6 Utility Maximization

As in Chapter 2 we fix a strictly concave, differentiable utility function $U : \ ]0, \infty[\ \to \mathbb{R}$, satisfying the Inada conditions (29). We have to impose the following additional regularity condition in order to obtain satisfactory duality results.

Definition 6.1. *([66], compare also [13]): The asymptotic elasticity of the utility function $U : \ ]0, \infty[\ \to \mathbb{R}$ is defined as*

$$AE(U) = \limsup_{x \to \infty} \frac{xU'(x)}{U(x)}. \tag{222}$$

We say that U has reasonable asymptotic elasticity *if $AE(U) < 1$.*

For example, for $U(x) = \frac{x^\gamma}{\gamma}$, where $\gamma \in \ ]-\infty, 1[\backslash\{0\}$, we have $AE(U) = \gamma < 1$. We note that, for an increasing concave function U we always have $AE(U) \leqslant 1$. A typical example of a function U for which $AE(U) = 1$ is $U(x) = \frac{x}{\log(x)}$, for x sufficiently large, as one verifies by calculating (222).

We again denote by V the conjugate function

$$V(y) = \sup\{U(x) - xy : x > 0\}, \qquad y > 0,$$

and refer to [66, Corollary 6.1] for equivalent reformulations of the asymptotic elasticity condition (222) in terms of V.

We adopt the setting of Chapter 4 where we considered a continuous price process $(S_t)_{0 \leqslant t \leqslant T}$ which satisfies condition (CPS^λ) of existence of λ-consistent price systems, for all $0 < \lambda < 1$. We again assume throughout this section that the underlying filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leqslant t \leqslant T}, \mathbb{P})$ is Brownian so that every $(\mathcal{F}_t)_{0 \leqslant t \leqslant T}$-martingale is continuous.

We have established in Theorem 4.22 that the sets $\mathfrak{C}(1)$ and $\mathfrak{D}(1)$ satisfy the requirements of Proposition 3.1 in [66].

We therefore are verbatim in the setting of [66], Th 3.1 and 3.2]. For the convenience of the reader we restate the aspects of these theorems which are relevant in the present setting.

Theorem 6.2. *Suppose that the continuous, strictly positive process $S = (S_t)_{0 \leqslant t \leqslant T}$ satisfies condition $(CPS^{\lambda'})$, for each $0 < \lambda 1$. Fix $\lambda > 0$, and define the primal and dual value function as*

$$u(x) = \sup_{\varphi_T^0 \in \mathfrak{C}(x)} \mathbb{E}[U(\varphi_T^0)], \tag{223}$$

$$v(y) = \inf_{Z_T^0 \in \mathfrak{D}(y)} \mathbb{E}[V(Z_T^0)]. \tag{224}$$

Suppose that the utility function U has reasonable asymptotic elasticity (222) *and that $u(x) < \infty$, for some $x > 0$.*

(*i*) *The functions $u(x)$ and $v(y)$ are finitely valued, for all $x > 0, y > 0$, and mutually conjugate*

$$v(y) = \sup_{x>0}[u(x) - xy], \qquad u(x) = \inf_{y>0}[v(y) + xy].$$

The functions u and v are continuously differentiable and strictly concave (resp. convex) and satisfy

$$u'(0) = -v'(0) = \infty, \qquad u'(\infty) = v'(\infty) = 0.$$

(*ii*) *The optimizers $\hat{\varphi}^0_T(x)$ in* (223) *(resp. $\hat{Z}^0_T(y)$ in* (224)*) exist, are unique and take their values a.s. in $]0, \infty[$. If $x > 0$ and $y > 0$ are related by $u'(x) = y$ (or, equivalently, $x = -v'(y)$), then $\hat{\varphi}^0_T(x)$ and $\hat{Z}^0_T(y)$ are related by*

$$\hat{Z}^0_T(y) = U'(\hat{\varphi}^0_T(x)), \qquad \hat{\varphi}^0_T(x) = -V'(\hat{Z}^0_T(y)).$$

(*iii*) *For $x > 0$ and $y > 0$ such that $u'(x) = y$ we have*

$$xy = \mathbb{E}[\hat{\varphi}^0_T(x)\hat{Z}^0_T(y)].$$

In fact, the process $(\hat{Z}^0_t\hat{\varphi}^0_t + \hat{Z}^1_t\hat{\varphi}^1_t)_{0\leqslant t\leqslant T} = (\hat{Z}^0_t(\hat{\varphi}^0_t + \hat{\varphi}^1_t\tilde{S}_t))_{0\leqslant t\leqslant T}$ is a uniformly integrable martingale.

■

In order to find a candidate for the shadow price process corresponding to $\hat{Z}_T(y)$ as in section 2 above we crucially need the following result which again is taken from previous work of D. Kramkov and the present author ([66], Prop. 3.2)

Proposition 6.3. *Under the assumptions of Theorem 4.22 the subset $\mathcal{D}(1)$ of $\mathfrak{D}(1)$ (see Definition 4.16 and 4.20) satisfies*

$$\sup_{Z^0_T\in\mathcal{D}(1)} \mathbb{E}[gZ^0_T] = \sup_{Z^0_T\in\mathfrak{D}(1)} \mathbb{E}[gZ^0_T],$$

for each $g \in \mathcal{C}$, and is closed under countable convex combinations. Hence it follows from [66], Prop. 3.2) that

$$v(y) = \inf_{Z^0_T\in\mathcal{D}(1)} \mathbb{E}[V(yZ^0_T)], \qquad y > 0.$$

We thus may find a sequence $(Z^n)_{n=1}^\infty = ((Z^{0,n}_t, Z^{1,n}_t)_{0\leqslant t\leqslant T})_{n=1}^\infty$ of λ-consistent price systems such that

$$v(y) = \lim_{n\to\infty} \mathbb{E}[V(yZ^{0,n}_T)]. \tag{225}$$

■

In order to formulate the main result of this section (Theorem 6.5 below) in proper generality we still need the following notion which was introduced by C. Bender [4] as a generalisation of the notion of a continuous martingale. We shall see in chapter 8 below that fractional Brownian motion also enjoys this property. On the other hand, fractional Brownian motion is far from being a martingale.

Definition 6.4. *Let $X = (X_t)_{t \geqslant 0}$ be a real-valued continuous stochastic process and σ a finite stopping time. Set $\sigma_{+1} = \inf\{t > \sigma \mid X_t - X_\sigma > 0\}$ and $\sigma_{-1} = \inf\{t > \sigma \mid X_t - X_\sigma < 0\}$. Then we say that X satisfies the condition (TWC) of "two way crossing", if $\sigma_{+1} = \sigma_{-1}$ P-a.s.*

We can now formulate one of the main results of this lecture notes.

Theorem 6.5. *As in Theorem 6.2 suppose that the continuous, strictly positive process S satisfies condition $(CPS^{\lambda'})$, for each $0 < \lambda' < 1$. Suppose in addition that S satisfies the two way crossing property (TWC).*

Fix the transaction costs $0 < \lambda < 1$, as well as the utility function U having reasonable asymptotic elasticity, and suppose that the value function $u(x)$ in (223) is finite, for some $x > 0$. Fix $x > 0$ and let $y = u'(x)$.

Then the optimizer $\hat{Z}^0_T(y) \in \mathfrak{D}(y)$ in (224) is the terminal value of a super-martingale deflator $(\hat{Z}^0_t(y), \hat{Z}^1_t(y))_{0 \leqslant t \leqslant T}$ which, in fact, also is a local martingale.

We shall split the message of Theorem 6.5 into the two subsequent results which serve to clarify the role of the assumption (TWC). Clearly the two subsequent results imply Theorem 6.5.

Theorem 6.6. *Let S be a continuous process satisfying the assumptions of Theorem 6.2.*

Suppose in addition that the liquidation value process associated to the optimiser $\hat{\varphi}$ of (223)

$$V_t^{liq} := \hat{\varphi}^0_t + (1-\lambda)(\hat{\varphi}^1_t)_+ S_t - (\hat{\varphi}^1_t)_- S_t, \qquad 0 \leqslant t \leqslant T, \tag{226}$$

is almost surely strictly positive.

Then the assertion of Theorem 6.5 holds true, i.e. the dual optimizer $\hat{Z}^0_T \in \mathfrak{D}(y)$ is the terminal value of a local martingale $(\hat{Z}^0_t(y), \hat{Z}^1_t(y))_{0 \leqslant t \leqslant T} \in \mathcal{Z}^e$.

Proposition 6.7. *Under the assumptions of Theorem 6.5, the liquidation value process V^{liq} in (226) is strictly positive.*

Proof of the Proposition 6.7:
To show that (226) remains almost surely positive, we argue by contradiction. Define

$$\sigma = \inf\{t \in [0,T] \mid V_t^{liq}(\widehat{\varphi}) = 0\}, \tag{227}$$

and suppose that $P[\sigma < \infty] > 0$.

First observe that $V_\sigma^{liq}(\widehat{\varphi}) = 0$ on $\{\sigma < \infty\}$. Indeed, applying the product rule to (226) and noting that $\widehat{\varphi}$ has finite variation we obtain

$$dV_t^{liq}(\widehat{\varphi}) = \big((\widehat{\varphi}_t^1)^+(1-\lambda) - (\widehat{\varphi}_t^1)^-\big)dS_t + \big(d\widehat{\varphi}_t^0 + (1-\lambda)S_t d(\widehat{\varphi}_t^1)^+ - S_t d(\widehat{\varphi}_t^1)^-\big). \tag{228}$$

The first term is the increment of a continuous process while the second term is, by the self-financing condition under transaction costs, the increment of a non-increasing right continuous process. Hence $V_\sigma^{liq}(\widehat{\varphi}) = 0$ on the set $\{\sigma < \infty\}$.

So suppose that $V_\sigma^{liq}(\widehat{\varphi}) = 0$ on the set $\{\sigma < \infty\}$ with $P[\sigma < \infty] > 0$. We may and do assume that S "moves immediately after σ", i.e. $\sigma = \inf\{t > \sigma \mid S_t \neq S_\sigma\}$.

We shall repeatedly use the fact established in Theorem 2.10 of [21] (compare also Remark 2.13 of [21]) that the process

$$\widehat{V} = \big(\widehat{\varphi}_t^0 \hat{Z}_t^0 + \widehat{\varphi}_t^1 \hat{Z}_t^1\big)_{0 \leqslant t \leqslant T}$$

is a uniformly integrable P-martingale satisfying $\widehat{V}_T > 0$ almost surely.

Firstly, this implies that $\widehat{\varphi}_\sigma^1 \neq 0$ a.s. on $\{\sigma < \infty\}$. Indeed, otherwise $V_\sigma^{liq}(\widehat{\varphi}) = \widehat{V}_\sigma = 0$ on $\{\sigma < \infty\}$. As $\widehat{V}$ is a uniformly integrable martingale with strictly positive terminal value $\widehat{V}_T > 0$ we arrive at the desired contradiction.

Hence we have $\widehat{\varphi}_\sigma^1 \neq 0$ and may suppose w.l.o.g. that $\widehat{\varphi}_\sigma^1 > 0$ on $\{\sigma < \infty\}$. Next, we show that we cannot have $\hat{S}_\sigma = (1-\lambda)S_\sigma$ with strictly positive probability on $\{\sigma < \infty\}$. Indeed, this again would imply that $V_\sigma^{liq}(\widehat{\varphi}) = \widehat{V}_\sigma = 0$ on this set which yields a contradiction as in the previous paragraph.

Hence we assume that $\hat{S}_\sigma > (1-\lambda)S_\sigma$ on $\{\sigma < \infty\}$. This implies that the utility-optimizing agent defined by $\widehat{\varphi}$ cannot sell stock at time σ as well as for some time after σ. Note, however, that the optimizing agent may very well buy stock. But we shall see that this is not to her advantage.

Define the stopping time ϱ_n as the first time after σ when one of the following events happens

(i) $\hat{S}_t - (1-\lambda)S_t < \frac{1}{2}\big(\hat{S}_\sigma - (1-\lambda)S_\sigma\big)$ or

(ii) $S_t < S_\sigma - \frac{1}{n}$.

By the hypothesis of (TWC) of "two way crossing", we conclude that, a.s. on $\{\sigma < \infty\}$, we have that ϱ_n decreases to σ and that we have $S_{\varrho_n} = S_\sigma - \frac{1}{n}$, for n large enough. Choose n large enough such that $S_{\varrho_n} = S_\sigma - \frac{1}{n}$ on a subset of $\{\sigma < \infty\}$ of positive measure. Then $V^{liq}_{\varrho_n}(\widehat{\varphi})$ is strictly negative on this set which will give the desired contradiction. Indeed, the assumption $\widehat{\varphi}^1_\sigma > 0$ implies that the agent suffers a strict loss from this position because $S_{\varrho_n} < S_\sigma$. The condition (i) makes sure that the agent cannot have sold stock between σ and ϱ_n. The agent may have bought additional stock during the interval $[\![\sigma, \varrho_n]\!]$. However, this cannot result in a positive effect either as the subsequent calculation reveals, which holds true on $\{S_{\varrho_n} = S_\sigma - \frac{1}{n}\}$.

$$V^{liq}_{\varrho_n}(\widehat{\varphi}) \leqslant V^{liq}_\sigma(\widehat{\varphi}) + \widehat{\varphi}^1_\sigma(1-\lambda)(S_{\varrho_n} - S_\sigma) - \int_\sigma^{\varrho_n} \big(S_u - (1-\lambda)S_{\varrho_n}\big) d\widehat{\varphi}^{1,\uparrow}_u < 0.$$

This contradiction finishes the proof of the Proposition 6.7.

Proof of Theorem 6.6: To alleviate notation we drop the variables x and y. We denote by $\hat{\varphi} = (\hat{\varphi}^0_t, \hat{\varphi}^1_t)_{0\leqslant t\leqslant T}$ the primal optimizer which exists by Theorem 6.2. To show that the dual optimizer $\hat{Z}$ is a local martingale is, of course, a *local* result. Hence we are free to use stopping arguments throughout the subsequent proof.

By Proposition 6.7 we have that

$$V^{liq}_t = \hat{\varphi}^0_t + (1-\lambda)(\hat{\varphi}^1_t)_+ S_t - (\hat{\varphi}^1_t)_- S_t$$

is strictly positive.

Let $(Z^n)_{n=1}^\infty = ((Z^{0,n}_t, Z^{1,n}_t)_{0\leqslant t\leqslant T})$ be a maximising sequence of λ-consistent price systems as in (225). By passing to a localising sequence of stopping times we may assume that all processes Z^n are true martingales and that S is bounded from above and bounded away from zero. By passing to convex combinations we may suppose that $(Z^n)_{n=1}^\infty$ Fatou-converges to a càdlàg super-martingale deflator $\hat{Z} = (\hat{Z}^0_t, \hat{Z}^1_t)_{0\leqslant t\leqslant T}$ as in the proof of Theorem 4.22. Its terminal value $\hat{Z}^0_T$ is the unique dual optimizer by Theorem 6.2 (ii).

We claim that $\hat{Z}$ is a local martingale. Indeed, suppose that this is not the case. Then we obtain for the limiting super-martingale $\hat{Z}$ the Doob-Meyer decomposition

$$d\hat{Z}^0_t = d\hat{M}^0_t - d\hat{A}^0_t, \tag{229}$$

$$d\hat{Z}^1_t = d\hat{M}^1_t - d\hat{A}^1_t, \tag{230}$$

where the processes $\hat{A}^0$ and $\hat{A}^1$ are non-decreasing. We have to show that they vanish. We start by showing that these two processes are aligned in the following sense.

Claim:

$$(1-\lambda)S_t d\hat{A}^0_t \leqslant d\hat{A}^1_t \leqslant S_t d\hat{A}^0_t. \tag{231}$$

In other words $\frac{d\hat{A}^1_t}{d\hat{A}^0_t}$ takes values in the bid ask spread $[(1-\lambda)S_t, S_t]$. The precise meaning of (231) is given in integral form by

$$\int_\sigma^\tau (1-\lambda)S_t d\hat{A}^0_t \leqslant \int_\sigma^\tau d\hat{A}^1_t \leqslant S_t d\hat{A}^0_t \tag{232}$$

for all stopping times $\sigma \leqslant \tau$ such that the above integrals converge. The proof of (232) is postponed to the subsequent lemma.

Now consider the process $\hat{V}_t = \hat{\varphi}^0_t \hat{Z}^0_t + \hat{\varphi}^1_t \hat{Z}^1_t$ which by Theorem 6.2 is a true martingale.

We then have, using Itô and the fact that $\hat{\varphi}$ is of finite variation,

$$d\hat{V}_t = \hat{\varphi}^0_t(d\hat{M}^0_t - d\hat{A}^0_t) + \hat{\varphi}^1_t(d\hat{M}^1_t - d\hat{A}^1_t) + \hat{Z}^0_t d\hat{\varphi}^0_t + \hat{Z}^1_t d\hat{\varphi}^1_t.$$

The sum of the last two terms yields a non-positive increment by the self-financing condition. The terms involving $d\hat{M}^0_t$ and $d\hat{M}^1_t$ are local martingale increments. We therefore may write symbolically

$$0 = \mathbb{E}[d\hat{V}_t] \leqslant -\hat{\varphi}^0_t d\hat{A}^0_t - \hat{\varphi}^1_t d\hat{A}^1_t. \tag{233}$$

The precise meaning of this differential inequality is that

$$0 = \mathbb{E}\left[\int_\sigma^\tau d\hat{V}_t\right] \leqslant -\mathbb{E}\left[\int_\sigma^\tau \left(\hat{\varphi}^0_t d\hat{A}^0_t + \hat{\varphi}^1_t d\hat{A}^1_t\right)\right],$$

for all $[0,T]$-valued stopping times $\sigma \leqslant \tau$ such that the above expectations are finite. To estimate the increment (233) we distinguish the cases $\hat{\varphi}^1_t \geqslant 0$ and $\hat{\varphi}^1_t \leqslant 0$. In the latter case $\hat{\varphi}^1_t \leqslant 0$, we deduce from (233) and (231)

$$\begin{aligned}\mathbb{E}[d\hat{V}_t] &\leqslant -\hat{\varphi}^0_t d\hat{A}^0_t - \hat{\varphi}^1_t d\hat{A}^1_t \\ &\leqslant -\hat{\varphi}^0_t d\hat{A}^0_t + (\hat{\varphi}^1_t)_- S_t d\hat{A}^0_t \\ &= -\left[\hat{\varphi}^0_t - (\hat{\varphi}^1_t)_- S_t\right] d\hat{A}^0_t.\end{aligned}$$

The term in the bracket is the liquidation value V_t^{liq} and therefore strictly positive. Hence $\mathbb{E}[d\hat{V}_t] = 0$ implies that $d\hat{A}^0_t = 0$. Finally, from (231) we also obtain that $d\hat{A}^1_t = 0$.

In the case $\hat{\varphi}^1_t \geqslant 0$ we estimate

$$\begin{aligned}\mathbb{E}[d\hat{V}_t] &\leqslant -\hat{\varphi}^0_t d\hat{A}^0_t - \hat{\varphi}^1_t d\hat{A}^1_t \\ &\leqslant -\hat{\varphi}^0_t d\hat{A}^0_t - (\hat{\varphi}^1_t)_+ (1-\lambda) S_t d\hat{A}^0_t \\ &\leqslant -\left[\hat{\varphi}^0_t - (\hat{\varphi}^1_t)_+ (1-\lambda) S_t\right] d\hat{A}^0_t\end{aligned}$$

so that, again, we may conclude that $d\hat{A}^0_t = d\hat{A}^1_t = 0$.

By (229) and (230) we have shown that $\hat{Z}^0$ and $\hat{Z}^1$ are equal to the local martingales $\hat{M}^0$ and $\hat{M}^1$. ■

Lemma 6.8. *In the setting of Theorem 6.6, let $\varepsilon > 0$ and $\sigma \leqslant \tau$ be "convenient" XXX stopping times such that $(1-\varepsilon) \leqslant \frac{S_\tau}{S_\sigma} \leqslant 1+\varepsilon$. Then*

$$\begin{aligned}(1-\varepsilon)(1-\lambda)S_\sigma \mathbb{E}\left[\hat{A}^0_\tau - \hat{A}^0_\sigma | \mathcal{F}_\sigma\right] &\leqslant \mathbb{E}\left[\hat{A}^1_\tau - \hat{A}^1_\sigma | \mathcal{F}_\sigma\right] \\ &\leqslant (1+\varepsilon) S_\sigma \mathbb{E}\left[\hat{A}^0_\tau - \hat{A}^0_\sigma | \mathcal{F}_\sigma\right].\end{aligned} \tag{234}$$

Proof. We may and do assume that the stopping times σ and τ are such that we have

$$\hat{Z}_\sigma = \mathbb{P} - \lim_{n \mapsto \infty} Z^n_\sigma \quad \text{and} \quad \hat{Z}_\tau = \mathbb{P} - \lim_{n \mapsto \infty} Z^n_\tau, \tag{235}$$

the limits taken in probability.

Indeed **??** holds true iff it holds true for $\sigma^h = (\sigma + h)(T-h)$ and $\tau^h = (\tau+h) \wedge (T-h)$ instead of σ and τ, for every $h > 0$. This easily follows from the right-continuity and the uniform integrability of the process $\hat{A}$.

For all but at most countably many $h > 0$ we must have the $\hat{Z}$ is a.s. continuous at time σ^h and τ^h. This implies that the Fatou-limit $\hat{Z}$ of $(Z^n)^\infty_{n=1}$ then satisfies

$$\hat{Z}_{\sigma^h} = \mathbb{P} - \lim_{n \mapsto \infty} Z^n_{\sigma^h} \quad \text{and} \quad \hat{Z}_{\tau^h} = \mathbb{P} - \lim Z^n_{\tau^h}, \tag{236}$$

as one easily verifies (compare [20] for much more on this topic).

Summing up, there is no loss of generality in assuming (235). This implies . . .

□

$$\lim_{n\to\infty} \left(Z^{0,n}_\tau - Z^{0,n}_\sigma\right) = \left(\hat{Z}^0_\tau - \hat{Z}^0_\sigma\right) = \left(\hat{M}^0_\tau - \hat{M}^0_\sigma\right) - \left(\hat{A}^0_\tau - \hat{A}^0_\sigma\right), \tag{237}$$

$$\lim_{n\to\infty} \left(Z^{1,n}_\tau - Z^{1,n}_\sigma\right) = \left(\hat{Z}^1_\tau - \hat{Z}^1_\sigma\right) = \left(\hat{M}^1_\tau - \hat{M}^1_\sigma\right) - \left(\hat{A}^1_\tau - \hat{A}^1_\sigma\right). \tag{238}$$

We then have that

$$\lim_{C\to\infty} \lim_{n\to\infty} \mathbb{E}\left[\left(Z^{0,n}_\tau - Z^{0,n}_\sigma\right) \mathbb{1}_{\{Z^{0,n}_\tau - Z^{0,n}_\sigma \geqslant C\}} | \mathcal{F}_\sigma\right] = \mathbb{E}\left[\hat{A}^0_\tau - \hat{A}^0_\sigma | \mathcal{F}_\sigma\right], \tag{239}$$

holds true a.s., and similarly

$$\lim_{C\to\infty} \lim_{n\to\infty} \mathbb{E}\left[\left(Z^{1,n}_\tau - Z^{1,n}_\sigma\right) \mathbb{1}_{\{Z^{1,n}_\tau - Z^{1,n}_\sigma \geqslant C\}} | \mathcal{F}_\sigma\right] = \mathbb{E}\left[\hat{A}^1_\tau - \hat{A}^1_\sigma | \mathcal{F}_\sigma\right]. \tag{240}$$

Indeed, we have

$$\begin{aligned}0 &= \mathbb{E}\left[Z_\tau^{0,n} - Z_\sigma^{0,n} | \mathcal{F}_\sigma\right] \\ &= \mathbb{E}\left[\left(Z_\tau^{0,n} - Z_\sigma^{0,n}\right) \mathbb{1}_{\{Z_\tau^{0,n} - Z_\sigma^{0,n} \geqslant C\}} | \mathcal{F}_\sigma\right] \\ &+ \mathbb{E}\left[\left(Z_\tau^{0,n} - Z_\sigma^{0,n}\right) \mathbb{1}_{\{Z_\tau^{0,n} - Z_\sigma^{0,n} < C\}} | \mathcal{F}_\sigma\right].\end{aligned}$$

Note that

$$\begin{aligned}&\lim_{C\to\infty}\lim_{n\to\infty} \mathbb{E}\left[\left(Z_\tau^{0,n} - Z_\sigma^{0,n}\right) \mathbb{1}_{\{Z_\tau^{0,n} - Z_\sigma^{0,n} < C\}} | \mathcal{F}_\sigma\right] \\ &= \mathbb{E}\left[\hat{Z}_\tau^0 - \hat{Z}_\sigma^0 | \mathcal{F}_\sigma\right] = -\mathbb{E}\left[\hat{A}_\tau^0 - \hat{A}_\sigma^0 | \mathcal{F}_\sigma\right],\end{aligned}$$

where the last equality follows from (237). We thus have shown (239) and (240) follows analogously.

We even obtain from (239) and (240) that

$$\lim_{C\to\infty}\lim_{n\to\infty} \mathbb{E}\left[Z_\tau^{0,n} \mathbb{1}_{\{Z_\tau^{0,n} \geqslant C\}} | \mathcal{F}_\sigma\right] = \mathbb{E}\left[\hat{A}_\tau^0 - \hat{A}_\sigma^0 | \mathcal{F}_\sigma\right] \tag{241}$$

and

$$\lim_{C\to\infty}\lim_{n\to\infty} \mathbb{E}\left[Z_\tau^{1,n} \mathbb{1}_{\{Z_\tau^{1,n} \geqslant C\}} | \mathcal{F}_\sigma\right] = \mathbb{E}\left[\hat{A}_\tau^1 - \hat{A}_\sigma^1 | \mathcal{F}_\sigma\right] \tag{242}$$

Indeed, the sequence $(Z_\sigma^{0,n})_{n=1}^\infty$ converges a.s. to $\hat{Z}_\sigma^0$ so that by Egoroff's theorem it converges uniformly on sets of measure bigger than $1-\varepsilon$. Therefore the terms involving $Z_\sigma^{0,n}$ in (239) disappear in the limit $C \to \infty$.

Finally, observe that

$$\frac{Z_\tau^{1,n}}{Z_\tau^{0,n}} \in [(1-\lambda)S_\tau, S_\tau] \subseteq [(1-\varepsilon)(1-\lambda)S_\sigma, (1+\varepsilon)S_\sigma].$$

Conditioning again on $\mathcal{F}_\sigma$ this implies on the one hand

$$\lim_{C\to\infty}\lim_{n\to\infty} \mathbb{E}\left[Z_\tau^{1,n} \mathbb{1}_{\{Z_\tau^{0,n} \geqslant C\}} | \mathcal{F}_\sigma\right] = \mathbb{E}\left[\hat{A}_\tau^1 - \hat{A}_\sigma^1 | \mathcal{F}_\sigma\right]$$

and on the other hand

$$\begin{aligned}\frac{\mathbb{E}[\hat{A}_\tau^1 - \hat{A}_\sigma^1 | \mathcal{F}_\sigma]}{\mathbb{E}[\hat{A}_\tau^0 - \hat{A}_\sigma^0 | \mathcal{F}_\sigma]} &= \lim_{C\to\infty}\lim_{n\to\infty} \frac{\mathbb{E}[Z_\tau^{1,n} \mathbb{1}_{\{Z_\tau^{0,n} \geqslant C\}} | \mathcal{F}_\sigma]}{\mathbb{E}[Z_\tau^{0,n} \mathbb{1}_{\{Z_\tau^{0,n} \geqslant C\}} | \mathcal{F}_\sigma]} \\ &\in [(1-\varepsilon)(1-\lambda)S_\sigma, (1+\varepsilon)S_\sigma],\end{aligned}$$

which is assertion (234). ∎

Then

A Appendix

In section 1 we have used a number of elementary results from linear algebra. In particular, this includes the following facts:

- The bipolar set of a closed, convex set in $\mathbb{R}^d$ containing the origin is the set itself.
- A set containing the origin is polyhedral iff its polar is polyhedral.
- The projection of a polyhedral cone is again a polyhedral cone.

For the convenience of the reader we provide proofs and present the underlying theory in a rather self-contained way in this appendix.

Let E be a vector space over the real numbers with finite dimension d and E' its dual. The space E then is isomorphic to $\mathbb{R}^d$ and we will use this fact in some of the discussion below, in which case we will denote the origin by $0 \in \mathbb{R}^d$ and the canonical basis by $\{e_1, \dots, e_d\}$.

A.1 Polar sets

We start with some basic definitions following [87] and [33]; shorter introductions to the geometry of convex sets can be found in [32] and [37]. For any set $A \subseteq E$, the smallest closed convex set containing A is called the *closed convex hull* of A, i.e. $\overline{\text{conv}}(A)$ is the intersection of all closed convex sets containing A. A closed convex set $C \subseteq E$ is called a *closed convex cone* if $\lambda a \in C$ for every a in C and $\lambda \geqslant 0$. The *closed convex cone generated by a set* $W \subseteq E$ is the closure of the convex cone

$$\text{cone}(W) := \left\{ \sum_{i \in I} \mu_i w_i : \; w_i \in W, \; \mu_i \geqslant 0 \right\},$$

where I is finite. It is the smallest closed convex cone containing W. We define $\text{cone}(\varnothing) := \{0\}$. The following properties of cones can be checked easily:

- Every closed convex cone contains the origin.
- The intersection of two closed convex cones is again a closed convex cone.

For a set $A \subseteq E$ we define the *polar* A° of E as

$$A^\circ := \{ y \in E' : \langle x, y \rangle \leqslant 1, \text{ for all } x \in A \}.$$

If A is a cone, we may equivalently define A° as

$$A^\circ = \{y \in E' : \langle x, y\rangle \leqslant 0, \text{ for all } x \in A\}.$$

If A is a linear space, we even may equivalently define A° as the annihilator

$$A^\circ = \{y \in E' : \langle x, y\rangle = 0, \text{ for all } x \in A\}.$$

The *Minkowski sum* of two sets $A, B \subseteq E$ is defined as the set

$$A + B := \{a + b,\ a \in A,\ b \in B\}.$$

It is easy to verify that, for any two sets $A \subseteq B \subseteq E$, we have $A^\circ \supseteq B^\circ$. If $C_1, C_2 \subseteq E$ are cones, then $(C_1 + C_2)^\circ = C_1^\circ \cap C_2^\circ$. Note that the polar of a cone is a closed convex cone.

The following theorem is a version of the celebrated Hahn-Banach theorem. The proof presented here can be found in [82]; for a more general discussion see for example [76].

Proposition A.1 (Bipolar Theorem)**.** *For a set $A \subseteq E$ the* bipolar $A^{\circ\circ} = (A^\circ)^\circ$ *equals the closed convex hull of $A \cup \{0\}$.*

Proof: Let $B = \mathrm{conv}(A \cup \{0\})$. Since $B \supseteq A$ we have $B^\circ \subseteq A^\circ$.

On the other hand, let $y \in A^\circ$ and $M \in \mathfrak{N}$ and pick $\lambda_i \in [0,1]$, for $1 \leqslant i \leqslant M$, such that $\sum_{i=1}^M \lambda_i = 1$. Then we have, for any $a_i \in A \cup \{0\}$:

$$1 \geqslant \sum_{i=1}^M \lambda_i \langle y, a_i\rangle = \sum_{i=1}^M \langle y, \lambda_i a_i\rangle = \langle \sum_{i=1}^M \lambda_i a_i, y\rangle.$$

Every $x \in B$ can be written as $x = \sum_{i=1}^M \lambda_i a_i$. It follows that $B^\circ \supseteq A^\circ$ and hence $A^\circ = B^\circ$.

We will now prove that $B^{\circ\circ} = \overline{B}$ which finishes the proof. Let $x \in \overline{B}$. For any $y \in B^\circ$ we have $\langle x, y\rangle \leqslant 1$ by definition and continuity, from which it follows that $x \in B^{\circ\circ}$ and therefore $\overline{B} \subseteq B^{\circ\circ}$. Conversely, assume $x_1 \notin \overline{B}$. Then there exists an $y \in E'$ and a constant c such that $\langle x, y\rangle \leqslant c$, for $x \in B$, and $\langle x_1, y\rangle > c$ (this follows from the Hahn-Banach theorem in its version as separating hyperplanes theorem, see for example [82]).

Because $0 \in B$ we have $c \geqslant 0$. We can even assume $c > 0$. It follows that $\langle x, y/c\rangle \leqslant 1$, for $x \in B$, and thus $y/c \in B^\circ$. But from $\langle x_1, y/c\rangle > 1$ we see that $x_1 \notin B^{\circ\circ}$. ∎

Corollary A.2. *If $C \subseteq E$ is a closed convex cone then $C^{\circ\circ} = C$.*

A.2 Polyhedral sets

We will now introduce the concept of *polyhedral sets*, which can be defined in two distinct ways. The first definition builds a polyhedron "from inside": Let V and W be two finite sets in E. The Minkowski sum of $\mathrm{conv}(V)$ and the cone generated by W

$$P = \mathrm{conv}(V) + \mathrm{cone}(W)$$

is called a *V-polyhedron*, where the name comes from the fact that such a polyhedron is defined using its vertices. Note that P is *closed.*

Polyhedral sets can also be built "from outside". A set $P \subseteq E$ is called an *H-polyhedron*, if it can be expressed as the finite intersection of closed halfspaces, that is

$$P = \bigcap_{i=1}^{N} \{x \in E : \langle x, y_i \rangle \leqslant c_i\},$$

for some elements $y_i \in E'$, and some constants c_i, $i \in \{1, \dots N\}$. As a subset of $\mathbb{R}^d$ such a polyhedron can be written as

$$P = P(A, z) := \left\{x \in \mathbb{R}^d : Ax \leqslant z\right\} \qquad \text{for some } A \in \mathbb{R}^{N \times d}, z \in \mathbb{R}^N.$$

Note that an H-polyhedron with all $c_i = 0$, i.e. of the form $P(A, 0)$, is in fact a closed convex cone: we shall encounter such *polyhedral cones* quite often.

These two distinct characterizations for polyhedral sets are useful for calculations and will play an important part in the following discussion. As we will verify below, the notions of V- and H-polyhedral sets are equivalent.

Our first Lemma deals with the projection of H-cones. The proof and a more thorough discussion can be found in [87].

Proposition A.3. *A projection of an H-cone along any coordinate directions e_k, $1 \leqslant k \leqslant d$, is again an H-cone. More specifically, if C is an H-cone in $\mathbb{R}^d$, then so is its elimination cone $elim_k(C) := \{x + \mu e_k : x \in C, \mu \in \mathbb{R}\}$ and its projection cone $proj_k(C) := elim_k(C) \cap \{x \in \mathbb{R}^d : \langle x, e_k \rangle = 0\}$.*

Proof: Note that it suffices to show that the set $\mathrm{elim}_k(C)$ is an H-cone, for any k, because the projection cone is the intersection of the elimination cone with the two halfspaces $\{x \in \mathbb{R}^d : \langle x, e_k \rangle \leqslant 0\}$ and $\{x \in \mathbb{R}^d : \langle x, -e_k \rangle \leqslant 0\}$.

Suppose that $C = P(A, 0)$ and $a_1, a_2, \dots, a_N$ are the row vectors of A. We will construct a new matrix A^k such that $\mathrm{elim}_k(C) = P(A^k, 0)$.
Claim: $A^k = \{a_i : a_{ik} = 0\} \cup \{a_{ik}a_j - a_{jk}a_i : a_{ik} > 0, a_{jk} < 0\}$
If $x \in C$ then $Ax \leqslant 0$. But then we also have $A^k x \leqslant 0$, because A^k consists of nonnegative linear combinations of rows of A. Therefore $C \subseteq P(A^k, 0)$. As

the k^{th} component of A^k is zero by construction, we even have $\mathrm{elim}_k(C) \subseteq P(A^k, 0)$.

On the other hand, let $x \in P(A^k, 0)$. We want to show that there is a $\mu \in \mathbb{R}$ such that $x - \mu e_k \in C$, i.e. $A(x - \mu e_k) \leqslant 0$. Writing these equations out, we obtain the inequalities $a_j x - a_{jk}\mu \leqslant 0$, or

$$\begin{aligned} \mu &\geqslant \tfrac{a_i x}{a_{ik}}, \quad \text{if } a_{ik} > 0, \\ \mu &\leqslant \tfrac{a_j x}{a_{jk}}, \quad \text{if } a_{jk} < 0. \end{aligned}$$

Such a μ exists, because if $a_{ik} > 0$ and $a_{jk} < 0$, then $(a_{ik}a_j - a_{jk}a_i)x \leqslant 0$, since $x \in P(A^k, 0)$, which can be written as

$$\frac{a_i x}{a_{ik}} \leqslant \frac{a_j x}{a_{jk}}.$$

It follows that $P(A^k, 0) \subseteq \mathrm{elim}_k(C)$, finishing the proof. ∎

Proposition A.4. *Every V-polyhedron is an H-polyhedron and vice versa.*

We split the proposition into two claims for the two directions, which we prove independently.

Claim: Every V-polyhedron is an H-polyhedron.

Remark: Proving the claim directly turns out to be rather tedious, due to the difficulty of manipulating the necessary sets. There is, however, an elegant proof using *homogenization*: Every polyhedron in d-dimensional space can be regarded as a polyhedral cone in dimension $d + 1$. The equivalence between V-cones and H-cones is easier to show. The direct proof uses Fourier-Motzkin elimination to calculate the sets explicitly. It can be found, together with the indirect proof given here, in [87].

Proof: By mapping a point $x \in \mathbb{R}^d$ to $\left(\begin{smallmatrix} 1 \\ x \end{smallmatrix}\right) \in \mathbb{R}^{d+1}$ we associate with every polyhedral set P in $\mathbb{R}^d$ a cone in $\mathbb{R}^{d+1}$ in the following way: If $P = P(A, z)$ is a H-polyhedral set, define

$$C(P) := P\left(\left(\begin{smallmatrix} -1 & 0 \\ -z & A \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)\right).$$

Conversely, if $C \in \mathbb{R}^{d+1}$ is an arbitrary H-cone, then $\{x \in \mathbb{R}^d : \left(\begin{smallmatrix} 1 \\ x \end{smallmatrix}\right) \in C\}$ is a (possibly empty) H-polyhedral set.

On the other hand, if $P = \mathrm{conv}(V) + \mathrm{cone}(W)$ is a V-polyhedral set for some finite sets V and W, we define

$$C(P) := \mathrm{cone}\left(\begin{smallmatrix} 1 & 0 \\ V & W \end{smallmatrix}\right),$$

that is, we add a zeroth coordinate to the vectors in V and W before generating the cone, namely 1 and 0, respectively. As before, a straightforward calculation shows that if C is a V-cone in $\mathbb{R}^{d+1}$, then $\{x \in \mathbb{R}^d : \left(\begin{smallmatrix}1\\x\end{smallmatrix}\right) \in C\}$ is a V-polyhedral set in $\mathbb{R}^d$.

If we can now show that every V-cone is an H-cone we are done, since every V-polyhedral set in $\mathbb{R}^d$ can be identified with a V-cone in $\mathbb{R}^{d+1}$ and the H-cone then translated back to the H-polyhedral set. Consider thus a V-cone, which can be written as

$$C = \left\{ x \in \mathbb{R}^d : \exists \lambda_i \geqslant 0 : x = \sum_i \lambda_i w_i,\ w_i \in W \right\},$$

or equivalently as

$$C = \left\{ (x, \lambda) \in \mathbb{R}^{d+n} : \lambda_i \geqslant 0, x = \sum_i \lambda_i w_i,\ w_i \in W \right\},$$

the latter set being an H-cone in $\mathbb{R}^{d+n}$. By successsively projecting the cone onto the hyperplanes for which the k^{th} coordinate equals zero, for $d < k \leqslant d+n$, we obtain a cone in $\mathbb{R}^d$ since we already showed that such a projection of an H-cone is again an H-cone. This finishes the proof of the claim.

The second part of the equivalence can also be shown directly or via homogenization, but we will give a third proof, which makes use of an elegant induction argument. For a thorough discussion of these concepts (and the proof of the following claim) see also [33].

Claim: Every H-polyhedron is a V-polyhedron.

Proof: Let P be an intersection of finitely many closed halfspaces in $\mathbb{R}^d$. We may assume w.l.g. that the dimension of P is d and will prove the claim by induction on d. If $d = 1$, then P is a halfline or a closed interval and the claim is clear. For $d \geqslant 2$ we will show that every point in P can be represented as the convex combination $a = (1 - \lambda)b + \lambda c,\ \ 0 \leqslant \lambda \leqslant 1$, where b and c belong to two distinct facets F and G of P respectively, i.e.

$$F = \operatorname{conv}(V_F) + \operatorname{cone}(W_F) \quad \text{and} \quad G = \operatorname{conv}(V_G) + \operatorname{cone}(W_G),$$

for some finite sets V_F, W_F, V_G, W_G. This suffices to prove the claim since the Minkowski sum of two V-polyhedral sets is again a V-polyhedral set.

Since every facet has dimension $d - 1$, we know that the boundaries of P are polyhedral sets. Let a be any point in the interior of P. Then there is some line l through a that intersects two facets of P, which is not parallel to any of the generating hyperplanes and intersects them in distinct points. Since a must lie between two such intersection points it is the linear combination of finitely many elements of V-polyhedral sets and because a was an arbitrary point in the interior of P, it follows that P itself is V-polyhedral. ■

The next proof can also be found in [33], along with other constructive results regarding polyhedra.

Proposition A.5. *Let $A \subseteq E$ be a polyhedral set. Then its polar A° also is a polyhedral set.*

Proof: We show that the polar of a V-polyhedron is an H-polyhedron, which we calculate explicitly. Let therefore A be of the form

$$A = \mathrm{conv}(V) + \mathrm{cone}(W) = \mathrm{conv}(\{v_1, \dots, v_N\}) + \mathrm{cone}(\{w_1, \dots, w_K\}),$$

for some finite sets V and W. By definition, we have

$$\begin{aligned} A^\circ &= \left\{ y \in E' : \left\langle \sum_{i=1}^N \lambda_i v_i + \sum_{j=1}^K \mu_j w_j, y \right\rangle \leqslant 1, \lambda_i \geqslant 0, \mu_j \geqslant 0, \sum \lambda_i = 1 \right\} \\ &= \left\{ y \in E' : \sum_i \lambda_i \langle v_i, y \rangle + \sum_j \mu_j \langle w_j, y \rangle \leqslant 1, \lambda_i \geqslant 0, \mu_j \geqslant 0, \sum \lambda_i = 1 \right\}. \end{aligned}$$

We therefore find that

$$A^\circ = \bigcap_{i=1}^N \{y \in E' : \langle v_i, y \rangle \leqslant 1\} \cap \bigcap_{j=1}^K \{y \in E' : \langle w_j, y \rangle \leqslant 0\},$$

which is an H-polyhedron. ■

Corollary A.6. *A convex, closed set containing the origin is polyhedral iff its polar is so too.*

Proof: This follows immediately from the previous proposition and the bipolar theorem, since then A° is polyhedral and $A = A^{\circ\circ}$. ■

B The Legendre Transformation

Definition B.1. *Let $u : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ be a concave upper semi-continuous function and $D = \mathrm{int}\{u > -\infty\} \neq \varnothing$ its domain, which we assume to be non-empty. The* conjugate *v of u is the function*

$$v(y) := \sup\{u(x) - xy,\ x \in \mathbb{R}\}.$$

The function v is the Legendre transform of $-u(-x)$ and is therefore convex rather than concave.[4] Given the conjugate function v, the original function u can be recovered via the transformation

$$u(x) := \inf\{v(y) + xy,\ y \in \mathbb{R}\}.$$

From these definitions it is immediately clear that for every $(x, y) \in \mathbb{R}^2$ we have *Fenchel's inequality*:

$$u(x) - v(y) \leqslant xy. \tag{243}$$

Note that equality holds when the supremum (respectively the infimum) in the above definitions is attained for the corresponding values of x and y.

Definition B.2. *The* subdifferential $\partial v(y_0)$ *of a convex function* v *at* y_0 *is the set of* $x \in \mathbb{R}$ *such that*

$$v(y) \geqslant v(y_0) + x \cdot (y - y_0), \quad \text{for all } y \in \mathbb{R}.$$

For a concave function u *we define the* superdifferential $\partial u(x_0)$ *of at* x_0 *equivalently as the set of* $y \in \mathbb{R}$ *satisfying*

$$u(x) \leqslant u(x_0) + y \cdot (x - x_0), \quad \text{for all } x \in \mathbb{R}.$$

If $\partial u(x_0)$ *consists of one single element* y*, then* u *is* differentiable at x_0 *and* $\nabla u(x_0) = y$*. Equivalently if* $\partial v(y_0)$ *consists of one single element* x*, then* v *is* differentiable at y_0 *and* $\nabla v(y_0) = x$*.*

Our first duality result links the super- and subdifferential of the conjugate functions u and v:

Proposition B.3. *The superdifferential* $\partial u(x_0)$ *contains* y_0 *iff* $-x_0 \in \partial v(y_0)$*.*

Proof: Let y_0 be in $\partial u(x_0)$. Then we have, for every x,

$$\begin{aligned} u(x) &\leqslant u(x_0) + y_0(x - x_0) \\ u(x) - y_0 x &\leqslant u(x_0) - y_0 x_0. \end{aligned}$$

Since this also holds for the supremum and using Fenchel's inequality on the right hand side, we obtain for every y in $\mathbb{R}$

$$\begin{aligned} v(y_0) &\leqslant u(x_0) - x_0 y_0 \leqslant v(y) + x_0 y - x_0 y_0 \\ v(y_0) &\leqslant v(y) + x_0(y - y_0), \end{aligned}$$

[4] In fact, the classical duality theory considers the (convex) function $-u(-x)$ to obtain a perfectly symmetric setting.

which is exactly the requirement for $-x_0$ to be in the subdifferential $\partial v(y_0)$. The other direction can be proved analogously. ■

There is an important duality regarding the smoothness and the strict concavity of the dual functions u and v. The following proof can be found in [45].

Proposition B.4. *The following are equivalent:*

1. $u : D \to \mathbb{R}$ *is strictly concave.*

2. v *is differentiable on the interior of its domain.*

Proof: $(i) \Rightarrow (ii)$. Suppose that (ii) fails, i.e. there is some y such that $\partial v(y)$ contains two distinct points, and call them $-x_1$ and $-x_2$. We may suppose that $x_1 < x_2$. This is equivalent to the requirement that $y \in \partial u(x_1) \cap \partial u(x_2)$. For $i = 1, 2$ we have

$$u(x_i) - v(y) = x_i y$$

and using Fenchel's inequality, we get (for $0 \leqslant \lambda \leqslant 1$):

$$\lambda u(x_1) + (1-\lambda)u(x_2) - v(y) = y \cdot (\lambda x_1 + (1-\lambda)x_2) \tag{244}$$

$$\geqslant u(\lambda x_1 + (1-\lambda)x_2) - v(y). \tag{245}$$

But this implies that u is affine on $[x_1, x_2]$, a contradiction since u is strictly concave. Therefore $\partial v(y)$ must be single-valued for all $y \in \text{int dom }(v)$, i.e. v is continuously differentiable.

$(ii) \Rightarrow (i)$. Suppose that there are two distinct points x_1 and x_2 such that u is affine on the line segment $[x_1, x_2]$. If we set $x := \frac{1}{2}(x_1 + x_2)$, there is an y such that $\nabla v(y) = x$, i.e. $y \in \partial u(x)$. We can write

$$0 = u(x) - v(y) - xy = \frac{1}{2}\sum_{i=1}^{2}\left[u(x_i) - v(y) - yx_i\right],$$

which implies (using Fenchel's inequality), that each of the terms in the bracket on the right hand side must vanish. We therefore have $y \in \partial u(x_1) \cap \partial u(x_2)$, i.e. $\partial v(y)$ contains more than one point, which contradicts the assumption that v is differentiable. ■

An economic visualization

As a side step – which may be safely skipped – let us try to give an economic "interpretation", or rather "visualization" of the Legendre transform V of the utility function U. Instead of interpreting U as a function which maps money to happiness, it seems more feasible to interpret U as a production function.

Keeping in mind that we only do a rather hypothetical mind experiment, suppose that you own a gold mine. You have the choice to invest x Euros into the (infrastructure of the) gold mine which will result in a production of $U(x)$ kilos of gold. You only can make this investment decision once and then the story is finished.

Now suppose that gold is traded at a price of y^{-1} Euros per kilo of gold or, equivalently, y as the price of one Euro in terms of kilos of gold. What is your optimal investment into the gold mine? Clearly you should invest the amount $\hat{x}$ of Euros for which the marginal production $U'(\hat{x})$ (of kilos of gold) per invested Euro equals the market price y, i.e. $\hat{x}$ is determined by $U'(\hat{x}) = y$. Hence the net value $V(y)$ of your gold mine in terms of kilos of gold equals $V(y) = \sup_x(U(x) - xy) = U(\hat{x}) - \hat{x}y$.

Indeed, it is optimal to borrow $\hat{x}$ Euros, invest them into the mine so that it produces $U(\hat{x})$ many kilos of gold. Subsequently you sell $\hat{x}y$ many of those kilos of gold to obtain $\hat{x}$ Euros which you use to pay back your loan. In this way you end up with a net result of $U(\hat{x}) - \hat{x}y$ kilos of gold.

Summing up $V(y)$ equals the net value of your gold mine in terms of kilos of gold, provided that the price of a kilo of gold equals y^{-1} and that you invest optimally.

Let us next try to interpret the inversion formula

$$U(x) = \inf_y (V(y) + xy).$$

Suppose that you have given the gold mine to a friend, whom we might call the "devil" and he promises to give you in exchange for the mine its net value in gold, i.e. $V(y)$ many kilos of gold if the market price of one kilo of gold turns out to be y^{-1}. Fix $y > 0$. If you own an initial capital of x Euros and want to transform all your wealth, i.e. the claims to the devil plus the x Euros, into gold, the total amount of kilos of gold then equals

$$V(y) + xy.$$

Fix your initial capital x Euros. If the devil is able to manipulate the market he might choose it in such a way that your resulting position in gold is minimized, i.e.

$$V(y) + xy \mapsto min!, \qquad y > 0.$$

Again, the optimal $\hat{y}$ is determined by $V'(\hat{y}) = -x$. The duality relation

$$V(x) = \inf_y V(y) + xy = V(haty + x\hat{y}$$

thus may interpreted that – if the devil really does a choice of y which is least favourable for you, – you will earn the same amount of gold as if you could have done by keeping the mine and investing your x Euros directly into the mine.

We close this “visualisation” of the duality relation between U and V by stressing once more that the fictitious posession of a gold mine has, of course, no practical economic relevance and was presented purely for didactic reasons.