2

Lévy's Theorem Concerning Positiveness of Transition Probabilities

Paul Lévy investigated "the allure" of the sample paths of general Markov processes $\xi = \{\xi_t\}_{t \geq 0}$ with denumerable state space S by using the properties of the so-called i-intervals, that is the sets $I(i) = \{t \colon \xi_t = i\}$. Lévy's study concludes with a very fine property of the transition probabilities $p_{ij}(t)$ of ξ, known as the *Lévy dichotomy*:

> *for any pair* (i, j) *of states and* $t \in (0, +\infty), p_{ij}(t)$ *is either identically zero or everywhere strictly positive.*

(See P. Lévy (1951, 1958).)

D.G. Kendall, introducing a classification for Markovian theorems in the spirit of the swallow/deep classification of Kingman, pointed out that the Lévy dichotomy belongs to the class of theorems relying on the Chapman–Kolmogorov equations (see D.G. Kendall and E.F. Harding (1973), p. 37).

D.G. Austin proved Lévy's property by a probabilistic argument, using the right separability of the process and Lebesgue's theorem on differentiation of monotone functions. Another proof, more analytic, was latter given by D. Ornstein (see K.L. Chung (1967) for details on these results). Recently, K.L. Chung (1988)) proved Lévy's theorem by using some information from the corresponding Q-matrix: he assumes the states are stable.

In this section we shall show that Lévy's theorem has an expression in terms of directed cycles or circuits, when the state space is at most a countable set and the process admits an invariant probability distribution $\pi = (\pi_i, i \in S)$. Our approach relies on the circuit representation theory exposed in Part I according to which, for each t, the transition probabilities $p_{ij}(t)$ are completely determined by a class $\{\mathscr{C}(t), w_c(t)\}$, where $\mathscr{C}(t)$ and

$w_c(t)$ denote, respectively, a collection of directed circuits occurring in the graph of $(p_{ij}(t), i, j \in S)$ and strictly positive numbers. Specifically, the $p_{ij}(t)$'s are expressed as

$$\pi_i p_{ij}(t) = \sum_{c \in \mathscr{C}(t)} w_c(t) J_c(i,j), \qquad w_c(t) > 0, \quad t \geq 0, \quad i, j \in S,$$

where J_c is the passage function associated with c. Throughout this chapter the circuits will be considered to have distinct points (except for the terminals). Then for $t > 0$

$$w(i,j,t) \equiv \pi_i p_{ij}(t) > 0$$
$$\textit{if and only if } (i,j) \textit{ is an edge of some circuit } c \in \mathscr{C}(t).$$

Accordingly, we may say that Lévy's theorem expresses a qualitative property of the process ξ. This will then inspire a circuit version of Lévy's theorem according to which the representative circuits are time-invariant solutions to the circuit generating equations

$$\sum_j w(i,j,t) = \sum_k w(k,i,t), \qquad i \in S, \quad t > 0.$$

Finally, we shall discuss a physical interpretation of Lévy's theorem when the elements of $\mathscr{C}(t)$ are considered resistive (electric) circuits, the $\pi_i, i \in S$, represent node (time-invariant) currents and the $w(i,j,t), i, j \in S$, are branch currents.

2.1 Lévy's Theorem in Terms of Circuits

Given a countable set S, let $P = \{P(t), t \geq 0\}$ be any homogeneous stochastic standard transition-matrix function with $P(t) = (p_{ij}(t), i, j \in S)$. Assume P defines an irreducible positive-recurrent Markov process $\xi = \{\xi_t, t \geq 0\}$ on a probability space $(\Omega, \mathscr{K}, \mathbb{P})$. Suppose further that $P(t), t > 0$, is of bounded degree (that is, for any $i \in S$ there are finitely many states j and k such that $p_{ij}(t) > 0$ and $p_{ki}(t) > 0$). For any $t > 0$ consider the discrete t-skeleton $\Xi_t = \{\xi_{nt}, n \geq 0\}$ of ξ, that is, the S-state Markov chain whose transition probability matrix is $P(t)$. The above assumptions on P imply that any skeleton-chain Ξ_t is an irreducible aperiodic positive-recurrent Markov chain.

Now we shall appeal to the circuit representation Theorems 3.3.1 and 5.5.2 of Part I according to which, there exists a probabilistic algorithm providing a unique circuit representation $\{\mathscr{C}_t, w_c(t)\}$ for each $P(t)$, that is,

$$\pi_i p_{ij}(t) = \sum_{c \in \mathscr{C}_t} w_c(t) J_c(i,j), \qquad t \geq 0, \quad i, j \in S, \tag{2.1.1}$$

where $\pi = (\pi_i, i \in S)$ denotes the invariant probability distribution of $P(t), t > 0, \mathscr{C}_t$ is the collection of the directed circuits occurring on almost all the trajectories of $\Xi_t, t > 0$, and $w_c(t), c \in \mathscr{C}_t$, are the cycle skipping rates defined by Theorem 3.2.1. Then the $w_c(t)$'s are strictly positive on $(0, +\infty)$.

On the other hand, if we suppose that ξ is reversible, that is, for each $t > 0$ the condition $\pi_i p_{ij}(t) = \pi_j p_{ji}(t)$ is satisfied for all $i, j \in S$, we may apply the deterministic algorithm of Theorem 3.4.2 for defining a circuit representation $(\mathscr{C}(t), \tilde{w}_c(t))$ of each $P(t)$ with all $\tilde{w}_c(t) > 0$ on $(0, +\infty)$. As already mentioned we shall consider directed circuits (with distinct points except for the terminals) as representatives. Furthermore, we shall distinguish the probabilistic collection of representative circuits from the deterministic ones using the notation $\mathscr{C}_t$ for the first and $\mathscr{C}(t)$ for the second ones. Also, the theorems quoted below belong to Part I. Denote by sgn x the *signum*, that is, the function on $[0, +\infty)$ defined as sgn $x = 1$ if $x > 0$, and sgn $x = 0$ if $x = 0$.

We are now in a position to apply to Lévy's property the argument of the circuit decomposition above, and to show that this property has an expression in terms of the directed circuits.

Theorem 2.1.1. *Let S be any finite set. Then for any S-state irreducible Markov process $\xi = \{\xi_t\}_{t \geq 0}$ defined either by a standard matrix function $P(t) = (p_{ij}(t), i, j \in S), t \geq 0$, or by a probabilistic or deterministic collection of directed circuits and weights, the following statements are equivalent:*

- *(i) Lévy's property: for any pair (i, j) of states, the* $\operatorname{sgn}(p_{ij}(t))$ *is time invariant on* $(0, +\infty)$.
- *(ii)* Arcset $\mathscr{C}(t) =$ Arcset $\mathscr{C}(s)$, *for all $t, s > 0$ and for all the deterministic classes $\mathscr{C}(t)$ and $\mathscr{C}(s)$ of directed circuits occurring in Theorem* 4.2.1 *when representing Ξ_t, and Ξ_s, respectively, where* Arcset $\mathscr{C}(u)$ *denotes the set of all directed edges of the circuits of $\mathscr{C}(u), u > 0$.*
- *(iii) $\mathscr{C}_t = \mathscr{C}_s$, for all $t, s > 0$, where $\mathscr{C}_t$ and $\mathscr{C}_s$ denote the unique probabilistic classes of directed circuits occurring in Theorem* 4.1.1 *when representing Ξ_t and Ξ_s, respectively.*

If S is countable, then the above equivalence is valid for reversible processes. In any case, we always have (i) $\Leftrightarrow$ (iii).

Proof. First, consider that S is a finite set. The equivalence (i) $\Leftrightarrow$ (ii) follows immediately. Let us prove that (iii) $\Rightarrow$ (i). Consider $t_0 > 0$. Then for any pair (i, j) of states we have

$$p_{ij}(t_0) = \sum_{c \in \mathscr{C}_{t_0}} \frac{1}{\pi_i} w_c(t_0) J_c(i, j), \tag{2.1.2}$$

where $\pi = (\pi_i, i \in S)$ is the invariant probability distribution of ξ and

$w_c(t_0), c \in \mathscr{C}_{t_0}$, are the cycle skipping rates (introduced by Theorem 3.2.1). If $p_{ij}(t_0) > 0$, it follows from (2.1.2) that there is at least one circuit $c_0 \in \mathscr{C}_{t_0}$ such that $w_{c_0}(t_0) > 0$ and $J_{c_0}(i,j) = 1$. Then, by hypothesis $c_0 \in \mathscr{C}_t$ for all $t > 0$. As a consequence, the $p_{ij}(\cdot)$, written as in (2.1.2), will be strictly positive on $(0, +\infty)$. Therefore (iii) $\Rightarrow$ (i).

To prove that (i) $\Rightarrow$ (iii) we first note that the Chapman–Kolmogorov equations and standardness imply that $\mathscr{C}_s \subseteq \mathscr{C}_t$ for $s \leq t$. It remains to show the converse inclusion. Let c be a circuit of $\mathscr{C}_t$, that is, $c = (i_1, \ldots, i_k, i_1)$ has the points $i_1, \ldots, i_k$ distinct from each other when $k > 1$ and

$$p_{i_1 i_2}(t) p_{i_2 i_3}(t) \cdot \ldots \cdot p_{i_k i_1}(t) > 0.$$

Then, from hypothesis (i) we have

$$p_{i_1 i_2}(s) p_{i_2 i_3}(s) \cdot \ldots \cdot p_{i_k i_1}(s) > 0.$$

Therefore $c \in \mathscr{C}_s$, so that $\mathscr{C}_s \equiv \mathscr{C}_t$ for all $s, t > 0$.

Finally, for the countable state space case we have to appeal to the representation Theorems 3.3.1 and 3.4.2, and to repeat the above reasoning. The proof is complete. □

As an immediate consequence of Theorem 2.1.1, the circuit decomposition (2.1.1), or the cycle decomposition (5.5.2) of Chapter 5 (Part I) should be written in terms of a single class $\mathscr{C} \equiv \mathscr{C}_t$, independent of the parameter-value $t > 0$, that is,

$$\pi_i p_{ij}(t) = \sum_{c \in \mathscr{C}} w_c(t) J_c(i,j), \qquad t \geq 0, \quad i, j \in S.$$

Accordingly, $(\mathscr{C}, w_c(t))_{t \geq 0}$ will be the probabilistic circuit (cycle) representation of ξ.

2.2 Physical Interpretation of the Weighted Circuits Representing a Markov Process

One of the physical phenomena which can be modeled by a circuit process is certainly that of a continuous electrical current flowing through a resistive network. Accordingly, the circuits and the positive circuit-weights representing a recurrent Markov process should be interpreted in terms of electric networks. Then certain stochastic properties of circuit processes may have analogues in some physical laws of electric networks.

Let S be a finite set and $\xi = \{\xi_t\}_{t \geq 0}$ be an irreducible reversible Markov process whose transition matrix function and invariant probability distribution are $P(t) = (p_{ij}(t), i, j \in S)$ and $\pi = (\pi_i, i \in S)$, respectively. Denote by $\mathscr{C}_0$ the collection of all the directed circuits with distinct points (except for the terminals) occurring in the graph of $P(t)$. Since $\mathscr{C}_0$ is symmetric, we may write it as the union $\mathscr{C} \cup \mathscr{C}_-$ of two collections of directed circuits in S such that $\mathscr{C}_-$ contains the reversed circuits of those of $\mathscr{C}$.

Then the probabilistic circuit representation Theorem 4.1.1 and Lévy's theorem enable us to write the equations

$$\pi_i p_{ij}(t) = \sum_{c \in \mathscr{C}} w_c(t) J_c(i,j) + \sum_{w_{c_-} \in \mathscr{C}_-} w_{c_-}(t) J_{c_-}(i,j), \qquad (2.2.1)$$

for any $i, j \in S$, $t > 0$, where the $w_c(t)$'s and $w_{c_-}(t)$'s denote the cycle skipping rates for all the circuits c and c_- with period greater than 2 and the halves of the skipping rates for all the circuits c with periods 1 and 2. The passage functions J_c and J_{c_-} occurring in (2.2.1) are those introduced by Definition 1.2.2 of Part I.

Consider $w(i,j,t) \equiv \sum_{c \in \mathscr{C}} w_c(t) J_c(i,j)$. Then, applying Theorem 1.3.1 of Part II, we have

$$\tfrac{1}{2}\pi_i = \textstyle\sum_j w(i,j,t) = \sum_k w(k,i,t), \qquad i \in S, \quad t > 0. \qquad (2.2.2)$$

If we relate each circuit $c \in \mathscr{C}_0$ with a resistive circuit, we may interpret the $w(i,j,t), i,j \in S$, as a branch current flowing at time t from node i to node j. Suppose Ohm's law is obeyed. Then equations (2.2.1) express Kirchhoff's current law for the resistive network associated with $\mathscr{C}$.

Invoking the Lévy theorem in terms of circuits, equations (2.2.2) may be interpreted in the electrical setting above as follows: if at some moment $t > 0$ there exist currents $w_c(t)$ flowing through certain electric circuits c according to the law of a circuit Markov process, then this happens at any time and with the same circuits. But, using an argument from the electrical context, the same conclusion arises as follows. The time invariance of the node currents $\pi_i, i \in S$, and the equilibrium Kirchhoff equations (2.2.2) enable one to write

$$\sum_j w(j,i,t-\Delta t) = \sum_k w(i,k,t+\Delta t) = \tfrac{1}{2}\pi_i, \quad i \in S,\ t > 0. \qquad (2.2.3)$$

Then, π being strictly positive at the points of every circuit $c = (i_1, \ldots, i_s, i_1)$ at any time $t > 0$, the existence of a branch current $w(i_k, i_{k+1}, t - \Delta t)$ requires the existence of $w(i_{k+1}, i_{k+2}, t + \Delta t)$, and vice versa. Therefore the time invariance of the node currents π_i and the equilibrium equations (2.2.3) require the existence of the branch currents $w(j,i,t-\Delta t) > 0$ and $w(i,k,t+\Delta t) > 0$ entering and leaving i. Then the collection $\mathscr{C}_t$ of electrical circuits through which the current flows at time $t > 0$ should be time-invariant, and this is in good agreement with Lévy's theorem.

In general, when interpreting a circuit Markov process, the diffusion of electrical currents through the corresponding resistive network can be replaced by the diffusion of any type of energy whose motion obeys rules similar to the Kirchhoff current law. For instance, relations (2.2.2) have a mechanical analogue as long as Kirchhoff current law has a full analogy in Newton's law of classical mechanics. To review briefly some basic mechanical elements of a mechanical system, we can recall any free-body diagram where a body is accelerated by a net force which equals, according to Newton's law, the derivative of the momentum. This equality becomes, when replacing, respectively, forces, velocity, friction, mass, and displacement by currents, voltage, resistor, capacitor, and flux, formally equivalent to Kirchhoff's current law. The previous analogy enables us to consider circuit processes associated to mechanical systems which obey Newton's laws. For instance, let us observe the motion of a satellite at finitely many points $i_1, i_2, \ldots, i_m$ of certain time-invariant overlapping closed orbits c (where Newton's laws are always obeyed). Then the passages of the satellite at time $t > 0$ through the points $i_1, i_2, \ldots, i_m$ under the traction forces $w_c(t)$, follow a Markovian trajectory of a circuit process with transition matrix function

$$\tilde{p}_{ij}(t) = \frac{w(i,j,t)}{\tilde{\pi}_i} \qquad \text{for all} \quad t > 0 \quad \text{and} \quad i, j \in \{i_1, i_2, \ldots, i_m\},$$

where $w(i,j,t) \equiv \sum_c w_c(t) J_c(i,j)$ and $\tilde{\pi}_i \equiv \sum_j w(i,j,t)$. When a trajectory correction is necessary at some instant of time, this will correspond to a perturbation of either the Markov property or strict stationarity. Then we have to change the stochastic model into another circuit process where the corrected orbits will play the rôle of the new representative circuits for the process.