state space in different phases. Finally, by these results we have approximated the solution of the equations for two-species electrolytes, showing clearly that phase separation arises when the free-energy is not convex. However, we were not able to obtain the solution when the mixture is asymmetric and the ion diameter is small. Further improvements to the numerical solver are needed in order to treat successfully any kind of electrolyte.

A. Convex Functions

The brief theoretical introduction about convex functions and their conjugates is adapted from 13; another introductory book on convex analysis is 2.

A.1. Definition

Definition A.1. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. The set

$$\operatorname{epi} f := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid y \geq f(x)\}$$

is called *epigraph* of f.

Definition A.2. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be *convex* if $\operatorname{epi} f$ is convex as a subset of $\mathbb{R}^{n+1}$.

It is possible to define convexity for a finite function defined on a subset S of $\mathbb{R}^n$, but with the definition given above it is no longer necessary to keep track of the function domains. Moreover, a function $g : S \to \overline{\mathbb{R}}$ can be extended to $\mathbb{R}^n$ by taking

$$\tilde{g}(x) = \begin{cases} g(x) & \text{if } x \in S \\ +\infty & \text{if } x \notin S \end{cases};$$

then $\operatorname{epi} g = \operatorname{epi} \tilde{g}$ as subsets of $\mathbb{R}^{n+1}$ and g is convex as a function on S if and only if $\tilde{g}$ is convex according to Definition A.2.

Remark A.3. In general $\operatorname{conv} \operatorname{epi} f \neq \operatorname{epi} \operatorname{conv} f$; consider for example a gaussian function.

Definition A.4. Let f be a convex function. The projection of $\operatorname{epi} f$ on $\mathbb{R}^n$ is called the *effective domain* of f and is denoted by $\operatorname{dom} f$.

It is immediately observed that

$$\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$$

and that $\operatorname{dom} f$ is a convex subset of $\mathbb{R}^n$, since it is the image of $\operatorname{epi} f$ under a linear transformation, i.e. the projection on $\mathbb{R}^n$.

Definition A.5. A convex function f is said to be *proper* if $f(x) < +\infty$ for at least one $x \in \mathbb{R}^n$ and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$, otherwise it is called *improper*.

Thus a proper function is finite and convex on the not-empty convex set $\operatorname{dom} f$, whereas it is $+\infty$ elsewhere.

Definition A.6. The *convex hull* of a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, indicated by $\operatorname{conv} f$, is the greatest convex function majorized by f or, equivalently, the function whose epigraph is the convex hull of $\operatorname{epi} f$.

Proposition A.7. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convex such that $f(x_0) = -\infty$ for a certain $x_0 \in \mathbb{R}^n$. Then $f = -\infty$ on $\operatorname{int}\operatorname{dom} f$.*

Proof. Let $x_1 \in \operatorname{int}\operatorname{dom} f$. Because x_1 is an internal point of $\operatorname{dom} f$, the point $x_1 + \varepsilon(x_1 - x_0)$ lies in $\operatorname{dom} f$ for $\varepsilon > 0$ small enough; we will denote such a point by the symbol x_2. Let $x_\lambda := (1-\lambda)x_0 + \lambda x_2$; for a certain $\lambda_1 \in]0,1[$ we have that $x_{\lambda_1} = x_1$. Because $\operatorname{epi} f$ is convex, we must have that $f(x_1) = f(x_{\lambda_1}) \leq (1-\lambda_1)f(x_0) + \lambda_1 f(x_2) = -\infty$ and thus $f(x_1) = -\infty$. □

A.2. Semicontinuous hulls

Definition A.8. Given $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, we indicate with the symbol $\underline{f}$ the *lower semicontinuous hull* of f, i.e. the greatest lower semicontinuous function majorized by f; similarly, we denote by $\overline{f}$ the *upper semicontinuous hull* of f, i.e. the smallest upper semicontinuous function minorized by f.

The semicontinuous hulls have an important characterization, presented in the following Theorem.

Theorem A.9. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and let $x_0 \in \mathbb{R}^n$. We then have*

$$\underline{f}(x_0) = \sup_{V \in \mathcal{U}(x_0)} \inf_{x \in V} f(x),$$

where $\mathcal{U}(x_0)$ is the collection of all neighbourhoods of x_0. Moreover, fixed $U \in \mathcal{U}(x_0)$, we have that

$$\underline{f}(x_0) = \sup_{V \in \mathcal{U}(x_0) | V \subseteq U} \inf_{x \in V} f(x).$$

The same characterization holds for the upper semicontinuous hull.

Thanks to the characterization presented, it is easy to highlight the local nature of the semicontinuous hull operation.

Lemma A.10. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be such that $f \geq g$ on a neighbourhood U of x_0. Then $\underline{f}(x_0) \geq \underline{g}(x_0)$, $\overline{f}(x_0) \geq \overline{g}(x_0)$ and $\underline{(\overline{f})}(x_0) \geq \underline{(\overline{g})}(x_0)$.*

Proof. By Theorem A.9 we have that

$$\begin{aligned} \underline{f}(x_0) &= \sup_{V \in \mathcal{U}(x_0) | V \subseteq U} \inf_{x \in V} f(x) \\ &\geq \sup_{V \in \mathcal{U}(x_0) | V \subseteq U} \inf_{x \in V} g(x) \\ &= \underline{g}(x_0); \end{aligned}$$

the same proof can be used for the inequality involving the upper semicontinuous hull. In order to prove the third inequality, observe that by definition of neighbourhood we can find an open set V such that $x_0 \in V \subset U$; we obviously have that $f \geq g$ on V. The open set V is a neighbourhood of any point $x \in V$ and thus by the second part of the lemma we have $\overline{f} \geq \overline{g}$ on V; by the first part we then have $\underline{(\overline{f})}(x_0) \geq \underline{(\overline{g})}(x_0)$. □

Corollary A.11. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be such that $f = g$ on a neighbourhood U of x_0. Then $\underline{f}(x_0) = \underline{g}(x_0)$, $\overline{f}(x_0) = \overline{g}(x_0)$ and $\underline{(\overline{f})}(x_0) = \underline{(\overline{g})}(x_0)$.*

A.3. Closure

The notion of lower semicontinuity in the case of a convex function is often called closedness; this is due to the following result.

Proposition A.12. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, not necessarily convex. Then the following conditions are equivalent:*

(i) $\operatorname{epi} f$ *is closed as a subset of* $\mathbb{R}^{n+1}$*;*

(ii) f is lower semicontinuous in $\mathbb{R}^n$*.*

Proof. Suppose (i) true. Consider two sequences $x_n \in \mathbb{R}^n$ e $y_n \in \mathbb{R}$ such that $\lim x_n = x$, $\lim y_n = y$ and $y_n \geq f(x_n)$ for all n; this means (x_n, y_n) is a sequence of points of $\operatorname{epi} f$ tending in $\mathbb{R}^{n+1}$ to (x, y). Because f is lower semicontinuous $f(x) \leq \lim f(x_n) \leq \lim y_n = y$ and thus $(x, y) \in \operatorname{epi} f$.

Suppose (ii) true. In order to demonstrate (i) is sufficient to show that the set $L_\alpha := \{x \in \mathbb{R}^n \mid f(x) \leq \alpha\}$ is closed for all $\alpha \in \mathbb{R}$. Let $x_n \in L_\alpha$ a sequence tending to $x \in \mathbb{R}^n$ and consider the sequence (x_n, α). This sequence is in $\operatorname{epi} f$, its limit is still in $\operatorname{epi} f$ for closedness and thus $x \in L_\alpha$. □

Fact A.13. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then* $\operatorname{epi} \underline{f} = \operatorname{cl} \operatorname{epi} f$.

We can now define the notion of closedness of convex functions.

Definition A.14. Let f be a convex function. If $f(x) > -\infty$ for all $x \in \mathbb{R}^n$, the *closure* of f is defined to be the lower semicontinuous hull of f and is indicated by $\operatorname{cl} f$. Otherwise, $\operatorname{cl} f$ is defined as the constant function $-\infty$.

Definition A.15. A convex function f is said to be *closed* if $\operatorname{cl} f = f$.

For proper functions (and the constant $+\infty$), closedness is equivalent to lower semicontinuity (or closedness of the epigraph); moreover, it can be proved that if f is proper then it can differ from $\operatorname{cl} f$ only on the boundary of $\operatorname{dom} f$ (on open subsets of $\operatorname{dom} f$ the convex function f is even continuous). The exception made in the definition for functions assuming the value $-\infty$ is needed to make Theorem B.12 valid even when the function is improper and limits the possibile improper closed functions to the constants $+\infty$ e $-\infty$.

Example A.16. The following function is convex and proper, but not closed.

$$f(x) = \begin{cases} +\infty & \text{if } x < 0 \\ 1 & \text{if } x = 0 \\ 0 & \text{if } x > 0 \end{cases}$$

Remark A.17. If f is a closed convex function, we do not necessarily have that $\operatorname{dom} f$ is a closed set. Consider for example

$$f(x) = \begin{cases} \frac{1}{1-|x|} & \text{if } x \in]-1,1[\\ +\infty & \text{elsewhere} \end{cases};$$

we have that f is a closed convex function, but that $\operatorname{dom} f =]-1,1[$.

A.4. Subdifferentiability

Definition A.18. Given $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convex and $x \in \mathbb{R}^n$, a vector $\xi \in \mathbb{R}^n$ is called a *subgradient* of f at x if the affine function $z \mapsto f(x) + \langle \xi, z - x \rangle$ is majorized by f. The collection of all subgradients of f at x is called the subdifferential of f at x and is denoted by $\partial f(x)$.

The subdifferential is a local notion, as shown in the next result.

Lemma A.19. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex functions such that $f = g$ on a neighbourhood U of $x_0 \in \mathbb{R}^n$. Then $\partial f(x_0) = \partial g(x_0)$.*

Proof. By symmetry of the enunciate, it suffices to prove that if $\xi \in \partial f(x_0)$ then $\xi \in \partial g(x_0)$. We will suppose $\xi \notin \partial g(x_0)$ and show that this takes us to a contradiction. If $\xi \notin \partial g(x_0)$ then there exists $\tilde{x} \in \mathbb{R}^n$ such that

$$g(x_0) + \langle \xi, \tilde{x} - x_0 \rangle > g(\tilde{x}). \tag{A.1}$$

By convexity of g the segment with extremes $(x_0, g(x_0))$ and $(\tilde{x}, g(\tilde{x}))$ lies in $\operatorname{epi} g$, i.e. for all $x \in \mathbb{R}^n$ on the segment with extremes $\tilde{x}$ and x_0 we have by (A.1) that

$$\begin{aligned} g(x) &\leq g(x_0) + \frac{g(\tilde{x}) - g(x_0)}{\|\tilde{x} - x_0\|} \|x - x_0\| \\ &< f(x_0) + \frac{\langle \xi, \tilde{x} - x_0 \rangle}{\|\tilde{x} - x_0\|} \|x - x_0\| \\ &= f(x_0) + \langle \xi, x - x_0 \rangle , \end{aligned}$$

where $\|x\| = \langle x, x \rangle$ and the last equality holds because $x - x_0$ and $\tilde{x} - x_0$ lies on the same line through the origin. Because U is a neighbourhood of x_0, we can find $x \in U$ which lies on the segment with extremes $\tilde{x}$ and x_0 ; for such a point we have

$$f(x) = g(x) < f(x_0) + \langle \xi, x - x_0 \rangle ,$$

in contradiction with the hypothesis $\xi \in \partial f(x_0)$. □

Proposition A.20. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ not necessarily convex and let $x, \xi \in \mathbb{R}^n$ be such that $f(x) + \langle \xi, z - x \rangle \leq f(z)$ for every $z \in \mathbb{R}^n$. Then $\xi \in \partial \operatorname{conv} f(x)$.*

Proof. The function $z \mapsto f(x) + \langle \xi, z - x \rangle$ is affine and thus convex; by definition of convex hull we then have that $f(x) + \langle \xi, z - x \rangle \leq \operatorname{conv} f(z)$ for every $z \in \mathbb{R}^n$. Because $\operatorname{conv} f(x) \leq f(x)$, we conclude that $\operatorname{conv} f(x) + \langle \xi, z - x \rangle \leq \operatorname{conv} f(z)$ for every $z \in \mathbb{R}^n$, i.e. $\xi \in \partial \operatorname{conv} f(x)$. □

Proposition A.21. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ a proper convex function and let $x \in \mathbb{R}^n$. We have that*

(i) $\partial f(x)$ is a closed convex set;

(ii) if $x \notin \operatorname{dom} f$ then $\partial f(x) = \emptyset$;

(iii) if $x \in \operatorname{int} \operatorname{dom} f$ then $\partial f(x) \neq \emptyset$.

Proof. Let $\xi_1, \xi_2 \in \partial f(x)$ and $0 \leq \lambda \leq 1$; we will show that $\lambda \xi_1 + (1 - \lambda)\xi_2 \in \partial f(x)$. We have indeed for any $z \in \mathbb{R}^n$ that

$$\begin{aligned} f(x) + \langle \lambda \xi_1 + (1-\lambda)\xi_2, z - x \rangle &= \lambda \left[f(x) + \langle \xi_1, z - x \rangle \right] + \\ &\quad (1 - \lambda) \left[f(x) + \langle \xi_2, z - x \rangle \right] \\ &\leq \lambda f(z) + (1 - \lambda) f(z) = f(z). \end{aligned}$$

Let now ξ_n a sequence in $\partial f(x)$ such that $\xi_n \to \bar{\xi}$; we will show that $\bar{\xi} \in \partial f(x)$. For any $z \in \mathbb{R}^n$ and for every n it holds $f(x) + \langle \xi_n, z - x \rangle \leq f(z)$; by limiting, we immediately obtain that $f(x) + \langle \bar{\xi}, z - x \rangle \leq f(z)$ for every z.

Suppose $x \notin \operatorname{dom} f$ and let $x_0 \in \operatorname{dom} f$, which exists because f is proper. By hypothesis there exists $\xi \in \mathbb{R}^n$ such that $\xi \in \partial f(x)$; we then have

$$+\infty = f(x) + \langle \xi, x_0 - x \rangle \leq f(x_0) < +\infty,$$

which is contradictory.

Suppose $x \in \operatorname{int} \operatorname{dom} f$. We have that $\operatorname{epi} f$ is a convex set and $(x, f(x)) \in \partial \operatorname{epi} f$; by the known properties of convex sets, there is an hyperplane through $(x, f(x))$ such that $\operatorname{epi} f$ lies entirely in one of the semispaces in which $\mathbb{R}^{n+1}$ results divided. This hyperplane cannot be of the form $\{(z, y) \in \mathbb{R}^{n+1} \mid \langle \xi, z - x \rangle = 0\}$ because x is an internal point of $\operatorname{dom} f$; then it has to be the graph of an affine function through $(x, f(x))$ majorized by f, and thus $\partial f(x) \neq \emptyset$. □

Remark A.22. In the case $x \in \partial \operatorname{dom} f$ it may be possible that $\partial f(x) = \emptyset$. Consider for example $f(x) = -\sqrt{x}$ for $x \geq 0$ and $+\infty$ elsewhere; we have $\partial f(0) = \emptyset$.

Theorem A.23. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ a proper convex function and let $x_0 \in \mathbb{R}^n$ such that $\partial f(x_0) \neq \emptyset$. Then $\operatorname{cl} f(x_0) = f(x_0)$.*

Proof. Because $\partial f(x_0) \neq \emptyset$, there is an affine function h such that $h(x_0) = f(x_0)$ and $h \leq f$. Because f is proper, we have $\operatorname{cl} f = \underline{f}$ and thus, being h a lower semicontinuous function, we have $h \leq \operatorname{cl} f \leq f$; the thesis follows from evaluating the last inequality in x_0. □

Corollary A.24. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ a proper convex function. Then $\operatorname{cl} f = f$ on $\operatorname{int} \operatorname{dom} f$.*

Proof. By joint application of Proposition A.21 and Theorem A.23. □

Theorem A.25. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ a proper convex function. Then f is continuous on any open convex set contained in $\operatorname{dom} f$.*

A.5. Subdifferentiability: the one-dimensional case

Now we study in detail and give the proofs for the one dimensional case.

Definition A.26. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$. For every $x_1, x_2 \in \mathbb{R}$ we define

$$R_f(x_1, x_2) := \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Theorem A.27. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ convex and $x_1, x_2, x_3 \in \operatorname{dom} f$ such that $x_1 < x_2 < x_3$. Then*

$$R_f(x_1, x_2) \leq R_f(x_1, x_3) \leq R_f(x_2, x_3).$$

Proof. Because f is convex, the points of the segment joining $(x_1, f(x_1))$ and $(x_3, f(x_3))$ are in $\operatorname{epi} f$; the slope of the line containing this segment is $R_f(x_1, x_3)$. Thus, for all $x_1 \leq x \leq x_3$, we have

$$f(x_1) + R_f(x_1, x_3)(x - x_1) = f(x_3) + R_f(x_1, x_3)(x - x_3) \geq f(x);$$

in particular

$$f(x_1) + R_f(x_1, x_3)(x_2 - x_1) = f(x_3) + R_f(x_1, x_3)(x_2 - x_3) \geq f(x_2),$$

from which the thesis immediately follows. □

Corollary A.28. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ convex. Let $x_1, x_2, x_3, x_4 \in \operatorname{dom} f$ such that $x_1 < x_2 < x_3 < x_4$. Then $R_f(x_1, x_2) \leq R_f(x_3, x_4)$.*

Corollary A.29. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ convex. Fix $x_1 \in \operatorname{dom} f$; then $R_f(x_1, x_2)$ is a increasing function in x_2 on $\mathbb{R}$.*

Proof. When $x_2 \in \operatorname{dom} f$ it is sufficient to apply Theorem A.27 by observing that $R_f(x, x') = R_f(x', x)$ for all $x, x' \in \operatorname{dom} f$; otherwise, $R(x_1 x_2)$ is $+\infty$ when $x_2 > x_1$ and $-\infty$ when $x_2 < x_1$, and thus globally increasing. □

Theorem A.30. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ convex and let $x_0 \in \operatorname{int}\operatorname{dom} f$, where $\operatorname{int}\operatorname{dom} f$ is the interior of $\operatorname{dom} f$; then the left and right derivatives of f in x_0, denoted by $f'_-(x_0)$ and $f'_+(x_0)$, exist, are finite and satisfy $f'_-(x_0) \leq f'_+(x_0)$.*

Proof. By definition

$$f'_+(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0^+} R_f(x_0, x).$$

This limits exists because the $R_f(x_0, \cdot)$ is monotone by Corollary A.29; it is finite because for all $x > x_0$ we have $R_f(x_0, x) \leq R_f(x_0, x_0 - \varepsilon) < +\infty$ where $\varepsilon > 0$ is such that $x_0 - \varepsilon \in \operatorname{dom} f$ (such a ε exists because x_0 is an internal point of $\operatorname{dom} f$). The same reasoning can be applied to $f'_-(x_0)$, while the inequality $f'_-(x_0) \leq f'_+(x_0)$ follows by limiting the inequality $R_f(x_0, x_0 - h) \leq R_f(x_0, x_0 + h)$ for $h \to 0^+$. □

Remark A.31. Because $\operatorname{dom} f$ is a convex set of $\mathbb{R}$, it is an interval or a half-line; for $x_0 \in \partial \operatorname{dom} f$ only one of $f'_-(x_0)$ and $f'_+(x_0)$ exists, by the same reasoning of Theorem A.30, but it can be also $+\infty$ or $-\infty$.

Corollary A.32. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ convex. Then f is continuous on $\operatorname{int}\operatorname{dom} f$.*

Remark A.33. The finiteness of $f'_-(x_0)$ and $f'_+(x_0)$ is necessary; for $x_0 \in \partial \operatorname{dom} f$, if $f'_-(x_0)$ ($f'_+(x_0)$) is $+\infty$ ($-\infty$), then f is not necessarily left (right) continuous.

Corollary A.34. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ convex and let $x_1, x_2 \in \operatorname{int}\operatorname{dom} f$ such that $x_1 < x_2$. Then*

$$f'_-(x_1) \leq f'_+(x_1) \leq f'_-(x_2) \leq f'_+(x_2).$$

Proof. By Theorem A.30, it suffices to prove $f'_+(x_1) \leq f'_-(x_2)$, i.e.

$$f'_+(x_1) = \lim_{h \to 0^+} R_f(x_1, x_1 + h) \leq \lim_{h \to 0^+} R_f(x_2, x_2 - h) = f'_-(x_2);$$

this is true by limiting the inequality obtained by Corollary A.28 (which is applicable because for h sufficiently small we have $x_1 + h < x_2 - h$). □

Proposition A.35. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ convex and $x_0 \in \operatorname{dom} f$. Then*

(i) if $f'_-(x_0)$ is finite, then $f'_-(x_0) = \min \partial f(x_0)$;

(ii) if $f'_+(x_0)$ is finite, then $f'_+(x_0) = \max \partial f(x_0)$;

(iii) if $x_0 \in \operatorname{int}\operatorname{dom} f$, then $\partial f(x_0) = [f'_-(x_0), f'_+(x_0)]$;

(iv) if x_0 is the left extreme of $\operatorname{dom} f$, then $\partial f(x_0) =]-\infty, f'_+(x_0)]$ if $f'_+(x_0)$ is finite and it is $\emptyset$ otherwise;

(v) if x_0 is the right extreme of $\operatorname{dom} f$, then $\partial f(x_0) = [f'_-(x_0), +\infty[$ if $f'_-(x_0)$ is finite and it is $\emptyset$ otherwise.

Proof. We first prove that $f'_-(x_0) \in \partial f(x_0)$, i.e.

$$f(x_0) + f'_-(x_0)(x - x_0) \leq f(x) \qquad \forall x \in \mathbb{R}; \tag{A.2}$$

this means proving (the case $x = x_0$ is trivial)

$$\begin{cases} f'_-(x_0) \leq R(x_0, x) & \forall x > x_0 \\ f'_-(x_0) \geq R(x, x_0) & \forall x < x_0 \end{cases},$$

which follows immediately from Corollary A.29. If $x_0 \in \partial \operatorname{dom} f$, equation (A.2) is trivial on one of the half-lines originating from x_0, while on the other Corollary A.29 is again applicable.

We now prove that for every $\xi < f'_-(x_0)$ we have $\xi \notin \partial f(x_0)$. As seen before, we have that $R(x_0, x) \nearrow f'_-(x_0)$ for $x \to x_0^-$; by definition of limit we can find $\tilde{x} < x_0$ such that $\xi < R(x_0, \tilde{x}) \leq f'_-(x_0)$. Thus

$$f(\tilde{x}) = f(x_0) + R(x_0, \tilde{x})(\tilde{x} - x_0) < f(x_0) + \xi(\tilde{x} - x_0),$$

which means $\xi \notin \partial f(x_0)$.

The same reasoning can also be applied to (ii).

In order to demonstrate (iii), observe that if $x_0 \in \operatorname{int} \operatorname{dom} f$ then, thanks to Theorem A.30, we can conclude by (i) and (ii) that $f'_-(x_0) = \min \partial f(x_0)$ and $f'_+(x_0) = \max \partial f(x_0)$. We can now conclude thanks to Proposition (A.21).

In order to prove (iv), suppose initially that if $f'_+(x_0)$ exists finite. Then (ii) is applicable and all that is left to prove is that for all $\xi \leq f'_+(x_0)$ we have $f(x_0) + \xi(x - x_0) \leq f(x)$. This is true because for $x > x_0$ we have $\xi(x - x_0) \leq f'_+(x_0)(x - x_0)$, while for $x < x_0$ we have $f(x) = +\infty$. If $f'_+(x_0)$ is not finite, it is $-\infty$ and by a reasoning similar to the one used in the proof of (i) we can show that $\xi \notin \partial f(x_0)\ \forall \xi \in \mathbb{R}$. The proof of (v) is similar. □

A.6. A characterization of the convex hull

Lemma A.36. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then*

$$\operatorname{conv} f(x) = \inf_{\{(\alpha_i, x_i)\}_i \in C(x)} \sum_{i \in I} \alpha_i f(x_i),$$

where $C(x)$ is the collection of the finite sets $\{(\alpha_i, x_i)\}_{i \in I}$ such that x is convex combination of the points x_i weighted by coefficients $\alpha_i \in [0, 1]$. In the sum $\sum_{i \in I} \alpha_i f(x_i)$ the conventions $+\infty - \infty = +\infty$ and $0 \times \pm\infty = 0$ are used.

Proof. Fixed $x \in \mathbb{R}^n$, for every set I with the required properties we must have $\operatorname{conv} f(x) \leq \sum_{i \in I} \alpha_i f(x_i)$ by convexity of the epigraph and thanks to the the special conventions used in the sum; thus we have $\operatorname{conv} f(x) \leq \inf \sum_{i \in I} \alpha_i f(x_i)$ for every x. In order to prove the other inequality, it suffices to show that the right side, which we will denote by $g(x)$, is

convex and that $g \leq f$, because by definition conv f is the largest of the convex functions majorized by f. The fact that $g \leq f$ is trivial because among the sets on which the infimus is taken there is $\{(1, x)\}$; in order to prove convexity, let then $(x_1, y_1), (x_2, y_2) \in \operatorname{epi} g$. Being $y_1 \geq g(x_1)$, for any $\varepsilon > 0$ by the properties of the infimum we can find I_1 such that $\sum_{i \in I_1} \alpha_i f(x_i) < y_1 + \varepsilon$; similarly we can find I_2 such that $\sum_{i \in I_2} \alpha_i f(x_i) < y_2 + \varepsilon$. By weighting the $(x_i)_{i \in I_1}$ with weights $(\lambda \alpha_i)_{i \in I_1}$ and the $(x_i)_{i \in I_2}$ with weights $((1-\lambda)\alpha_i)_{i \in I_2}$ we obtain a convex combination of $\lambda x_1 + (1 - \lambda)x_2$; we then have

$$\begin{aligned} g(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda \sum_{i \in I_1} \alpha_i f(x_i) + (1-\lambda) \sum_{i \in I_2} \alpha_i f(x_i) \\ &< \lambda y_1 + (1-\lambda) y_2 + \varepsilon. \end{aligned}$$

Because $\varepsilon > 0$ is arbitrary, we then have $\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) \in \operatorname{epi} g$, i.e. g is convex. □

Corollary A.37. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then* $\operatorname{dom} \operatorname{conv} f = \operatorname{conv}\{x \in \mathbb{R}^n \mid f(x) < +\infty\}$.

Corollary A.38. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then f is convex if and only if for any $x_1, x_2 \in \mathbb{R}^n$ and $\alpha \in [0, 1]$ we have that*

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2),$$

where the conventions established in Lemma A.36 are again used.

Proof. Necessity follows immediately from Lemma A.36. We will now prove sufficiency. Fix $x \in \mathbb{R}^n$ and $\{(\alpha_i, x_i)\}_{i=1,\ldots,N}$ such that x is convex combination of the points x_i weighted by coefficients $\alpha_i \in [0, 1]$. We will now prove that $f(x) \leq \sum_{i=1}^N \alpha_i f(x_i)$; by Lemma A.36 we can then conclude that $f(x) \leq \operatorname{conv} f(x)$ and thus $f(x) = \operatorname{conv} f(x)$. We will proceed by induction on N. The case $N = 1$ is trivial; suppose then that the property holds for a certain N and consider the case $N + 1$. We can decompose x as the convex combination with coefficient α_1 of x_1 and $\tilde{x}_2 := \frac{1}{1-\alpha_1} \sum_{i=2}^N \alpha_i x_i$; we can then apply to $\tilde{x}_2$ the induction hypothesis, obtaining that $f(\tilde{x}_2) \leq \frac{1}{1-\alpha_1} \sum_{i=2}^N \alpha_i f(x_i)$. Applying the hypothesis we can then conclude by observing that

$$\begin{aligned} f(x) &= f(\alpha_1 x_1 + (1-\alpha_1)\tilde{x}_2) \\ &\leq \alpha_1 f(x_1) + (1-\alpha_1) f(\tilde{x}_2) \\ &\leq \sum_{i=1}^N \alpha_i f(x_i). \end{aligned}$$

□

Corollary A.39. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ convex. Then $f + g$ is convex.*

Proof. Fix $x_1, x_2 \in \mathbb{R}^n$, $\alpha \in [0, 1]$ and denote by the symbol x the convex combination $\alpha x_1 + (1 - \alpha)x_2$. By applying Corollary A.38 to the convex functions f and g we obtain

that

$$\begin{aligned}(f+g)(x) &= f(x)+g(x)\\ &\leq \alpha f(x_1)+(1-\alpha)f(x_2)+\alpha g(x_1)+(1-\alpha)g(x_2)\\ &\leq \alpha(f+g)(x_1)+(1-\alpha)(f+g)(x_2);\end{aligned}$$

we can then conclude by applying again Corollary A.38. □

Corollary A.40. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then* $\operatorname{conv} f + \operatorname{conv} g \leq \operatorname{conv}(f+g)$.

Proof. Because $\operatorname{conv} f + \operatorname{conv} g \leq f+g$, after observing that by Corollary A.39 $\operatorname{conv} f + \operatorname{conv} g$ is convex, we can immediately conclude by definition of convex hull. □

Remark A.41. Equality does not hold in general. Consider for example the functions

$$f(x) = \begin{cases} 0 & \text{if } x \in [0,1] \\ 1 & \text{elsewhere} \end{cases}$$

and

$$g(x) = \begin{cases} 1 & \text{if } x \in [0,1] \\ 0 & \text{elsewhere} \end{cases}.$$

We have that $\operatorname{conv} f = \operatorname{conv} g \equiv 0$, while $\operatorname{conv}(f+g) = f+g \equiv 1$.

Theorem A.42 (Carathéodory)**.** *Let $\Omega \subseteq \mathbb{R}^n$ and $x \in \operatorname{conv}\Omega$. Then there is $m \leq n+1$ such that x can be expressed as a convex combination of m elements of Ω.*

Proof. Because $x \in \operatorname{conv}\Omega$ there are $x_1, \ldots, x_m \in \Omega$ and $\alpha_1, \ldots, \alpha_m \in [0,1]$ such that $\sum_{i=1}^m \alpha_i x_i = x$. We can suppose that m is the minimal number of points for which such a representation is possible; we thus also have that $\alpha_i > 0$ for all i. We will now prove that $m \leq n+1$. If it were $m > n+1$, the vectors $\{x_i - x_1\}_{i=2,\ldots,m}$ would be linearly dependent, i.e. there would be some scalars $\{\lambda_i\}_{i=2,\ldots,m}$ such that $\sum_{i=2}^m \lambda_i(x_i - x_1) = 0$ and $\lambda_i > 0$ for at least one i (at least one coefficient is not zero by linear dependence; if it is negative, we can change the sign of all λ_i obtaing a strictly positive coefficient). Defining $\lambda_1 = -\sum_{i=2}^m \lambda_i$, we then have

$$\sum_{i=1}^m \lambda_i x_i = 0 \text{ and } \sum_{i=1}^m \lambda_i = 0.$$

Then for every $\beta \in \mathbb{R}$ we have

$$\begin{aligned}x &= \sum_{i=1}^m \alpha_i x_i + \beta \sum_{i=1}^m \lambda_i x_i \\ &= \sum_{i=1}^m (\alpha_i + \beta\lambda_i) x_i;\end{aligned}$$

observe that we have $\sum_{i=1}^{m}(\alpha_i + \beta\lambda_i) = 1$. We now want to choose a $\beta \in \mathbb{R}$ such that $\alpha_i + \beta\lambda_i \geq 0$ for every i; by taking $\beta \leq 0$ the required property is true for every i such that $\lambda_i \leq 0$. In order to make it hold also for all i such that $\lambda_i > 0$ (as we have seen there is at least one such i), we need also to satisfy $\beta \geq -\frac{\alpha_i}{\lambda_i}$ for all such i; being $\frac{\alpha_i}{\lambda_i} \geq 0$ this is possible. We can in particular take $\beta = -\max_{i:\lambda_i>0} \frac{\alpha_i}{\lambda_i}$; with this choice at least one of the m coefficients $\alpha_i + \beta\lambda_i$ is zero. We have thus expressed x a convex combination of less than m elements of Ω, in contradiction with the minimality of m. □

Corollary A.43. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then*

$$\operatorname{conv} f(x) = \inf_{\{(\alpha_i,x_i)\}_i \in C_{n+2}(x)} \sum_{i=1}^{n+2} \alpha_i f(x_i),$$

where $C_{n+2}(x)$ is the collection of the sets $\{(\alpha_i, x_i)\}_{i\in\{1,\dots,n+2\}}$ such that x is the convex combination of the points x_i weighted by coefficients $\alpha_i \in [0, 1]$. In the sum $\sum_{i=1}^{n+2} \alpha_i f(x_i)$ the conventions $+\infty - \infty = +\infty$ and $0 \times \pm\infty = 0$ are used.

Proof. By Lemma A.36, we have to prove that

$$\inf_{\{(\alpha_i,x_i)\}_i \in C(x)} \sum_{i\in I} \alpha_i f(x_i) = \inf_{\{(\alpha_i,x_i)\}_i \in C_{n+2}(x)} \sum_{i=1}^{n+2} \alpha_i f(x_i).$$

The inequality $\leq$ is easily proved because $C_{n+2}(x) \subset C(x)$; to prove the other direction we will show that for every element of $C(x)$ there is an element of $C_{n+2}(x)$ for which the value of the sum is lower equal. Let $\{(\alpha_i, x_i)\}_{i\in I} \in C(x)$. Consider the point $(x, y) \in \mathbb{R}^{n+1}$, where $y = \sum_{i\in I} \alpha_i f(x_i)$; we have that $(x, y) = \sum_{i\in I} \alpha_i (x_i, f(x_i))$. By Theorem A.42 we can then find $J \subseteq I$ such that $|J| = n + 2$ and $(x, y) = \sum_{i\in J} \alpha_i (x_i, f(x_i))$; then $\{(\alpha_i, x_i)\}_{i\in J} \in C_{n+2}(x)$ and the values of the two sums are equal. □

A.7. Convex hulls of compact sets

Lemma A.44. *Let $\Omega \subset \mathbb{R}^n$ a bounded set. Then* $\operatorname{conv}\Omega$ *is a bounded set.*

Proof. Let $x \in \operatorname{conv}\Omega$; we can then write it as

$$x = \sum_{i=1}^{N} \alpha_i x_i,$$

where $x_i \in \Omega$ and $\alpha_i \in [0, 1]$ for every $i \in \{1, \dots N\}$ and $\sum_{i=1}^{N} \alpha_i = 1$. Because Ω is bounded, there exists $R \in \mathbb{R}$ such that $\|y\| \leq R$ for every $y \in \Omega$; we thus have

$$\|x\| \leq \sum_{i=1}^{N} \alpha_i \|x_i\| \leq \sum_{i=1}^{N} \alpha_i R = R,$$

which, by generality of x, means that $\operatorname{conv}\Omega$ is bounded. □

Example A.45. If Ω is closed, $\operatorname{conv}\Omega$ is not in general closed. Take for example $\Omega \subset \mathbb{R}^2$ as the epigraph of a gaussian function; we have that $\operatorname{conv}\Omega$ is the open set $\{(x,y) \in \mathbb{R}^2 \mid y > 0\}$.

Corollary A.46. *If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is finite on a bounded set $\Omega \subseteq \mathbb{R}^n$ and $+\infty$ elsewhere, then $\operatorname{dom}\operatorname{conv} f = \operatorname{conv}\Omega$ is a bounded set.*

Proof. By Corollary A.37 we have that $\operatorname{dom}\operatorname{conv} f = \operatorname{conv}\Omega$; it is then sufficient to apply Lemma A.44. □

Lemma A.47. *Let $\Omega \subset \mathbb{R}^n$ a compact set. Then $\operatorname{conv}\Omega$ is a compact set.*

Proof. Consider a sequence of points $x^{(N)} \in \operatorname{conv}\Omega$; by Theorem A.42 we can write them as

$$x^{(N)} = \sum_{i=1}^{n+1} \alpha_i^{(N)} x_i^{(N)},$$

where $\alpha_i^{(N)} \in [0,1]$ and $x_i^{(N)} \in \Omega$ for every N and every $i = 1, \dots, n+1$. We then have that $(\alpha_1^{(N)}, \dots, \alpha_{n+1}^{(N)}, x_1^{(N)}, \dots, x_{n+1}^{(N)}) \in [0,1]^{n+1} \times \Omega^{n+1}$, which is a compact subset of $\mathbb{R}^{2(n+1)}$; there is thus a subsequence $(\alpha_1^{(k_N)}, \dots, \alpha_{n+1}^{(k_N)}, x_1^{(k_N)}, \dots, x_{n+1}^{(k_N)})$ converging to a certain $(\bar\alpha_1, \dots, \bar\alpha_{n+1}, \bar x_1, \dots, \bar x_{n+1}) \in [0,1]^{n+1} \times \Omega^{n+1}$. By the properties of limits, we have that $\sum_{i=1}^{n+1} \alpha_i^{(k_N)} = 1$ tends to $\sum_{i=1}^{n+1} \bar\alpha_i$; we thus have that the point $\bar x := \sum_{i=1}^{n+1} \bar\alpha_i \bar x_i$ is a convex combination of points of Ω. Again by the properties of limits, $x^{(k_N)}$ converges then to $\bar x \in \operatorname{conv}\Omega$; thus every sequence in $\operatorname{conv}\Omega$ admits a subsequence converging to an element of Ω, i.e. $\operatorname{conv}\Omega$ is compact. □

Corollary A.48. *If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is finite on a compact set $\Omega \subset \mathbb{R}^n$ and $+\infty$ elsewhere, then $\operatorname{dom}\operatorname{conv} f = \operatorname{conv}\Omega$ is a compact set.*

Proof. By Corollary A.37 we have that $\operatorname{dom}\operatorname{conv} f = \operatorname{conv}\Omega$; it is then sufficient to apply Lemma A.47. □

A.8. Closedness of convex hulls

Theorem A.49. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ lower semicontinuous. Suppose that $\Omega := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$ is compact and that there exists $L \in \mathbb{R}$ such that $f \geq L$; then $\operatorname{conv} f$ is closed.*

Proof. Because $f \geq L$, being the constant L a convex function we have that $\operatorname{conv} f \geq L$; then proving that $\operatorname{conv} f$ is closed is equivalent to prove that $\operatorname{conv} f$ is lower semicontinuous, i.e. to prove that the set $S_\gamma := \{x \in \mathbb{R}^n \mid \operatorname{conv} f(x) \leq \gamma\}$ is closed for every $\gamma \in \mathbb{R}$. Consider a sequence of points $x^{(N)} \in S_\gamma$ converging to $\bar x \in \mathbb{R}^n$; we need to prove that $\bar x \in S_\gamma$. Being $L \leq \operatorname{conv} f(x^{(N)}) \leq \gamma$ and thus finite, by the characterization of Corollary A.43 and the properties of the infimum, for every N we can find $\alpha_i^{(N)} \in [0,1]$

and $x_i^{(N)} \in \mathbb{R}^n$ with $i \in \{1, \dots, n+2\}$ such that $\sum_{i=1}^{n+2} \alpha_i^{(N)} = 1$, $\sum_{i=1}^{n+2} \alpha_i^{(N)} x_i^{(N)} = x^{(N)}$ and

$$\operatorname{conv} f(x^{(N)}) + \varepsilon > \sum_{i=1}^{n+2} \alpha_i^{(N)} f(x_i^{(N)}) > \operatorname{conv} f(x^{(N)}). \tag{A.3}$$

By compactness of $[0,1]^{n+2} \times \Omega^{n+2}$ we can extract a subsequence (which again for simplicity will be denoted by the same symbol) such that $\alpha_i^{(N)} \to \bar{\alpha}_i \in [0,1]$ and $x_i^{(N)} \to \bar{x}_i \in \Omega$; moreover, by the properties of limits, we have that $\sum_{i=1}^{n+2} \bar{\alpha}_i = 1$ and $\bar{x} = \sum_{i=1}^{n+2} \bar{\alpha}_i x_i$. If the function

$$\begin{aligned} [0,1]^{n+2} \times \Omega^{n+2} &\to \mathbb{R} \\ (\alpha_i, x_i) &\mapsto \sum_{i=1}^{n+2} \alpha_i f(x_i) \end{aligned} \tag{A.4}$$

is lower semicontiuous, we could apply Corollary A.43 and inequality (A.49) to obtain

$$\operatorname{conv} f(\bar{x}) \leq \sum_{i=1}^{n+2} \bar{\alpha}_i f(\bar{x}_i) \leq \liminf_{N \to \infty} \sum_{i=1}^{n+2} \alpha_i^{(N)} f(x_i^{(N)}) \leq \gamma + \varepsilon;$$

by generality of ε we could then obtain $\operatorname{conv} f(\bar{x}) \leq \varepsilon$ as needed.

We will now prove the lower semicontinuity of (A.4); because the sum of lower semicontinuous functions is lower semicontinuous, it is sufficient to prove that the function $(\alpha, x) \mapsto \alpha f(x)$ is lower semicontinuous. Let $(x_N, \alpha_N) \to (\bar{x}, \bar{\alpha})$ such that $\alpha_N f(x_N) \leq \gamma$ for every N. We have $\alpha_N, \bar{\alpha} \in [0,1]$; suppose $\bar{\alpha} \neq 0$. By convergence, for N sufficiently large and $\varepsilon < \bar{\alpha}$ we have that $\alpha_N > \bar{\alpha} - \varepsilon > 0$; then have $f(x_N) \leq \frac{\gamma}{\alpha_N} < \frac{\gamma}{\bar{\alpha}-\varepsilon}$; by lower semicontinuity of f we then have $f(\bar{x}) \leq \frac{\gamma}{\bar{\alpha}-\varepsilon}$ and thus $\bar{\alpha} f(\bar{x}) \leq \frac{\bar{\alpha}}{\bar{\alpha}-\varepsilon}\gamma$ for every $\varepsilon < \bar{\alpha}$, which means $\bar{\alpha} f(\bar{x}) \leq \gamma$. If instead $\bar{\alpha} = 0$, we have to prove that $\gamma \geq 0$; if there is an N such that $\alpha_N = 0$ or $\alpha_N \neq 0$ and $f(x_N) \geq 0$ we have finished. Otherwise, by convergence for every $M > 0$ we can find N such that $0 < \alpha_N < \frac{1}{L+M}$; we then have $-L \leq f(x_N) < 0$ and thus $\gamma \geq \alpha_N f(x_N) \geq -\frac{L}{L+M}$, which by generality of M implies $\gamma \geq 0$. □

Now we present several situation which can occur when the hypothesis of Theorem A.49 do not hold.

Example A.50. A convex function which is not lower semicontinuous even if it is uniformly continuous when restricted to its domain (because $\operatorname{dom} f$ is open, f cannot be closed).

$$f(x) = \begin{cases} 0 & \text{if } x \in]0,1[\\ +\infty & \text{elsewhere} \end{cases}$$

Example A.51. A function which is uniformly continuous on $\mathbb{R}$, but whose convex hull is nowhere finite.

$$f(x) = -|x|$$

Example A.52. A lower semicontinuous function whose convex hull is not lower semicontinuous.

$$f(x) = \begin{cases} \tan x & \text{if } -\frac{\pi}{2} < x < \frac{\pi}{2} \\ -\infty & \text{if } x < -\frac{\pi}{2} \\ +\infty & \text{if } x > \frac{\pi}{2} \end{cases}$$

The function f is lower semicontinuous, but its convex hull $\operatorname{conv} f$ is $-\infty$ on the *open* set $]-\infty, \pi/2[$ and $+\infty$ elsewhere, and thus not lower semicontinuous.

Example A.53. A convex, proper and closed function which is not continuous.

$$f(x,y) = \begin{cases} \frac{y^2}{2x} & \text{if } x > 0 \\ 0 & \text{if } x = 0 \wedge y = 0 \\ +\infty & \text{elsewhere} \end{cases}$$

The function f is convex, proper and closed, but it is not continuous in $(0,0)$, point on the boundary of $\operatorname{dom} f$. This can be seen, for example, approaching the origin along the curve $x = \frac{y^2}{2\alpha}$ with $\alpha \neq 0$: on this curve f is α constantly.

Example A.54. A function finite on a bounded set and continuous on it whose convex hull is not continuous on its domain and not even lower semicontinuous.

$$\begin{aligned} f(x,y) &= \begin{cases} \exp\left(-\frac{1}{1-|y|}x\right) & \text{if } (x,y) \in [0,1]\times]-1,1[\\ +\infty & \text{elsewhere} \end{cases} \\ \operatorname{conv} f(x,y) &= \begin{cases} 1 & \text{if } x = 0 \text{ and } y \in]-1,1[\\ 0 & \text{if } (x,y) \in]0,1]\times]-1,1[\\ +\infty & \text{elsewhere} \end{cases} \end{aligned}$$

Example A.55. A function finite on a close unbounded set and continuous on it (and thus lower semicontinuous on $\mathbb{R}^n$) whose convex hull is again not even lower semicontinuous on its domain.

$$\begin{aligned} f(x,y) &= \begin{cases} \exp\left(-(1+|y|)x\right) & \text{if } x \in [0,1] \\ +\infty & \text{elsewhere} \end{cases} \\ \operatorname{conv} f(x,y) &= \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \in]0,1] \\ +\infty & \text{elsewhere} \end{cases} \end{aligned}$$

A.9. Two more results on convex hulls

Theorem A.56. *Let $\Omega_N = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$ finite and let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ finite on Ω_N. Then*

(i) $\operatorname{conv} f_{\Omega_N}$ *is closed;*

(ii) $\partial \operatorname{conv} f_{\Omega_N}(x) \neq \emptyset$ *for every* $x \in \Omega_N$.

Proof. By the characterization of Theorem A.36 we have that

$$\operatorname{conv} f_{\Omega_N}(x) = \inf_{\alpha} \sum_{i=1}^{N} \alpha_i f(x_i),$$

where the infimum is taken on the set

$$\left\{ \alpha \in [0,1]^N \mid \sum_{i=1}^{N} \alpha_i x_i = x \wedge \sum_{i=1}^{N} \alpha_i = 1 \right\}.$$

Then by Corollary 19.1.2 of [13] (pag. 173) the function $\operatorname{conv} f_{\Omega_N}$ is polyhedral, i.e. $\operatorname{epi} \operatorname{conv} f_{\Omega_N}$ is polyhedral. Since $\operatorname{conv} f_{\Omega_N}(x)$ is finite when $x \in \Omega_N$, from Corollary 19.1.2 we get the first part of the thesis and from Theorem 23.10 of [13] (pag. 226) we get its second part. □

Theorem A.57. *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be finite on a closed convex set* Ω *and* $+\infty$ *elsewhere. Let* $x_0 \in \partial\Omega$ *and let* h *an hyperplane such that* $x_0 \in h$ *and that* Ω *lies entirely on one side of* h. *Then* $(\operatorname{conv} f)\,|_h = \operatorname{conv}(f|_h)$.

Proof. Because $x_0 \in \Omega$, by Theorem A.42 we can write it as

$$x_0 = \sum_{i=1}^{m} \alpha_i x_i,$$

where $m \leq n+1$, $\alpha_i \in]0,1[$, $x_i \in \Omega$ and $\sum_{i=1}^{m} \alpha_i = 1$. Suppose that $x_k \notin h$ for a certain index k; without loss of generality we can suppose $k = m$. If $m = 1$ this is obviously contradictory because $x_0 = x_m$. If $m \geq 2$, we then have

$$x_0 = \alpha_m x_m + (1 - \alpha_m)\tilde{x},$$

where

$$\tilde{x} = \sum_{i=1}^{m-1} \frac{\alpha_i}{1 - \alpha_m} x_i;$$

because Ω is convex, we must have that $\tilde{x} \in \Omega$. However, because $\alpha_m > 0$, the point $\tilde{x}$ must lie on the opposite side of h with respect to x_m: by hypothesis $\tilde{x}$ and x_m cannot be both in Ω and thus we have a contradiction.

Then x_0 can only be expressed as a convex combination of points of Ω if those points all lie on h; by the characterization of Corollary A.43 we then have the thesis. □

B. The Legendre-Fenchel transformation

B.1. Definition and basic properties

Definition B.1. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. The function

$$\begin{aligned} f^* : \mathbb{R}^n &\to \overline{\mathbb{R}} \\ \xi &\mapsto \sup_{x\in\mathbb{R}^n} \left[\langle x, \xi\rangle - f(x)\right] = -\inf_{x\in\mathbb{R}^n} \left[f(x) - \langle x, \xi\rangle\right], \end{aligned}$$

where $\langle\cdot,\cdot\rangle$ is the scalar product of $\mathbb{R}^n$, is called the *Legendre-Fenchel transform*, or *conjugate*, of f.

Remark B.2. We have $f^*(0) = -\inf_{x\in\mathbb{R}^n} f(x)$.

Proposition B.3. *Let $\{f_i\}_{i\in I}$ an arbitrary collection of closed convex functions. The function $f(x) := \sup_{i\in I}[f_i(x)]$ is then a closed convex function.*

Proof. If the collection is empty, then we have that $f(x) = -\infty$, which is a closed convex function; if $f_i(x) = -\infty$ for a certain $i \in I$, then it is possible to remove f_i from the collection without affecting f. It is easily verifiable, even for generic functions f_i, that $\operatorname{epi} f = \bigcap_{i\in I} \operatorname{epi} f_i$; for the ipothesis $\operatorname{epi} f_i$ are convex and closed and thus their (arbitrary) intersection is also convex and closed. □

Corollary B.4. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, not necessarily convex. Then f^* is a closed convex function.*

Proof. By definition f^* is the pointwise supremum of a collection of affine functions (which may include the constants $+\infty$ e $-\infty$); because these are closed convex functions, f^* is also convex and closed by the previous definition. □

Proposition B.5. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ such that $f \leq g$. Then $f^* \geq g^*$.*

Proof. It follows trivially from the Definition B.1. □

Remark B.6 (Geometric construction). Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, let $\xi \in \mathbb{R}^n$ and let h be the hyperplane through the origin having equation $y = \langle \xi, x\rangle$. Then, by Definition B.1, $f^*(\xi)$ is the supremum value of the signed distance along the y-axis between the graph of f and h; for an example in $\mathbb{R}$ see Figure B.1.

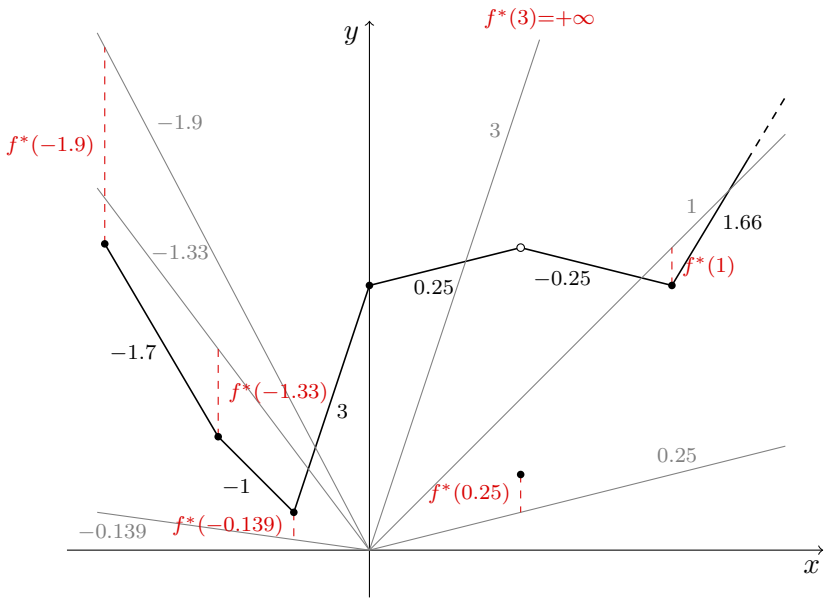

Figure B.1.: Geometric construction of the Legendre-Fenchel conjugate in $\mathbb{R}$. The function f and its slopes are drawn in black, the lines through the origin and their slopes are drawn in grey, while the distances representing the values of f^* are drawn in red; $f^*(3)$ is $+\infty$ because the distance between the line with slope 3 and f grows indefinetely for $x \to +\infty$.

B.2. Conjugation of convex functions

Lemma B.7. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, non necessarily convex. Then an affine function $h_{\xi,\mu}(x) := \langle \xi, x \rangle - \mu$, where $\xi \in \mathbb{R}^n$ and $\mu \in \mathbb{R}$, is majorized by f if and only if $(\xi, \mu) \in \operatorname{epi} f^*$.*

Proof. Suppose that $h_{\xi,\mu}$ is majorized by f. Then for all $x \in \mathbb{R}^n$ we have that $f(x) \geq h_{\xi,\mu}(x)$ and thus $\mu \geq \langle \xi, x \rangle - f(x)$; by definition of f^*, we conclude that $\mu \geq f^*(\xi)$, i.e.$(\xi, \mu) \in \operatorname{epi} f^*$. The other direction is similar. □

In the conjugate of f is thus coded the information necessary to determine the collection of affine functions which are majorized by f. It is possible to demonstrate that if f is convex then it coincides with the pointwise supremum of that collection; f^* is thus another rappresentation of the function f and it should be possible to find an inverse transform capable to reconstruct f starting from f^*: it will be shown that this inverse transform is the conjugate itself.

Lemma B.8. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a closed convex function. Then f is the pointwise supremum of the collection of all affine functions majorized by f.*

Proof. If f is improper, then the thesis is trivially true. If f is proper, then $\operatorname{epi} f$ is a non-empty closed convex subset of $\mathbb{R}^{n+1}$. By the known properties of convex sets, $\operatorname{epi} f$

is then the intersection of all closed half-spaces of $\mathbb{R}^{n+1}$ containing it. The closed half-spaces of $\mathbb{R}^{n+1}$ can be of one of three forms: $y \geq h(x)$, $y \leq h(x)$ or $\overline{h}(x) \leq 0$, where h and $\overline{h}$ are affine functions $\mathbb{R}^n \to \mathbb{R}$ and $\overline{h}$ is not constantly 0; we note that the half-spaces of the first form can also be indicated by $\operatorname{epi} h$. The collection whose intersection is $\operatorname{epi} f$ cannot contain half-spaces of the second form and must contain at least one half-space of the first form because f is proper; the half-spaces of the first form containing $\operatorname{epi} f$ are the ones characterized by an affine function majorized by f. Thus if we prove that by not including the half-spaces of the third form in the intersection we still obtain $\operatorname{epi} f$, we can conclude that $\operatorname{epi} f = \bigcap_{h \text{ affine},\, h \leq f} \operatorname{epi} h$ and thus f is the pointwise supremum of all h.

Let $V = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid \overline{h}(x) \leq 0\}$ where $\overline{h}$ is a non-null affine function. If V is to be not influent on the intersection, we have to find for each $(x_0, y_0) \notin V$ an affine function $h \leq f$ such that $(x_0, y_0) \notin \operatorname{epi} h$. As we have highlighted before, if f is proper there is at least one affine function $\tilde{h} \leq f$. Fixed $\lambda > 0$, we define the function $h_\lambda(x) := \lambda \overline{h}(x) + \tilde{h}(x)$, which, as a sum of two affine function, is itself affine. For every $x \in \operatorname{dom} f$ we have that $\overline{h}(x) \leq 0$ and $\tilde{h}(x) \leq f(x)$ and thus $h_\lambda \leq f$ on $\operatorname{dom} f$; the same is trivially true also outside $\operatorname{dom} f$, where $f(x) = +\infty$. Because $\overline{h}(x_0) > 0$, we can take $\bar{\lambda}$ sufficiently large to ensure that $h_{\bar{\lambda}}(x_0) > y_0$ and thus $(x_0, y_0) \notin \operatorname{epi} h_{\bar{\lambda}}$; $h_{\bar{\lambda}}$ is then the affine function needed. □

Lemma B.9. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, not necessarily convex. Then $f^* = \left(\underline{f}\right)^*$.*

Proof. For Lemma B.7 it suffices to show that the collections of affine functions which are majorized respectively by f and by $\underline{f}$ are the same. Let g be an affine function. If $g \leq \underline{f}$ then $g \leq f$, because by definition of $\underline{f}$ we have $\underline{f} \leq f$. If $g \leq f$, we have $g \leq \underline{f}$ because an affine function is lower semicontinuous and $\underline{f}$ is the greatest among all the lower semicontinuous functions majorized by f. □

Corollary B.10. *Let f be a convex function. Then $f^* = (\operatorname{cl} f)^*$.*

Definition B.11. If $f(x) > -\infty$ for all $x \in \mathbb{R}^n$, by definition we have $\operatorname{cl} f = \underline{f}$ and thus Lemma B.9 applies; otherwise, we have that $\operatorname{cl} f$ is the constant function $-\infty$. If there is a point x_0 such that $f(x_0) = -\infty$, then no affine function can be majorized by f and thus by Lemma B.7 we have $\operatorname{epi} f^* = \emptyset$; this means f^* is the constant function $+\infty$, which is the conjugate of the constant function $-\infty$.

Theorem B.12. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function. Then $\operatorname{cl} f = f^{**}$.*

Proof. In Corollary B.10, it will be proved that $f^* = (\operatorname{cl} f)^*$; we can then assume without loss of generality that f is closed. Because trivially $(+\infty)^* = -\infty$ and $(-\infty)^* = +\infty$, the result is immediate if f is improper. If f is proper and closed, Lemma B.8 says that

$$f(x) = \sup_{h \text{ affine},\, h \leq f} h(x),$$

which, by Lemma B.7, becomes

$$f(x) = \sup_{(\xi, \mu) \in \operatorname{epi} f^*} \left(\langle \xi, x \rangle - \mu\right).$$

Because $h_{\xi,\mu_1} \geq h_{\xi,\mu_2}$ when $\mu_1 \leq \mu_2$ and because, fixed ξ, the minimum value of μ such that $(\xi, \mu) \in \operatorname{epi} f^*$ is $f^*(\xi)$, the supremum simplifies to

$$f(x) = \sup_{\xi \in \mathbb{R}^n} \left[\langle \xi, x \rangle - f^*(\xi) \right],$$

which is equal by definition to $f^{**}(x)$. □

Corollary B.13. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be closed convex functions. Then $f = g$ if and only if $f^* = g^*$.*

We can now immediately prove a converse to Lemma B.7.

Lemma B.14. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function. Then an affine function $h_{x,y}(\xi) := \langle x, \xi \rangle - y$, where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$, is majorized by f^* if and only if $(x, y) \in \operatorname{epi} \operatorname{cl} f$.*

B.3. Conjugation of general functions

Lemma B.15. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then $f^* = (\operatorname{conv} f)^*$.*

Proof. We again apply Lemma B.7 and show that, given g an affine function, we have $g \leq f$ if and only if $g \leq \operatorname{conv} f$.If $g \leq \operatorname{conv} f$ then $g \leq f$, because by definition of convex hull $\operatorname{conv} f \leq f$. If $g \leq f$, we have $g \leq \operatorname{conv} f$ because affine functions are convex. □

Corollary B.16. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then $f^* = [\operatorname{cl}(\operatorname{conv} f)]^*$.*

Proof. It follows from Lemma B.15 and Corollary B.10. □

Theorem B.17. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, not necessarily convex. Then $f^{**} = \operatorname{cl}(\operatorname{conv} f)$.*

Proof. It follows directly by Lemma B.16 and Theorem B.12. □

B.4. Multi-dimensional conjugation

An important property of the Legendre-Fenchel conjugation is that the general transformation in dimension n can be reduced to a sequence of one-dimensional transformations; this property sets it apart from the convex hull operation and makes a double conjugation a more practical option for computing convex hulls in more than one dimension. To simplify notation, given $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and an index $i \in \{1, \dots n\}$, we use the symbol f^{*i} for the transformation along the ith-dimension, i.e.

$$\begin{aligned} f^{*i} : \mathbb{R}^n &\to \overline{\mathbb{R}} \\ (x_1, \dots, x_{i-1}, \xi_i, x_{i+1}, \dots, x_n) &\mapsto \sup_{x_i \in \mathbb{R}} \left[x_i \xi_i - f(x_1, \dots, x_i, \dots x_n) \right] \\ &\quad = \left[f(x_1, \dots, x_{i-1}, \cdot, x_i, \dots x_n) \right]^* (\xi_i). \end{aligned}$$

We then have the following theorem.

Theorem B.18. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then*

$$f^* = \left(- \left(\cdots \left(- \left(-f^{*1} \right)^{*2} \right)^{*3} \cdots \right)^{*(n-1)} \right)^{*n}.$$

Proof. It can be shown easily through factorization of the supremum operator:

$$\begin{aligned} f^*(\xi) &= \sup_{x \in \mathbb{R}^n} \left[\langle x, \xi \rangle - f(x) \right] \\ &= \sup_{x_n \in \mathbb{R}} \left[x_n \xi_n + \sup_{x_{n-1} \in \mathbb{R}} \left[x_{n-1} \xi_{n-1} + \cdots \right. \right. \\ &\quad \left. \left. \cdots \sup_{x_3 \in \mathbb{R}} \left[x_3 \xi_3 + \sup_{x_2 \in \mathbb{R}} \left[x_2 \xi_2 + \sup_{x_1 \in \mathbb{R}} \left[x_1 \xi_1 - f(x) \right] \right] \right] \right] \right]. \end{aligned}$$

□

B.5. Subdifferentials

Lemma B.19. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function. Then for all $x_0, \xi_0 \in \mathbb{R}^n$ the following propositions are true:*

(i) if $\xi_0 \in \partial f(x_0)$ then $\xi_0 \in \partial \left(\operatorname{cl} f \right)(x_0)$;

(ii) if $\xi_0 \in \partial \left(\operatorname{cl} f \right)(x_0)$ and $\operatorname{cl} f(x_0) = f(x_0)$ then $\xi_0 \in \partial f(x_0)$;

(iii) $\xi_0 \in \partial \left(\operatorname{cl} f \right)(x_0)$ if and only if $x_0 \in \partial f^*(\xi_0)$;

Proof. Suppose $\xi_0 \in \partial f(x_0)$; we then know that $h_1(x) = f(x_0) + \langle \xi_0, x - x_0 \rangle$ is majorized by f and we have to prove that $h_2(x) = \operatorname{cl} f(x_0) + \langle \xi_0, x - x_0 \rangle$ is majorized by $\operatorname{cl} f$. Because $\operatorname{cl} f \leq f$, we have that $h_2 \leq h_1$; then, because we have supposed that $h_1 \leq f$, we also have $h_2 \leq f$. Because h_2 is a lower semicontinuous, we then have $h_2 \leq \operatorname{cl} f$ by definition of closure. Suppose now that $\xi_0 \in \partial \left(\operatorname{cl} f \right)(x_0)$ and $\operatorname{cl} f(x_0) = f(x_0)$; we then know that $h_1 = h_2 \leq \operatorname{cl} f$ and we have to prove that $h_1 \leq f$, which is true because $\operatorname{cl} f \leq f$.

By application of Lemma B.7 and Lemma B.10we have

$$\begin{aligned} \xi_0 \in \partial \left(\operatorname{cl} f \right)(x_0) &\iff \operatorname{cl} f(x_0) + \langle \xi_0, \cdot - x_0 \rangle \leq \operatorname{cl} f \\ &\iff (\xi_0, \langle \xi_0, x_0 \rangle - \operatorname{cl} f(x_0)) \in \operatorname{epi} \left(\operatorname{cl} f \right)^* \\ &\iff \langle \xi_0, x_0 \rangle - \operatorname{cl} f(x_0) \geq \left(\operatorname{cl} f \right)^* (\xi_0) = f^*(\xi_0), \end{aligned}$$

while by application of Lemma B.14 we have

$$\begin{aligned} x_0 \in \partial f^*(\xi_0) &\iff f^*(\xi_0) + \langle x_0, \cdot - \xi_0 \rangle \leq f^* \\ &\iff (x_0, \langle x_0, \xi_0 \rangle - f^*(\xi_0)) \in \operatorname{epi} \operatorname{cl} f \\ &\iff \langle x_0, \xi_0 \rangle - f^*(\xi_0) \geq \operatorname{cl} f(x_0). \end{aligned}$$

□

Corollary B.20. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function and let $x_0, \xi_0 \in \mathbb{R}^n$ such that $\xi_0 \in \partial f(x_0)$. Then $x_0 \in \partial f^*(\xi_0)$.*

Remark B.21. The converse is not true. Consider for example

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x > 0 \\ +\infty & \text{if } x < 0 \end{cases}.$$

f is convex, but not closed; Moreover, we have that $\partial f(0) = \emptyset$. The conjugate of f is

$$f^*(\xi) = \begin{cases} 0 & \text{if } \xi \leq 0 \\ +\infty & \text{if } \xi > 0 \end{cases}.$$

Then, for any $\xi \leq 0$, we have $0 \in \partial f^*(\xi)$, but not $\xi \in \partial f(0)$. Nonetheless, we can prove that this problem can only arise on $\partial \operatorname{dom} f$, as shown in the next proposition.

Corollary B.22. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper convex function. Then for all $x_0 \in \operatorname{int} \operatorname{dom} f$ and for all $\xi_0 \in \mathbb{R}^n$ we have $\xi_0 \in \partial f(x_0)$ if and only if $x_0 \in \partial f^*(\xi_0)$.*

Proof. By joint application of Lemma B.19 and Corollary A.24. □

Corollary B.23. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a closed convex function. Then for all $x_0, \xi_0 \in \mathbb{R}^n$ we have $\xi_0 \in \partial f(x_0)$ if and only if $x_0 \in \partial f^*(\xi_0)$.*

Corollary B.24. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper convex function. Then for all $x_0, \xi_0 \in \mathbb{R}^n$ the following propositions are true:*

(i) if $\xi_0 \in \partial f(x_0)$ then $\xi_0 \in \operatorname{dom} f^*$;

(ii) if $x_0 \in \partial f^*(\xi_0)$ and $\operatorname{cl} f(x_0) = f(x_0)$ then $x_0 \in \operatorname{dom} f$;

Proof. It follows directly from the joint application of Lemma B.19 and Proposition A.21. □

Remark B.25. Converse propositions such as "if $\xi_0 \in \operatorname{dom} f^*$ then there exists $\tilde{x} \in \mathbb{R}^n$ such that $\xi_0 \in \partial f(\tilde{x})$" are not true in general; if the one presented were for example true, we would then have by Corollary B.20 that $\tilde{x} \in \partial f^*(\xi_0)$, i.e. $\partial f^*(\xi_0) \neq \emptyset$ which is not necessarily true, as seen in Remark A.22.

Theorem B.26. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function and let $x_0 \in \operatorname{dom} f$ and $\xi_0 \in \mathbb{R}^n$. Then*

$$f^*(\xi_0) = \langle x_0, \xi_0 \rangle - f(x_0)$$

if and only if

$$\xi_0 \in \partial f(x_0).$$

Proof. Suppose $\xi_0 \in \partial f(x_0)$; then

$$f(x_0) + \langle \xi_0, x - x_0 \rangle \leq f(x) \qquad \forall x \in \mathbb{R}^n.$$

This means

$$\langle \xi_0, x \rangle - f(x) \leq \langle \xi_0, x_0 \rangle - f(x_0) \qquad \forall x \in \mathbb{R}^n,$$

from which follows that

$$f^*(\xi_0) = \sup_{x \in \mathbb{R}^n} [\langle x, \xi_0 \rangle - f(x)] = \langle \xi_0, x_0 \rangle - f(x_0).$$

Suppose that $f^*(\xi_0) = \langle x_0, \xi_0 \rangle - f(x_0)$; by definition of conjugation we then have

$$\langle x, \xi_0 \rangle - f(x) \leq \langle x_0, \xi_0 \rangle - f(x_0)$$

for all $x \in \mathbb{R}^n$. By rearranging we obtain that

$$\langle x - x_0, \xi_0 \rangle + f(x_0) \leq f(x)$$

for all $x \in \mathbb{R}^n$, i.e. $\xi_0 \in \partial f(x_0)$. □

Corollary B.27. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex functions such that $f(x_0) = g(x_0)$ for a certain $x_0 \in \mathbb{R}^n$. Then for all $\xi \in \partial f(x_0) \cap \partial g(x_0)$ we have that $f^*(\xi) = g^*(\xi)$.*

Corollary B.28. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex functions such that $f = g$ on a neighbourhood U of a point $x_0 \in \mathbb{R}^n$. Then $f^* = g^*$ on $\partial f(x_0)$.*

Proof. It follows directly from Theorem B.26 and Lemma A.19. □

Corollary B.29. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex functions such that $f = g$ on an open set $\Omega \subseteq \mathbb{R}^n$. Then $f^* = g^*$ on $\bigcup_{x \in \Omega} \partial f(x)$.*

Bibliography

[1] A. M. Andrew. Another efficient algorithm for convex hulls in two dimensions. *Information Processing Letters*, 9(5):216–219, December 1979. doi: 10.1016/0020-0190(72)90045-2.

[2] Heinz H. Bauschke and Patrick L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, May 2011. ISBN 978-1441994660.

[3] Lucilla Corrias. Fast Legendre–Fenchel Transform and Applications to Hamilton–Jacobi Equations and Conservation Laws. *SIAM J. Numer. Anal.*, 33(4):1534–1558, 1996. doi: 10.1137/S0036142993260208.

[4] Alexandre Ern, Rémi Joubaud, and Tony Lelièvre. Mathematical study of non-ideal electrostatic correlations in equilibrium electrolytes. *Nonlinearity*, 25(6):1635–1652, June 2012. doi: 10.1088/0951-7715/25/6/1635.

[5] R. L. Graham. An efficient algorithm for determining the convex hull of a finite planar set. *Information Processing Letters*, 1(4):132–133, June 1972. doi: 10.1016/0020-0190(72)90045-2.

[6] Philippe Helluy and Hélène Mathis. Pressure laws and fast Legendre transform. *Math. Models Methods Appl. Sci.*, 21(4):745–775, April 2011. doi: 10.1142/S0218202511005209.

[7] Jean-Baptiste Hiriart-Urruty. Lipschitz r-continuity of the approximate subdifferential of a convex function. *Math. Scand.*, 47:123–134, 1980.

[8] Rémi Joubaud. *Mathematical and numerical modelling of fluids at nanometric scales*. PhD thesis, Université Paris-Est, November 2012.

[9] Yves Lucet. A fast computational algorithm for the Legendre-Fenchel transform. *Comput. Optim. Appl.*, 6(1):27–57, July 1996. doi: 10.1007/BF00248008.

[10] Yves Lucet. Faster than the Fast Legendre Transform, the Linear-time Legendre Transform. *Numer Algor*, 16(2):171–185, March 1997. doi: 10.1023/A:1019191114493.

[11] James Clerk Maxwell. On the Dynamical Evidence of the Molecular Constitution of Bodies. *Nature*, 11:357–359, March 1875. doi: 10.1038/011357a0.

[12] F. P. Preparata and S. J. Hong. Convex hulls of finite sets of points in two and three dimensions. *Communications of the ACM*, 20(2):87–93, February 1977. doi: 10.1145/359423.359430.

[13] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970. ISBN 0-691-08069-0.

[14] Jean Serra. *Image Analysis and Mathematical Morphology*. Academic Press, January 1983. ISBN 978-0126372427.