probability p_{ij} depending only on the points $\vec{q}_i$ and $\vec{q}_j$. This continues. We see that there is an induced Markov chain on the $\vec{q}_i$, depending on which one $\vec{z}_n(t)$ is near. The interested reader may consult [FW, §6.6 p. 198].

Unfortunately, the Freidlin-Wentzell theory is technical, and even more unfortunately, some of the most important things to learn aren't the theorems, but the techniques brought forth in the proofs. We have tried to keep the conditions for the theorems strong, so as to make the proofs simpler and more easily understood. We weaken the conditions through an excessive number of exercises. We strongly recommend that, on your first perusal of this chapter, you just read Exercises 6.23, 6.34, 6.41, and 6.42, and completely ignore all other exercises. The reason for including so many technical exercises is that we need many similar results for the applications chapters, and proving these results is not difficult, once the basics have been understood. However, it is time- and page-consuming to detail the arguments.

We recommend that you begin your study by reading §6.1, including the assumptions but not the exercises, through the paragraph following the statement of Theorem 6.17. Then skip to §6.4 and read the entire section. Then go back to §6.1 and read the lemmas leading up to the proofs of Theorems 6.15 and 6.17, again skipping the exercises, and read the proofs, too. Continue to §6.2. Then, if you are motivated, try reading the entire chapter, and test your prowess on the Exercises.

6.1. The Exit Problem

The first problem we consider in this chapter is the classic exit problem. This problem was the first considered by Freidlin and Wentzell using large deviations. The question is: If a process is given sufficient time to escape from a given set G that contains an "attracting point," how long would it take, and what path would it follow on its escape route? Our setup basically follows that of the general process theory in $\mathbb{R}^d$. The process $\vec{z}_n$ we are interested in is defined by its generator

$$L_n f(\vec{x}) = \sum_{i=1}^{k} n\lambda_i(\vec{x}) \left(f\left(\vec{x} + \frac{\vec{e}_i}{n}\right) - f(\vec{x}) \right). \tag{6.3}$$

We start with a (non-empty) set G in $\mathbb{R}^d$. The set of paths that exit G at time zero is

$$S \stackrel{\triangle}{=} \{\vec{r} : \vec{r}(-T) = \vec{q},\ \vec{r}(t) \in G^o \text{ for } -T < t < 0,\ \vec{r}(0) \in \partial G,\ T > 0\}, \tag{6.4}$$

where $T = \infty$ is acceptable. Note that since none of the quantities of interest depend explicitly on time, we may, whenever convenient, shift everything in time; the only effect this may have is to slightly confuse the reader. For a set of paths S' we use the notation

$$I^*(S') \stackrel{\triangle}{=} \inf \left\{ I_{-T}^0(\vec{r}) : \vec{r} \in S' \right\}, \tag{6.5}$$

where T depends on $\vec{r}$. The following assumptions are used in order to smooth the presentation (by avoiding technicalities), and will be relaxed, mostly through exercises. Figure 6.7 pictorially shows what most of the assumptions mean.

Assumption 6.1. G is bounded in $\mathbb{R}^d$ with $G = \overline{G^o}$; that is, G consists of the closure of its interior.

Assumption 6.2. The log-jump rates $\log \lambda_i(\vec{x})$ are uniformly Lipschitz continuous on a compact set whose interior contains $\overline{G}$.

That is, the jump rates are defined, are smooth, and are bounded away from zero on a set that includes the closure of G in its interior.

Assumption 6.3. The boundary ∂G of G is smooth (Definition A.33), and at every point $\vec{x}$ in ∂G there is an open exterior cone $c(\vec{x})$ so that, for some $\varepsilon > 0$,

$$c(\vec{x}) \cap \{\vec{y} : |\vec{y} - \vec{x}| < \varepsilon\} \cap G = \emptyset.$$

That is, we can find a small neighborhood outside G so that each point in this neighborhood is connected to $\vec{x}$ by a straight line that does not re-enter G.

As in (5.7) we define a vector field

$$\vec{f}(\vec{x}) \stackrel{\triangle}{=} \sum_{i=1}^{k} \lambda_i(\vec{x})\vec{e}_i.$$

Recall that $\vec{f}(\vec{x})$ defines $\vec{z}_\infty(t)$, since $d\vec{z}_\infty(t)/dt = \vec{f}(\vec{z}_\infty(t))$. Let $\vec{n}(\vec{x})$ denote the unit outward normal to the set G, which exists since the boundary is smooth. The next assumption guarantees that whenever $\vec{z}_\infty$ starts on the boundary of G, it will immediately go into the interior; therefore, if $\vec{z}_\infty$ starts inside G, it never escapes G.

Assumption 6.4. There is a $\delta > 0$ so that, at every point $\vec{x}$ in ∂G,

$$\langle \vec{f}(\vec{x}), \vec{n}(\vec{x}) \rangle \leq -\delta < 0.$$

Assumption 6.5. There is a point $\vec{q} \in G^o$ and, for every $\varepsilon > 0$, there is a $T < \infty$ such that, uniformly over $\vec{x} \in G$, with $\vec{z}_\infty(0) = \vec{x}$,

$$|\vec{z}_\infty(t) - \vec{q}| < \varepsilon \text{ for all } t > T.$$

Assumption 6.6. There exists a unique path $\vec{r}^{\,*}(t)$ and corresponding time T^* (possibly infinite!) such that

$$I^*(S) \stackrel{\triangle}{=} \inf_{\vec{r} \in S} I^0_{-T}(\vec{r}) = I^0_{-T^*}(\vec{r}^{\,*}).$$

Remark. Note that, in our search for an optimal path, we could include in S paths that leave G before time 0, as long as they start at $\vec{q}$ and are on the boundary at $t = 0$. To see that this would change neither I^* nor $\vec{r}^{\,*}$, assume $\vec{r}^{\,*}(t_1) \in \partial G$ for some time $t_1 > -T^*$ before 0. Since ℓ is non-negative, the definition of $\vec{r}^{\,*}$ and

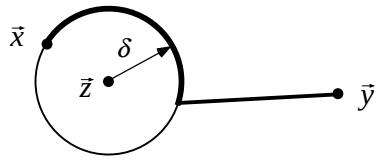

Figure 6.24. The dark line is a path satisfying (6.11).

This would be desirable when the assumptions do not hold near $\vec{z}$ (usually $\vec{q}$), for example if the rates go to zero near $\vec{z}$. Fix some $\eta > 0$, and for each $\vec{x}, \vec{y} \in N_\varepsilon(\vec{z})$, $\delta < \varepsilon < \eta$ construct a path $\vec{r}$ with $\vec{r}(0) = \vec{x},\ \vec{r}(\varepsilon) = \vec{y}$ that satisfies both (6.9) and (6.11) for a constant $C(\delta)$ independent of ε. Assume (6.8) and that $\log \lambda_i$ are bounded on

$$\left\{\vec{x} : \vec{x} \in \overline{N}_\eta(\vec{z}),\ \vec{x} \notin N_{\delta/2}(\vec{z})\right\}.$$

If $\log \lambda_i$ are bounded on $N_\eta(\vec{z})$ then the constant $C(\delta)$ does not depend on δ. On the other hand, if we only assume that

$$\sup_{\delta \le |\vec{x}-\vec{z}| \le 1} \delta(\lambda_i(\vec{x}) + |\log \lambda_i(\vec{x})|) \to 0 \quad \text{as} \quad \delta \to 0 \tag{6.12}$$

for each i, then $C(\delta)$ satisfies $\delta C(\delta) \to 0$ as $\delta \to 0$. Hint: for the last part, examine the proof of Lemma 5.32. ♠

Thus it is "cheap" for the process to move between close points. Note that in the one-dimensional case we need special arguments when some rates are zero at $\vec{q}$. Here is an important consequence.

Exercise 6.25. Assume $\log \lambda_i(\vec{x})$ are bounded and continuous. Fix a path $\vec{r}$ and define $\vec{r}_\alpha(t) \stackrel{\triangle}{=} \vec{r}(\alpha t)$: note that $\vec{r}_\alpha$ traverses the same path as $\vec{r}$, except that the time it takes to do so is different. For each T, the function $I_0^{T/\alpha}(\vec{r}_\alpha)$ is continuous in α near $\alpha = 1$. Hint: by Lemmas 5.17 and 5.32, if $|\vec{y}|$ is large then $\ell(\vec{x}, \vec{y})$ is bounded above and below by a constant times $|\vec{y}| \log |\vec{y}|$. As in the proof of Theorem 5.35 show that the set where $\vec{r}\,'$ is large can be ignored. Now invoke Exercise 5.34. ♠

The next lemma shows that the "price" of avoiding a neighborhood of the stable point $\vec{q}$ increases linearly in time. This is the main ingredient in reducing the infinite-time problem to a bounded time interval. For a given positive T and ε (d is Euclidean distance), consider the set of paths that avoid q:

$$S_T(\vec{x}) \stackrel{\triangle}{=} \left\{\vec{r} \in D^d[0,T] : \inf_{0 \le t \le T} |\vec{r}(t) - \vec{q}| > \varepsilon, \right.$$
$$\left. \vec{r}(0) = \vec{x},\ \inf_{0 \le t \le T} d(\vec{r}(t), \partial G) > 0 \right\}. \tag{6.13}$$

G^o at, say, the j^{th} attempt (i.e., at time τ_{2j}) decreases exponentially (in n) with rate roughly $I^*(S)$.

Lemma 6.32. *Under Assumptions 6.1–6.5, for each $\varepsilon > 0$ small enough there is a $T < \infty$, depending only on ε, and there are positive functions C_e and C_o such that if n is large enough then for $k = 1, 2, \ldots$*

$$\mathbb{P}_{\vec{x}}\left(\tau_{j+1} - \tau_j > kT \;\middle|\; \vec{z}_n(\tau_j) \in G^o\right) \le \begin{cases} \left(1 - e^{-nC_e(\varepsilon)}\right)^k & j \text{ even}, \\ e^{-nC_o(\varepsilon)k} & j \text{ odd}. \end{cases} \tag{6.15}$$

Moreover, $C_e(\varepsilon) = O(\varepsilon)$.

Proof. For even j, note that if $\vec{z}_n(\tau_j) \in G^o$ then necessarily $\vec{z}_n(\tau_j) \in \overline{N_{\varepsilon_s}(\vec{q})}$, so

$$\begin{aligned}
&\mathbb{P}_{\vec{x}}\left(\tau_{j+1} - \tau_j > kT \;\middle|\; \vec{z}_n(\tau_j) \in G^o\right) \\
&\qquad = \mathbb{P}_{\vec{x}}\left(\vec{z}_n(t) \in N_\varepsilon(\vec{q}),\ \tau_j < t \le \tau_j + kT \;\middle|\; \vec{z}_n(\tau_j) \in G^o\right) \\
&\qquad \le \sup_{|\vec{z}-\vec{q}|\le\varepsilon_s} \mathbb{P}_{\vec{x}}\left(\vec{z}_n(t) \in N_\varepsilon(\vec{q}),\ \tau_j < t \le \tau_j + kT \;\middle|\; \vec{z}_n(\tau_j) = \vec{z}\right) \\
&\qquad = \sup_{|\vec{z}-\vec{q}|\le\varepsilon_s} \mathbb{P}_{\vec{z}}\left(\vec{z}_n(t) \in N_\varepsilon(\vec{q}),\ 0 < t \le kT\right)
\end{aligned}$$

by the (strong) Markov property—Theorems A.137–A.138. Using the Markov property again, conditioning and then replacing the conditioning by maximization as above,

$$\begin{aligned}
&\mathbb{P}_{\vec{z}}\left(\vec{z}_n(t) \in N_\varepsilon(\vec{q}),\ 0 < t \le kT\right) \\
&\qquad \le \mathbb{E}_{\vec{z}}\Big\{\mathbf{1}\left[\vec{z}_n(t) \in N_\varepsilon(\vec{q}),\ 0 < t \le (k-1)T\right] \\
&\qquad\qquad \times \sup_{\vec{y}\in N_\varepsilon(\vec{q})} \mathbb{E}_{\vec{y}}\mathbf{1}\left[\vec{z}_n(t) \in N_\varepsilon(\vec{q}),\ 0 \le t \le T\right]\Big\}.
\end{aligned}$$

Now for each $\vec{y}$ in $N_\varepsilon(\vec{q})$, let $\vec{z} \stackrel{\triangle}{=} \vec{y} + 2\varepsilon(\vec{y} - \vec{q})|\vec{y} - \vec{q}|^{-1}$, so that $\vec{z} \notin N_{2\varepsilon}(\vec{q})$. Let $\vec{r}$ be the path constructed in Lemma 6.21, with $\vec{r}(0) = \vec{y}$ and $\vec{r}(\varepsilon) = \vec{z}$, and let S_ε be the set of paths that remain within ε of $\vec{r}$ during $[0, T]$. Then

$$\begin{aligned}
\mathbb{E}_{\vec{y}}\mathbf{1}\left[\vec{z}_n(t) \in N_\varepsilon(\vec{q}),\ 0 \le t \le T\right] &\le 1 - \mathbb{P}_{\vec{y}}\left(|\vec{z}_n(t) - \vec{r}(t)| < \varepsilon,\ 0 \le t \le T\right) \\
&\le 1 - e^{-nI^*_{\vec{y}}(S_\varepsilon)+O(n)} \\
&\le 1 - e^{-nC\cdot 2\varepsilon},
\end{aligned}$$

where the second inequality holds uniformly in $\vec{y}$ in $N_\varepsilon(\vec{q})$ by the lower bound, and the last inequality holds by (6.9) of Lemma 6.21, where C can be chosen independently of $\vec{y}$ (how?). Thus with $C_e(\varepsilon) \stackrel{\triangle}{=} C \cdot 2\varepsilon$,

Corollary 5.65 the upper bound estimate is uniform in n for $\vec{z}$ in the compact set $\overline{N_{2\varepsilon}}(\vec{q})$. Thus, for all n large enough and all such $\vec{z}$

$$\mathbb{P}_{\vec{z}}(\vec{z}_n \in F_T) \leq e^{-n(I^*_{\vec{z}}(F_T)-\delta)}.$$

But Lemma 6.21 proves that, for ε small enough,

$$I^*_{\vec{z}}(F_T) \geq I^* - \delta$$

for all T (proof by contradiction is easy). We now bound the first term on the right-hand-side of (6.21), by choosing T large. Using Lemma 6.32, let $\tilde{T}$ be large enough so that

$$\mathbb{P}_{\vec{z}}\left(\tau_0 \geq k\tilde{T}\right) \leq e^{-nC_o(\varepsilon)k}.$$

Thus choosing k so large that $C_0(\varepsilon)k > I^* + 1$, this term is indeed negligible when $T = T'' \stackrel{\triangle}{=} k\tilde{T}$ (note that we might have $T > T^*$, but this does not affect the result).

For the final upper bound (6.20), by the last calculation it suffices to estimate the probability of $\vec{z}_n \in F_2$, where

$$F_2 \stackrel{\triangle}{=} F_{T''} \cap \left\{\vec{r} : \sup_{-t_r \leq t \leq 0} \left|\vec{r}(t+t_r) - \vec{r}^{\,*}(t)\right| \geq \eta\right\},$$

where t_r is the first time $\vec{r}$ leaves G^o. Since $\vec{r}^{\,*}$ is the unique minimizer of I over S, the path $\{\vec{r}^{\,*}(t),\ -T'' \leq t \leq 0\}$ is also the unique minimizer of $I^0_{-T''}$, among paths starting at $\vec{r}^{\,*}(-T'')$ (recall that if $t < -T^*$ then by definition $\vec{r}^{\,*}(t) = \vec{q}$). But, as discussed in the introduction to this chapter, using the fact that I_0^T is a good rate function (lower semicontinuous, with compact level sets) or, using Proposition 5.46, we obtain that for any $\eta > 0$, there is a δ small enough so that

$$I^*_{\vec{r}^{\,*}(-T'')}(F_2) \geq I_{\vec{r}^{\,*}(-T'')}(\vec{r}^{\,*}) + 4\delta,$$

and this holds for any $T'' \geq T_1$, where δ is independent of T'' and ε (Exercise 6.38). Note that F_2 contains paths on $[0, T]$, where $T > T^*$ is possible: we use $T < \infty$, since by definition $\vec{r}^{\,*}$ is optimal among all paths of all lengths. Now choose ε small, and increase T'' if necessary, so that $\vec{r}^{\,*}(-T'')$ is in $N_\varepsilon(\vec{q})$. Applying Lemma 6.21 to both sides of the last inequality, we obtain (increasing T'' if necessary)

$$I^*_{\vec{z}}(F_2) \geq I^* + 4\delta - \delta = I^* + 3\delta$$

for all $\vec{z}$ in the compact set $\overline{N_{2\varepsilon}}(\vec{q})$. The large deviations upper bound of Corollary 5.65 now gives

$$\mathbb{P}_{\vec{z}}\left(\vec{z}_n \in F_2\right) \leq e^{-n(I^*+2\delta)}$$

for n large enough. The uniformity in $\vec{z}$ over $\overline{N_{2\varepsilon}}(\vec{q})$ follows from the continuity of $I^*_{\vec{x}}(F_2)$ in $\vec{x}$, which is established just like the continuity of $I^*_{\vec{x}}(F)$. Finally, since by decreasing T'' the probability on the left-hand side of (6.20) decreases, we conclude that T'' can also be chosen arbitrarily—it need not be large. ■

The proof of Lemma 6.36 does not extend to $d = 1$; if the optimal path r^* exits G^o, say, to the right of q while $z_n(\tau_j)$ is to the left of q, then the only way for the process to exit G^o at τ_{j+1} is at the left boundary point. Most likely this will be done in the optimal way of exiting to the left of q; there is no reason the cost should be close to I^*. The probability of exit by the optimal path is obviously zero in this case. However, the basic idea is still valid:

Exercise 6.38. a. Lemma 6.36 holds in the case $d = 1$. Hint: by choosing ε small we can guarantee that the ratio of the probability of exiting N_{ε_δ} in the "right" direction compared to the "wrong" direction is bounded below (actually, diverges). b. Show that δ (in the 4δ argument) can indeed be chosen independently of T'' and of ε, provided ε is small. Hint: the first claim follows from lower semicontinuity. For the second, assume ε much smaller than η and bound the cost to go from the ε ball to the η ball. ♠

The methods of Lemma 6.36 allow us also to compute the probability of exiting the boundary through a prescribed subset of the boundary, which does not include the optimal exit point $\vec{r}^*(0)$. Fix a set $H \subset \partial G$ and assume that it is a "nice" set, that is, H is the closure of its relative interior. Now recall (6.4) and define

$$S_H \stackrel{\triangle}{=} S \cap \{\vec{r} : \vec{r}(-T) = \vec{q},\ \vec{r}(0) \in H,\ T > 0\}. \tag{6.22}$$

We shall say that $\vec{z}_n$ exited G^o at τ_{j+1} and exit was through H if the straight line between $\vec{z}_n(\tau_{j+1})$ and $\vec{z}_n(\tau_{j+1}^-)$, the position just before exiting, crosses H.

Exercise 6.39. Assume $H \subset \partial G$ is the closure of its relative interior. Under the conditions of Lemma 6.36, for any $\delta > 0$ small enough, n large enough, and uniformly in compact subsets of G^o,

$$e^{-n(I^*(S_H)+\delta)} \le \mathbb{P}_{\vec{x}}\left(\vec{z}_n(\tau_{j+1}) \notin G^o,\ \text{exit was through } H \ \middle|\ \vec{z}_n(\tau_j) \in G^o\right) \le e^{-n(I^*(S_H)-\delta)}.$$

Extend also to $d = 1$ as in Exercise 6.38. Hint: just as in the proof of Lemma 6.36, the idea is to show that exit must occur in a bounded time, and then apply the theory of Chapter 5. For the lower bound note that if $V(\vec{x}^*) = \inf_{\vec{x}\in H} V(\vec{x})$ then there is a $\vec{y} \in H^o$ near $\vec{x}^*$ with $V(\vec{y})$ near $V(\vec{x}^*)$. Use Lemma 6.21 for approximations and continuity. ♠

Exercise 6.40. The purpose of this exercise is to show that Lemma 6.36 holds without even assuming stability of $\vec{q}$, and under very weak conditions. The key condition is that it is not likely for the process to remain away from $\vec{q}$ but inside G for a long time. Recall the notation (6.17). Let $d \ge 2$. Assume that $\log \lambda_i$ are Lipschitz continuous in each G_i, that the $\vec{e}_i$ span $\mathbb{R}^d$ (6.8) and that the rates do not increase or decrease too fast, so that (6.12) holds. Fix δ and choose $\vec{r}_\delta$ in S so that $\vec{r}_\delta(-T_\delta) = \vec{q}$, $\vec{r}_\delta(0) \in \partial G$ and

$$I^0_{-T_\delta}(\vec{r}_\delta) < I^*(S) + \delta.$$

(i) If for some $C > 0$ we can extend $\vec{r}_\delta$ for all ε small enough up to time u^ε so that $\vec{r}_\delta(u^\varepsilon)$ is away from the boundary of G, but not too far, or more precisely

$$N_\varepsilon(\vec{r}_\delta(u^\varepsilon)) \cap G = \emptyset \quad \text{and} \quad |\vec{r}_\delta(u^\varepsilon) - \vec{r}_\delta(0)| < C\varepsilon,$$

and if $\log \lambda_i$ are Lipschitz continuous also in a ε-neighborhood of $\{\vec{r}_\delta(t) : -T_\delta \le t \le u^\varepsilon\}$, then the lower bounds (6.18)–(6.19) hold, with $\vec{r}_\delta$ replacing $\vec{r}^{\,*}$.

(ii) Assume that if the process starts near $\vec{q}$, then it either gets closer or escapes G altogether. More precisely, assume that for some $\eta > \varepsilon > 0$, there is a $T = T(\varepsilon, \eta)$ so that if $\vec{z}_\infty(0) \in N_{2\varepsilon}(\vec{q}) \cap N^C_{\varepsilon/2}(\vec{q})$ then for some $t \le T$ either $\vec{z}_\infty(t) \in N_{\varepsilon/2}$ or $\vec{z}_\infty(t) \notin G^\eta$. Then the upper bound in (6.18) holds. If in addition $\vec{r}^{\,*}$ exists and is unique, then (6.20) holds.

(iii) The results of (i)–(ii) hold in the case $d = 1$, in the form of Exercise 6.38, provided that for ε small enough, the process can exit the ε-ball in the desired direction, i.e.,

$$\mathbb{P}_x \left(|z_n(\tau_j) - r^*(0)| < |q - r^*(0)| \;\middle|\; z_n(\tau_{j-1}) \in G^o \right) \ge e^{-n(\delta/2)}. \quad (6.23)$$

Hints: for (i), the proof of Lemma 6.36 extends. For (ii) note that for n large, $\vec{z}$ needs only be taken outside $N_\varepsilon(\vec{q})$ but in $N_{2\varepsilon}(\vec{q})$. For the last part of (iii) compute the probability of a sequence of jumps all in the right direction. ♠

Remark. The complicated condition concerning $T(\varepsilon, \eta)$ in Exercise 6.40, part (ii), is implied, for example, by the stability Assumption (6.5). It simply means that there are no limit cycles for $\vec{z}_\infty$ inside G. Equation (6.23) holds if the rates are bounded and (at least one of) the rates of the jumps in the direction from q to $r^*(0)$ satisfies, for some positive η, $\lambda_i(q) > 0$ and $\int_{q-\eta}^{q+\eta} |\log \lambda_i(x)|\, dx < \infty$, which holds, in particular, whenever $\lambda_i(x)$ is polynomial in x. On the other hand, if all jumps are of size one and the point $\vec{q}$ is absorbing [that is, $\lambda_i(\vec{q}) = 0$ for all i], then the condition does not hold. Indeed, in this case the lower bound cannot hold.

Lemmas 6.32–6.36 indicate that the segments of the processes between stopping times are almost identically distributed, and in a sense independent. We shall need the following more precise statements in the course of the proof of Theorem 6.17.

Exercise 6.41. Assume the estimates (6.15) of Lemma 6.32 and (6.16) of Exercise 6.34 (or 6.35) hold. Fix $\delta > 0$ and define $b_j \stackrel{\triangle}{=} \mathbf{1}\left[\tau_{2j} - \tau_{2j-2} \ge T\right]$—the event that the j^{th} "cycle" is longer than T. Let $\tilde{b}_1, \tilde{b}_2, \ldots$ be a sequence of Bernoulli random variables with $\mathbb{P}\left(\tilde{b}_j = 0\right) = e^{-n\delta}$. There is a positive T so that, if $\tau_{2K} < \infty$,

$$\mathbb{P}_{\vec{x}} \left(\sum_{j=1}^{K} b_j \le M \right) \le \mathbb{P}_{\vec{x}} \left(\sum_{j=1}^{K} \tilde{b}_j \le M \right)$$

for all M and all n large enough. Hint: use (6.16) to bound the probability that $b_j = 1$. Use induction on K, condition on a sum up to $K-1$ and on the position at τ_{2K-1}, and use the estimates (6.15) of Lemma 6.32. Remember the position is discrete, for each n, and use the Markov property. ♠

Exercise 6.42. Assume the estimates (6.15) of Lemma 6.32 and (6.16) of Exercise 6.34 (or 6.35) hold. Fix $\delta > 0$ and define $c_j \stackrel{\triangle}{=} \tau_{2j} - \tau_{2j-2}$—the length of the j^{th} "cycle." Let $\tilde{c}_1, \tilde{c}_2, \ldots$ be a sequence of i.i.d. exponential random variables with mean $e^{n\delta}$. There is a positive T so that, if $\tau_{2K} < \infty$,

$$\mathbb{P}_{\vec{x}}\left(\sum_{j=1}^{K} c_j \geq M\right) \leq \mathbb{P}_{\vec{x}}\left(\sum_{j=1}^{K} (\tilde{c}_j + 2T) \geq M\right)$$

for all M and all n large enough. Hint: let first $g_1, g_2, \ldots$ be geometric random variables with mean $e^{n\delta}$. For integer i use

$$\mathbb{P}_{\vec{x}}\left(\sum_{j=1}^{K} c_j \geq M\right) \leq \sum_{i} \mathbb{P}\left(M - iT \leq \sum_{j=1}^{K} c_j \leq M - iT + T,\ c_K \geq iT\right)$$

and now use the same conditioning as in Exercise 6.41. Finally, replace the geometric random variables with $T+$ exponential ones. ♠

All the ingredients are now in place to prove the main results of this section. For convenience, let us restate each theorem before its proof.

Theorem 6.15. *Under Assumptions 6.1–6.6, for each finite $T < T^*$ and $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}_{\vec{x}}\left(\sup_{\tau_n(G)-T\leq t\leq \tau_n(G)} |\vec{z}_n(t) - \vec{r}^{\,*}(t - \tau_n(G))| < \varepsilon,\ \tau_n(G) > T\right) = 1.$$

The rate of convergence is exponential, uniformly over $\vec{x}$ in any compact subset of G^o.

Proof. We provide the proof in the case $d \geq 2$. The calculations are validated for the case $d = 1$ in Exercise 6.44.

First note that $\tau_n(G)$ is almost surely finite (use a coupling argument as in Examples 4.8–4.9 to show that, eventually, there will be a sequence of jumps in the, say, $\vec{e}_i$ direction, uninterrupted by other jumps, that will cause an exit, independently of the starting position). On the other hand, by Kurtz's Theorem 5.3, for any finite time T, the probability that the process exits G before T (and thus does not follow $\vec{z}_\infty$) $\mathbb{P}_{\vec{x}}(\tau_n(G) \leq T) \to 0$, uniformly in $\vec{x}$ in any compact subset of G^o. Therefore we can (and shall) ignore the stipulation $\{\tau_n(G) > T\}$.

We use the notation of Lemma 6.36. Fix some $T < T^*$. Since we want to show that events of the type (6.19) dominate, we may as well increase the ε of (6.19) to

equal η and set $\eta = \varepsilon$. So, we define a sequence of events, and for convenience we suppress the dependence on n and the condition $t \geq \tau_{2j-1}$.

$$R_{2j} \stackrel{\triangle}{=} \left\{ \omega : \sup_{\tau_{2j}-T \leq t \leq \tau_{2j}} |\vec{z}_n(t) - \vec{r}^{\,*}(t-\tau_{2j})| < \varepsilon \quad \text{and} \quad \vec{z}_n(\tau_{2j}) \notin G^o \right\}.$$

R_{2j} is the "desired" event, that the process exits near $\vec{r}^{\,*}$, and does so at τ_{2j}. Now let W_{2j} denote the event that the exit is "wild" in that it wanders away from $\vec{r}^{\,*}$.

$$W_{2j} \stackrel{\triangle}{=} \left\{ \omega : \sup_{\tau_{2j}-T \leq t \leq \tau_{2j}} |\vec{z}_n(t) - \vec{r}^{\,*}(t-\tau_{2j})| \geq \varepsilon \quad \text{and} \quad \vec{z}_n(\tau_{2j}) \notin G^o \right\}.$$

Note that $R_{2j} \bigcup W_{2j} = \left\{ \omega : \vec{z}_n(\tau_{2j}) \notin G^o \right\}$. Recall that we adopted the convention that $\tau_j = \infty$ if $\tau_k = \tau_n(G)$ for some $k < j$. The event we are interested in is the union of the R_{2j}, and clearly

$$R \stackrel{\triangle}{=} \bigcup_{j>0} R_{2j} = \bigcup_{j>0} \left\{ R_{2j} \cap \{\tau_{2j-1} < \infty\} \right\}.$$

Therefore

$$\begin{aligned}
\mathbb{P}_{\vec{x}}(R) &= \mathbb{P}_{\vec{x}} \left(\bigcup_{j>0} \left\{ R_{2j} \cap \{\tau_{2j-1} < \infty\} \right\} \right) \\
&= \sum_{j=1}^{\infty} \mathbb{P}_{\vec{x}} \left(R_{2j} \cap \{\tau_{2j-1} < \infty\} \right) \\
&= \sum_{j=1}^{\infty} \mathbb{P}_{\vec{x}} \left(R_{2j} \mid \tau_{2j-1} < \infty \right) \mathbb{P}_{\vec{x}} \left(\tau_{2j-1} < \infty \right) \\
&\geq e^{-n(I^*+\delta)} \sum_{j=1}^{\infty} \mathbb{P}_{\vec{x}} \left(\tau_{2j-1} < \infty \right)
\end{aligned}$$

for all n large, by (6.19) of Lemma 6.36. This estimate shows also that the last sum is finite. Similarly (with the obvious definition of W),

$$\begin{aligned}
\mathbb{P}_{\vec{x}}(W) &= \mathbb{P}_{\vec{x}} \left(\bigcup_{j>0} \left\{ W_{2j} \cap \{\tau_{2j-1} < \infty\} \right\} \right) \\
&= \sum_{j=1}^{\infty} \mathbb{P}_{\vec{x}} \left(W_{2j} \mid \tau_{2j-1} < \infty \right) \mathbb{P}_{\vec{x}} \left(\tau_{2j-1} < \infty \right) \\
&\leq e^{-n\delta} e^{-n(I^*+\delta)} \sum_{j=1}^{\infty} \mathbb{P}_{\vec{x}} \left(\tau_{2j-1} < \infty \right)
\end{aligned}$$

for all n large, by (6.20) of Lemma 6.36. Combining the two estimates

$$\mathbb{P}_{\vec{x}}(R) + \mathbb{P}_{\vec{x}}(W) \leq \mathbb{P}_{\vec{x}}(R)\left(1 + e^{-n\delta}\right) \tag{6.24}$$

for all n large. Now let

$$R_0 \stackrel{\triangle}{=} \{\omega \,:\, \tau_G = \tau_0\}\,.$$

That is, R_0 is the set of paths that escape right away, before entering $N_{\varepsilon_s}(\vec{q})$ at all. Then it is an easy consequence of Assumption 6.4 and Kurtz's Theorem that, uniformly over $\vec{x}$ in any compact subset of G^o, we have a constant η such that

$$\mathbb{P}_{\vec{x}}(R_0) < e^{-n\eta} \tag{6.25}$$

for all large enough n. Since $\tau_n(G) < \infty$ a.s. and since the sets R_0, R, and W are disjoint,

$$\mathbb{P}_{\vec{x}}\left(R_0 \bigcup R \bigcup W\right) = \mathbb{P}_{\vec{x}}(R_0) + \mathbb{P}_{\vec{x}}(R) + \mathbb{P}_{\vec{x}}(W) = 1,$$

so that (6.24) and (6.25) imply that, for large n,

$$\mathbb{P}_{\vec{x}}(R) \geq 1 - 2e^{-n(\eta\wedge\delta)} \to 1 \text{ as } n \to \infty$$

so that convergence is exponential. Finally, the uniformity in $\vec{x}$ follows since the estimates in Lemma 6.36 are uniform. ■

Exercise 6.43. Extend Theorem 6.15 to $T \geq T^*$. Hint: since T is finite, only the case $T^* < \infty$ needs be treated. Use the extension to the definition of $\vec{r}^{\,*}$ for large (negative) times and the estimates on the probability of remaining near $\vec{q}$. ♠

Exercise 6.44. Prove Theorem 6.15 when $d = 1$. Hint: see Exercise 6.38. ♠

We can now characterize the relative probability of exiting the set G through a specified part H of the boundary ∂G. Recall the definition (6.22) of S_H and the following discussion of the meaning of exiting G^o through H.

Corollary 6.45. *Assume $H \subset \partial G$ is the closure of its relative interior. Under Assumptions 6.1–6.5, uniformly over $\vec{x}$ in any compact subset of G^o,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{x}}\big(\vec{z}_n(t) \textit{ exits } G^o \textit{ through } H\big) = I^*(S) - I^*(S_H).$$

Suppose now that all the assumptions hold in a larger set $\tilde{G}$, but $\vec{q} \in \tilde{G}^o \setminus G^o$. Then for each $\vec{x} \in G^o$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{x}}\big(\vec{z}_n(t) \textit{ exits } G^o \textit{ through } H\big)$$
$$= \inf\left\{I^0_{-T}(\vec{r}) : \vec{r}(-T) = \vec{x},\ \vec{r}(0) \in H,\ T > 0\right\}.$$

We take T as defined in Exercise 6.41. Now

$$\begin{aligned}\mathbb{P}_{\vec{x}}\left(\tau_n(G) < Te^{n(I^*-3\delta)}\right) &= \sum_{i=1}^{\infty} \mathbb{P}_{\vec{x}}\left(j_G = i,\ \tau_n(G) < TK_n e^{-n\delta}\right) \\ &\leq \sum_{i=1}^{K_n-1} \mathbb{P}_{\vec{x}}\left(\vec{z}_n(\tau_{2i}) \notin G^o,\ \vec{z}_n(\tau_{2i-1}) \in G^o\right) \\ &\quad + \sum_{i=K_n}^{\infty} \mathbb{P}_{\vec{x}}\left(\tau_n(G) < TK_n e^{-n\delta},\ j_G = i\right).\end{aligned}$$

For the first summand, if $d \geq 2$ Lemma 6.36 implies

$$\begin{aligned}\mathbb{P}_{\vec{x}}\left(\vec{z}_n(\tau_{2i}) \notin G^o,\ \vec{z}_n(\tau_{2i-1}) \in G^o\right) &\leq \mathbb{P}_{\vec{x}}\left(\vec{z}_n(\tau_{2i}) \notin G^o \;\middle|\; \vec{z}_n(\tau_{2i-1}) \in G^o\right) \\ &\leq e^{-n(I^*-\delta)}.\end{aligned}$$

(If $d = 1$, condition on τ_{2i-2} and use Exercise 6.38 to obtain this estimate.) Now, for any non-negative random variables $\sigma_1, \sigma_2, \ldots$ and any $T > 0$, $\sigma_i \geq T\mathbf{1}[\sigma_i \geq T]$, so that for any K and A,

$$\mathbb{P}\left(\sum_{i=1}^{K} \sigma_i < A\right) \leq \mathbb{P}\left(T\sum_{i=1}^{K} \mathbf{1}[\sigma_i \geq T] < A\right).$$

Therefore, using the notation and results of Exercise 6.41, for all n large, and for $i \geq K_n$,

$$\begin{aligned}\mathbb{P}_{\vec{x}}\left(\tau_n(G) < TK_n e^{-n\delta},\ j_G = i\right) &\leq \mathbb{P}_{\vec{x}}\left(T\sum_{j=1}^{i} \tilde{b}_j \leq TK_n e^{-n\delta}\right) \\ &\leq \mathbb{P}_{\vec{x}}\left(\frac{1}{i}\sum_{j=1}^{i} \tilde{b}_j \leq \frac{1}{3}\right).\end{aligned}$$

This puts us in a position to use large deviations, with n fixed, and i becoming large. Note that the rate function $\ell_p(a)$ for Bernoulli random variables is monotone in p (the probability of a one) when a is far from the mean (consult Exercise 1.17). So, with $p = p(n)$,

$$\mathbb{P}_{\vec{x}}\left(\frac{1}{i}\sum_{j=1}^{i} \tilde{b}_j \leq \frac{1}{3}\right) \leq e^{-i\ell_{p(n)}(1/3)} \leq e^{-i\ell},$$

where $\ell = \ell_{2/3}(1/3) > 0$ is the value of the rate function when $p = 2/3$ at $a = 1/3$. Putting the estimates together, for every $T > 0$ we have

$$\mathbb{P}_{\vec{x}}\left(\tau_n(G) < Te^{n(I^*-3\delta)}\right) \leq K_n e^{-n(I^*-\delta)} + \frac{e^{-\ell K_n}}{1-e^{-\ell}}.$$

Therefore, since the second term decreases much faster than the first, we obtain

$$\mathbb{P}_{\vec{x}}\left(\frac{\log \tau_n(G)}{n} \geq I^* - 3\delta\right) \geq 1 - \frac{e^{-n\delta}}{2}.$$

This gives a lower bound on $\tau_n(G)$ for part (i) of the theorem. This also provides us with the desired lower bound on the expectation in (ii).

To get an upper bound on $\tau_n(G)$ in (i), define T to be the larger of the T defined in Lemma 6.32 and the T defined in Exercise 6.42 (note that this T makes the conclusion of both hold). Fix $s \geq 2\delta$ and let now $K_n \stackrel{\triangle}{=} e^{n(I^*+s)}$. As in the first part of the proof (and with the same caveat in the case $d = 1$), using now Exercise 6.42 and the notation therein,

$$\begin{aligned}
&\mathbb{P}_{\vec{x}}\left(\tau_n(G) > Te^{n(I^*+2s)}\right) \\
&\quad \leq \mathbb{P}_{\vec{x}}\left(j_G \geq K_n\right) + \sum_{i=1}^{K_n} \mathbb{P}_{\vec{x}}\left(\sum_{j=1}^{i} c_j \geq Te^{n(I^*+2s)},\ \tau_{2i} < \infty\right) \\
&\quad \leq \mathbb{P}_{\vec{x}}\left(\vec{z}_n(\tau_0) \in G^o\right) \prod_{i=1}^{K_n} \mathbb{P}_{\vec{x}}\left(\vec{z}_n(\tau_{2i}) \in G^o \mid \tau_{2i-1} < \infty\right) \\
&\qquad\qquad + \sum_{i=1}^{K_n} \mathbb{P}_{\vec{x}}\left(\frac{1}{i}\sum_{j=1}^{i} \tilde{c}_j \geq \frac{Te^{n(I^*+2s)}}{i} - 2T\right) \\
&\quad \leq \left(1 - e^{-n(I^*+\delta)}\right)^{K_n} + \sum_{i=1}^{K_n} \exp\left(-i\,\ell_n\left(\frac{Te^{n(I^*+2s)}}{i} - 2T\right)\right), \quad (6.26)
\end{aligned}$$

where the first term of the last line of (6.26) is obtained by Lemma 6.32. The second term is obtained from the large deviations upper bound for fixed n, so that ℓ_n is the rate function for i.i.d. exponential random variables $\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_i$ with mean $e^{n\delta}$. Since $s \geq 2\delta$ and $i \leq K_n$, the argument of ℓ_n is larger than $Te^{2n\delta}$, which in turn is larger than the mean $e^{n\delta}$ of $\tilde{c}_j$. Thus ℓ_n is increasing in this region. Now using the scaling property of the rate function for exponential random variables (cf. Exercise 1.20), $\ell_n(\alpha) = \ell_1(\alpha e^{-n\delta})$, where ℓ_1 corresponds to mean one. Therefore,

$$\begin{aligned}
\ell_n\left(\frac{Te^{n(I^*+2s)}}{i} - 2T\right) &\geq \ell_1\left(Te^{n(2s-s-\delta)} - 2Te^{-n\delta}\right) \\
&\geq C_1 e^{n(s-\delta)} - C_2
\end{aligned}$$

since ℓ_1 is convex, and $C_1 > 0$ since ℓ_1 is increasing in the range of interest.

Therefore we have for the second term in (6.26)

$$\begin{aligned}\sum_{i=1}^{K_n} \exp\left(-i\,\ell_n\left(\frac{Te^{n(I^*+2s)}}{i} - 2T\right)\right) &\le \sum_{i=1}^{K_n} \exp\left(-i\left(C_1 e^{n(s-\delta)} - C_2\right)\right) \\ &\le K_n \exp\left(-C_1 e^{n(s-\delta)} + C_2\right) \\ &= \exp\left(n(I^*+s) - C_1 e^{n(s-\delta)} + C_2\right)\end{aligned} \tag{6.27}$$

which decreases superexponentially fast with n, uniformly in $s \ge 2\delta$. As for the first term in (6.26), since

$$\lim_{x\to 0} (1-x)^{1/x} = \frac{1}{e}$$

we have, with $x_n = e^{-n(I^*+\delta)}$

$$\begin{aligned}\left(1 - e^{-n(I^*+\delta)}\right)^{K_n} &= (1-x_n)^{(\exp n(s-\delta))/x_n} \\ &\le (2/e)^{\exp n(s-\delta)}.\end{aligned} \tag{6.28}$$

Therefore, combining (6.26), (6.27), and (6.28), we see that

$$\begin{aligned}\mathbb{P}_{\vec{x}}\left(\tau_n(G) \ge Te^{n(I^*+2s)}\right) &\le (2/e)^{\exp(n(s-\delta))} \\ &\quad + \exp\left(n(I^*+s) - C_1 e^{n(s-\delta)} + C_2\right)\end{aligned} \tag{6.29}$$

and the upper bound in (i) is established.

To obtain the upper bound for (ii) note that the bound (6.29) is uniform in $s \ge 2\delta$. Let $\alpha = Te^{n(I^*+4\delta)}$. Clearly

$$\mathbb{E}_{\vec{x}}\tau_n(G) \le \alpha + \int_\alpha^\infty \mathbb{P}_{\vec{x}}\left(\tau_n(G) \ge t\right)\, dt.$$

With the change of variable $t = Te^{n(I^*+2s)}$, we get $dt = 2nTe^{n(I^*+2s)}\, ds$, and using (6.29) we easily find positive numbers C_3 and C_4 so that

$$\begin{aligned}\int_\alpha^\infty \mathbb{P}_{\vec{x}}\left(\tau_n(G) \ge t\right)\, dt &\le 2nTe^{nI^*} \int_{2\delta}^\infty \mathbb{P}_{\vec{x}}\left(\tau_n(G) \ge Te^{n(I^*+2s)}\right) e^{2ns}\, ds \\ &\le 2nTe^{nI^*} \int_{2\delta}^\infty \left((2/e)^{\exp n(s-\delta)} + e^{(C_4 - C_3\exp(ns))}\right) ds,\end{aligned}$$

which converges to zero as $n \to \infty$, and the theorem is established. ∎