2 Utility Maximization under Transaction Costs: the Case of Finite Ω

In this section we again adopt the simple setting of a finite filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbb{P})$. In addition to the ingredients of the previous section, i.e. the stock price process $S = (S_t)_{t=0}^T$ and the level of transaction costs $0 \leqslant \lambda < 1$, we also fix a utility function

$$U : D \to \mathbb{R}. \tag{28}$$

The domain D of U will be either $D =]0, \infty[$ or $D =]-\infty, \infty[$, and U is supposed to be a concave, $\mathbb{R}$-valued (hence continuous), increasing function on D. We also assume that U is strictly concave and differentiable on the interior of D. This assumption is not very essential but avoids to speak about subgradients instead of derivatives and allows for the uniqueness of solutions. More importantly, we assume that U satisfies the Inada conditions

$$\lim_{x \searrow x_0} U'(x) = \infty, \quad \lim_{x \nearrow \infty} U'(x) = 0, \tag{29}$$

where $x_0 \in \{-\infty, 0\}$ denotes the left boundary of D.

Remark 2.1. *Some widely studied examples for utility functions include:*

- $U(x) = \log(x)$,
- $U(x) = \frac{x^{1/2}}{1/2}$ *or, more generally,* $U(x) = \frac{x^\gamma}{\gamma}$, *for* $\gamma \in]0, 1[$,
- $U(x) = \frac{x^{-1}}{-1}$ *or, more generally,* $U(x) = \frac{x^\gamma}{\gamma}$, *for* $\gamma \in]-\infty, 0[$,
- $U(x) = -\exp(-x)$, *or, more generally,* $U(x) = -\exp(-\mu x)$, *for* $\mu > 0$.

The first three examples pertain to the domain $D =]0, \infty[$, while the second pertains to $D =]-\infty, \infty[$.

We also fix an initial endowment $x \in D$, denoted in units of bond. The aim is to find a trading strategy $(\varphi_t^0, \varphi_t^1)_{t=-1}^T$ maximizing expected utility of terminal wealth (measured in units of bond). More formally, we consider the optimization problem

$$(P_x) \qquad \begin{aligned} &\mathbb{E}[U(x + \varphi_T^0)] \to \max! \\ &\varphi_T^0 \in \mathcal{C}^\lambda \end{aligned} \tag{30}$$

In (P_x) the random variables φ_T^0 run through the elements of C^λ, i.e. such that there is a self-financing trading strategy $(\varphi_t^0, \varphi_t^1)_{t=-1}^T$, starting at $\varphi_{-1}^0, \varphi_{-1}^1 = (0, 0)$.

The interpretation is that an agent, whose preferences are modeled by the utility function U, starts with x units of bond (and no holdings in stock). She then trades at times $t = 0, \dots, T-1$, and at terminal date T she liquidates her position in stock so that $\varphi_T^1 = 0$ (this equality constraint clearly is equivalent to the inequality constraint $\varphi_T^1 \geqslant 0$ when solving the problem (P_x)). She then evaluates the performance of her trading strategy in terms of the expected utility of her final holdings φ_T^0 in bond.

Of course, we could formulate the utility maximization problem in greater generality. For example, we could consider initial endowments (x, y) in bonds as well as in stocks, instead of restricting to the case $y = 0$. We also could replace the requirement $\varphi_T^1 \geqslant 0$ by introducing a utility function $\mathcal{U}(x, y)$ defined on an appropriate domain $D \subseteq \mathbb{R}^2$ and consider

$$(P_{x,y}) \qquad \mathbb{E}[\mathcal{U}(\varphi_T^0, \varphi_T^1)] \to \max!$$

where $(\varphi_T^0, \varphi_T^1)$ runs through all terminal values of trading strategies $(\varphi_t^0, \varphi_t^1)_{t=-1}^T$ starting at $(\varphi_{-1}^0, \varphi_{-1}^1) = (x, y)$.

Note that (28) corresponds to the two-dimensional utility function

$$\mathcal{U}(x, y) = \begin{cases} U(x), & \text{if } y \geqslant 0, \\ -\infty, & \text{if } y < 0. \end{cases} \tag{31}$$

We refer to [19] and [47] for a thorough treatment of such a more general framework. For the present purposes we prefer, however, to remain in the realm of problem (30) as this allows for easier and crisper formulations of the results.

Using (28) and Corollary 1.12, we can reformulate (P_x) as a concave maximization problem under linear constraints:

$$(P_x) \qquad \mathbb{E}[U(x + \varphi_T^0)] \to \max! \qquad \varphi_T^0 \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}), \tag{32}$$

$$\mathbb{E}_Q[\varphi_T^0] \leqslant 0, \qquad Q \in \mathcal{M}^\lambda. \tag{33}$$

As $\mathcal{M}^\lambda$ is a compact polyhedron we can replace the infinitely many constraints (33) by finitely many: it is sufficient that (33) holds true for the extreme points $(Q^1, \dots, Q^M)$ of $\mathcal{M}^\lambda$.

We now are precisely in the well-known situation of utility optimization as in the frictionless case, which in the present setting reduces to a concave optimization problem on the finite-dimensional vector space $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ under linear constraints. Proceeding as in ([22, section 3.2]) we obtain the following basic duality result, where V denotes the conjugate function (the Legendre transform up to the choice of signs) of U

$$V(y) = \sup_{x \in D}\{U(x) - xy\}, \qquad y > 0. \tag{34}$$

Theorem 2.2. *(compare [22, Th. 3.2.1]): Fix $0 \leqslant \lambda < 1$ and suppose that in the above setting the (NA^λ) condition is satisfied for some fixed $0 \leqslant \lambda < 1$.*

Denote by u and v the value functions

$$\begin{aligned} u(x) &= \sup\left\{\mathbb{E}[U(x+\varphi_T^0)] : (\varphi_T^0, \varphi_T^1) \in \mathcal{A}^\lambda, \varphi_T^1 \geqslant 0\right\} \\ &= \sup\{\mathbb{E}[U(x+\varphi_T^0)] : \varphi_T^0 \in \mathcal{C}^\lambda\}, \qquad x \in D. \end{aligned} \tag{35}$$

$$\begin{aligned} v(y) &= \inf\{\mathbb{E}[V(y\tfrac{dQ}{d\mathbb{P}})] : Q \in \mathcal{M}^\lambda\} \\ &= \inf\{\mathbb{E}[V(Z_T^0)] : Z_T = (Z_T^0, Z_T^1) \in \mathcal{B}^\lambda, \mathbb{E}[Z_T^0] = y\}, \quad y > 0. \end{aligned} \tag{36}$$

Then the following statements hold true:

(i) The value functions $u(x)$ and $v(y)$ are mutually conjugate, and the indirect utility function $u : D \to \mathbb{R}$ is smooth, concave, increasing, and satisfies the Inada conditions (29).

(ii) For $x \in D$ and $y > 0$ such that $u'(x) = y$, the optimizers $\hat{\varphi}_T^0 = \hat{\varphi}_T^0(x) \in \mathcal{C}^\lambda$ and $\hat{Q} = \hat{Q}(y) \in \mathcal{M}^\lambda$ in (35) and (36) exist, are unique, and satisfy

$$x + \hat{\varphi}_T^0 = I\left(y\frac{d\hat{Q}}{d\mathbb{P}}\right), \quad y\frac{d\hat{Q}}{d\mathbb{P}} = U'\left(x+\hat{\varphi}_T^0\right), \tag{37}$$

where $I = (U')^{-1} = -V'$ denotes the "inverse" function. The measure $\hat{Q}$ is equivalent to $\mathbb{P}$, i.e. $\hat{Q}$ assigns a strictly positive mass to each $\omega \in \Omega$.

(iii) The following formulae for u' and v' hold true

$$u'(x) = \mathbb{E}_{\mathbb{P}}\left[U'\left(x+\hat{\varphi}_T^0(x)\right)\right], \qquad v'(y) = \mathbb{E}_{\hat{Q}(y)}\left[V'\left(y\frac{d\hat{Q}(y)}{d\mathbb{P}}\right)\right], \tag{38}$$

$$x\ u'(x) = \mathbb{E}_{\mathbb{P}}\left[\left(x+\hat{\varphi}_T^0(x)\right)U'\left(x+\hat{\varphi}_T^0(x)\right)\right], \quad y\ v'(y) = \mathbb{E}_{\mathbb{P}}\left[y\frac{d\hat{Q}(y)}{d\mathbb{P}}V'\left(y\frac{d\hat{Q}(y)}{d\mathbb{P}}\right)\right]. \tag{39}$$

Proof: We follow the reasoning of [22, section 3.2]. Denote by $\{\omega_1, \ldots, \omega_N\}$ the elements of Ω. We may identify a function $\varphi^0 \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ with the vector $(\xi_n)_{n=1}^N = (\varphi^0(\omega_n))_{n=1}^N \in \mathbb{R}^N$.

Denote by $Q^1, \ldots, Q^M$ the extremal points of the compact polyhedron $\mathcal{M}^\lambda$ and, for $1 \leqslant m \leqslant M$, by $(q_n^m)_{n=1}^N = (Q^m[\omega_n])_{n=1}^N$ the weights of Q^m. We may write the Lagrangian for the problem (32) as

$$\begin{aligned} L(\xi_1, \ldots, \xi_N, \eta_1, \ldots, \eta_M) &= \sum_{n=1}^N p_n U(\xi_n) - \sum_{m=1}^M \eta_m \left(\sum_{n=1}^N q_n^m \xi_n - x\right) = \\ &\sum_{n=1}^N p_n \left(U(\xi_n) - \sum_{m=1}^M \frac{\eta_m q_n^m}{p_n}\xi_n\right) + x\sum_{m=1}^M \eta_m. \end{aligned}$$

Here x is the initial endowment in bonds, which will be fixed in the sequel. The variables ξ_n vary in $\mathbb{R}$, the variables η_m in $\mathbb{R}_+$. Our aim is to find the (hopefully uniquely existing) saddle point $(\hat{\xi}_1, \dots, \hat{\xi}_N, \hat{\eta}_1, \dots, \hat{\eta}_M)$ of L which will give the primal optimizer via $x + \hat{\varphi}^0_T(\omega_n) := \hat{\xi}_n$, as well as the dual optimizer via $y\hat{Q} = \sum_{m=1}^M \hat{\eta}_m Q^m$, where $y = \sum_{m=1}^M \hat{\eta}_m$ so that $\hat{Q} \in \mathcal{M}^\lambda$.

In order to do so we shall consider $\max_\xi \min_\eta L(\xi, \eta)$ as well as $\min_\eta \max_\xi L(\xi, \eta)$. Define

$$\begin{aligned}\Phi(\xi_1, \dots, \xi_N) &= \inf_{\eta_1, \dots, \eta_M} L(\xi_1, \dots, \xi_N, \eta_1, \dots, \eta_M) \\ &= \inf_{y>0, Q \in \mathcal{M}^\lambda} \left\{ \sum_{n=1}^N p_n \left(U(\xi_n) - \frac{y q_n}{p_n} \xi_n \right) + yx \right\}\end{aligned}$$

Again the relation between $(\eta_1, \dots, \eta_M)$ and $y > 0$ and $Q \in \mathcal{M}^\lambda$ is given via $y = \sum_{m=1}^M \eta_m$ and $Q = \sum_{m=1}^M \frac{\eta_m}{y} Q^m$, where we denote by q_n the weights $q_n = Q[\omega_n]$.

Note that $\Phi(\xi_1, \dots, \xi_N)$ equals the target functional (30) if $(\xi_1, \dots, \xi_N)$ is admissible, i.e. satisfies (33), and $-\infty$ otherwise. Identifying the elements $\varphi^0 \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ with $(\xi_1, \dots, \xi_N) \in \mathbb{R}^N$, this may be written as

$$\Phi(\varphi^0) = \begin{cases} \mathbb{E}[U(\varphi^0)], \text{ if } \mathbb{E}_Q[\varphi^0] \leqslant x \text{ for all } Q \in \mathcal{M}^\lambda \\ -\infty, \text{ otherwise.} \end{cases} \tag{40}$$

Let us now pass from the max min to the min max: identifying $(\eta_1, \dots, \eta_M)$ with (y, Q) as above, define

$$\begin{aligned}\Psi(y, Q) &= \sup_{\xi_1, \dots, \xi_N} L(\xi_1, \dots, \xi_N, y, Q) \\ &= \sup_{\xi_1, \dots, \xi_N} \sum_{n=1}^N p_n \left(U(\xi_n) - y \tfrac{q_n}{p_n} \xi_n \right) + xy \\ &= \sum_{n=1}^N p_n \sup_{\xi_n} \left(U(\xi_n) - y \tfrac{q_n}{p_n} \xi_n \right) + xy \\ &= \sum_{n=1}^N p_n V\left(y \tfrac{q_n}{p_n} \right) + xy \\ &= \mathbb{E}_{\mathbb{P}} \left[V\left(y \tfrac{dQ}{d\mathbb{P}} \right) \right] + xy.\end{aligned}$$

We have used above the definition (30) of the conjugate function V of U.

Defining

$$\Psi(y) = \inf_{Q \in \mathcal{M}^\lambda} \Psi(y, Q) \tag{41}$$

we infer from the compactness of $\mathcal{M}^\lambda$ that, for $y > 0$, there is a minimizer $\hat{Q}(y)$ in (41). From the strict convexity of V (which corresponds to the differentiability of U as we recall in the appendix) we infer, as in [22], section 3.2, that $\hat{Q}(y)$ is unique and $\hat{Q}(y)[\omega] > 0$, for each $\omega \in \Omega$.

Finally, we minimize $y \mapsto \Psi(y)$ to obtain the optimizer $\hat{y} = \hat{y}(x)$ by solving

$$\Psi'(\hat{y}) = 0. \tag{42}$$

Denoting by $v(y)$ the dual value function which is obtained from $\Psi(y)$ by dropping the term xy, i.e.

$$v(y) = \inf_{Q \in \mathcal{M}^\lambda} \mathbb{E}\left[V(y\tfrac{dQ}{d\mathbb{P}})\right],$$

we obtain from (42) the relation

$$v'(\hat{y}(x)) = -x.$$

The uniqueness of $\hat{y}(x)$ follows from the strict convexity of v which, in turn, is a consequence of the strict convexity of V (see Proposition B.4 of the appendix).

Turning back to the Lagrangian $L(\xi_1, \dots, \xi_N, y, Q)$, the first order conditions

$$\tfrac{\partial}{\partial \xi_n} L(\xi_1, \dots, \xi_N, y, Q)|_{\hat{\xi}_1, \dots, \hat{\xi}_N, \hat{y}, \hat{Q}} = 0 \tag{43}$$

for a saddle point yield the following equations for the primal optimizers $\hat{\xi}_1, \dots, \hat{\xi}_N$

$$U'(\hat{\xi}_n) = \hat{y}\tfrac{\hat{q}_n}{p_n}, \tag{44}$$

where $\hat{y} = \hat{y}(x)$ and $\hat{Q} = \hat{Q}(\hat{y}(x))$. By the Inada conditions (29), as well as the smoothness and strict concavity of U, equation (44) admits unique solutions $(\hat{\xi}_1, \dots, \hat{\xi}_N) = (\hat{\xi}_1(x), \dots, \hat{\xi}_N(x))$.

Summing up, we have found a unique saddle point $(\hat{\xi}_1, \dots, \hat{\xi}_N, \hat{y}, \hat{Q})$ of the Lagrangian L. Denoting by $\hat{L} = \hat{L}(x)$ the value

$$\hat{L} = L(\hat{\xi}_1, \dots, \hat{\xi}_N, \hat{y}, \hat{Q})$$

we infer from the concavity of L in $\xi_1, \dots, \xi_N$ and convexity in y and Q that

$$\max_{\xi} \min_{y,Q} L = \min_{y,Q} \max_{\xi} L = \hat{L}. \tag{45}$$

It follows from (40) that $\hat{L}$ is the optimal value of the primal problem (P_x) in (30), i.e.

$$u(x) = \sum_{n=1}^N p_n U(\hat{\xi}_n) = \hat{L}. \tag{46}$$

The second equality in (45) yields

$$\hat{L} = \Psi(\hat{y}) = v(\hat{y}) + x\hat{y}. \tag{47}$$

Equations (46) and (47), together with the concavity (resp. convexity) of u (resp. (v)) and $v'(\hat{y}) = -x$ are tantamount to the fact that the functions u and v are conjugate.

We thus have shown (i) of Theorem 2.2. The listed qualitative properties of u are straightforward to verify (compare [22], section 3.2). Item (ii) now follows from the above obtained existence and uniqueness of the saddle point $(\hat{\xi}_1, \dots, \hat{\xi}_N, \hat{y}, \hat{Q})$ and (iii) again is straightforward to check as in [22]. ∎

We remark that in the above proof we did not apply an abstract minimax theorem guaranteeing the existence of a saddle point of the Lagrangian. Rather we directly found the saddle point by using the first order conditions, very much as we did in high school: differentiate and set the derivative to zero! The assumptions of the theorem are designed in such a way to make sure that this method yields a unique solution.

We now adapt the idea of *market completion* as developed in [53] to the present setting. Fix the initial endowment $x \in D$, and $y = u'(x)$. Define a frictionless financial market, denoted by AS, in the following way. For each fixed $\omega \in \Omega$, the *Arrow security* AS^ω, paying $AS^\omega_T = \mathbb{1}_\omega$ units of bond at time $t = T$, is traded (without transaction costs) at time $t = 0$ at price $AS^\omega_0 := \hat{Q}(y)[\omega]$. In other words, AS^ω pays one unit of bond at time T if ω turns out at time T to be the true state of the world, and zero otherwise. We define, for each $\omega \in \Omega$, the price *process* of AS^ω as the $\hat{Q}(y)$-martingale

$$AS^\omega_t = \mathbb{E}_{\hat{Q}(y)}[\mathbb{1}_\omega | \mathcal{F}_t], \qquad t = 0, \dots, T.$$

The set $\mathcal{C}^A$, where A stands for *Kenneth Arrow*, of claims attainable at price zero in this complete, frictionless market equals the half-space of $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$

$$\mathcal{C}^A = H_{\hat{Q}(y)} = \{\varphi^0_T \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) : \mathbb{E}_{\hat{Q}(y)}[\varphi^0_T] \leqslant 0\}. \tag{48}$$

Indeed, every $\varphi^0_T \in H_{\hat{Q}(y)}$ may trivially be written as a linear combination of Arrow securities

$$\begin{aligned}\varphi^0_T &= \sum_{\omega \in \Omega} \varphi^0_T(\omega) \mathbb{1}_{\{\omega\}} \\ &= \sum_{\omega \in \Omega} \varphi^0_T(\omega) AS^\omega_T(\omega)\end{aligned}$$

which may be purchased at time $t = 0$ at price

$$\sum_{\omega \in \Omega} \varphi_T^0(\omega) AS_0^\omega(\omega) = \mathbb{E}_{\hat{Q}(y)}[\varphi_T^0] \leqslant 0.$$

The Arrow securities AS^ω are quite different from the original process $S = (S_t)_{t=0}^T$ or, more precisely, the process of bid-ask intervals $([(1-\lambda)S_t, S_t])_{t=0}^T$. But we know from the fact that $\hat{Q}(y) \in \mathcal{M}^\lambda$ that

$$\mathcal{C}^\lambda \subseteq \mathcal{C}^A = H_{\hat{Q}(y)}. \tag{49}$$

In prose: the contingent claims φ_T^0 attainable at price 0 in the market S under transaction costs λ are a subset of the contingent claims φ_T^0 attainable at price zero in the frictionless Arrow market AS.

The message of the next theorem is the following: although the complete, frictionless market AS offers better terms of trade than S (under transaction costs λ), the economic agent modeled by (28) will choose as her terminal wealth the same optimizer $\hat{\varphi}_T^0 \in \mathcal{C}^\lambda$, although she can choose in the bigger set $\mathcal{C}^A$.

Theorem 2.3. *Fix $S = (S_t)_{t=0}^T$, transaction costs $0 \leqslant \lambda < 1$ such that (NA^λ) is satisfied, as well as $U : D \to \mathbb{R}$ verifying* (29) *and $x \in D$. Using the notation of Theorem 2.2, let $y = u'(x)$ and denote by $\hat{Q}(y) \in \mathcal{M}^\lambda$ the dual optimizer in* (36).

Define the optimization problem

$$(P_x^A) \qquad \begin{aligned} &\mathbb{E}[U(x + \varphi_T^0)] \to \max! \\ &\mathbb{E}_{\hat{Q}(y)}[\varphi_T^0] \leqslant 0, \end{aligned} \tag{50}$$

where φ_T^0 ranges through all D-valued, $\mathcal{F}_T$-measurable functions.

The optimizer $\hat{\varphi}_T^0(x)$ of the above problem exists, is unique, and coincides with the optimizer of problem (P_x) defined in (30).

Proof: As $\hat{Q}(y) \in \mathcal{M}^\lambda$ we have that $\hat{Q}(y)|_{\mathcal{C}^\lambda} \leqslant 0$ so that $\hat{Q}(y)|_{x+\mathcal{C}^\lambda} \leqslant x$. It follows from (50) that in (P_x^A) we optimize over a larger set than in (P_x).

Denote by $\hat{\varphi}_T^0 = \hat{\varphi}_T^0(x)$ the optimizer of (P_x) which uniquely exists by Theorem 2.2. Denote by $\hat{y} = \hat{y}(x)$ the corresponding Lagrange multiplier $\hat{y} = u'(x)$. We shall now show that $\hat{Q}(\hat{y})$ induces the *marginal utility pricing functional.*

Fix $1 \leqslant k \leqslant N$ and consider the variation functional corresponding to ω_k

$$\begin{aligned} v_k(h) &= \mathbb{E}\left[U(\hat{\varphi}_T^0 + h\mathbb{1}_{\omega_k})\right] \\ &= \sum_{\substack{n=1 \\ n \neq k}}^{N} p_n U(\hat{\xi}_n) + p_k U(\hat{\xi}_k + h), \qquad h \in \mathbb{R}. \end{aligned}$$

The function v_k is strictly concave and its derivative at $h = 0$ satisfies by (44)

$$v_k'(0) = p_k \hat{y} \tfrac{\hat{q}_k}{p_k} = \hat{y}\hat{q}_k.$$

Let $\zeta \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}), \zeta \neq 0$ be such that $\mathbb{E}_{\hat{Q}}(\zeta) = 0$. The variation functional v_ζ

$$v_\zeta(h) = \mathbb{E}\left[U(\hat{\varphi}_T^0 + h\zeta)\right] = \left(\sum_{k=1}^N p_k \zeta(\omega_k) v_k(h),\right) \qquad h \in \mathbb{R},$$

$$= \sum_{k=1}^N p_k U(\hat{\xi}_k + h\zeta_k)$$

has as derivative

$$v_\zeta'(h) = \sum_{k=1}^N p_k U'(\hat{\xi}_k + h\zeta_k)\zeta_k.$$

Hence

$$\begin{aligned} v_\zeta'(0) &= \sum_{k=1}^N p_k \underbrace{U'(\hat{\xi}_k)}_{=\hat{y}\frac{\hat{q}_k}{p_k}} \zeta_k = \sum_{k=1}^N \hat{y}\hat{q}_k\xi_k \\ &= \hat{y}\mathbb{E}_{\hat{Q}}[\xi] \\ &= 0 \end{aligned}$$

The function $h \mapsto v_\zeta(h)$ is strictly concave and therefore attains its unique maximum at $h = 0$.

Hence, for every $\varphi_T^0 \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}), \varphi_T^0 \neq \hat{\varphi}_T^0$ such that $\mathbb{E}_{\hat{Q}}[\varphi_T^0] = x$ we have

$$\mathbb{E}[U(\varphi_T^0)] < \mathbb{E}[U(\hat{\varphi}_T^0)].$$

Indeed, it suffices to apply the previous argument to $\zeta = \varphi_T^0 - \hat{\varphi}_T^0$. Finally, by the monotonicity of U, the same inequality holds true for all $\varphi_T^0 \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}_{\hat{Q}}[\varphi_T^0] < x$.

The proof of Theorem 2.3 now is complete. ■

In the above formulation of Theorems 2.2 and 2.3 we have obtained the unique primal optimizer $\hat{\varphi}_T^0$ only in terms of the final holdings in bonds; similarly the unique dual optimizer $\hat{Q}$ is given in terms of a probability measure which corresponds to a one dimensional density $Z^0 = \frac{d\hat{Q}}{d\mathbb{P}}$. What are the "full" versions of these optimizers in terms of $(\varphi_T^0, \varphi_T^1) \in \mathcal{A}^\lambda$, i.e., in terms of bond and stock, resp. $(Z^0, Z^1) \in \mathcal{B}^\lambda$ which is an $\mathbb{R}^2_+$-valued martingale? As regards the former, we mentioned already that it is economically obvious (and easily checked mathematically) that the unique optimizer $(\varphi^0_T, \varphi^1_T) \in \mathcal{A}^\lambda$ corresponding to $\hat{\varphi}^0_T \in \mathcal{C}^\lambda$ in (35) simply is $(\varphi^0_T, \varphi^1_T) = (\hat{\varphi}^0_T, 0)$, i.e. the optimal holding in stock at terminal date T is zero. As regards the optimizer $(Z^0, Z^1) \in \mathcal{B}^\lambda$ in (36) corresponding to the optimizer $\hat{Q} \in \mathcal{M}^\lambda$ the situation is slightly more tricky. By the definition (14) of $\mathcal{D}^\lambda$, for given $\hat{Z}^0 \in \mathcal{D}^\lambda$ there is $\hat{Z}^1 \in L^1_+(\Omega, \mathcal{F}, \mathbb{P})$ such that $(\hat{Z}^0, \hat{Z}^1) \in \mathcal{B}^\lambda$. But this $\hat{Z}^1$ need not be unique, even in very regular situations as shown by the subsequent easy example. Hence the "shadow price process" $(\tilde{S}_t)_{t=0}^T = \left(\frac{\hat{Z}^1_t}{\hat{Z}^0_t}\right)_{t=0}^T$ need not be unique. The terminology "shadow price" will be explained below, and will be formally defined in 2.7.

Example 2.4. In the above setting suppose that $(S_t)_{t=0}^T$ is a martingale under the measure $\mathbb{P}$. Then it is economically obvious (and easily checked) that it is optimal not to trade at all (even under transaction costs $\lambda = 0$). More formally, we obtain $u(x) = U(x), v(y) = V(y)$ and, for $x \in D$, the unique optimizers in Theorem 2.2 are given by $\hat{\varphi}^0_T \equiv 0$, and $\hat{Q} = \mathbb{P}$, as well as $\hat{y} = U'(x)$. For the optimal shadow price process $\tilde{S}$ we may take $\tilde{S} = S$. But this choice is not unique. In fact, we may take any $\mathbb{P}$-martingale $\tilde{S} = (\tilde{S}_t)_{t=0}^T$ taking values in the bid-ask spread $([(1-\lambda)S_t, S_t])_{t=0}^T$.

In the setting of Theorem 2.2 let $(\hat{Z}^0_T, \hat{Z}^1_T)$ be an optimizer of (36) and denote by $\hat{\tilde{S}}$ the process

$$\hat{\tilde{S}}_t = \frac{\mathbb{E}[\hat{Z}^1_T | \mathcal{F}_t]}{\mathbb{E}[\hat{Z}^0_T | \mathcal{F}_t]}, \qquad t = 0, \ldots, T,$$

which is a martingale under $\hat{Q}(y)$. We shall now justify why we have called this process a *shadow price process* for S under transaction costs λ.

Fix $x \in D$ and $y = u'(x)$. To alleviate notation we write $\tilde{S} = (\tilde{S}_t)_{t=0}^T$ for $\hat{\tilde{S}}(y)$ and Q for $\hat{Q}(y)$. Denote by $\mathcal{C}^{\tilde{S}}$ the cone of random variables φ^0_T dominated by a contingent claim of the form $(H \cdot \tilde{S})_T$, i.e.

$$\mathcal{C}^{\tilde{S}} = \{\varphi^0_T \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) : \varphi^0_T \leqslant (H \cdot \tilde{S})_T, \text{ for some } H \in \mathcal{P}\}.$$

Here we use standard notation from the frictionless theory. The letter $\mathcal{P}$ denotes the space of *predictable* $\mathbb{R}$-valued trading strategies $(H_t)_{t=1}^T$, i.e. H_t is $\mathcal{F}_{t-1}$-measurable, and $(H \cdot \tilde{S})_T$ denotes the stochastic integral

$$(H \cdot \tilde{S})_T = \sum_{t=1}^T H_t(\tilde{S}_t - \tilde{S}_{t-1}). \tag{51}$$

In prose: $\mathcal{C}^{\tilde{S}}$ denotes the cone of random variables φ_T^0 which can be super-replicated in the financial market $\tilde{S}$ without transaction costs and with zero initial endowment.

Lemma 2.5. *Using the above notation and assuming that S satisfies (NA^λ) we have*

$$\mathcal{C}^\lambda \subseteq \mathcal{C}^{\tilde{S}} \subseteq \mathcal{C}^A \tag{52}$$

Proof: The first inclusion was already shown in the proof of the Fundamental Theorem 1.8; it corresponds to the fact that trading without transaction costs on $\tilde{S}$ yields better terms of trade than trading on S under transaction costs λ.

As regards the second inclusion note that, for $(H \cdot \tilde{S})_T$ as in (51), we have

$$\mathbb{E}_Q[(H \cdot \tilde{S})_T] = 0,$$

whence $(H \cdot \tilde{S})_T$ belongs to $\mathcal{C}^A$ by (48). As $\mathcal{C}^A$ also contains the negative orthant $L^\infty(\Omega, \mathcal{F}, \mathbb{P}; -\mathbb{R}^2_+)$ we obtain

$$\mathcal{C}^{\tilde{S}} \subseteq \mathcal{C}^A. \qquad \blacksquare$$

Corollary 2.6. *Using the above notation and assuming that S satisfies (NA^λ), the optimization problem*

$$(P_x^{\tilde{S}}) \qquad \mathbb{E}[U(x + \varphi_T^0)] \to \max! \tag{53}$$

$$\varphi_T^0 \leqslant (H \cdot \tilde{S})_T, \quad \textit{for some } H \in \mathcal{P}. \tag{54}$$

has the same unique optimizer $\hat{\varphi}_T^0$ as the problem (P_x) defined in (30) as well as the problem (P_x^A) defined in (50).

If the λ-self-financing process $(\hat{\varphi}_t^0, \hat{\varphi}_t^1)_{t=0}^T$, starting at zero is a maximizer for problem (P_x) then

$$\hat{H}_t = \hat{\varphi}_{t-1}^1, \qquad t = 1, \ldots, T$$

defines a maximizer for problem $(P_x^{\tilde{S}})$ and we have

$$(\hat{H} \cdot \tilde{S})_T = \sum_{t=1}^T \hat{H}_t(\tilde{S}_t - \tilde{S}_{t-1}) = \hat{\varphi}_T^0 \tag{55}$$

and more generally,

$$\begin{aligned}(\hat{H} \cdot \tilde{S})_t &= \hat{\varphi}_t^0 + \hat{\varphi}_t^1 \tilde{S}_t \\ &= \hat{\varphi}_{t-1}^0 + \hat{\varphi}_{t-1}^1 \tilde{S}_t, \qquad t = 1, \ldots, T.\end{aligned} \tag{56}$$

Proof: The first part follows form (52) and Theorem 2.3.

As regards the second part, let us verify (56) by induction. Rewrite these equations as

$$(\hat{H} \cdot \tilde{S})_t = \hat{\varphi}^0_{t-1} + \hat{\varphi}^1_{t-1}\tilde{S}_t + a_t \tag{57}$$

$$(\hat{H} \cdot \tilde{S})_t = \hat{\varphi}^0_t + \hat{\varphi}^1_t\tilde{S}_t + b_t \tag{58}$$

We have to show that the elements $a_t, b_t \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ are all zero. Obviously $a_0 = 0$. As inductive hypothesis assume that $0 = a_0 \leqslant b_0 = a_1 \leqslant \ldots \leqslant b_{t-1} = a_t$. We claim that $a_t \leqslant b_t$. Indeed, $(\hat{\varphi}^0_t, \hat{\varphi}_t{}^1)$ is obtained from $(\hat{\varphi}^0_{t-1}, \hat{\varphi}^1_{t-1})$ by trading at price S_t or $(1-\lambda)S_t$, depending on whether $\hat{\varphi}_t{}^1 - \hat{\varphi}^1_{t-1} \geqslant 0$ or $\hat{\varphi}_t{}^1 - \hat{\varphi}^1_{t-1} \leqslant 0$. As $\tilde{S}_t$ takes values in $[(1-\lambda)S_t, S_t]$ we get in either case

$$(\hat{\varphi}_t{}^1 - \hat{\varphi}^1_{t-1})\tilde{S}_t + (\hat{\varphi}^0_t - \hat{\varphi}^0_{t-1}) \leqslant 0,$$

which gives $a_t \leqslant b_t$.

To complete the inductive step we have to show that $b_t = a_{t+1}$, i.e.

$$(\hat{H} \cdot \tilde{S})_{t+1} - (\hat{H} \cdot \tilde{S})_t = \hat{\varphi}_t{}^1(\tilde{S}_{t+1} - \tilde{S}_t).$$

As the left hand side equals $\hat{H}_{t+1}(\tilde{S}_{t+1} - \tilde{S}_t)$ this follows from the definition $\hat{H}_{t+1} = \hat{\varphi}^1_t$.

Having completed the inductive step we conclude that $b_T \geqslant 0$. We have to show that $b_T = 0$. If this were not the case we would have

$$\begin{aligned}\mathbb{E}\Big[U(x + (\hat{H} \cdot \tilde{S})_T)\Big] &= \mathbb{E}\left[U(x + \hat{\varphi}^0_T + b_T)\right] \\ &> \mathbb{E}\left[U(x + \hat{\varphi}^0_T)\right],\end{aligned}$$

which contradicts the first part of the corollary, showing (55) and (56). ■

Here is the economic interpretation of the above argument: whenever $\hat{\varphi}^1_t - \hat{\varphi}^1_{t-1} \neq 0$ we must have that $\tilde{S}_t$ equals either the bid or the ask price $(1-\lambda)S_t$, resp. S_t, depending on the sign of $\hat{\varphi}^1_t - \hat{\varphi}^1_{t-1}$. More formally

$$\left\{\hat{\varphi}^1_t - \hat{\varphi}^1_{t-1} > 0\right\} \subseteq \left\{\tilde{S}_t = S_t\right\}, \tag{59}$$

$$\left\{\hat{\varphi}^1_t - \hat{\varphi}^1_{t-1} < 0\right\} \subseteq \left\{\tilde{S}_t = (1-\lambda)S_t\right\}, \qquad t = 0, \ldots, T. \tag{60}$$

The predictable process $(\hat{H}_t)^T_{t=1}$ denotes the holdings of stock during the intervals $(]t-1, t])^T_{t=1}$. Inclusion (59) indicates that the utility maximizing agent, trading optimally in the frictionless market $\tilde{S}$, only increases her investment in stock when $\tilde{S}$ equals the ask price S. Inclusion (60) indicates the analogous result for the case of decreasing the investment in stock. The inclusions pertain to $\mathcal{F}_{t-1}$-measurable sets, i.e. to investment decisions done at time $t-1$, where $t-1$ range from 0 to T. The reader may check that, defining $\hat{H}_0 = \hat{H}_{T+1} = 0$, this reasoning also extends to the trading decisions at time $t = 0$ and $t = T + 1$.

The reader may wonder why we index the process H by $(H_t)_{t=0}^{T+1}$, while y is indexed by $(y_t)_{t=-1}^{T}$. As regards H, this is the usual definition of a *predictable* process from the frictionless theory (where H_{T+1} plays no role). The reason why we shift the indexation for t by 1 will be discussed in the more general continuous time setting in section 4.

One may also turn the point of view around and start from a process $\tilde{S}$ (obtained, e.g., from an educated guess) such that the associated (frictionless) optimizer $\hat{\varphi}_t^1 = \hat{H}_{t+1}$ satisfies (59) and (60), and deduce from the solution of $(P_x^{\tilde{S}})$ the solution of (P_x). In fact, this idea will turn out to work very nicely in the applications (see section 3 below).

Here is a formal definition [48].

Definition 2.7. *Fix a process* $(S_t)_{t=0}^T$ *and* $0 \leqslant \lambda < 1$ *such that* (NA^λ) *is satisfied, as well as a utility function* U *and an initial endowment* $x \in D$ *as above. In addition, suppose that* $\tilde{S} = (\tilde{S}_t)_{t=0}^T$ *is an adapted process defined on* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbb{P})$, *taking its values in the bid-ask spread* $([(1-\lambda)S_t, S_t])_{t=0}^T$. *We call* $\tilde{S}$ *a* shadow price process *for* S *if there is an optimizer* $(\hat{H}_t)_{t=1}^T$ *for the frictionless market* $\tilde{S}$, *i.e.*

$$\mathbb{E}_{\mathbb{P}}\left[U\left(x + (\hat{H}\cdot\tilde{S})_T\right)\right] = \sup\left\{\mathbb{E}_{\mathbb{P}}\left[U(x + (H\cdot\tilde{S})_T)\right] : H \in \mathcal{P}\right\},$$

such that

$$\left\{\Delta\hat{H}_t > 0\right\} \subseteq \{\tilde{S}_{t-1} = S_{t-1}\}, \qquad t = 1,\dots,T, \tag{61}$$

$$\left\{\Delta\hat{H}_t < 0\right\} \subseteq \{\tilde{S}_{t-1} = (1-\lambda)S_{t-1}\}, \qquad t = 1,\dots,T. \tag{62}$$

Theorem 2.8. *Suppose that* $\tilde{S}$ *is a shadow price for* S, *and let* $\hat{H}, U, x$, *and* $0 \leqslant \lambda < 1$ *be as in Definition 2.7.*

Then we obtain an optimal (in the sense of (30)*) trading strategy* $(\hat{\varphi}_t^0, \hat{\varphi}_t^1)_{t=-1}^T$ *in the market* S *under transaction costs* λ *via the identification* $\hat{\varphi}_{-1}^0 = \hat{\varphi}_{-1}^1 = 0$ *and*

$$\hat{\varphi}_{t-1}^1 = \hat{H}_t, \qquad t = 1,\dots,T, \tag{63}$$

$$\hat{\varphi}_{t-1}^0 = -\hat{\varphi}_{t-1}^1\tilde{S}_{t-1} + (\hat{H}\cdot\tilde{S})_{t-1}, \qquad t = 1,\dots,T, \tag{64}$$

as well as $\hat{\varphi}_T^1 = 0, \hat{\varphi}_T^0 = (\hat{H}\cdot\tilde{S})_T$.

Proof: Again the proof reduces to the economically obvious fact that trading in the frictionless market $\tilde{S}$ yields better terms of trade than in the market S under transaction costs λ. This is formalized by the first inclusion in Lemma 2.5. Hence (61) and (62) imply that the frictionless trading strategy $(\hat{H}_t)_{t=1}^T$ can be transformed into a trading strategy $(\hat{\varphi}_t^0, \hat{\varphi}_t^1)_{t=-1}^T$ via (63) and (64). ■

Remark 2.9. *In the above analysis the notion of the Legendre transform played a central role.*

As a side step – which may be safely skipped without missing any mathematical content – let us try to give an economic "interpretation", or rather "visualisation" of the conjugate function V

$$V(y) = \sup_x (U(x) - xy). \tag{65}$$

Instead of interpreting U as a function which maps money to happiness, it seems more feasible for the present purpose to interpret U as a production function.

We shall only give a hypothetical mind experiment which is silly form a realistic point of view: 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, then take the resulting kilos of gold, and then the story is finished. In other words, the gold mine is a machine turning money into gold. The monotonicity and concavity of U correspond to the "law of diminishing returns".

Now suppose that gold is traded at a price of y^{-1} Euros for one kilo of gold or, equivalently, y is 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 of $\hat{x}$ Euros for which the marginal production $U'(\hat{x})$ of kilos of gold per invested Euro equals the market price y of one Euro in terms of gold, i.e. $\hat{x}$ is determined by $U'(\hat{x}) = y$.

Given the price y, we thus may interpret the conjugate function (65) as the net value $V(y)$ of your gold mine in terms of kilos of gold: it equals $V(y) = \sup_x(U(x) - xy) = U(\hat{x}) - \hat{x}y$. Indeed, starting from an initial capital of 0 Euros it is optimal for you to borrow $\hat{x}$ Euros and 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 the 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} Euros 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 of x Euros. If the devil is able to manipulate the market, then he might be evil and choose the price y 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 $\hat{y}$ (i.e. the meanest choice of the devil) is determined by the first order condition $V'(\hat{y}) = -x$. The duality relation

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

thus may interpreted in the following way: if the devil does the choice of y which is least favourable for you, then you will earn the same amount of gold as if you would have done by keeping the mine and investing your x Euros directly into the mine. In both cases the result equals $U(x)$ kilos of gold.

Next we try to visualize the theme of Theorem 2.2: we not only consider the utility function U, but also the financial market S under transaction costs λ. In this variant of the above story you invest into the goldmine at time T to transform an investment of ξ units of Euros into $U(\xi)$ many kilos of gold. At time $t = 0$ you start with an initial capital of x Euros and you are allowed to trade in the financial market S under transaction costs λ by choosing a trading strategy φ. This will result in a random variable of $x + \varphi^0_T(\omega)$ Euros which you can transform into $U(x + \varphi^0_T(\omega))$ kilos of gold. Passing to the optimal strategy $\hat{\varphi}^0_T$ you therefore obtain $U(x + \hat{\varphi}^0_T(\omega))$ many kilos of gold if ω turns out to be the true state of the world. In average this will yield $u(x) = \mathbb{E}_{\mathbb{P}}[U(x + \hat{\varphi}^0_T)]$ many kilos of gold. We thus may consider the indirect utility function $u(x)$ as a machine which turns the original wealth x into $u(x)$ many expected kilos of gold, provided you invest optimally into the financial market S and subsequently into the gold mine also in an optimal way.

We now pass again to the dual problem, i.e., to the devil to whom you have given your gold mine. Fix your initial wealth x and first regard $u(x)$ simply as a utility function as in the first part of this remark.

We may define the conjugate function

$$v(y) = \sup_{\xi}(u(\xi) - \xi y) \tag{66}$$

and interpret it as the net value of the gold mine, denoted in expected kilos of gold, if the price y of Euro versus gold equals y at time $t = 0$. Indeed the argument works exactly as in the first part of this remark where again we interpret u as a machine turning money into gold (measured in expectation and assuming that you trade optimally). In particular we get for the "devilish" price $\hat{y}$ at time $t = 0$, given by $\hat{y} = u'(x)$, that the devil gives you at time $t = 0$ precisely the amount of $v(\hat{y})$ kilos of gold such that $v(\hat{y}) + x\hat{y}$ equals $u(x)$, i.e. the expected kilos of gold which you could obtain by trading optimally and investing into the gold mine at time T.

But this time there is an additional feature: the devil will also do something more subtle. He offers you, alternatively, to pay $V(y(\omega))$ many kilos of gold as recompensation for leaving him the goldmine. The payment now depends on the prize $y(\omega)$ of one Euro in terms of gold at time T which may depend on the random element ω and which is only revealed at time T. The function V now is the conjugate function of the original utility function U as defined in (65).

The main message of Theorem 2.2 can be resumed in prose as follows

(a) there is a choice of "devilish" prices $\hat{y}(\omega)$ given by the marginal utility of the optimal terminal wealth

$$\hat{y}(\omega) = U'(x + \hat{\varphi}^0_T(\omega)), \quad \omega \in \Omega.$$

(b) There is a probability measure $\hat{Q}$ on Ω such that

$$\hat{y}(\omega) = \hat{y}\frac{d\hat{Q}}{d\mathbb{P}}(\omega), \text{ where } \quad \hat{y} \quad \text{is the optimizer in} \quad (66).$$

It follows that $\sum_{\omega} \hat{y}(\omega)\mathbb{P}(\omega) = \hat{y}$, i.e., $\hat{y}$ is the $\mathbb{P}$-average of the prizes $\hat{y}(\omega)$.

(c) The formula

$$v(\hat{y}) = \mathbb{E}_{\mathbb{P}}\Big[V\big(\hat{y}(\omega)\big)\Big] = \mathbb{E}_{\mathbb{P}}\Big[V\Big(\hat{y}\frac{d\hat{Q}}{d\mathbb{P}}(\omega)\Big)\Big]$$

now has the interpretation that the devil gives you (in average) the same amount of gold, namely $v(\hat{y})$ many kilos, independently of whether you do the deal with him at time $t = 0$ or $t = T$.

(d) If you choose any strategy φ we have the inequality

$$\mathbb{E}_{\hat{Q}}[\varphi_T^0] \leqslant \mathbb{E}_{\hat{Q}}[\hat{\varphi}_T^0] = x$$

as $\hat{Q}$ is a λ-consistent price system. Hence

$$\mathbb{E}_{\mathbb{P}}\big[\big(x + \varphi_T^0(\omega)\big)\hat{y}(\omega)\big] \leqslant \mathbb{E}_{\mathbb{P}}\big[\big(x + \hat{\varphi}_T^0(\omega)\big)\hat{y}(\omega)\big] \tag{67}$$

which may be interpreted in the following way: if you accept the devil's offer to get the amount of $V(\hat{y}(\omega))$ kilos of gold at time T, you cannot improve your expected result by changing from $\hat{\varphi}$ to some other trading strategy φ, while the devil remains his choice of prices $\hat{y}(\omega)$ unchanged.

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