A Hands-on Approach to Optimal Stopping

6th School of Probability and Stochastic Processes, Guanajuato, Mexico, 31st August - 4th September, 2009
by

Andreas E. Kyprianou,
Department of Mathematical Sciences,
University of Bath,
Claverton Down,
Bath, BA2 7AY.
a.kyprianou@bath.ac.uk

These notes give a more detailed account of the material discussed in lectures. They are largely based on material found in books and research papers. In the former category, we have relied heavily on the following two texts:

G. Peskir and A.N. Shiryaev (2006) Optimal stopping and free-boundary problems. Lectures in Mathematics, ETH Zurich, Birkhäuser.

A.E. Kyprianou (2006) Introductory Lectures on fluctuations of Lévy process with applications. Universitext, Springer.

I have also made use of earlier lecture notes of my colleague A.M.G. Cox at
`http://www.maths.bath.ac.uk/ mapamgc/OptimalStopping.htm`

1 Some aspects of the theory of optimal stopping in discrete time

Suppose that $(G_n)_{n\geq 0}$ is a sequence of adapted non-negative random variables on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$, and interpret G_n as the gain, or reward we receive for stopping at time n. Define $\mathcal{M}_\tau$ as the set of almost surely finite stopping times which are greater than τ (which itself may be a stopping time). We are interested in solving the optimal stopping problem

$$V = \sup_{\tau \in \mathcal{M}_0} \mathbb{E} G_\tau.$$

That means we are interested in both the stopping time which optimizes the expectation (if it exists) as well as the optimal value V. There is an obvious assumption we should make here before we start our analysis. Namely

$$\mathbb{E}\left[\sup_n G_n\right] < \infty. \tag{1.1}$$

Essentially this ensures that all expectations are uniformly finite.

It seems sensible that the decision to stop must in some sense be based on the future expectation of the return given what has been seen to date. So if stopping has not occurred by step n then one should look at all the quantities $\mathbb{E}(G_\tau|\mathcal{F}_n)$ for stopping times $\tau \in \mathcal{M}_n$, which can be thought of as the potential returns if not stopping at that step, and compare them against the gain G_n which would be incurred if one stopped at that step. In particular one would like to look at the quantity "$\sup_{\tau\in\mathcal{M}_n} \mathbb{E}(G_\tau|\mathcal{F}_n)$" however there are an uncountable number of stopping times in $\mathcal{M}_n$ and each of the objects in the supremum is a random variable. It is therefore not entirely clear what one means by "$\sup_{\tau\in\mathcal{M}_n} \mathbb{E}(G_\tau|\mathcal{F}_n)$". To get resolve this issue, we need to introduce the concept of an **essential supremum**.

1.1 Essential supremum

This section is entirely devoted to the following theorem which gives us the definition and functionality of the concept of an essential supreumum of random variables.

Theorem 1.1 *Let $\{Z_\alpha; \alpha \in I\}$ be a collection of real-valued random variables in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with I an arbitrary index set. Then there exists a countable subset $J \subseteq I$ such the random variable $Z^* : \Omega \to \mathbb{R} \cup \{-\infty, \infty\}$ defined by*

$$Z^* = \sup_{\alpha \in J} Z_\alpha$$

satisfies

(i) $\mathbb{P}(Z_\alpha \leq Z^*) = 1, \quad \forall \alpha \in I,$

(ii) if $Y : \Omega \to \bar{\mathbb{R}}$ is another random variable satisfying (i), then

$$\mathbb{P}(Z^* \leq Y) = 1.$$

We call Z^* the **essential supremum** *of $\{Z_\alpha; \alpha \in I\}$, and write*

$$Z^* = \operatorname*{esssup}_{\alpha \in I} Z_\alpha.$$

It is defined uniquely $\mathbb{P}$-almost surely.

Proof. Mapping by a suitable mapping $f : [-\infty, \infty] \to [-1, 1]$ (eg. $f(x) = (2/\pi) \arctan x$), we may assume all Z_α are bounded in $[-1, 1]$. Let $\mathcal{C}$ be the set of countable subsets of I. Then by monotinicity

$$[-1, 1] \ni a := \sup_{C \in \mathcal{C}} \mathbb{E}\left[\sup_{\alpha \in C} Z_\alpha\right] = \sup_{n \geq 1} \mathbb{E}\left[\sup_{\alpha \in C_n} Z_\alpha\right]$$

for some suitable sequence $C_n \subseteq \mathcal{C}$.[1] However, $\bigcup_{n \geq 1} C_n$ is a countable set, so we may define the random variable

$$Z^* = \sup_{\alpha \in \bigcup_n C_n} Z_\alpha$$

and thus we may take $J = \bigcup_n C_n$.

We now need to confirm that Z^* respects the properties (i) and (ii). Note that the first property is clearly true for all $\alpha \in J$. Suppose that $\beta \in I \backslash J$ and $\mathbb{P}(Z_\beta > Z^*) > 0$. Then, on the one hand, because of the latter positive probability, we have $a < \mathbb{E}(Z_\beta \vee Z^*)$ and on the other hand

$$\mathbb{E}(Z_\beta \vee Z^*) = \mathbb{E}\left[\sup_{\alpha \in J \cup \{\beta\}} Z_\alpha\right] \leq \sup_{C \in \mathcal{C}} \mathbb{E}\left[\sup_{\alpha \in C} Z_\alpha\right] = a,$$

where the inequality follows on account of the fact that $J \cup \{\beta\}$ is countable. We therefore reach a contradiction. The property (i) thus follows.

The second property is a straightforward consequence of the fact that if for all $\alpha \in I$, $Z_\alpha \leq Y$ almost surely, then in particular for all $\alpha \in J$, $Z_\alpha \leq Y$ and hence by the countability of J,

$$Z^* = \sup_{\alpha \in J} Z_\alpha \leq Y$$

almost surely. ■

The following corollary is an important feature that will be of most interest to us with regard to the definition of the essential supremum.

[1] Note that this statement can be derived from the fact that if a_n is a sequence of positive numbers then there exists an increasing subsequence α_n such that $\sup\{a_n\} = \lim_{n \uparrow \infty} \alpha_n = \sup_n \alpha_n$.

Corollary 1.1 *If the family* $\{Z_\alpha : \alpha \in I\}$ *has the lattice property that for any* α, β *in* I *there exists a* $\gamma \in I$ *such that* $Z_\alpha \vee Z_\beta \leq Z_\gamma$ *in the* $\mathbb{P}$*-almost sure sense, then the countable subset* $J = \{\alpha_0, \alpha_1, \alpha_2, \cdots\}$ *may be chosen so that*

$$Z^* = \lim_{n\uparrow\infty} Z_{\alpha_n}$$

where $Z_{\alpha_0} \leq Z_{\alpha_1} \leq Z_{\alpha_2} \leq \cdots$ *in the* $\mathbb{P}$*-almost sure sense.*

Proof. Suppose that the countable set in the previous theorem is $J = \{\alpha_0, \alpha_1, \alpha_2, \cdots, \cdots\}$. Then we replace it by a new sequence $J^* = \{\alpha_0^*, \alpha_1^*, \alpha_2^*, \cdots\}$ where $\alpha_0^* = \alpha_0$ and the remaining elements in the sequence are inductively chosen so that, making use of the lattice property, α_{n+1}^* is such that $Z_{\alpha_{n+1}^*} \geq Z_{\alpha_n^*} \vee Z_{\alpha_n}$ almost surely.

Since $Z_{\alpha_{n+1}^*} \geq Z_{\alpha_n}$ almost surely and since $\{Z_{\alpha_n^*}\}$ is an increasing sequence whose elements which are each almost surely bounded by Z^*, it is now obvious that

$$Z^* \geq \lim_{n\uparrow\infty} Z_{\alpha_n^*} = \sup_{\alpha\in J^*} Z_\alpha \geq \sup_{\alpha\in J} Z_\alpha = Z^*$$

in the almost sure sense. That is to say $Z^* = \lim_{n\uparrow\infty} Z_{\alpha_n^*}$ almost surely. ■

1.2 Solution to the optimal stopping problem

Now that we understand how to define an essential supremum (ie a "supremum over random variables") we may proceed to look at putting our intuition into rigour and showing how the process

$$S_n = \operatorname*{esssup}_{\tau\in\mathcal{M}_n} \mathbb{E}\left[G_\tau|\mathcal{F}_n\right] \tag{1.2}$$

(the best gain one can expect by not stopping at time n) can be used in comparison with G_n (the gain obtained by stopping at time n) to produce an optimal stopping strategy. Note that since $\tau = n$ is a stopping time in $\mathcal{M}_n$ it is trivial to see that $S_n \geq G_n$ almost surely. It is therefore conceivable that the optimal strategy is to stop at the first time n such that $S_n = G_n$ as otherwise, when $S_n > G_n$ not stopping would appear to have greater "value". The process $\{S_n : n \geq 0\}$ is called the Snell envelope; named after Laurie Snell, it "envelopes" the gain process $\{G_n : n \geq 0\}$.

Theorem 1.2 *Fix* $n \in \{0, 1, 2, \cdots\}$*. Suppose* (1.1) *holds. Let*

$$\tau_n = \inf\{k \geq n : S_k = G_k\}.$$

and assume that $\mathbb{P}(\tau_n < \infty) = 1$*. Consider the optimal stopping problem:*

$$V_n = \sup_{\tau\in\mathcal{M}_n} \mathbb{E}G_\tau. \tag{1.3}$$

Then,

(i) $S_n = \max\{G_n, \mathbb{E}(S_{n+1}|\mathcal{F}_n)\}$;

(ii) *The stopped process* $\{S_{k\wedge\tau_n} : k \geq n\}$ *is a martingale;*

(iii) $V_n = \mathbb{E}S_n$ *and the stopping time* τ_n *is optimal in* (1.3);

(iv) *The process* $\{S_k : k \geq n\}$ *is the smallest supermartingale which dominates* $\{G_k : k \geq n\}$.

Proof. We begin by showing that $S_n \geq \mathbb{E}[S_{n+1}|\mathcal{F}_n]$. Suppose that $\sigma_1, \sigma_2 \in \mathcal{M}_{n+1}$ and let $A = \{\mathbb{E}[G_{\sigma_1}|\mathcal{F}_{n+1}] \geq \mathbb{E}[G_{\sigma_2}|\mathcal{F}_{n+1}]\} \in \mathcal{F}_{n+1}$. Then we can define a stopping time[2]

$$\sigma_3 = \sigma_1 \mathbf{1}_A + \sigma_2 \mathbf{1}_{A^C} \in \mathcal{M}_{n+1}.$$

Hence:

$$\begin{aligned}
\mathbb{E}[G_{\sigma_3}|\mathcal{F}_{n+1}] &= \mathbf{1}_A \mathbb{E}[G_{\sigma_1}|\mathcal{F}_{n+1}] + \mathbf{1}_{A^C}\mathbb{E}[G_{\sigma_2}|\mathcal{F}_{n+1}] \\
&= \max\left\{\mathbb{E}[G_{\sigma_1}|\mathcal{F}_{n+1}], \mathbb{E}[G_{\sigma_2}|\mathcal{F}_{n+1}]\right\}.
\end{aligned}$$

This shows that the random variables $\{\mathbb{E}(G_\sigma|\mathcal{F}_{n+1}) : \sigma \in \mathcal{M}_{n+1}\}$ have the lattice property. Now, by Corollary 1.1, we know there exists a sequence of stopping times $\sigma_k \in \mathcal{M}_{n+1}$ such that

$$S_{n+1} = \lim_{k\to\infty} \mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]$$

where $\mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]$ is a sequence of increasing random variables. Note that since $\mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]$ is $\mathcal{F}_{n+1}$-measurable, it follows that, as an increasing limit, S_{n+1} is also $\mathcal{F}_{n+1}$ measurable. So thanks to dominated (or indeed monotone) convergence

$$\begin{aligned}
\mathbb{E}[S_{n+1}|\mathcal{F}_n] &= \mathbb{E}\left[\lim_{k\to\infty} \mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]|\mathcal{F}_n\right] \\
&= \lim_{k\to\infty} \mathbb{E}\left[\mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]|\mathcal{F}_n\right] \\
&= \lim_{k\to\infty} \mathbb{E}[G_{\sigma_k}|\mathcal{F}_n] \\
&\leq S_n \qquad (1.4)
\end{aligned}$$

where the final inequality follows by virtue of the definition of S_n. Recalling that $S_n \geq G_n$ we have shown that

$$S_n \geq \max\{G_n, \mathbb{E}(S_{n+1}|\mathcal{F}_n)\}$$

On the other hand, for all $\tau \in \mathcal{M}_n$, we get,

$$\begin{aligned}
\mathbb{E}[G_\tau|\mathcal{F}_n] &= G_n \mathbf{1}_{\{\tau=n\}} + \mathbf{1}_{\{\tau>n\}}\mathbb{E}[G_\tau|\mathcal{F}_n] \\
&= G_n \mathbf{1}_{\{\tau=n\}} + \mathbf{1}_{\{\tau>n\}}\mathbb{E}[G_{\tau\vee(n+1)}|\mathcal{F}_n] \\
&= G_n \mathbf{1}_{\{\tau=n\}} + \mathbf{1}_{\{\tau>n\}}\mathbb{E}[\mathbb{E}[G_{\tau\vee(n+1)}|\mathcal{F}_{n+1}]|\mathcal{F}_n] \\
&\leq G_n \mathbf{1}_{\{\tau=n\}} + \mathbf{1}_{\{\tau>n\}}\mathbb{E}[S_{n+1}|\mathcal{F}_n] \\
&\leq \max\{G_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]\},
\end{aligned}$$

[2] It is a simple exercise to check that this is a stopping time.

so that $S_n \le \max\{G_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]\}$. We have thus established property (i) that

$$S_n = \max\{G_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]\}.$$

It is now immediately obvious that $S_n \ge \mathbb{E}[S_{n+1}|\mathcal{F}_n]$ showing that it is a supermartingale (which, as earlier shown, dominates the gain). It is also obvious that that if $n \le k < \tau_n$ then by definition of τ_n, $S_k = \mathbb{E}[S_{k+1}|\mathcal{F}_k]$ a.s.. Since $S_{\tau_n} = G_{\tau_n}$ it follows that

$$S_{k\wedge\tau_n} = \mathbb{E}[S_{(k+1)\wedge\tau_n}|\mathcal{F}_k], \quad \forall k \ge n, \tag{1.5}$$

showing property (ii). In particular this implies that for $N > n$,

$$S_n = S_{n\wedge\tau_n} = \mathbb{E}[S_{\tau_n\wedge N}|\mathcal{F}_n] = \mathbb{E}[G_{\tau_n}\mathbf{1}_{\{\tau_n<N\}} + S_N\mathbf{1}_{\{\tau_n\ge N\}}|\mathcal{F}_n].$$

Now note that for $N > n$,

$$S_N = \operatorname*{esssup}_{\tau\in\mathcal{M}_N} \mathbb{E}(G_\tau|\mathcal{F}_N) \le \mathbb{E}(\sup_n G_n|\mathcal{F}_N)$$

and therefore

$$\begin{aligned}
\lim_{N\uparrow\infty} \mathbb{E}(S_N\mathbf{1}_{\{\tau_n\ge N\}}|\mathcal{F}_n) &\le \lim_{N\uparrow\infty} \mathbb{E}\mathbb{E}(\sup_n G_n|\mathcal{F}_N)\mathbf{1}_{\{\tau_n\ge N\}}|\mathcal{F}_n) \\
&\le \lim_{N\uparrow\infty} \mathbb{E}(\sup_n G_n\mathbf{1}_{\{\tau_n\ge N\}}|\mathcal{F}_n) \\
&= 0
\end{aligned}$$

where we have used dominated convergence and the assumption that $\mathbb{E}(\sup_n G_n) < \infty$,and hence that $\mathbb{E}(\sup_n G_n|\mathcal{F}_n) < \infty$ almost surely, together with the assumption that . $\mathbb{P}(\tau_n < \infty)$ to conclude that

$$S_n = \mathbb{E}[G_{\tau_n}|\mathcal{F}_n].$$

Thus, on the one hand,

$$\mathbb{E}[S_n] = \mathbb{E}[S_{\tau_n}] = \mathbb{E}[G_{\tau_n}] \le \sup_{\tau\in\mathcal{M}_n} \mathbb{E}[G_\tau].$$

However, on the other hand, as $S_n \ge G_n$ and S_n is a supermartingale we have

$$\sup_{\tau\in\mathcal{M}_n} \mathbb{E}[G_\tau] \le \sup_{\tau\in\mathcal{M}_n} \mathbb{E}[S_\tau] \le \mathbb{E}[S_n].$$

This gives us conclusion (iii) of the theorem.

To prove (iv) Suppose $(U_k)_{k\ge n}$ is also a supermartingale dominating $(G_k)_{k\ge n}$. Then

$$S_k = \mathbb{E}[G_{\tau_k}|\mathcal{F}_k] \le \mathbb{E}[U_{\tau_k}|\mathcal{F}_k] \le U_k,$$

where the final inequality follows from conditional Fatou, applied to $U_{m\wedge\tau_k}$. ■

1.3 Markovian setting

Now consider a time-homogeneous Markov chain, $X = (X_n)_{n\geq 0}$, $X_n \in E$ for some measurable space $(E, \mathcal{B})$, defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P}_x)$, with $\mathbb{P}_x(X_0 = x) = 1$. Recall that the Markov property states that for all positive, continuous and uniformly bounded functions f, we have that

$$\mathbb{E}(f(X_{n+k})|\mathcal{F}_n) = \mathbb{E}(f(X_{n+k})|X_n) = \mathbb{E}_x(f(\widetilde{X}_k))\Big|_{x=X_n}$$

where $\widetilde{X}$ is an independent copy of X. Further, for simplicity, we will assume that $(\Omega, \mathcal{F}) = (E^{\mathbb{Z}_+}, \mathcal{B}^{\mathbb{Z}_+})$, so that the shift operators $\theta_n : \Omega \to \Omega$ can be defined by $f \circ \theta_n(w.) = f(w._{+n})$ for functionals f of the canonical process $w.$, and $\mathcal{F}_0$ is trivial. We will typically use the shift operator on stopping times. For example of $\tau \in \mathcal{M}_0$ then $\tau \circ \theta_n \in \mathcal{M}_n$. In particular we could take for example $\tau_k^B := \inf\{n \geq k : X_n \in B\}$ for some appropriate $B \in \mathcal{B}$. In that case

$$\tau_0^B \circ \theta_k = \inf\{m \geq 0 : X_{m+k} \in B\} = \inf\{n \geq k : X_n \in B\} - k = \tau_k^B - k.$$

Of course, this is just a special case of the setting considered in the previous section, so can we say anything extra? Consider an infinite horizon problem

$$V(x) = \sup_{\tau \in \mathcal{M}_0} \mathbb{E}_x G(X_\tau) \tag{1.6}$$

where we note that we have replaced the general process G_n by a process depending only on the current state of the Markov chain, $G(X_n)$, where it is usual to take G as a continuous function. Assume that there is an optimal strategy for the problem. The next result shows that we can make a more explicit connection between the Snell envelope:

Lemma 1.1 *Suppose that $\mathbb{P}_x(\tau_0 < \infty) = 1$ for all $x \in E$. Let S_n be the Snell envelope as given by* (1.2)*. Then we have:*

$$S_n = V(X_n), \tag{1.7}$$

$\mathbb{P}_x$-a.s., for all $x \in E, n \geq 0$.

Proof. We first note that

$$\begin{aligned} S_n &= \operatorname*{esssup}_{\tau\in\mathcal{M}_n} \mathbb{E}(G(X_\tau)|\mathcal{F}_n) \\ &\geq \operatorname*{esssup}_{\tau\in\mathcal{M}_0} \mathbb{E}(G(X_{n+\tau\circ\theta_n})|\mathcal{F}_n) \\ &= \operatorname*{esssup}_{\tau\in\mathcal{M}_0} \mathbb{E}_x(G(\widetilde{X}_\tau))\Big|_{x=X_n} \\ &= V(X_n), \end{aligned}$$

where $\widetilde{X}$ is an independent copy of X. The inequality is a result of the fact that for $\tau \in \mathcal{M}_0$, $n + \tau \circ \theta_n$ belongs to $\mathcal{M}_n$ but not all stopping times in $\mathcal{M}_n$ can be written in this way. We also note that for all $x \in E$

$$\begin{aligned} V(x) &\geq \sup_{\tau \in \mathcal{M}_1} \mathbb{E}_x[G(X_\tau)] \text{ as } \mathcal{M}_1 \subseteq \mathcal{M}_0 \\ &\geq \sup_{\tau \in \mathcal{M}_0} \mathbb{E}_x[G(X_{1+\tau\circ\theta_1})] \\ &= \sup_{\tau \in \mathcal{M}_0} \mathbb{E}_x[\mathbb{E}[G(X_{1+\tau\circ\theta_1})|\mathcal{F}_1]] \\ &= \sup_{\tau \in \mathcal{M}_0} \mathbb{E}_x\left[\mathbb{E}_x G(\widetilde{X}_\tau)]\Big|_{x=X_1} \right] \\ &\geq \mathbb{E}_x\left[\mathbb{E}_x G(\widetilde{X}_{\tau_0})]\Big|_{x=X_1} \right] \\ &= \mathbb{E}_x\left[\left\{ \sup_{\tau \in \mathcal{M}_0} \mathbb{E}_x[G(\widetilde{X}_\tau)] \right\}\Big|_{x=X_1} \right] \\ &\geq \mathbb{E}_x[V(X_1)]. \end{aligned}$$

where we have used the optimality of the stopping time τ_0 from (1.2). This implies that $\{V(X_n) : n \geq 0\}$ is a supermartingale. Indeed note that

$$\mathbb{E}(V(X_{n+1})|\mathcal{F}_n) = \mathbb{E}_x(V(X_1))|_{x=X_n} \leq V(X_n)$$

It is also clear from its definition that $V(x) \geq \mathbb{E}_x(G(X_\tau))$ for any stopping time in $\mathcal{M}_0$. If in particular we take the stopping time $\tau = 0$ then we observe that for all $x \in E$, $V(x) \geq G(x)$. This means that $\{V(X_n) : n \geq 0\}$ is a supermartingale dominating the gain and hence by Theorem 1.2 (iv) it follows that $V(X_n) \geq S_n$. Combined with our observation at the beginning of the proof we deduce the claim of the lemma. ■

We are now in a position to state an analogue of Theorem 1.2. Given the value function V, we define the continuation region:

$$C = \{x \in E : V(x) > G(x)\}$$

and the stopping region

$$D = \{x \in E : V(x) = G(x)\}.$$

The candidate optimal stopping time is then

$$\tau_D = \inf\{n \geq 0 : X_n \in D\}. \tag{1.8}$$

Theorem 1.3 *Suppose that $\mathbb{P}_x(\tau_D < \infty) = 1$ for all $x \in E$. Then*

(i) the stopped process $\{V(X_{n\wedge\tau_D}) : n \geq 0\}$ is a $\mathbb{P}_x$-martingale for every $x \in E$;

(ii) the stopping time τ_D is optimal in (1.6);

(iii) the process $\{V(X_n) : n \geq 0\}$ *is the smallest supermartingale dominating* $\{n \geq 0 : G(X_n)$.

Proof. This follows directly from Theorem 1.2, having made the identification $S_n = V(X_n)$ in (1.7) ■

We conclude by noting that the requirement that $\mathbb{P}_x(\tau_n < \infty) = 1$ for all $x \in E$ may appear to be an unnatural requirement. Indeed in the Markovian setting the process X may simply never visit the domain D with positive probability on account of D consisting of transient states. It is possible however to adapt many of the arguments above to accommodate for the adjustment that $\mathcal{M}$ contains optimal stopping times which may be infinite-valued with positive probability. In that case one needs to be more specific about the the meaning of $\lim_{n\uparrow\infty} G_n$. We shall see later in our examples that we will need to be able to accommodate such stopping times and indeed we shall give an operational sense to the above limit.

2 Continuous time and a change of perspective

Given what we have seen for discrete time, one would expect that one may make a direct translation to the case of continuous time stochastic processes. Indeed this is possible, although one needs to be somewhat more careful about path continuity of the underlying (super)martingales as well as some other measure-theoretic considerations.

We could easily fill some additional pages laying out this theory, however, at this point it is worth stepping back to the discrete time case and asking oneself how one would realistically go about solving a given optimal stopping problem. Let us consider just the Markovian case which, in principle, should be slightly easier. In this case we note from Theorem 1.3 that the optimal stopping time, τ_D, turns out to require prior knowledge of the value function $V(x)$. This is potentially a tautology as one could argue that in fact since $V(x) = \mathbb{E}_x(G(X_{\tau_D}))$ one needs prior knowledge of τ_D in order to know $V(x)$. This dependency makes it hard to see how one might systematically approach any Markovian optimal stopping problem.

This brings us to a technique known as "guess and verify". This is an approach in which one uses intuition to guess a solution to an optimal stopping problem and then verify that it is indeed optimal by checking a number of sufficient conditions that are inspired by the probabilistic features that we would expect to see given what we have learnt in Theorem 1.3.

We need a platform on which to exemplify this "hands-on" method of guess and verify. The optimal stopping problems we consider in this chapter will be of the form

$$v(x) = \sup_{\tau \in \mathcal{T}} \mathbb{E}_x(\mathrm{e}^{-q\tau} G(X_\tau)) \tag{2.1}$$

where $X = \{X_t : t \geq 0\}$ under $\mathbb{P}_x$ a Lévy process started from $X_0 = x$. Further, G is a non-negative measurable function, $q \geq 0$ and $\mathcal{T}$ is a family of stopping times with respect to the natural filtration generated by X, $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$. We recall below the definition of a Lévy process.

Definition 2.1 (Lévy Process) *A process $X = \{X_t : t \geq 0\}$ is said to be a Lévy process if it possesses the following properties:*

(i) *The paths of X are almost surely right continuous with left limits.*

(ii) $X_0 = 0$ *almost surely.*

(iii) *For $0 \leq s \leq t$, $X_t - X_s$ is equal in distribution to X_{t-s}.*

(iv) *For $0 \leq s \leq t$, $X_t - X_s$ is independent of $\{X_u : u \leq s\}$.*

A Lévy process started from $X_0 = x$ is simply defined as $x + X_t$ for $t \geq 0$.

Note that the properties of stationary and independent increments implies that a Lévy process is a Markov process. Thanks to almost sure right continuity of paths, one may show in addition that Lévy processes are also Strong Markov processes.

Henceforth we keep to the standard notation that $(X, \mathbb{P}_x)$ is a Lévy processes issued from $x \in \mathbb{R}$. For convenience we shall write $\mathbb{P}$ in place of $\mathbb{P}_0$. Here we give sufficient conditions with which one may verify that a candidate solution solves the optimal stopping problem (2.1). In effect we will be considering the gain to be a function of the two dimensional Markov process (t, X_t) (ie. a function of space and time).

2.1 Sufficient Conditions for Optimality

Lemma 2.1 *Consider the optimal stopping problem (2.1) for $q \geq 0$ under the assumption that for all $x \in \mathbb{R}$,*

$$\mathbb{P}_x(\textit{there exists } \lim_{t\uparrow\infty} \mathrm{e}^{-qt} G(X_t) < \infty) = 1. \tag{2.2}$$

Suppose that $\tau^* \in \mathcal{T}$ is a candidate optimal strategy for the optimal stopping problem (2.1) and let $v^*(x) = \mathbb{E}_x(\mathrm{e}^{-q\tau^*} G(X_{\tau^*}))$. Then the pair (v^*, τ^*) is a solution if

(i) *$v^*(x) \geq G(x)$ for all $x \in \mathbb{R}$,*

(ii) *the process $\{\mathrm{e}^{-qt} v^*(X_t) : t \geq 0\}$ is a right continuous supermartingale.*

Proof. The definition of v^* implies that

$$\sup_{\tau\in\mathcal{T}} \mathbb{E}_x(\mathrm{e}^{-q\tau} G(X_\tau)) \geq v^*(x)$$

for all $x \in \mathbb{R}$. On the other hand, property (ii) together with Doob's Optional Stopping Theorem[3] imply that for all $t \geq 0$, $x \in \mathbb{R}$ and $\sigma \in \mathcal{T}$,

$$v^*(x) \geq \mathbb{E}_x(\mathrm{e}^{-q(t\wedge\sigma)} v^*(X_{t\wedge\sigma}))$$

and hence by property (i), Fatou's Lemma, the non-negativity of G and assumption (2.2)

$$\begin{aligned} v^*(x) &\geq \liminf_{t\uparrow\infty} \mathbb{E}_x(\mathrm{e}^{-q(t\wedge\sigma)} G(X_{t\wedge\sigma})) \\ &\geq \mathbb{E}_x(\liminf_{t\uparrow\infty} \mathrm{e}^{-q(t\wedge\sigma)} G(X_{t\wedge\sigma})) \\ &= \mathbb{E}_x(\mathrm{e}^{-q\sigma} G(X_\sigma)). \end{aligned}$$

As $\sigma \in \mathcal{T}$ is arbitrary, it follows that for all $x \in \mathbb{R}$

$$v^*(x) \geq \sup_{\tau\in\mathcal{T}} \mathbb{E}_x(\mathrm{e}^{-q\tau} G(X_\tau)).$$

In conclusion it must hold that

$$v^*(x) = \sup_{\tau\in\mathcal{T}} \mathbb{E}_x(\mathrm{e}^{-q\tau} G(X_\tau))$$

[3] Right continuity of paths is implicitly used here.

for all $x \in \mathbb{R}$. ■

When G is a monotone function and $q > 0$, a reasonable class of candidate solutions that one may consider in conjunction with the previous lemma are those based on first passage times over a specified threshold. That is, either first passage above a given constant in the case that G is monotone increasing or first passage below a given constant in the case that G is monotone decreasing. An intuitive justification may be given as follows.

Suppose that G is monotone increasing. In order to optimise the value $G(X_\tau)$ one should stop at some time τ for which X_τ is large. On the other hand, this should not happen after too much time on account of the exponential discounting. This suggests that there is a threshold, which may depend on time, over which one should stop X in order to maximise the expected discounted gain. Suppose however, that by time $t > 0$ one has not reached this threshold. Then, by the Markov property, given $X_t = x$, any stopping time τ which depends only on the continuation of the path of X from the space-time point (x, t) would yield an expected gain $\mathrm{e}^{-qt}\mathbb{E}_x(\mathrm{e}^{-q\tau}G(X_\tau))$. The optimisation of this expression over the latter class of stopping times is essentially the same procedure as in the original problem (2.1). Note that since X is a Markov process, there is nothing to be gained by considering stopping times which take account of the history of the process $\{X_s : s < t\}$. These arguments suggest that threshold should not vary with time and hence a candidate for the optimal strategy takes the form

$$\tau_y^+ = \inf\{t > 0 : X_t \in A\},$$

where $A = [y, \infty)$ or (y, ∞) for some $y \in \mathbb{R}$. Similar reasoning applies when G is monotone decreasing.

It is quite difficult to make the above intuition rigorous. Indeed in many situations, it may simply be wrong as montonicity is not the only issue at stake for a threshold strategy to be optimal. Optimal stopping is not a categorical theory in the sense that, aside from the general theorems we have addressed in the previous section, the specific nature of optimal strategies for individual problems can vary tremendously in their nature. Experience shows that there are families of problems which exhibit similarities in their solutions and the method of solution. However, for every example of a principle which appears to lie behind several optimal stopping problems, it is never too difficult to construct another optimal stopping problem for which this principle is not applicable.

We therefore continue our exposition by looking at two concrete examples of optimal stopping problems which can be solved using similar ideas and which appeal to the intuition given above; however, we also add the additional warning that this is but a small taste of the techniques and types of mathematical structure that one can expect to see.

2.2 The McKean Optimal Stopping Problem

This optimal stopping problem is given by

$$v(x) = \sup_{\tau \in \mathcal{T}} \mathbb{E}_x(\mathrm{e}^{-q\tau}(K - \mathrm{e}^{X_\tau})^+), \tag{2.3}$$

where $q > 0$ and $\mathcal{T}$ is all $\mathbb{F}$-stopping times. The solution to this optimal stopping problem was first considered by [9] for the case that X is linear Brownian motion in the context of the optimal time to sell a risky asset for a fixed price K and in the presence of discounting, where the value of the risky asset follows the dynamics of an exponential Brownian motion.

In [5] a solution to a discrete-time analogue of (2.3) was obtained. In that case, the process X is replaced by a random walk. Some years later and again within the context of the the optimal time to sell a risky asset (the pricing of an American put), a number of authors dealt with the solution to (2.3) for a variety of special classes of Lévy processes.[4] Below we give the solution to (2.3) as presented in [11]. The proof we shall give here comes however from Alili and Kyprianou (2005) and remains close in nature to the random walk proofs of [5].

Theorem 2.1 *The solution to (2.3) under the stated assumption is given by*

$$v(x) = \frac{\mathbb{E}\left(\left(K\mathbb{E}\left(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}\right) - \mathrm{e}^{x+\underline{X}_{\mathbf{e}_q}}\right)^+\right)}{\mathbb{E}\left(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}\right)}$$

and the optimal stopping time is given by

$$\tau^* = \inf\{t > 0 : X_t < x^*\}$$

where

$$x^* = \log K\mathbb{E}\left(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}\right).$$

Here, the symbol $\mathbf{e}_q$ denotes an independent random variable which is independent of X and exponentially distributed. Further, $\underline{X}_t = \inf_{s\leq t} X_s$.

Before proving this theorem, we need to establish and auxiliary lemma which will help us establish the value of stopping on a first passage strategy. To this end, define for $x \in \mathbb{R}$ the first passage times

$$\tau_x^+ := \inf\{t > 0 : X_t > x\} \text{ and } \tau_x^- := \inf\{t > 0 : X_t < x\}$$

and let $\overline{X}_t = \sup_{s\leq t} X_s$.

Lemma 2.2 *For all $q > 0$, $\beta \geq 0$ and $x \geq 0$ we have*

$$\mathbb{E}\left(e^{-q\tau_x^+ - \beta X_{\tau_x^+}}\mathbf{1}_{(\tau_x^+<\infty)}\right) = \frac{\mathbb{E}\left(e^{-\beta\overline{X}_{\mathbf{e}_q}}\mathbf{1}_{(\overline{X}_{\mathbf{e}_q}>x)}\right)}{\mathbb{E}\left(e^{-\beta\overline{X}_{\mathbf{e}_q}}\right)}. \tag{2.4}$$

[4] [6] dealt with the case of bounded variation spectrally positive Lévy processes; Boyarchenko and Levendorskiĭ (2002) handled a class of tempered stable processes; [3] covers the case of spectrally negative processes; [2] deal with spectrally negative Lévy processes again; [1] look at Lévy processes which have phase-type jumps and [4] again for the spectrally negative case.

Proof. First, assume that q, β, $x > 0$ and note that

$$\begin{aligned}
&\mathbb{E}\left(e^{-\beta\overline{X}_{\mathbf{e}_q}}\mathbf{1}_{(\overline{X}_{\mathbf{e}_q}>x)}\right)\\
&= \mathbb{E}\left(e^{-\beta\overline{X}_{\mathbf{e}_q}}\mathbf{1}_{(\tau_x^+<\mathbf{e}_q)}\right)\\
&= \mathbb{E}\left(\mathbf{1}_{(\tau_x^+<\mathbf{e}_q)}e^{-\beta X_{\tau_x^+}}\mathbb{E}\left(e^{-\beta(\overline{X}_{\mathbf{e}_q}-X_{\tau_x^+})}\Big|\,\mathcal{F}_{\tau_x^+}\right)\right).
\end{aligned}$$

Now, conditionally on $\mathcal{F}_{\tau_x^+}$ and on the event $\{\tau_x^+ < \mathbf{e}_q\}$ the random variables $\overline{X}_{\mathbf{e}_q} - X_{\tau_x^+}$ and $\overline{X}_{\mathbf{e}_q}$ have the same distribution thanks to the lack of memory property of $\mathbf{e}_q$ and the strong Markov property. Hence, we have the factorization

$$\mathbb{E}\left(e^{-\beta\overline{X}_{\mathbf{e}_q}}\mathbf{1}_{(\overline{X}_{\mathbf{e}_q}>x)}\right) = \mathbb{E}\left(e^{-q\tau_x^+-\beta X_{\tau_x^+}}\right)\mathbb{E}\left(e^{-\beta\overline{X}_{\mathbf{e}_q}}\right).$$

The case that β or x are equal to zero can be achieved by taking limits on both sides of the above equality. ■

Proof of Theorem 2.1. First note that the assumption (2.2) is trivially satisfied. In view of the remarks following Lemma 2.1 let us define the bounded functions

$$v_y(x) = \mathbb{E}_x\left(\mathrm{e}^{-q\tau_y^-}(K - \mathrm{e}^{X_{\tau_y^-}})^+\right). \tag{2.5}$$

We shall show that the solution to (2.3) is of the form (2.5) for a suitable choice of $y \le \log K$ by using Lemma 2.1.

By replacing X by $-X$ in the previous lemma, we get the following analogous result for first passage into the negative half line.

$$\mathbb{E}_x\left(\mathrm{e}^{-\alpha\tau_y^-+\beta X_{\tau_y^-}}\mathbf{1}_{(\tau_y^-<\infty)}\right) = \mathrm{e}^{\beta x}\frac{\mathbb{E}(\mathrm{e}^{\beta\underline{X}_{\mathbf{e}_\alpha}}\mathbf{1}_{(-\underline{X}_{\mathbf{e}_\alpha}>x-y)})}{\mathbb{E}(\mathrm{e}^{\beta\underline{X}_{\mathbf{e}_\alpha}})} \tag{2.6}$$

for $\alpha, \beta \ge 0$ and $x - y \ge 0$ and hence it follows that

$$v_y(x) = \frac{\mathbb{E}\left((K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}) - \mathrm{e}^{x+\underline{X}_{\mathbf{e}_q}})\mathbf{1}_{(-\underline{X}_{\mathbf{e}_q}>x-y)}\right)}{\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}})}. \tag{2.7}$$

Lower bound (i). The lower bound $v_y(x) \ge (K - \mathrm{e}^x)^+$ is respected if and only if $v_y(x) \ge 0$ and $v_y(x) \ge (K - \mathrm{e}^x)$. From (2.5) we see that $v_y(x) \ge 0$ always holds. On the other hand, a straightforward manipulation shows that

$$v_y(x) = (K - \mathrm{e}^x) + \frac{\mathbb{E}\left((\mathrm{e}^{x+\underline{X}_{\mathbf{e}_q}} - K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}))\mathbf{1}_{(-\underline{X}_{\mathbf{e}_q}\le x-y)}\right)}{\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}})}. \tag{2.8}$$

From (2.8) we see that a sufficient condition that $v_y(x) \ge (K - \mathrm{e}^x)$ is that

$$\mathrm{e}^y \ge K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}). \tag{2.9}$$

Supermartingale property (ii). On the event $\{t < \mathbf{e}_q\}$ the identity $\underline{X}_{\mathbf{e}_q} = \underline{X}_t \wedge (X_t + I)$ holds where conditionally on $\mathcal{F}_t$, I has the same distribution as $\underline{X}_{\mathbf{e}_q}$. In particular it follows that on $\{t < \mathbf{e}_q\}$, $\underline{X}_{\mathbf{e}_q} \leq X_t + I$. If

$$\mathrm{e}^y \leq K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}) \tag{2.10}$$

then for $x \in \mathbb{R}$

$$\begin{aligned} v_y(x) &\geq \frac{\mathbb{E}\left(\mathbf{1}_{(t<\mathbf{e}_q)}\mathbb{E}\left((K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}) - \mathrm{e}^{x+X_t+I})\mathbf{1}_{(-(X_t+I)>x-y)}\middle|\,\mathcal{F}_t\right)\right)}{\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}})} \\ &\geq \mathbb{E}\left(\mathrm{e}^{-qt}v_y(x+X_t)\right) \\ &= \mathbb{E}_x\left(\mathrm{e}^{-qt}v_y(X_t)\right). \end{aligned}$$

Note that what we have thrown away from the expectation in the first line is positive on account of (2.10) thereby justifying the inequality. Stationary independent increments now imply that for $0 \leq s \leq t < \infty$,

$$\mathbb{E}(\mathrm{e}^{-rt}v_y(X_t)|\mathcal{F}_s) = \mathrm{e}^{-rs}\mathbb{E}_{X_s}(\mathrm{e}^{-r(t-s)}v_y(X_{t-s})) \leq \mathrm{e}^{-rs}v_y(X_s) \tag{2.11}$$

showing that $\{\mathrm{e}^{-qt}v_y(X_t) : t \geq 0\}$ is a $\mathbb{P}_x$-supermartingale. Right continuity of its paths follow from the right continuity of the paths of X and right continuity of v_y which can be seen from (2.8).

To conclude, we see then that it would be sufficient to take $y = \log K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}})$ in order to satisfy conditions (i) and (ii) of Lemma 2.1 and establish a solution to (2.3). ■

2.3 Smooth Fit versus Continuous Fit

It is clear that the solution to (2.3) is lower bounded by the gain function G and further is equal to the gain function on the domain on which the distribution of X_{τ^*} is concentrated. It turns out that there are different ways in which the function v "fits" on to the gain function G according to certain path properties of the underlying Lévy process. The McKean optimal stopping problem provides a good example of where a dichotomy appears in this respect. We say that there is *continuous fit* at x^* if the left and right limit points of v at x^* exist and are equal. In addition, if the left and right derivatives of v exist at the boundary x^* and are equal then we say that there is *smooth fit* at x^*. The remainder of this section is devoted to explaining the dichotomy of smooth and continuous fit in (2.3).

Consider again the McKean optimal stopping problem. The following Theorem is again taken from Alili and Kyprianou (2005).

Theorem 2.2 *The function $v(\log y)$ is convex in $y > 0$ and in particular there is continuous fit of v at x^*. The right derivative at x^* is given by $v'(x^*+) = -\mathrm{e}^{x^*} + K\mathbb{P}(\underline{X}_{\mathbf{e}_q} = 0)$.*

Proof. Note that for a fixed stopping time $\tau \in \mathcal{T}$ the expression $\mathbb{E}(\mathrm{e}^{-q\tau}(K - \mathrm{e}^{x+X_\tau})^+)$ is convex in e^x as the same is true of the function $(K - c\mathrm{e}^x)^+$ where $c > 0$ is a constant. Further, since taking the supremum is a subadditive operation, it can easily be deduced that $v(\log y)$ is a convex function in y. In particular v is continuous.

Next we establish necessary and sufficient conditions for smooth fit. Since $v(x) = K - \mathrm{e}^x$ for all $x < x^*$, and hence $v'(x^*-) = -\mathrm{e}^{x^*}$, we are required to show that $v'(x^*+) = -\mathrm{e}^{x^*}$ for smooth fit. Starting from (2.7) and recalling that $\mathrm{e}^{x^*} = K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}})$, we have

$$\begin{aligned} v(x) &= -K\mathbb{E}\left((\mathrm{e}^{x-x^*+\underline{X}_{\mathbf{e}_q}} - 1)\mathbf{1}_{(-\underline{X}_{\mathbf{e}_q} > x-x^*)}\right) \\ &= -K(\mathrm{e}^{x-x^*} - 1)\mathbb{E}\left(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}\mathbf{1}_{(-\underline{X}_{\mathbf{e}_q} > x-x^*)}\right) \\ &\quad -K\mathbb{E}\left((\mathrm{e}^{\underline{X}_{\mathbf{e}_q}} - 1)\mathbf{1}_{(-\underline{X}_{\mathbf{e}_q} > x-x^*)}\right). \end{aligned}$$

From the last equality we may then write

$$\begin{aligned} \frac{v(x) - (K - \mathrm{e}^{x^*})}{x - x^*} &= \frac{v(x) + K(\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}) - 1)}{x - x^*} \\ &= -K\frac{(\mathrm{e}^{x-x^*} - 1)}{x - x^*}\mathbb{E}\left(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}\mathbf{1}_{(-\underline{X}_{\mathbf{e}_q} > x-x^*)}\right) \\ &\quad +K\frac{\mathbb{E}\left((\mathrm{e}^{\underline{X}_{\mathbf{e}_q}} - 1)\mathbf{1}_{(-\underline{X}_{\mathbf{e}_q} \le x-x^*)}\right)}{x - x^*}. \end{aligned}$$

To simplify notations let us call A_x and B_x the last two terms, respectively. It is clear that

$$\lim_{x \downarrow x^*} A_x = -K\mathbb{E}\left(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}\mathbf{1}_{(-\underline{X}_{\mathbf{e}_q} > 0)}\right). \tag{2.12}$$

On the other hand, we have that

$$\begin{aligned} B_x &= K\frac{\mathbb{E}\left((\mathrm{e}^{\underline{X}_{\mathbf{e}_q}} - 1)\mathbf{1}_{(0 < -\underline{X}_{\mathbf{e}_q} \le x-x^*)}\right)}{x - x^*} \\ &= K\int_{(0,x-x^*]} \frac{\mathrm{e}^{-z} - 1}{x - x^*}\mathbb{P}(-\underline{X}_{\mathbf{e}_q} \in \mathrm{d}z) \\ &= K\frac{\mathrm{e}^{x^*-x} - 1}{x - x^*}\mathbb{P}(0 < -\underline{X}_{\mathbf{e}_q} \le x - x^*) \\ &\quad + \frac{K}{x - x^*}\int_0^{x-x^*} \mathrm{e}^{-z}\mathbb{P}(0 < -\underline{X}_{\mathbf{e}_q} \le z)\mathrm{d}z, \end{aligned}$$

where in the first equality we have removed the possible atom at zero from the expectation by noting that $\exp\{\underline{X}_{\mathbf{e}_q}\} - 1 = 0$ on $\{\underline{X}_{\mathbf{e}_q} = 0\}$. This leads to $\lim_{x\downarrow x^*} B_x = 0$. Using the expression for e^{x^*} we see that $v'(x^*+) = -\mathrm{e}^{x^*} +$

$K\mathbb{P}(-\underline{X}_{\mathbf{e}_q} = 0)$ which equals $-\mathrm{e}^{x^*}$ if and only if $\mathbb{P}(-\underline{X}_{\mathbf{e}_q} = 0) = 0$; in other words, if and only if 0 is regular for $(-\infty, 0)$. ■

Let us now discuss intuitively the dichotomy of continuous and smooth fit as a mathematical principle. The first thing we need to address is whether the atom in the distribution of $-\underline{X}_{\mathbf{e}_q}$ ever occurs (otherwise the previous theorem would say there can only ever be smooth pasting). It is quite easy to construct examples for which the atom is present. Indeed, only only needs to consider Lévy processes of the form

$$X_t = ct + \sum_{i=1}^{N_t} \xi_i$$

where $c > 0$, $\{N_t : t \geq 0\}$ is a Poisson process and $\{\xi_i : i \geq 1\}$ are independent and identically distributed random variables. For such processes, if we denote by T_1 the time of the first jump, then we can even lower estimate

$$\mathbb{P}(-\underline{X}_{\mathbf{e}_q} = 0) \geq \mathbb{P}(T_1 > \mathbf{e}_q) > 0.$$

Compound Poisson processes with positive drift are not the only examples of Lévy processes which have this property. A full examination of such Lévy processes is beyond this text, however we note that they are more generally known as Lévy processes for which 0 is irregular for $(-\infty, 0)$ or just *irregular* for short. (A process which is not irregular will be called *regular*). Moreover, the property of irregularity of 0 for $(-\infty, 0)$ is equivalent to the property that

$$\mathbb{P}(\tau_0^- > 0) = 1.$$

(It is important to note that the only alternative value this probability can take is 0 which is a consequence of Blumenthal's $0-1$ law for Markov processes).

Reconsidering the general expression (2.5) for the candidate value function v_y where $y \leq \log K$ (before optimizing the value of y) we see the relevance of irregularity. Indeed, when $x < y$ we have that $\mathbb{P}_x(\tau_y^- = 0) = 1$ and $v_y(x) = (K - e^x)^+$ showing that $v_y(y-) = (K - e^y)$. On the other hand, if $x \geq y$

$$v_y(y+) = \mathbb{E}_y\left(\mathrm{e}^{-q\tau_y^-}(K - \mathrm{e}^{X_{\tau_y^-}})^+\right) = \mathbb{E}\left(\mathrm{e}^{-q\tau_0^-}(K - \mathrm{e}^{y+X_{\tau_0^-}})^+\right).$$

The right hand side is equal to $(K - e^y)$ when there is regularity. However, when there is irregularity, the expectation on the right hand side may be greater or less than $(K - e^y)$ thereby destroying the property of continuity of the function v_y. This is also clear from (2.8) when we think of irregularity in terms of the possible atom at zero in the distribution of $-\underline{X}_{\mathbf{e}_q}$. Indeed note that

$$v_y(y+) = (K - \mathrm{e}^y) + \frac{(\mathrm{e}^y - K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}}))}{\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}})}\mathbb{P}(-\underline{X}_{\mathbf{e}_q} = 0)$$

showing there is a positive discontinuity at y when $e^y > K\mathbb{E}(\mathrm{e}^{\underline{X}_{\mathbf{e}_q}})$ and a negative discontinuity when this strict inequality is reversed.

Figs. 1 and 2 sketch what one can expect to see in the shape of v_y by perturbing the value y about x^* for the cases of regular and irregular Lévy processes. With these diagrams in mind we may now intuitively understand the appearance of smooth or continuous fit as a principle via the following reasoning.

For the case 0 **is irregular for** $(-\infty, 0)$ **for** X**.** In general v_y has a discontinuity at y. When $y < x^*$, thanks to (2.8) we know the function v_y does not upper bound the gain function $(K - e^x)^+$ due to a negative discontinuity at y and hence τ_y^- is not an admissible strategy in this regime of y. On the other hand, from (2.8) and (2.9) if $y \geq x^*$, v_y upper bounds the gain function. Again from (2.8) we see that there is a discontinuity in v_y at y when $y > x^*$ and continuity when $y = x^*$. By bringing y down to x^* it turns out that the function v_y is pointwise optimised. Here then we experience *a principle of continuous fit* and from Theorem 2.2 it transpires there is no smooth fit.

For the case 0 **is regular for** $(-\infty, 0)$ **for** X**.** All curves v_y are continuous. There is in general however a discontinuity in the first derivative of v_y at the point y. When $y < x^*$ the function v_y cannot upper bound the gain function $(K - e^y)^+$ as $v_y'(y+) < v_y'(y-)$ and hence τ_y^- is not an admissible strategy in this regime of y. As before, if $y \geq x^*$, v_y upper bounds the gain function. There is a discontinuity in v_y' at y if $y > x^*$ and otherwise it is smooth when $y = x^*$. It turns out this time that by bringing y down to x^* the gradient $v_y'(y+)$ becomes equal to $v_y'(y-)$ and the function v_y is pointwise optimised. We experience then in this case *a principle of smooth fit* instead.

Whilst the understanding that smooth fit appears in the solutions of optimal stopping problems as a principle dates back to [10], the idea that continuous fit appears in certain classes of optimal stopping problems *as a principle* first appeared for the first time only recently in the work of [13, 14]).

2.4 An interesting peculiarity

Suppose that X is a compound Poisson process with two sided jumps which have no atoms (this excludes the possibility that X can jump exactly onto a prescribed point). For this class of processes we note that

$$\inf\{t > 0 : X_t < y\} = \inf\{t \geq 0 : X_t \leq y\}$$

$\mathbb{P}_x$- almost surely unless $y = x$. In that case, the stopping time on the left is strictly positive $\mathbb{P}_x$-almost surely and the stopping time on the right is zero $\mathbb{P}_x$-almost surely. In other words, for all $x, y \in \mathbb{R}$, the optimal stopping time on the left is $\mathbb{P}_x$-almost surely greater than the optimal stopping time on the right. Suppose we re-define $\tau_y^- = \inf\{t \geq 0 : X_t \leq y\}$ and take $\tau_{x^*}^-$ (under this new definition) as the candidate optimal stopping time to the McKean optimal stopping problem instead of the one given in Theorem 2.1. Revisiting

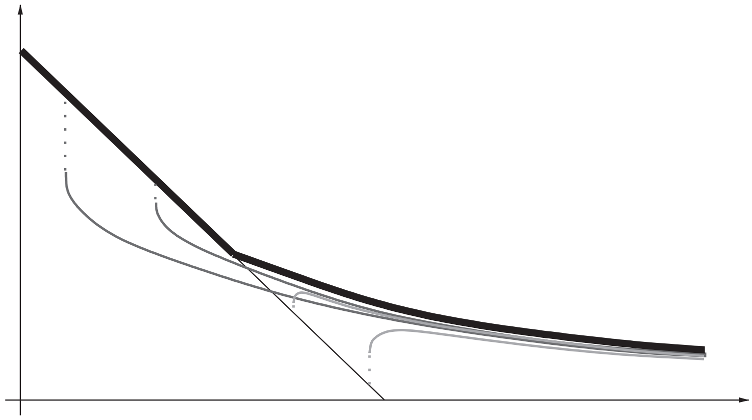

Figure 1: A sketch of the functions $v_y(\log x)$ for different values of y when X is irregular. Curves which do not upper bound the function $(K-x)^+$ correspond to examples of $v_y(\log x)$ with $y < x^*$. Curves which are lower bounded by $(K-x)^+$ correspond to examples of $v_y(\log x)$ with $y > x^*$. The unique curve which upper bounds the gain with continuous fit corresponds to $v_y(\log x)$ with $y = x^*$.

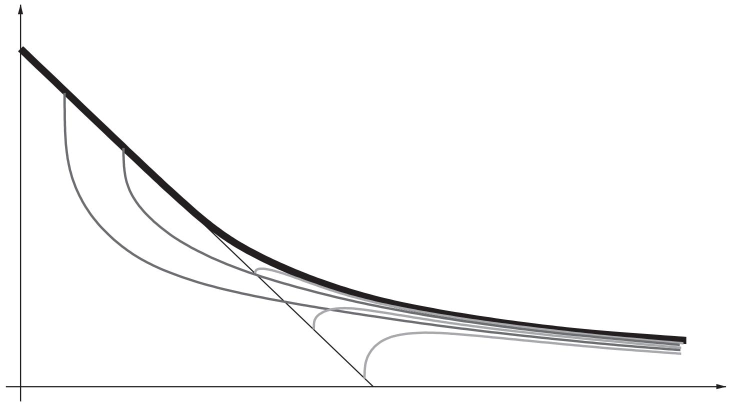

Figure 2: A sketch of the functions $v_y(\log x)$ for different values of y when X is regular. Curves which do not upper bound the function $(K-x)^+$ correspond to examples of $v_y(\log x)$ with $y < x^*$. Curves which are lower bounded by $(K-x)^+$ correspond to examples of $v_y(\log x)$ with $y > x^*$. The unique curve which upper bounds the gain with smooth fit corresponds to $v_y(\log x)$ with $y = x^*$.

the proof of Theorem 2.1 one easily finds that the value function is the same with this new stopping time. In showing this, one needs to start by making the strict inequality in (2.6) a weak inequality and working the consequence of this change through the computations. One may also perform the analysis concerning continuous fit with the re-defined stopping times τ_y^- and find that in general there is a discontinuity in v_y, however, unlike before the function v_y is now left continuous with a right limit at y instead of right continuous with a left limit at y. None the less when $y = x^*$, with either definition of $\tau_{x^*}^-$, the value function emerges as the same.

We therefore see that, although there is a unique value for the solution to the optimal stopping problem. The optimal strategy is not necessarily unique. Indeed we have found, at least in the compound Poisson case, that there is another optimal stopping time which is almost surely smaller than the optimal stopping found in the proof of Theorem 2.1.

2.5 The Novikov–Shiryaev Optimal Stopping Problem

The following family of optimal stopping problems was recently solved by [12] in an analogous random walk setting. Consider

$$v_n(x) = \sup_{\tau \in \mathcal{T}} \mathbb{E}_x(\mathrm{e}^{-q\tau}(X_\tau^+)^n), \tag{2.13}$$

where $\mathcal{T}$ is the set of $\mathbb{F}$-stopping times and it is assumed that X is any Lévy process, $q > 0$ and we may choose n to be any strictly positive integer. The solution we shall give is based on the arguments of [12][5]. We first need to introduce a special class of polynomials based on cumulants of specified random variables.

Recall that if a non-negative random variable Y has characteristic function $\phi(\theta) = \mathbb{E}(\mathrm{e}^{i\theta Y})$ then its cumulant generating function is defined by $\log\phi(\theta)$. If Y has up to n moments then it is possible to make a Taylor expansion of the cumulant generating function up to order n plus an error term. In that case, the coefficients $\{\kappa_1, ..., \kappa_n\}$ are called the first n *cumulants*. If the first n cumulants are finite, then they may be written it terms of the first n moments. For example,

$$\begin{aligned}
\kappa_1 &= \mu_1, \\
\kappa_2 &= \mu_2 - \mu_1^2, \\
\kappa_3 &= 2\mu_1^3 - 3\mu_1\mu_2 + \mu_3, \\
& \quad ...
\end{aligned}$$

where $\mu_1, \mu_2, ...$ are the first, second, third, etc. moments of Y.

For a concise overview of cumulant generating functions, the reader is referred to [8].

[5]The continuous time arguments are also given in Kyprianou and Surya ([7])

Definition 2.2 (Appell Polynomials) *Suppose that Y is a non-negative random variable with nth cumulant given by κ_n for $n = 1, 2, ...$ Then define the Appell polynomials iteratively as follows. Take $Q_0(x) = 1$ and assuming that $|\kappa_n| < \infty$ (equivalently, Y has an nth moment) given $Q_{n-1}(x)$ we define $Q_n(x)$ via*

$$\frac{\mathrm{d}}{\mathrm{d}x} Q_n(x) = n Q_{n-1}(x). \tag{2.14}$$

This defines Q_n up to a constant. To pin this constant down we insist that $\mathbb{E}(Q_n(Y)) = 0$. The first three Appell polynomials are given for example by

$$Q_0(x) = 1,\ Q_1(x) = x - \kappa_1,\ Q_2(x) = (x - \kappa_1)^2 - \kappa_2,$$

$$Q_3(x) = (x - \kappa_1)^3 - 3\kappa_2(x - \kappa_1) - \kappa_3,$$

under the assumption that $\kappa_3 < \infty$. See also Schoutens (2000) for further details of Appell polynomials.

In the following theorem, we shall work with the Appell polynomials generated by the random variable $Y = \overline{X}_{\mathbf{e}_q}$ where as usual, for each $t \in [0, \infty)$, $\overline{X}_t = \sup_{s \in [0,t]} X_s$ and $\mathbf{e}_q$ is an exponentially distributed random variable which is independent of X. As a condition to the next and several subsequent results we will see the condition $\mathbb{E}((X_1^+)^n) < \infty$. This turns out to be a sufficient condition under which $\mathbb{E}(\overline{X}_{\mathbf{e}_q}^n) < \infty$, however we will not give a proof of this fact.

Theorem 2.3 *Fix $n \in \{1, 2, ...\}$. Assume that $\mathbb{E}((X_1^+)^n) < \infty$. Then $Q_n(x)$ has finite coefficients and there exists $x_n^* \in [0, \infty)$ being the largest root of the equation $Q_n(x) = 0$. Let*

$$\tau_n^* = \inf\{t \geq 0 : X_t \geq x_n^*\}.$$

Then τ_n^* is an optimal strategy to (2.13). Further,

$$v_n(x) = \mathbb{E}_x(Q_n(\overline{X}_{\mathbf{e}_q}) \mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq x_n^*)}).$$

Similarly to the McKean optimal stopping problem, we can establish a necessary and sufficient criterion for the occurrence of smooth fit. Once again, it boils down to the underlying path regularity.

Theorem 2.4 *For each $n = 1, 2, ...$ the solution to the optimal stopping problem in Theorem 9.6 is convex, in particular exhibiting continuous fit at x_n^*, and*

$$v_n'(x_n^*-) = v_n'(x_n^*+) - Q_n'(x_n^*)\mathbb{P}(\overline{X}_{\mathbf{e}_q} = 0).$$

Hence there is smooth fit at x_n^* if and only if 0 is regular for $(0, \infty)$ for X.

The proofs of the last two theorems require some preliminary results given in the following lemmas.

Lemma 2.3 (Mean value property) *Fix $n \in \{1, 2, ...\}$ Suppose that Y is a non-negative random variable satisfying $\mathbb{E}(Y^n) < \infty$. Then if Q_n is the nth Appell polynomial generated by Y, we have that*

$$\mathbb{E}(Q_n(x + Y)) = x^n$$

for all $x \in \mathbb{R}$.

Proof. Note the result is trivially true for $n = 1$. Next suppose the result is true for Q_{n-1}. Then using dominated convergence we have from (2.14) that

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathbb{E}(Q_n(x + Y)) = \mathbb{E}\left(\frac{\mathrm{d}}{\mathrm{d}x}Q_n(x + Y)\right) = n\mathbb{E}(Q_{n-1}(x + Y)) = nx^{n-1}.$$

Solving together with the requirement that $\mathbb{E}(Q_n(Y)) = 0$ we have the required result. ■

Lemma 2.4 (Fluctuation identity) *Fix $n \in \{1, 2, ...\}$ and suppose that $\mathbb{E}((X_1^+)^n) < \infty$. Then for all $a > 0$ and $x \in \mathbb{R}$,*

$$\mathbb{E}_x(\mathrm{e}^{-qT_a^+} X^n_{T_a^+} \mathbf{1}_{(T_a^+ < \infty)}) = \mathbb{E}_x(Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq a)}),$$

where $T_a^+ = \inf\{t \geq 0 : X_t \geq a\}$.

Proof. On the event $\{T_a^+ < \mathbf{e}_q\}$ we have that $\overline{X}_{\mathbf{e}_q} = X_{T_a^+} + S$ where S is independent of $\mathcal{F}_{T_a^+}$ and has the same distribution as $\overline{X}_{\mathbf{e}_q}$. It follows that

$$\mathbb{E}_x(Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq a)} | \mathcal{F}_{T_a^+}) = \mathbf{1}_{(T_a^+ < \mathbf{e}_q)} h(X_{T_a^+}),$$

where $h(x) = \mathbb{E}_x(Q_n(\overline{X}_{\mathbf{e}_q})) = x^n$ and the last equality follows from Lemma 2.3 with $Y = \overline{X}_{\mathbf{e}_q}$. We see, by taking expectations again in the previous calculation, that

$$\mathbb{E}_x(Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq a)}) = \mathbb{E}_x(e^{-qT_a^+} X^n_{T_a^+} \mathbf{1}_{(T_a^+ < \infty)})$$

as required. ■

Lemma 2.5 (Largest positive root) *Fix $n \in \{1, 2, ...\}$. Suppose that $\mathbb{E}((X_1^+)^n) < \infty$ and that Q_n is generated by $\overline{X}_{\mathbf{e}_q}$. Then Q_n has a unique strictly positive root x_n^* such that $Q_n(x)$ is negative on $[0, x_n^*)$ and positive and increasing on $[x_n^*, \infty)$.*

Proof. First note that the statement of the lemma is clearly true for $Q_1(x) = x - \kappa_1$. We proceed then by induction and assume that the result is true for Q_{n-1}.

The first step is to prove that $Q_n(0) \leq 0$. Let

$$\eta(a, n) = \mathbb{E}(e^{-qT_a^+} X^n_{T_a^+} \mathbf{1}_{(T_a^+ < \infty)}),$$

where $T_a^+ = \inf\{t \geq 0 : X_t \geq a\}$ and note that $\eta(a,n) \geq 0$ for all $a \geq 0$ and $n = 1, 2, ...$ On the other hand

$$\begin{aligned}
\eta(a,n) &= \mathbb{E}(Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq a)}) \\
&= -\mathbb{E}(Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} < a)}) \\
&= -\mathbb{P}(\overline{X}_{\mathbf{e}_q} < a)Q_n(0) \\
&\quad + \mathbb{E}((Q_n(0) - Q_n(\overline{X}_{\mathbf{e}_q}))\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} < a)}),
\end{aligned}$$

where the first equality follows from Lemma 2.4 and the second by Lemma 2.3. Since by definition

$$Q_n(x) = Q_n(0) + n\int_0^x Q_{n-1}(u)\mathrm{d}y \tag{2.15}$$

for all $x \geq 0$ we have the estimate

$$\left|\mathbb{E}_x((Q_n(0) - Q_n(\overline{X}_{\mathbf{e}_q}))\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} < a)})\right| \leq na \sup_{y\in[0,a]} |Q_{n-1}(y)|\mathbb{P}(\overline{X}_{\mathbf{e}_q} < a)$$

which tends to zero as $a \downarrow 0$. We have in conclusion that

$$0 \leq \eta(a,n) \leq -\mathbb{P}(\overline{X}_{\mathbf{e}_q} < a)[Q_n(0) + 0(a)]$$

and hence as a may be made arbitrarily small it necessarily follows that $Q_n(0) \leq 0$.

Under the induction hypothesis for Q_{n-1}, we see from (2.15), together with the fact that $Q_n(0) \leq 0$, that Q_n is negative and decreasing on $[0, x^*_{n-1})$. The point x^*_{n-1} corresponds to the minimum of Q_n thanks to the positivity and monotonicity of $Q_{n-1}(u)$ for $x > x^*_{n-1}$. In particular, $Q_n(x)$ tends to infinity from its minimum point and hence there must be a unique strictly positive root of the equation $Q_n(x) = 0$. ■

We are now ready to move to the proofs of the main theorems of this section.

Proof of Theorem 2.3. Fix $n \in \{1, 2, ...\}$. Thanks to the assumption $\mathbb{E}((X_1^+)^n) < \infty$ we have $\mathbb{E}(X_1) \in [-\infty, \infty)$ and hence the Strong Law of Large Numbers for Lévy processes[6] implies that (2.2) is automatically satisfied since $q > 0$. Indeed, $(X_t^+)^n$ grows no faster than Ct^n for some constant $C > 0$.

[6] The version of Strong Law of Large Numbers for Lévy processes that we need says that whenever $\mathbb{E}(X_1)$ is defined in $[-\infty, \infty]$ then $\lim_{t\uparrow\infty} X_t/t = \mathbb{E}(X_1)$.

Define

$$v_n^a(x) = \mathbb{E}_x(e^{-qT_a^+}(X_{T_a^+}^+)^n \mathbf{1}_{(T_a^+<\infty)}), \tag{2.16}$$

where as usual $T_a^+ = \inf\{t \geq 0 : X_t \geq a\}$. From Lemma 2.4 we know that

$$v_n^a(x) = \mathbb{E}_x(Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq a)}).$$

Again referring to the discussion following Lemma 2.1 we consider pairs (v_n^a, T_a^+) for $a > 0$ to be a class of candidate solutions to (2.13). Our goal then is to verify with the help of Lemma 2.1 that the candidate pair (v_n^a, T_a^+) solve (2.13) for some $a > 0$.

Lower bound (i). We need to prove that $v_n^a(x) \geq (x^+)^n$ for all $x \in \mathbb{R}$. Note that this statement is obvious for $x \in (-\infty, 0) \cup (a, \infty)$ just from the definition of v_n^a. Otherwise when $x \in (0, a)$ we have, using the mean value property in Lemma 2.3, that for all $x \in \mathbb{R}$,

$$\begin{aligned} v_n^a(x) &= \mathbb{E}_x(Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq a)}) \\ &= x^n - \mathbb{E}(Q_n(x + \overline{X}_{\mathbf{e}_q})\mathbf{1}_{(x+\overline{X}_{\mathbf{e}_q}<a)}). \end{aligned} \tag{2.17}$$

From Lemma 2.5 and specifically the fact that $Q_n(x) \leq 0$ on $(0, x_n^*]$ it follows that, provided

$$a \leq x_n^*,$$

we have in (2.17) that $v_n^a(x) \geq (x^+)^n$.

Supermartingale property (ii). Provided

$$a \geq x_n^*$$

we have almost surely that

$$Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq a)} \geq 0.$$

On the event that $\{\mathbf{e}_q > t\}$ we have $\overline{X}_{\mathbf{e}_q}$ is equal in distribution to $(X_t + S) \vee \overline{X}_t$ where S is independent of $\mathcal{F}_t$ and equal in distribution to $\overline{X}_{\mathbf{e}_q}$. In particular $\overline{X}_{\mathbf{e}_q} \geq X_t + S$. It now follows that

$$\begin{aligned} v_n^a(x) &\geq \mathbb{E}_x(\mathbf{1}_{(\mathbf{e}_q>t)}Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} \geq a)}) \\ &\geq \mathbb{E}_x(\mathbf{1}_{(\mathbf{e}_q>t)}\mathbb{E}_x(Q_n(X_t + S)\mathbf{1}_{(X_t+S\geq a)}|\mathcal{F}_t)) \\ &= \mathbb{E}_x(\mathrm{e}^{-qt}v_n^a(X_t)). \end{aligned}$$

From this inequality together with the Markov property, it is easily shown as in the McKean optimal stopping problem that $\{\mathrm{e}^{-qt}v_n^a(X_t) : t \geq 0\}$ is a supermartingale. Right continuity follows again from the right continuity of the paths of X together with the right continuity of v_n^a which is evident from (2.17).

We now see that the unique choice $a = x_n^*$ allows all the conditions of Lemma 2.1 to be satisfied thus giving the solution to (2.13). ■

Proof of Theorem 2.4. In a similar manner to the proof of Theorem 2.2 it is straighforward to prove that v is convex and hence continuous.

To establish when there is a smooth fit at x_n^* we compute as follows. For $x < x_n^*$,

$$\begin{aligned}\frac{v_n(x_n^*) - v_n(x^n)}{x_n^* - x} &= \frac{(x_n^*)^n - x^n}{x_n^* - x} + \frac{\mathbb{E}_x(Q_n(\overline{X}_{\mathbf{e}_q})\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} < x_n^*)})}{x_n^* - x} \\ &= \frac{(x_n^*)^n - x^n}{x_n^* - x} + \frac{\mathbb{E}_x((Q_n(\overline{X}_{\mathbf{e}_q}) - Q_n(x_n^*))\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} < x_n^*)})}{x_n^* - x},\end{aligned}$$

where the final equality follows because $Q_n(x_n^*) = 0$. Clearly

$$\lim_{x \uparrow x_n^*} \frac{(x_n^*)^n - x^n}{x_n^* - x} = v_n'(x_n^*+).$$

However,

$$\begin{aligned}&\frac{\mathbb{E}_x((Q_n(\overline{X}_{\mathbf{e}_q}) - Q_n(x_n^*))\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} < x_n^*)})}{x_n^* - x} \\ &= \frac{\mathbb{E}_x((Q_n(\overline{X}_{\mathbf{e}_q}) - Q_n(x))\mathbf{1}_{(x < \overline{X}_{\mathbf{e}_q} < x_n^*)})}{x_n^* - x} \\ &\quad - \frac{\mathbb{E}_x((Q_n(x_n^*) - Q_n(x))\mathbf{1}_{(\overline{X}_{\mathbf{e}_q} < x_n^*)})}{x_n^* - x} \qquad (2.18)\end{aligned}$$

where in the first term on the right-hand side we may restrict the expectation to $\{x < \overline{X}_{\mathbf{e}_q} < x_n^*\}$ as, under $\mathbb{P}_x$, the possible atom of $\overline{X}_{\mathbf{e}_q}$ at x gives zero mass to the expectation. Denote by A_x and B_x the two expressions on the right hand side of (2.18). We have that

$$\lim_{x \uparrow x_n^*} B_x = -Q_n'(x_n^*)\mathbb{P}(\overline{X}_{\mathbf{e}_q} = 0).$$

Integration by parts also gives

$$\begin{aligned}A_x &= \int_{(0, x_n^* - x)} \frac{Q_n(x+y) - Q_n(x)}{x_n^* - x}\mathbb{P}(\overline{X}_{\mathbf{e}_q} \in \mathrm{d}y) \\ &= \frac{Q_n(x_n^*) - Q_n(x)}{x_n^* - x}\mathbb{P}(\overline{X}_{\mathbf{e}_q} \in (0, x_n^* - x)) \\ &\quad - \frac{1}{x_n^* - x}\int_0^{x_n^* - x} \mathbb{P}(\overline{X}_{\mathbf{e}_q} \in (0, y])Q_n'(x+y)\mathrm{d}y.\end{aligned}$$

Hence it follows that

$$\lim_{x \uparrow x_n^*} A_x = 0.$$

In conclusion we have that

$$\lim_{x \uparrow x_n^*} \frac{v_n(x_n^*) - v_n(x)}{x_n^* - x} = v_n'(x_n^*+) - Q_n'(x_n^*)\mathbb{P}(\overline{X}_{\mathbf{e}_q} = 0)$$

which concludes the proof. ■

References

[1] Asmussen, S., Avram, F. and Pistorius, M. (2004) Russian and American put options under exponential phase-type Lévy models. *Stochast. Process. Appl.* **109**, 79–111.

[2] Avram, F., Chan, T. and Usabel, M. (2002a) On the valuation of constant barrier options under spectrally one-sided exponential Lévy models and Carr's approximation for American puts. *Stochast. Process. Appl.* **100**, 75–107.

[3] Chan, T. (2004) Some applications of Lévy processes in insurance and finance. *Finance* **25**, 71–94.

[4] Chesney, M. and Jeanblanc, M. (2004) Pricing American currency options in an exponential Lévy model. *Appl. Math. Finance* **11**, 207–225.

[5] Darling, D.A., Liggett, T. and Taylor, H.M. (1972) Optimal stopping for partial sums. *Ann. Math. Stat.* **43**, 1363–1368.

[6] Gerber, H.U. and Shiu, E.S.W. (1994) Martingale approach to pricing perpetual American options. *Astin Bull.* **24**, 195–220.

[7] Kyprianou, A.E. and Palmowski, Z. (2005) A martingale review of some fluctuation theory for spectrally negative Lévy processes. *Séminaire de Probabilités XXXVIII.* 16–29. Lecture Notes in Mathematics, Springer, Berlin Heidelberg New York.

[8] Lukacs, E. (1970) *Characteristic Functions.* Second edition, revised and enlarged. Hafner Publishing Co., New York.

[9] McKean, H. (1965) Appendix: A free boundary problem for the heat equation arising from a problem of mathematical economics. *Ind. Manag.Rev.* **6**, 32–39.

[10] Mikhalevich, V.S. (1958) Baysian choice between two hypotheses for the mean value of a normal process. *Visnik Kiiv. Univ.* **1**, 101–104.

[11] Mordecki, E. (2002) Optimal stopping and perpetual options for Lévy processes. *Finance Stoch.* **6**, 473–493.

[12] Novikov, A.A. (2004) Martingales and first-exit times for the Ornstein–Uhlenbeck process with jumps. *Theory Probab. Appl.* **48**, 288–303.

[13] Peskir, G. and Shiryaev, A.N. (2000), Sequential testing problems for Poisson processes. *Ann. Statist.* **28**, 837–859.

[14] Peskir, G. and Shiryaev, A.N. (2002) Solving the Poisson disorder problem. *Advances in Finance and Stochastics.* pp. 295–312, Springer, Berlin Heidelberg New York.