For Chapter 31

31-2. For the last assertion one may for each ω, view Q as a probability measure on $(\mathbf{D}([0,\infty),\Psi),\mathcal{H})$. Then Q_t is the distribution of the Ψ-valued random variable $\varphi \rightsquigarrow \varphi_t$ defined on the probability space $(\mathbf{D}([0,\infty),\Psi),\mathcal{H},Q)$. Since $\varphi_u \to \varphi_t$ as $u \searrow t$ and almost sure (in this case sure) convergence implies convergence in distribution, $Q_u \to Q_t$ as $u \searrow t$ (for each ω, not just the requested 'a.s.').

31-3. (i) for all $x \in \Psi$, $\mu_{x,t} \to \mu_{x,0} = \delta_x$ as $t \searrow 0$; (ii) for all Borel $A \subseteq \Psi$ and $t \geq 0$, the function $x \rightsquigarrow \mu_{x,t}(A)$ is measurable; (iii) for all Borel $A \subseteq \Psi$, $s,t \geq 0$, and $x \in \Psi$,

$$\mu_{x,s+t}(A) = \int_\Psi \mu_{y,s}(A)\,\mu_{x,t}(dy)\,.$$

31-9. Let R_t denote the distribution of the Lévy process at time t. Then

$$T_t f(x) = \int f(x+y)\,R_t(dy)\,.$$

Let R denote the distribution of the Lévy process. Then the corresponding Markov family $(Q^x : x \in \mathbb{R})$ is defined by

$$Q^x(B) = R\big(\{\varphi\colon [t \rightsquigarrow (x+\varphi_t)] \in B\}\big)\,.$$

31-15. $Gf(x) = \kappa \int (f(y) - f(x))\,Q(dy)$, in the notation of Example 1 of Chapter 30.

31-21. Suppose that Q_0 is an equilibrium distribution for $\widetilde{T}$. Then

$$Q_0 T_t = e^{-ct} \sum_{k=0}^{\infty} \frac{(ct)^k}{k!} Q_0 \widetilde{T}^k = e^{-ct} \sum_{k=0}^{\infty} \frac{(ct)^k}{k!} Q_0 = Q_0\,.$$

Hint: for converse: Use Problem 16.

31-23. *Hint:* Let f be the indicator function of the one-point set $\{y\}$ and use Theorem 14.

31-25.

$$\begin{aligned}
p_{00}(t) &= (q_{01}+q_{10})^{-1}\big[q_{10} + q_{01}\exp[-(q_{01}+q_{10})t]\big]\\
p_{01}(t) &= (q_{01}+q_{10})^{-1}\,q_{01}\big(1 - \exp[-(q_{01}+q_{10})t]\big)\\
p_{10}(t) &= (q_{01}+q_{10})^{-1}\,q_{10}\big(1 - \exp[-(q_{01}+q_{10})t]\big)\\
p_{12}(t) &= (q_{01}+q_{10})^{-1}\big(q_{01} + q_{10}\exp[-(q_{01}+q_{10})t]\big)
\end{aligned}$$

The limits at ∞ of both p_{00} and p_{10} are the same: $(q_{01}+q_{10})^{-1}q_{10}$, the value the equilibrium distribution assigns to $\{0\}$. The limits at ∞ of both p_{01} and p_{11} are the same: $(q_{01}+q_{10})^{-1}q_{01}$, the value the equilibrium distribution assigns to $\{1\}$.

31-28. The solution to Problem 23 involves applying Theorem 14 to the indicator functions of one-point sets. When the rates are unbounded, such functions may not be in the domain of the generator. For example, let the state space be $\mathbb{Z}^+$, let the transition rates q_{xy} have the property that $q_{x0} \to \infty$ as $x \to \infty$, and let f be the indicator function of $\{0\}$. Then f is not in the domain of the infinitesimal generator because the limit in the definition does not exist boundedly, and Theorem 14 does not apply. Nevertheless, it can be shown that (31.12) holds whenever the state space is countable, even in the case of unbounded rates.

31-29. Let M be the largest member of the support of ρ, x_0 the initial state, and U_n the time of the n^{th} jump. The construction ensures that $X_{U_n} \leq x_0 + (M-1)n$. Therefore, conditioned on $\mathcal{F}_{U_{n-1}}$, $U_n - U_{n-1}$ is exponential with mean at least $1/\gamma(x_0 + (M-1)(n-1))$. An inductive argument based on this fact shows that for each n, the distribution function of U_n is bounded above by the distribution function of the sum of n independent exponentially distributed random variables with means $1/\gamma x_0$, $1/(\gamma(x_0+M-1))$, … ,$1/(\gamma(x_0+(M-1)(n-1))$. Such a sum of exponentially distributed random variables diverges almost surely as $n \to \infty$ by the Kolmogorov Three-Series Theorem. It follows that $U_n \to \infty$ a.s. as $n \to \infty$.

31-36. $\frac{d\nu}{d\lambda}(y) = b(1-b)ce^{-(1-b)cy}$, $\nu\{\infty\} = 0$; equilibrium distribution assigns value $(1-b)b^x$ to x; jump-rate function is

$$x \rightsquigarrow \begin{cases} cb^{-1} & \text{if } x = 0 \\ c & \text{if } x > 0\,; \end{cases}$$

transition probabilities from x to $x-1$ equal 1 for $x > 0$ and from 0 to $x > 0$ equal $(1-b)b^{x-1}$; transition rates from x to $x-1$ equal c for $x > 0$ and from 0 to $x > 0$ equal $c(1-b)b^x$ and all others equal 0

For Chapter 32

32-1.

$$(\xi, \eta) \rightsquigarrow \begin{cases} b_x(\xi) & \text{if } \eta = \xi^x \\ d_x(\xi) & \text{if } \eta = {}_x\xi \\ j_{xy}(\xi) & \text{if } \eta = {}_x\xi^y \\ 0 & \text{otherwise} \end{cases}$$

32-7. For $\xi \in \Xi$, let $X^{(\xi)}$ be the process defined in the construction with initial state ξ. The discussion in the paragraph following the proof of Theorem 2 shows that for each time $t \geq 0$, the function $\xi \rightsquigarrow X_t^{(\xi)}$ is almost surely a continuous function. It follows from the Bounded Convergence Theorem that, for any continuous function $f\colon \Xi \to \mathbb{R}$, the function $\xi \rightsquigarrow E(f \circ X_t^{(\xi)})$ is continuous. Thus, the transition semigroup is Feller.

32-10. Here is one way to make a correct 'if and only if' statement: Let G and $G^{(k)}$ be as in the first sentence of Problem 9. The 'if' statement is: If $G^{(k)}f \to Gf$ pointwise as $k \to \infty$ for all $f \in \mathfrak{F}$, then $X_t^{(k)} \to X_t$ as $k \to \infty$, uniformly for t in bounded subsets of $[0,\infty)$ and for all choices of initial states $\xi^{(k)}$ and ξ such that $\xi^{(k)} \to \xi$. The 'only if' statement is: If there exists a function $f \in \mathfrak{F}$ and a state η such that $G^{(k)}f(\eta)$ does not converge to $Gf(\eta)$ as $k \to \infty$, then there exists a time $t > 0$ and a sequence of initial states $\xi^{(k)}$ converging to a state ξ as $k \to \infty$ such that $X_t^{(k)}$ does not converge to X_t as $k \to \infty$. (In this second statement, we may take $\xi^{(k)} = \xi = \eta$ for all k and let t be any sufficiently small positive time.)

To prove the 'if' statement, it is enough to show that for any site x and any time $t \geq 0$, there exists a nonnegative random variable K that is almost surely finite such that $X_s^{(k)}(x) = X_s(x)$ for all $k \geq K$ and $s \in [0, t]$. This last statement is a slightly stronger version of the statement made in the paragraph immediately following the proof of Theorem 2. To prove this stronger statement, first note that since $G^{(k)} f \to Gf$ for $f \in \mathfrak{F}$ as $k \to \infty$, each rate in the system with infinitesimal generator $G^{(k)}$ converges uniformly as $k \to \infty$ to the corresponding rate in the system with infinitesimal generator G. Now consider the construction of $X^{(k)}$ and X using the universal coupling. Let A be as in the statement following the proof of Theorem 2 and let K be large enough so that $\xi^{(k)}$ agrees with ξ at sites in A for $k \geq K$. We can also choose K large enough so that the rates of $G^{(k)}$ at sites in A are uniformly as close as we like to the corresponding rates of G when $k \geq K$. A simple modification of the proof of Theorem 2 shows that we can thereby make the probability arbitrarily close to 1 that the processes $X^{(k)}$ and X take the same values at x at all times in $[0, t]$. Further details are left to the reader.

The hypothesis in the 'only if' statement implies that there exists a site x such that at least one of the rates at x for the process with infinitesimal generator G is not the pointwise limit as $k \to \infty$ of the corresponding rates for the processes with infinitesimal generators $G^{(k)}$. It follows that there exist arbitrarily large integers k and a state η such that the process with infinitesimal generator $G^{(k)}$ and initial state η will not behave the same at the site x as the process with infinitesimal generator G and initial state η, at least for short time periods. Once again, the details are left to the reader.

32-13. (This problem is incorrectly stated in the book. The statement is not true for the contact process with threshold birth rates. Also, a stronger statement is proved for the contact process with sexual reproduction in Problem 12. So the problem should only be done for the contact process of Example 2.) For finite sets $A \subseteq \mathbb{Z}^d$, let

$$f_A(\xi) = \sum_{x \in A} \xi(x)\,.$$

Direct calculation shows that if ξ is a state with only finitely many occupied sites, then

$$Gf_A(\xi) \leq (1 - \delta) f_A(\xi)\,, \tag{7.22}$$

provided A is chosen large enough to include all x such that $\xi(x) = 1$.

Let ξ_0 be a state with only finitely many occupied sites, and let (X_t) be the interacting particle system with initial state ξ_0 and infinitesimal generator G. For each finite set $A \subseteq \mathbb{Z}^d$, define a random time σ_A by

$$\sigma_A = \inf\{t \geq 0\colon X_t(x) = 1 \text{ for some } x \notin A\}\,.$$

Also, let

$$\tau = \inf\{t \geq 0\colon X_t = \bar{0}\}\,.$$

Since the interacting particle system is a solution to the martingale problem for G, it follows from (7.22) and the Optional Sampling Theorem that for any time $t \geq 0$,

$$E(f_A(X_{t \wedge \sigma_A \wedge \tau})) - f_A(\xi_0) \leq E\Big(\int_0^{t \wedge \sigma_A \wedge \tau} (1 - \delta) f_A(X_s)\, ds\Big)\,.$$

Since $\delta > 1$, the integrand on the right side is bounded above by $(1-\delta)$ for all $s < \sigma_A$, so

$$E(f_A(X_{t\wedge\sigma_A\wedge\tau})) - f_A(\xi_0) \le (1-\delta)E(t\wedge\sigma_A\wedge\tau)\,,$$

from which it follows immediately that

$$f_A(\xi_0) \ge (\delta-1)E(t\wedge\sigma_A\wedge\tau)\,.$$

We leave it to the reader to check that $\sigma_A \nearrow \infty$ a.s. as $A \nearrow \mathbb{Z}^d$. Thus, after first letting $A \nearrow \mathbb{Z}^d$ and then letting $t \nearrow \infty$, we have by the Monotone Convergence Theorem that

$$\sum_{x\in\mathbb{Z}^d} \xi_0(x) \ge (\delta-1)E(\tau)\,.$$

Since ξ_0 has only finitely many occupied sites, the left side of this inequality is finite. It follows that τ has finite expectation, and hence that τ is finite almost surely, as desired.

32-16. It is easily checked that for each site x, the process $(X_t(x), t \ge 0)$ is a pure-jump Markov process with state space $\{0,1\}$, transition rates $q_{01} = 1$ and $q_{10} = 2^{|x|}$, and initial state 0. It follows from Problem 25 of Chapter 31 that

$$P[X_t(x) = 1] < 2^{-|x|}\,.$$

By the Borel Lemma, $\sum_x X_t(x)$ is finite a.s. Thus, for any fixed time t, the number of occupied sites at time t is finite a.s.

For the second part of the problem, we fix $t \in (0,\infty)$. We know from the previous part of the problem that at any given time s there are infinitely many vacant sites. Since the birth rates are all equal to 1 at vacant sites, it is not hard to show that, with probability 1, infinitely many births occur during every time interval of positive length. In particular, infinitely many births occur with probability 1 during the time interval $(0,t)$. Let

$$x_1 = \min\{x > 0\colon \text{ there is a birth at } x \text{ during } (0,t)\}\,.$$

Let U_1 be the time of the first birth at x_1 and V_1 the time of the first death at x_1.

We now proceed by induction. We assume that random sites $x_1,\dots,x_n$ have been defined for some $n \ge 1$, with corresponding random times $U_1,\dots,U_n$ and $V_1,\dots,V_n$, where for each $k = 2,\dots,n$, U_k is the time of the first birth at x_k after time U_{k-1}, and V_k is the time of the first death at x_k after time U_k. Note that $U_1 < U_2 < \cdots < U_n$. As part of the induction, we also assume that $U_n < t\wedge V_1\wedge\cdots\wedge V_n$. This assumption implies that the time interval $(U_n, t\wedge V_1\wedge\cdots\wedge U_n)$ has positive length, so the following random site is almost surely defined:

$$x_{n+1} = \min\{x > 0\colon \text{ there is a birth at } x \text{ during } (U_n, t\wedge V_1\wedge\cdots\wedge V_n)\,.$$

Let U_{n+1} be the time of the first birth at x_{n+1} after time U_n, and V_{n+1} the time of the first death at x_{n+1} after time U_{n+1}. Note that our construction ensures that $U_{n+1} < t\wedge V_1\wedge\cdots\wedge V_{n+1}$, as required by the assumption made in the inductive step.

Let $U = \lim_{n\to\infty} U_n$. Our construction of U shows that U is defined almost surely, and that when it is defined, $U \le t$. This construction also shows that $X_{U-}(x_n) = 1$ for all $n = 1,2,\dots$. Our construction of the process (X_t) shows that, with probability 1, at most one death can occur at time U, so infinitely many sites are occupied at time U, as desired.

32-18. $j_{xy}(\xi) = \rho\{y - x\}$ if $\xi(y) = 0$; $d_x(\xi) = \sum_y \xi(y)\rho\{y - x\}$; other rates are 0

32-24. There is an error in Example 8: in order for the if and only if statement at the end of the example to be true, one must assume that the death rate is bounded away from zero. Under that additional assumption, a solution to Problem 24 (with arbitrary finite range r) can be made as follows. First do the problem for the case in which the initial state ξ satisfies the property that $\xi(x) = 0$ for all sites x to the right of some site y. Deduce that y can be chosen so that the probability is arbitrarily close to 1 that the sites to the right of $-r$ remain vacant for all time. Prove similar statements for the case in which the initial state satisfies $\xi(x) = 0$ for all x to the left of some site y. Use these facts to show that $y > 0$ can be chosen so that if the initial state ξ satisfies $\xi(x) = 0$ for $-y \le x \le y$, then the probability is at least $\frac{1}{2}$ that $X_t^\xi \to \bar{0}$ as $t \to \infty$. Having chosen such a y, use the assumption on the death rates to show that for any initial state, the process almost surely spends an infinite amount of time in states η for which $\eta(x) = 0$ for $-y \le x \le y$. Use the strong Markov property to complete the proof.

For Chapter 33

33-2. *Hint:* Let $(\mathcal{F}_t\colon t \ge 0)$ denote the minimal filtration of the Wiener process W. Square both sides of (33.1) and then take expectations. Six terms result on the right side. The following calculation shows that one of them is equal to 0:

$$E\Big(Z_{n\varepsilon}a(Z_{n\varepsilon})(W_{(n+1)\varepsilon} - W_{n\varepsilon})\Big) = E\Big(E\big(Z_{n\varepsilon}a(Z_{n\varepsilon})(W_{(n+1)\varepsilon} - W_{n\varepsilon}) \;\big|\; \mathcal{F}_{n\varepsilon}\big)\Big)$$
$$= E\Big(Z_{n\varepsilon}a(Z_{n\varepsilon})E\big((W_{(n+1)\varepsilon} - W_{n\varepsilon}) \;\big|\; \mathcal{F}_{n\varepsilon}\big)\Big) = 0\,.$$

Similarly,

$$E\Big(b(Z_{n\varepsilon})\varepsilon a(Z_{n\varepsilon})(W_{(n+1)\varepsilon} - W_{n\varepsilon})\Big) = 0\,.$$

The following calculation is relevant for another of the six terms:

$$E\Big([a(Z_\varepsilon)]^2(W_{(n+1)\varepsilon} - W_{n\varepsilon})^2\Big) = E\Big(E\big([a(Z_\varepsilon)]^2(W_{(n+1)\varepsilon} - W_{n\varepsilon})^2 \;\big|\; \mathcal{F}_{n\varepsilon}\big)\Big)$$
$$= E\Big([a(Z_\varepsilon)]^2E\big((W_{(n+1)\varepsilon} - W_{n\varepsilon})^2 \;\big|\; \mathcal{F}_{n\varepsilon}\big)\Big) = \varepsilon E\Big([a(Z_\varepsilon)]^2\Big)\,.$$

33-5. yes

33-12. $d(e^{\alpha W}) = \alpha e^{\alpha W}\, dW + \frac{1}{2}\alpha^2 e^{\alpha W}\, dt$

33-15. Equation (33.18) is to be interpreted as an almost sure statement. In the following, we assume that the relevant properties of the Itô integral have been extended to allow for integrands like $\operatorname{sgn}(W_u - x)$.

Let

$$L_t^\delta = \frac{1}{2\delta}\int_0^t I_{[-\delta+x,\delta+x]}(W_u)\, du\,,$$

so that for each x,

$$|W_\cdot - x| - \int_0^\cdot \operatorname{sgn}(W_u - x)\, dW_u = L_\cdot(x) = \lim_{\delta\searrow 0} L_\cdot^\delta(x) \text{ i.p.}$$

In the notation of the example,

$$L_t^\delta(x) = f_\delta(W_t - x) - \int_0^t f'_\delta(W_u - x)\, dW_u$$

(see (33.16)).

Using Theorem 4 and the fact that $(a+b)^2 \le 2a^2 + 2b^2$ for real numbers a and b, we have

$$E\Big((L_t(x) - L_t^\delta(x))^2\Big) \le$$
$$2E\Big((|W_t - x| - f_\delta(W_t - x))^2\Big) + 2E\Big(\int_0^t (\operatorname{sgn}(W_u - x) - f'_\delta(W_u - x))^2\, du\Big)\,.$$

It follows that for any bounded Borel set B,

$$\lim_{\delta\searrow 0} \int_B E\Big((L_t(x) - L_t^\delta(x))^2\Big)\, dx = 0\,.$$

By the Fubini Theorem,

$$\lim_{\delta\searrow 0} \int_B \big(L_t(x) - L_t^\delta(x)\big)^2\, dx = 0 \text{ i.p.}$$

A simple application of the Cauchy-Schwarz Inequality then gives

$$\int_B L_t(x)\, dx = \lim_{\delta\searrow 0} \int_B L_t^\delta(x)\, dx \text{ i.p.} \tag{7.23}$$

Now suppose that B is a bounded interval in $\mathbb{R}$. Let

$$g^\delta(y) = \frac{1}{2\delta} \int_B I_{[-\delta+x,\delta+x]}(y)\, dx\,.$$

Clearly we have $0 \le g^\delta \le 1$ and $g^\delta \to g$ pointwise as $\delta \searrow 0$, where g is 1 on the interior of B, $1/2$ at the endpoints of B, and 0 elsewhere. By Dominated Convergence, applied to (7.23), we have

$$\int_B L_t(x)\, dx = \int_0^t g(W_u)\, du \text{ a.s.}$$

It is easy to deduce from this equation that

$$\int_B L_t(x)\, dx = \int_0^t I_B(W_u)\, du \text{ a.s.}$$

A standard argument using the Sierpiński Class Theorem shows that this last equation is valid for all Borel sets B, as claimed.

33-17. For $z \in \mathbb{R}$, let $Z^{(z)}$ denote the solution of (33.19) with initial state z, and let $(T_t, t \ge 0)$ denote the corresponding transition semigroup. Since $T_t f(z) = E(f \circ Z_t^{(z)})$, the Bounded Convergence Theorem implies that it is enough to show that for each $t \ge 0$ and $z \in \mathbb{R}$, $\lim_{y\to z} Z_t^{(y)} = Z_t^{(z)}$ a.s. In the proof of Theorem 7 it is shown that each random variable $Z_t^{(y)}$ is the limit in probability of random variables $Z_t^{(y,\varepsilon)}$ as $\varepsilon \searrow 0$. From the definitions it is apparent that $y \rightsquigarrow Z^{(y,\varepsilon)}$ is almost surely a continuous function for each $\varepsilon > 0$. Thus, it is enough to show that

$$\lim_{\varepsilon,\eta\searrow 0} \sup_{y\in\mathbb{R}} |Z^{(y,\varepsilon)} - Z^{(y,\eta)}| = 0\,. \tag{7.24}$$

Noting that the estimates used in the proof of Theorem 7 do not depend on the initial value y, we see that, with minor modifications, the argument in that proof can be used to give (7.24).

33-29. $Gf = \frac{1}{2}\Delta f$ for sufficiently nice functions $f\colon \mathbb{R}^d \to \mathbb{R}$. For bounded continuous functions f having bounded continuous first, second, and third partial derivatives, this fact can be proved by direct computation, using the second degree Taylor polynomial approximation of f with remainder.

For Appendix A

A-2. The derivative $x \rightsquigarrow 1 - \cos x$ is positive for $-2\pi < x < 0$ and also for $0 < x < 2\pi$. A theorem of calculus says that a continuous function on a closed interval that has a positive derivative at all interior points of that interval is strictly increasing on the closed interval. Therefore the given function is strictly increasing on the interval $[-2\pi, 0]$ and on the interval $[0, 2\pi]$. By the preceding problem it is strictly increasing on the interval $[-2\pi, 2\pi]$. (Notice that the argument can be extended to prove that the given function is strictly increasing on $\mathbb{R}$.)

For Appendix B

B-1. Proof that a closed subset of a compact set is compact. Let B be a closed subset of a compact set C, and let $\mathcal{O}$ be an open covering of B. Consider $\mathcal{O} \cup \{B^c\}$, the collection obtained by adjoining the complement of B to the collection $\mathcal{O}$. This collection is an open covering of C. It contains a finite subcovering of C. The members of $\mathcal{O}$ in this finite subcovering of C constitute a finite subcovering (from $\mathcal{O}$) of B.

B-5. The 'only if' part is trivial. We will prove the contrapositive of the 'if part', so suppose that the sequence does not converge to y. Then there exists $\varepsilon > 0$ and an infinite subsequence $(x_{n_k}\colon k = 1, 2, \dots)$ of (x_n) such that $\rho(x_{n_k}, y) > \varepsilon$ for all k. No further subsequence of this subsequence can converge to y because the distance between y and every member of that further subsequence would be greater than ε.

For Appendix C

C-5. Suppose that $x \in \partial B$. Case 1, $x \in B$: Every neighborhood of x contains a member of B —namely x itself. If some neighborhood did not contain a member of B^c, then x would be a member an open subset of that neighborhood which itself would be a subset of B. Hence x would belong to the interior of B and thus not to ∂B.

Case 2, $x \notin B$: Now we must show that every neighborhood of x contains a member of B. If there were some neighborhood lying entirely inside B^c, there would be an open subset of that neighborhood containing x and having the same property. The complement of that open set would be a closed set containing B and thus containing the closure of B. Therefore x would not belong to ∂B.

For the converse suppose that every neighborhood of x contains at least one point of B and least one point in B^c. First we observe that x cannot be a member of the interior of B, for, if it were, this interior would be a neighborhood of x that contains no member of B^c. To finish the proof we must show that x belongs to the closure of

B. If it did not, the complement of the closure of B would be a neighborhood of x containing no point of B, which is a contradiction.

C-6. *Hint:* Avoid doing work similar to that needed for the preceding problem.

C-9. $[a,b)$, both open and closed whether $b<\infty$ or $b=\infty$; $(a,b]$, neither open nor closed whether $a>-\infty$ or $a=-\infty$; $[a,b]$ closed but not open; (a,b) open but not closed whether a and b are finite or infinite; $[a,a]$ is only compact interval

C-10. Closure under arbitrary unions: clearly yes if all sets in the union belong to $\mathcal{O}$; if one of the sets in the union contains ∞ and has a complement that is a compact subset C of Ω, the union will contain ∞ and have a complement that is closed subset of the compact subset C of Ω. An appeal to Proposition 1 completes this portion of the proof.

Closure under finite intersections: clearly yes if one of the sets in the intersection does not contain ∞; if all do contain ∞, then so does the intersection and the complement of the intersection is the union of a finite number of compact subsets of Ω. The definition of compactness shows that a finite union of compact sets is compact.

Compactness: An open covering must have at least one set that contains ∞. Take any such set O. The remaining sets in the open covering cover the compact complement of O. Thus there is a finite subcovering of this complement. Adjoin O to this finite subcovering to obtain a finite subcovering of Ω^*.

C-14. The closed interval $[0,1]$ of $\mathbb{R}$ with the usual topology is not open in that topology, but it is an open subset of the topological space $[0,1]$ with the relative topology.

Now assume that $\Psi\in\mathcal{O}$ and that $O\subset\Psi$ is open in the relative topology on Ψ. Then $O=A\cap\Psi$ for some $A\in\mathcal{O}$. Hence, O, the intersection of two members of $\mathcal{O}$, is itself a member of $\mathcal{O}$.

For Appendix D

D-1. 30

D-2. $\frac{3}{11}$

D-14. According to Theorem 4 we only need prove that f is Riemann-Stieltjes integrable with respect to g, and for doing that, Proposition 2 says that we only need prove that f is bounded and fg' is Riemann integrable.

Suppose that f is unbounded. For each m there exists $x_m\in[a,b]$ such that $|f(x_m)|>m$. Let x denote a limit of a subsequence of (x_m). It cannot be that infinitely many members of the subsequence equal x. If infinitely many members are larger than x, then $f(x+)$ does not exist. If infinitely many members are smaller than x, then $f(x-)$ does not exist. Therefore the assumption that f is unbounded leads to a contradiction, and hence f is bounded.

For future use we show that for each $\delta>0$, there exists only finitely many x such that
$$f(x-)\vee f(x)\vee f(x+)>\delta+f(x-)\wedge f(x)\wedge f(x+)\,.$$
If there were infinitely many, then at the limit y of a convergence sequence of distinct such x, either $f(y+)$ or $f(y-)$ would fail to exist.

Turning to the proof of Riemann integrability of fg', we let $\varepsilon > 0$. For each $x \in [a,b]$ let J_x be an open interval in $[a,b]$ such that

- $x \in J_x$,
- $|f(y) - f(x+)| < \frac{\varepsilon}{4(b-a)}$ if $x < y \in J_x$,
- $|f(y) - f(x-)| < \frac{\varepsilon}{4(b-a)}$ if $x > y \in J_x$.

(Reminder: Intervals in $[a,b]$ including the endpoint a or b can be open in the relative topology of $[a,b]$. Alternatively, we could have let J_a and J_b be open intervals in $\mathbb{R}$ containing members outside the interval $[a,b]$.) Since $[a,b]$ is compact there exists a finite collection of intervals J_x whose union equals $[a,b]$. Let $\widehat{P}$ be the point partition of $[a,b]$ consisting of the endpoints of the intervals in this finite collection and the points midway between two consecutive endpoints.

For each point x for which

$$f(x-) \vee f(x) \vee f(x+) > \tfrac{\varepsilon}{4(b-a)} + f(x-) \wedge f(x) \wedge f(x+)\,,$$

of which there are only finitely many—say q —introduce a close interval $K_x \subseteq [a,b]$ containing x as an interior point and having length less than $\frac{\varepsilon}{4qs}$, where s denotes the supremum of $|f(x)g'(x)|$ for $x \in [a,b]$. Let P denote the point partition of $[a,b]$ obtained by adjoining the endpoints of each such K_x to $\widehat{P}$.

Consider any refinement P' of P. For any Riemann sum of fg' corresponding to P', the total contribution arising from intervals lying in the various K_x is less than $\varepsilon/4$. The contributions to any two such Riemann sums arising from other intervals differ by less than $3\varepsilon/4$. Thus any two Riemann sums of any refinement of P differ by less than ε.

Now a straightforward argument using a sequence of refinements corresponding to a decreasing sequence (ε_k) gives a Cauchy sequence of Riemann sums. Then the above argument can be used again to show that the limit of this Cauchy sequence is the value of the Riemann integral, and thus in particular, that the Riemann integral of fg' exists.

Comment: For those whose definition of Riemann integrals involves upper and lower integrals and sums rather than Riemann sums, the above argument can be shortened a bit. We have not adopted the 'upper-lower' approach because it does not generalize nicely to the Riemann-Stieltjes setting.

For Appendix E

E-4. We consider the real part of $\exp \circ \lambda$:

$$(\mathfrak{R} \circ \exp \circ \lambda) = (\exp \circ \mathfrak{R} \circ \lambda) \cdot (\cos \circ \mathfrak{I} \circ \lambda)\,.$$

Using the Product Rule and Chain Rule for $\mathbb{R}$-valued functions we obtain

$$\begin{aligned}(\mathfrak{R} \circ \beta)' &= (\exp \circ \mathfrak{R} \circ \lambda) \cdot (\mathfrak{R} \circ \lambda)' \cdot (\cos \circ \mathfrak{I} \circ \lambda) \\ &\qquad - (\exp \circ \mathfrak{R} \circ \lambda) \cdot (\sin \circ \mathfrak{I} \circ \lambda) \cdot (\mathfrak{I} \circ \lambda)' \\ &= (\mathfrak{R} \circ \lambda') \cdot (\mathfrak{R} \circ \exp \circ \lambda) - (\mathfrak{I} \circ \lambda') \cdot (\mathfrak{I} \circ \exp \circ \lambda) \\ &= \mathfrak{R} \circ \big(\lambda' \cdot (\exp \circ \lambda)\big)\,,\end{aligned}$$

as desired. We omit the similar calculation relevant for the imaginary part.

E-9. no