Elements of convex analysis

Hugo Touchette

Started: September 6, 2006; last compiled: October 23, 2013

Notes and extra references

- Updated 23/10/2013: Relation between $\mathrm{dom}\partial f$ and and $\mathrm{dom} f$ corrected (equal except possible at relative boundary points). Example of Rockafellar added. Results on concave points vs supporting line points added.
- See Appendix A of M. Costeniuc, R.S. Ellis, H. Touchette, and B. Turkington. The generalized canonical ensemble and its universal equivalence with the microcanonical ensemble. *J. Stat. Phys.*, 119:1283–1329, 2005. Available from: http://dx.doi.org/10.1007/s10955-005-4407-0.
- See pp. 1038-1042 of R. S. Ellis, K. Haven, and B. Turkington. Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. *J. Stat. Phys.*, 101:999–1064, 2000.
- S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.
- R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*, volume 317. Springer, New York, 1988. Available from: http://www.springerlink.com/content/978-3-540-62772-2.
- A. Bossavit. A course in convex analysis. 2003. Available from: http://butler.cc.tut.fi/~bossavit/BackupICM/CA.pdf.

1. Convex sets

Let A be a subset of $\mathbb{R}^n$.

- **Interior:** $\mathrm{int}(A)$
- **Closure:** $\mathrm{cl}(A)$
- **Relative interior:** $\mathrm{ri}(A)$. Interior of A relative to the smallest subspace containing A (defined technically as the interior relative to the affine hull of A). (Fig. 1) (VT, §4.8) (R, §6)
  - $\mathrm{int}(A)$ is the interior of A relative to $\mathbb{R}^n$.
  - $\mathrm{ri}(A) \subseteq A \subseteq \mathrm{cl}(A)$. (R, §6)
  - For $A \subseteq \mathbb{R}^n$, $\mathrm{ri}(A) = \mathrm{int}(A)$ if $\dim(A) = n$.

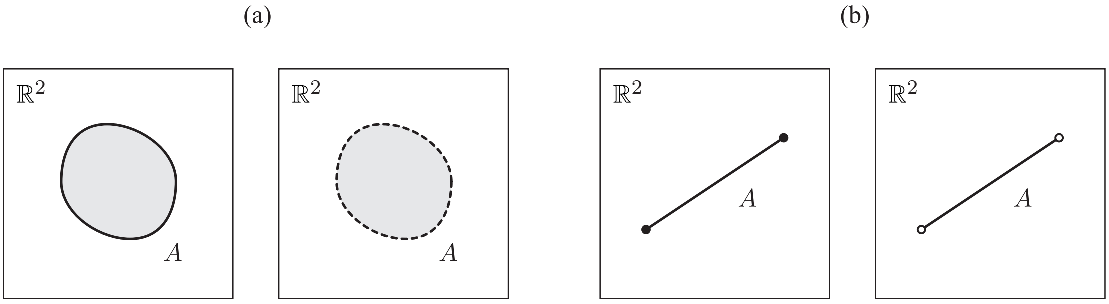

Figure 1: (a) $\text{int}(A) = \text{ri}(A)$. (b) $\text{int}(A) = \emptyset$ but $\text{ri}(A) \neq \emptyset$.

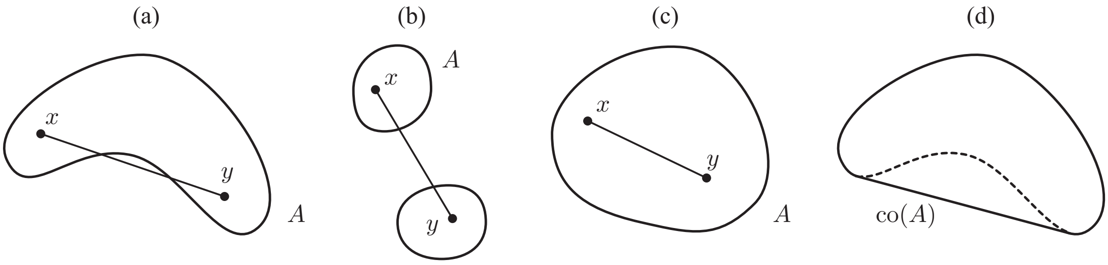

Figure 2: (a)-(b) Nonconvex sets. (c) Convex set. (d) Convex hull.

   - ri = int in 1D.

- **Convex set:** A is convex if $ax + (1-a)y \in A$ for all $x, y \in A$, $a \in [0,1]$. (Fig. 2) (B, §2)
   - Operations that preserve convexity: intersection, dilatation, addition, closure, linear transformations.
   - Convex sets are connected.
   - Convex sets have non-empty relative interiors.
- **Convex hull:** $\text{co}(A)$. Smallest convex set containing A.

---

2. Convex functions

Consider a function $f : X \to \mathbb{R}$, with $X \subseteq \mathbb{R}^n$.

- **Extended reals:** $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$
- **Extension of f:** (VT, §1.22)

$$\tilde{f}(x) = \begin{cases} f(x) & x \in X \\ \infty & x \notin X. \end{cases} \tag{1}$$

   - $\tilde{f}$ is a function of $\mathbb{R}^n$ to $\overline{\mathbb{R}}$.
   - One can always extend a function, so from now we consider only functions of $\mathbb{R}^n$ to $\overline{\mathbb{R}}$.
- **Effective domain:** $\text{dom}(f) = \{x \in \mathbb{R}^n : f(x) < \infty\}$. (VT, §5.11)

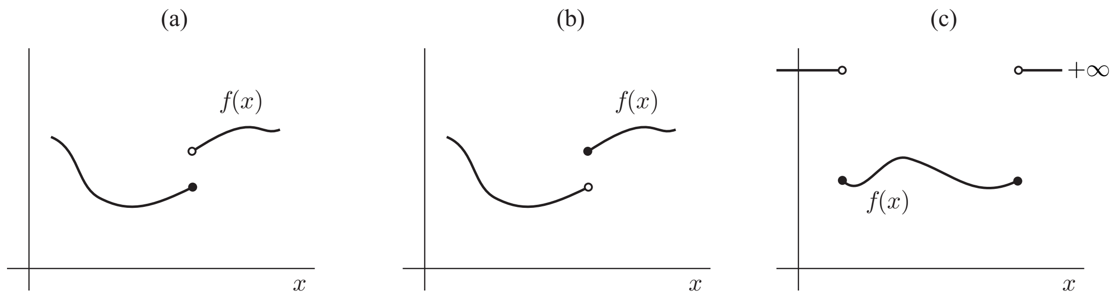

Figure 3: (a) Lower semi-continuous function. (A). (b) Upper semi-continuous function. (c) Lower semi-continuous, extended function.

- **Lower semi-continuity:** $f : X \to \overline{\mathbb{R}}$ is lower semi-continuous at $x_0 \in X$ if for each $k \in \mathbb{R}$, $k < f(x_0)$ there exists a neighborhood U of x_0 such that $f(U) > k$. (VT, §5.2)
    - Interpretation: function values near x_0 are either close to $f(x_0)$ or are greater than $f(x_0)$.
    - Graphical interpretation: if $f(x)$ is discontinuous at x_0, then $f(x_0)$ is on the lowest branch. (Fig. 3)
    - Equivalent definition: (VT, §5.7)
    $$\liminf_{x \to x_0} f(x) \geq f(x_0). \tag{2}$$
    - (Closed level sets) If f is lower semi-continuous, then $\{x \in X : f(x) \leq a\}$ is closed for all $a \in \mathbb{R}$. (Essential property for LDT.) (VT, §5.3)
    - If f is lower semi-continuous, then $\{x \in X : f(x) > a\}$ is open for all $a \in \mathbb{R}$. (VT, §5.3)
    - $f(x) = \sup_\lambda f_\lambda(x)$ is lower semi-continuous if the f_λ's are all lower semi-continuous. (VT, §5.4)
    - If f is lower semi-continuous on a compact space, then f assumes a minimum value (which may be $+\infty$). (Essential for LDT.) (VT, §5.4)
    - If f and g are lower semi-continuous, then so is λf, $\lambda > 0$, and $f + g$. (VT, §5.4)
    - A function is continuous if and only if it is both lower and upper semi-continuous.

- **Epigraph:** $\text{epi}(f) = \{(x, a) : f(x) \leq a, a \in \mathbb{R}\}$ (Fig. 4) (VT, §5.1)
    - $\text{epi}(f)$ is closed $\Leftrightarrow$ f is lower semi-continuous. (VT, §5.3)
    - From the greek "epi" meaning "upon" or "over".

- **Lower semi-continuous hull:** function $\overline{f}$ such that (Fig. 4) (VT, §5.5)
$$\text{epi}(\overline{f}) = \overline{\text{epi}(f)}. \tag{3}$$
    - $\overline{f}$ is the largest lower semi-continuous minorant of f, i.e., the largest lower semi-continuous function $g(x)$ such that $g(x) \leq f(x)$ for all $x \in \mathbb{R}^n$. (VT, §5.6)
    - If f is lower semi-continuous, then $f = \overline{f}$. (VT, §5.8)

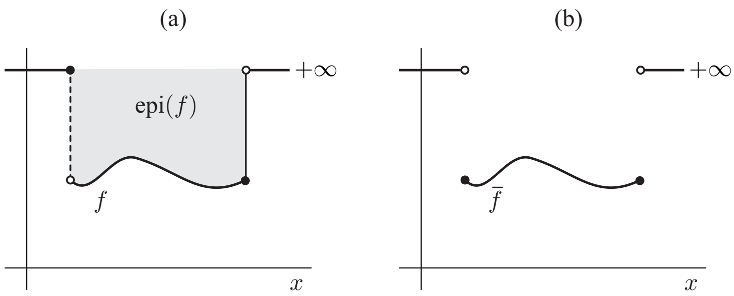

Figure 4: (a) epi(f). (b) Lower semi-continuous hull of f.

- **Subgradient:** $\alpha \in \mathbb{R}^n$ is said to be a subgradient of f at x_0 if (VT, §5.30)

$$f(x) \geq f(x_0) + \alpha \cdot (x - x_0) \tag{4}$$

for all $x \in \mathbb{R}^n$. (Fig. 5)

  - When the inequality is satisfied we also say that f has a supporting hyperplane at x_0 with gradient α.
  - A supporting hyperplane is said to be strictly supporting if the inequality is strict for all $x \neq x_0$.
  - If f is differentiable at $x_0 \in \text{dom}(f)$, then $\nabla f(x_0)$ is the unique subgradient of f at x_0.
  - In $\mathbb{R}$, we say that f has a supporting line with slope α.

- **Subdifferential:** Set of all subgradients of f at x_0: (VT, §5.30)

$$\partial f(x_0) = \{\alpha \in \mathbb{R}^n : f(x) \geq f(x_0) + \alpha \cdot (x - x_0), \forall x\}. \tag{5}$$

  - $\partial f(x_0)$ is a convex subset of $\mathbb{R}^n$.
  - $\partial f(x) = \{\nabla f(x)\}$ if f is differentiable at x.
  - If $f : \mathbb{R} \to \mathbb{R}$ is differentiable at x, then $\partial f(x) = \{f'(x)\}$.
  - $\text{dom}(\partial f) = \{x \in \mathbb{R}^n : \partial f(x) \neq \emptyset\}$.

- **Convex function:** f is convex if (VT, §5.9)

$$f(ax + (1-a)y) \leq af(x) + (1-a)f(y). \tag{6}$$

for all $x, y \in \mathbb{R}^n$ and $a \in [0, 1]$.

  - f is strictly convex if the inequality is strict for all $a \in (0, 1)$.
  - Proper convex function: $f \neq +\infty$. (VT, §5.11)
  - Improper convex function: $f(x) = -\infty$ for all $x \in \text{ri}(\text{dom}(f))$. If f is lower semi-continuous, then $\text{dom}(f)$ is closed, so that $f(x) = -\infty$ on $\text{dom}(f)$ in this case. (VT, §5.12)

- **Properties of convex functions:** Let f be a proper convex function. Then,

  - $\text{epi}(f)$ is convex. (VT, §5.10)

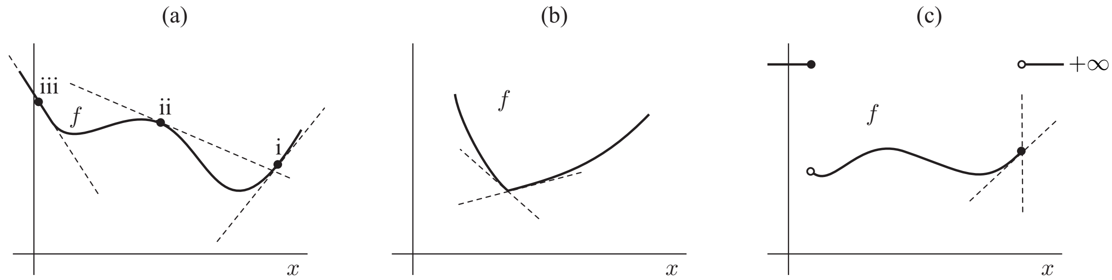

Figure 5: (a) (i) Point admitting a strict supporting line; (ii) point admitting no supporting line; (iii) non-strict supporting line. (b) $\partial f(x) = [f'_-, f'_+]$. (c) Supporting lines for boundary points: the left boundary point has no supporting lines, while the right boundary point has an infinite number of supporting lines with slope in $[f'_-, \infty)$.

- Convex level sets: f has convex level sets, i.e., $\{x : f(x) \leq a\}$ is a convex set for all $a \in \mathbb{R}$.
- $\text{dom}(f)$ is convex. (VT, §5.11)
- f has no isolated $(-\infty)$ singularities in its domain. (Fig. 6)
- $\text{ri}(\text{dom}(f)) \subseteq \text{dom}(\partial f) \subseteq \text{dom}(f)$. (R, §227)
  * This shows that $\partial f(x)$ is defined for all $x \in \text{dom} f$ except possibly at relative boundary points.
  * A proper convex function has supporting lines everywhere except possibly relative boundary points.
  * Example of convex function that is not subdifferentiable (in fact differentiable) everywhere: (R, §215)

$$f(x) = \begin{cases} -\sqrt{1 - |x|^2} & |x| \leq 1 \\ +\infty & \text{otherwise.} \end{cases} \tag{7}$$

  Then $\text{dom}\partial f = (-1, 1)$ but $\text{dom} f = [-1, 1]$.
- Continuity: f is continuous on $\text{int}(\text{dom}(f))$. (VT, §5.20)
- Relative continuity: The restriction of f to $\text{ri}(\text{dom}(f))$ is continuous. (VT, §5.23)
- Semi-continuity: f is lower semi-continuous at each point in $\text{ri}(\text{dom}(f))$.
- Subdifferential: f is everywhere subdifferentiable in its relative interior, i.e., $\partial f(x) \neq \emptyset$ for all $x \in \text{ri}(\text{dom}(f))$. (VT, §5.35)
- In $\mathbb{R}$, f has left- and right-derivatives everywhere in $\text{int}(\text{dom}(f))$.
- In $\mathbb{R}$, $\partial f(x) = [f'_+(x), f'_-(x)]$ for all $x \in \text{int}(\text{dom}(f))$.
- If $f : \mathbb{R} \to \mathbb{R}$ is convex, differentiable, then $f'(x)$ is monotonically increasing.
- $af(x) + b$, $a > 0$, is convex.
- Affinisation: $f(ax + b)$ is convex.
- Minimizers: f has no local minimum which is not a global minimum.
- Minimizers set: The set of minimizers of f is a convex set.

- **Other useful properties:**
  - Jensen's inequality: $f(E[X]) \leq E[f(X)]$, where $E[\cdot]$ denotes the expected value. (VT, §5.14)
  - Hessian: If f is twice continuously differentiable, then f is convex if and only if its Hessian is semi-definite (non-negative determinant). (VT, §5.29)
  - If $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable and $f''(x) > 0$, then f is convex. The converse does not hold (counterexample: $f(x) = x^4$). (VT, §1.11)
  - Convex superposition: $g(x) = \sum_i f_i(x)$ is convex if the $f_i(x)$'s are convex. (VT, §5.14)
  - Convex maximization: $g(x) = \sup_\lambda f_\lambda(x)$ is convex if $f_\lambda(x)$ is convex for all λ. Equivalently, $g(x) = \sup_y f(x, y)$ is convex if $f(x, y)$ is convex in x for all y.
  - Convex minimization: $g(x) = \inf_y f(x, y)$ is convex if $f(x, y)$ is jointly convex, i.e., convex as a "surface".
  - Pointwise limit: $f(x) = \lim_n f_n(x)$ is convex if f_n is convex for all n.
- **Convex hull:** (VT, §5.16)
$$\mathrm{co}(f)(x) = \inf\{a : (x, a) \in \mathrm{co}(\mathrm{epi}(f))\}. \tag{8}$$
  - $\mathrm{co}(f)$ is the largest convex minorant of f.
  - $\overline{\mathrm{co}(f)}$ is the largest lower semi-continuous, convex minorant of f.

---

3. Duality

- **Conjugate** or **dual function:** (VT, §6.1)
$$f^*(k) = \sup_{x \in \mathbb{R}^n} \{k \cdot x - f(x)\}. \tag{9}$$
- **Bipolar** or **double dual:**
$$f^{**}(x) = \sup_{k \in \mathbb{R}^n} \{k \cdot x - f^*(k)\} = (f^*)^*(x). \tag{10}$$
- **Properties:**
  - If $f \leq g$, then $f^* \geq g^*$. (VT, §6.3)
  - $(+\infty)^* = -\infty$.
  - If there is a point where f has the value $-\infty$, then $f^* = +\infty$. In this case, $f^{**} = -\infty$, and so f^{**} may not necessarily be equal to f.
  - $f^{**} \leq f$.
  - $(\inf_\lambda f_\lambda)^* = \sup_\lambda f_\lambda^*$.
  - $(\sup_\lambda f_\lambda)^* \leq \inf_\lambda f_\lambda^*$.
  - $(\lambda f)^*(k) = \lambda f^*(k/\lambda)$, $\lambda > 0$.
  - $(f + \lambda)^* = f^* + \lambda$.
  - $[f(x - y)]^*(k) = f^*(k) + k \cdot y$.
  - $\inf f(x) = -f^*(0)$.

- f^* is convex, lower semi-continuous. (VT, §6.8)
- f^{**} is convex, lower semi-continuous. (VT, §6.11)
- $f^{***} = f^*$.
- Fenchel's inequality: $f(x) + f^*(k) \geq k \cdot x$. (VT, §6.9)

- **Closure of f:** $\mathrm{cl}(f) = \bar{f}$ if f has nowhere the value $-\infty$; otherwise $\mathrm{cl}(f) = -\infty$. (VT, §6.13)
  - f is said to be closed when $\mathrm{cl}(f) = f$.

- **Duality:** (Fig. 6) See also (HT) for figures. (R, §23, 25)
  - $k \in \partial f(x) \Leftrightarrow f^*(k) = k \cdot x - f(x)$. (VT, §6.10)
  - $k \in \partial f^{**}(x) \Leftrightarrow x \in \partial f^*(k)$.
  - $k \in \partial f(x) \Leftrightarrow f(x) = f^{**}(x)$ except possibly at relative boundary points. (See Rockafellar's example).
  - $\partial f(x) \neq \emptyset$ $f(x) = f^{**}(x)$ except possibly at relative boundary points. (See Rockafellar's example).
  - $f^{**} = \mathrm{cl}(\mathrm{co}(f))$ in general; $f^{**} = \overline{\mathrm{co}(f)}$ if f is nowhere equal to $-\infty$. (VT, §6.15)
  - $f^{**} = \overline{f}$ if f is proper convex. (VT, §6.16)
  - $f^{**} = f$ if f is convex, lower semi-continuous or else $f = \pm\infty$. (VT, §6.18)
  - $\mathrm{dom} f \subseteq \mathrm{dom} f^{**}$.
    - Examples: f is not lower semi-continuous or f has a middle $+\infty$ (non-convex) part, i.e., $\mathrm{dom} f$ is not convex.
    - Corollary: If $f(x) < \infty$, then $f^{**}(x) < \infty$.
  - The map $f \to f^*$ is bijective for convex, lower semi-continuous functions. (VT, §6.19)
  - $f > f^{**}$ if $f \neq f^{**}$.
  - If f is nonconcave or affine somewhere, then f^* is non-differentiable somewhere.
  - If f is non-differentiable somewhere, then f^* has an affine region.
  - The dual is the same as the Legendre transform for strictly convex, differentiable functions.

- **Concave points vs supporting lines:**
  - Convex hull points: $\Gamma(f) = \{x : f(x) = f^{**}(x)\}$.
  - Concave points: $\Gamma(f) \cap \mathrm{dom} f$.
    The intersection with $\mathrm{dom} f$ comes from not wanting $+\infty$ points as concave.
  - Supporting line points: $C(f) = \{x : \partial f(x) \neq \emptyset\} = \mathrm{dom}\partial f$.
  - $C(f) = \Gamma(f) \cap \mathrm{dom}\partial f^{**} = \Gamma(f) \cap \mathrm{dom}\partial f$.
  - $\Gamma(f) \cap \mathrm{ri}(\mathrm{dom} f) \subseteq C(f) \subseteq \Gamma(f) \cap \mathrm{dom} f$.
    - *Proof*: Take $\Gamma(f)\cap$ of Rockafellar's inclusion result.
    - This shows that concave points are supporting line points except possibly at relative boundary points.

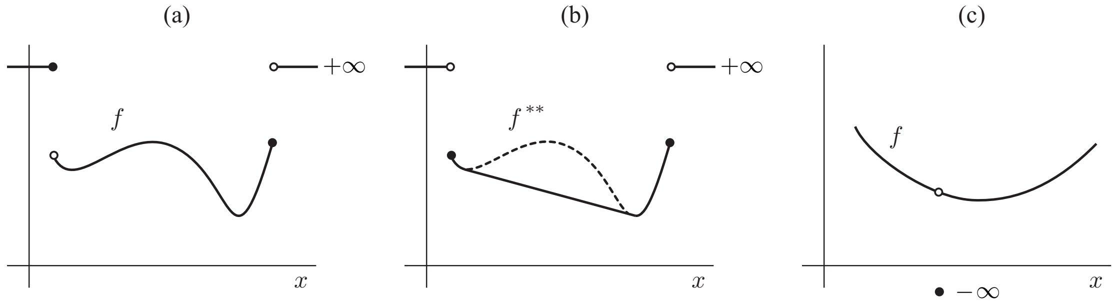

Figure 6: (a)-(b) f and its convex, lower semi-continuous hull. (c) f has the value $-\infty$ somewhere. Then $f^* = +\infty$, so that $f^{**} = -\infty$.

4. Optimization

- **Fenchel's duality Theorem:** Let f be a proper convex function and g be a proper concave function such that $\mathrm{ri}(\mathrm{dom}(f)) \cap \mathrm{ri}(\mathrm{dom}(g)) \neq \emptyset$. Then, (VT, §7.15)

$$\inf_{x \in \mathbb{R}^n} \{f(x) - g(x)\} = \max_{k \in \mathbb{R}^n} \{g^*(k) - f^*(k)\}.$$

  g^* is the dual defined for concave functions.

- **Constrained minimization:** Let C be a convex, non-empty subset of $\mathbb{R}^n$. Then, (VT, §7.16)

$$\inf_{x \in C} f(x) = \inf_{x \in \mathbb{R}^n} \{f(x) - g(x)\} = \max_{k \in \mathbb{R}^n} \{g^*(k) - f^*(k)\},$$

  where $g(x) = -\delta_C(x)$ (indicator function). Note that (VT, §5.15) (VT, §6.5)

$$\delta_C^*(k) = \sup_{x \in \mathbb{R}^n} \{k \cdot x - \delta_C(x)\} = \sup_{x \in C} k \cdot x.$$

References

[VT] J. van Tiel, *Convex Analysis: An Introductory Text*, John Wiley, New York, 1984.
Very good and concise introduction to the subject. The book starts with convex functions on $\mathbb{R}$ before it goes on to discuss convex functions on $\mathbb{R}^n$, which is very helpful for those who study convex analysis for the first time.

[R] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.
The definite reference on convex analysis. Not always easy to read, but a good source of information.

[B] D. P. Bertsakas, Lecture notes on convex analysis and optimization. Available on the MIT OpenCourse website.
Set of slides on convex optimization theory. The first few slides introduce (with no text) the basics of convex analysis. The book suggested for the course (written by Bertsekas) is another good reference.

[HT] H. Touchette, Legendre-Fenchel transforms in a nutshell. Unpublished report, 2005.
The basics of Legendre-Fenchel transforms (duals) for physicists with many figures.