M1 Internship Report

Léo Gayral

Synchronisation Problems on Markov Chains

Tutor:

Kouji Yano,
Department of Mathematics, Graduate School of Science,
Kyoto University,
kyano@math.kyoto-u.ac.jp

Graduate School of Science, Kyoto University

05/06/17 — 20/08/17

Contents

Undergoing of the Internship

I spent all my internship on the north campus of the Kyoto Graduate School of Science, in Japan. During these three months, the initial task given to me was to hold a weekly seminar, along the lines of a *groupe de lecture*. The first articles I presented are the reason I contacted Kouji Yano in the first place:

- [1] *Realization of an ergodic Markov chain as a random walk subject to a synchronizing road coloring*;
- [2] *Random walk in a finite directed graph subject to a road coloring*; and
- [3] *Random walk in a finite directed graph subject to a synchronizing road coloring*.

These three papers are quite complementary so I reordered their results in a way which felt more natural to me, focusing mostly on [2]. This problem roughly occupied the first third of my internship, and is summed up in the first section of the following report, as the starting point for a potential generalisation.

After this, Yano-sensei recommended me some more articles to read, notably:

- [4] *Tsirel'son's equation in discrete time* (Yor);
- [5] *Chaînes de Markov indexées par $-\mathbb{N}$: existence et comportement* (Brossard and Leuridan); and
- [6] *On the exchange of intersection and supremum of σ-fields in filtering theory* (Handel).

While the main questions of [4], [5], [6] and Yano's articles seem quite distinct, all these papers rely on a likewise model: a Markov chain with infinite past $(X_i)_{i\leq 0}$ and its asymptotic filtration $\mathcal{F}_{-\infty} = \bigcap_{n\leq 0} \sigma(X_i, i \leq n)$. It is to note that Yor's article is one of the inspirations of [2] and [5], so these problems are actually somehow related. For my last seminars, I roughly translated [5] (originally written in French) in English and presented the translation along with some lemmas of [4] it uses.

As these papers were getting too technical for my seminar, during the last third of the internship, I went back to Yano's articles and started working on a possible generalisation of their results in the case of a countable set space. This personal research will be explained in the latter parts of this report.

During my internship, I did not only held a weekly seminar but also attended a lot of seminars presented by other students. While their content was quite out of my league – the seminars were entirely done in Japanese and I lacked some prerequisites – the overall experience itself has been quite interesting.

Though unrelated to the subject of the internship, I also spent some time working on computer science as a warmup for the *Agrégation de mathématiques* – as I won't have any time to do so between the intership and the start of classes.

My thanks go to Kouji Yano and his students for welcoming me in their lab during this experience.

1 The Road Colouring Theorem

In this first section, we will study the random walks driven by a random colour introduced in Kouji Yano's articles. To contextualise the results, we will first briefly see the original – deterministic – road colouring problem on finite graphs. Then, we will introduce the notion of random walks driven by a random mapping, and how it can be related to the *a priori* more general notion of Markov chains. Finally, there will a more technical part, about the properties of these random walks up to the random road colouring theorem.

1.1 The Deterministic Road Colouring Problem

The canonical road colouring problem was first stated in the late 1970s in a purely deterministic framework. We say that a set of mappings $\Sigma_0 \subset \Sigma := \{\sigma : V \to V\}$ on a finite set V is *synchronising* if the composition semigroup generated by Σ_0 contains a constant application, ie there is some word $(n_0, \ldots, n_r) \in \Sigma_0^+$ such that $n_0 \circ \cdots \circ s_r$ is constant. The problem can be formulated as follows: given some oriented graph $G = (V, A)$ (allowing multiple edges with the same extremities), is there some partition $(\sigma_i)_{1 \leq i \leq d}$ of the edges A so that each σ_i naturally represents a mapping and that the set of mappings $\{\sigma_i, 1 \leq i \leq d\}$ is synchronising?

To understand this problem, let us consider a city map instead of an abstract graph, with the streets being the edges. At every intersection – vertex of the graph – each available road must be assigned a colour; globally, in order to induce a mapping, each colour must leave each crossroad exactly once. The colouration of the map is called *synchronising* if there exists some finite sequence of instructions – of colours – so that, starting from any crossroad of the city, following these directions will lead to the same destination. Such a sequence is called a *synchronising word*.

The previous example may seem a bit forced, as nobody would get *that* lost. However, such methods can be used in automata theory, where there should ideally be no human intervention on the machine. For the automaton to properly work, you need to reset it back to its initial state after usage. If a *hard reset* is not an option, this requires some knowledge about the current state of the machine. Thereby, if an error is detected in the data given to an automaton, it may be quicker to directly reset the machine instead of waiting to reach a known state – this is all the more true given that the input is erroneous and may have an unpredictable effect on the state. Therefore, the user may want to input a "reset key" – a synchronising word – so that the automaton goes back to its initial stage.

It follows quickly from the definitions that a graph admits some colouration if and only if it has a constant out-degree d. However, such colourations are not always synchronising. In Figure 1 below, we can see the two possible colourations of K_2: the left one uses two constant mappings, and thus is clearly synchronising, while the right one uses two permutations, and thus is clearly not. Similar constructions can be made on any complete graph K_n.

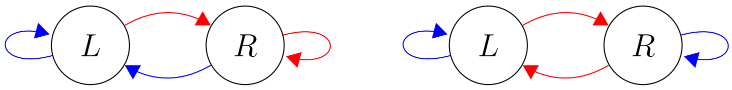

Figure 1: The two possible colourations of the complete graph K_2

In the late 2000s, in his well-known paper [7], Trahtman proved that any strongly-connected aperiodic graph G with constant out-degree d has a synchronising colouration.

1.2 Markov Chains and Random Colours

Now that we have introduced the deterministic ideas of road colouring, we will see how to use this model to define random-walks. As a colourable graph G can be entirely deduced from a colouration, we will tend to forget about G itself and focus on the colourations from now on; the former notion of colouration allows multiple colours to be associated to the same mapping of Σ, but two such colours play exactly the same behaviour in the transitions over V so we will merge these together, and represent a colouration as a subset $\Sigma_0 \subset \Sigma$ of our set of mappings.

Because a mapping $\sigma \in \Sigma$ corresponds to a deterministic transition law over V, a random colour N of law μ naturally induces random transitions and thus a random walk over V, and the support of its law $\mathrm{supp}(\mu)$ is a colouration as redefined above.

In what follows, we consider some arbitrary time interval $I \subset \mathbb{Z}$ which can always be brought back to $\mathbb{N}$ or $\mathbb{Z}$, depending on whether I has a minimum element (the walk has a *starting point* or an *initial law*) or not (the walk has an *infinite past*). The following definition is equivalent to the one on p. 261 in [2]:

Definition 1 (μ-Random Walk)**.** Consider some probability law $\mu \in \mathcal{P}(\Sigma)$ over the mappings on the (finite) set V. A random process $(X_n, N_n)_{n\in I}$ is called a *μ-random walk* if:

1. X_n are random variables over V;
2. $N_n \overset{\text{iid}}{\sim} \mu$ are independent random colours of common law μ;
3. the transitions in X are driven by N following $X_{n+1} = N_{n+1}X_n$; and
4. the transitions are independent of the past, so $N_n \perp\!\!\!\perp \sigma(X_i,\ N_i,\ i < n)$.

If $a = \inf I > -\infty$, then the law of X_a is called the *initial law* of the μ-random walk.

Lemma 1 (Induced Markov Chain)**.** *Consider a random colour law μ. Any mapping σ can be seen as the stochastic matrix $\sigma(y, x) = \delta_{\sigma x, y}$. We can define the transition matrix $P := \sum_{\sigma\in\Sigma} \mu(\sigma) \times \sigma$ induced by μ.*

If $(X_n, N_n)_{n\in I}$ is a μ-random walk, then $(X_n)_{n\in I}$ is a P-Markov chain.

The converse property is true, though a bit less forward to formulate. I give here my own construction of a law μ, which differs from the one stated on p. 772 in [1]. While their construction is somewhat minimal, my version has a maximal support and a briefer and more explicit construction, which will appear again in the last section.

Lemma 2 (Induced μ-Random Walk)**.** *Consider $(X_n)_{n\in I}$ a P-Markov chain, so that $\mathbb{P}(X_{n+1} = y|X_n = x) = P(y, x)$. We define the random colour law $\overline{\mu}$ by $\overline{\mu}(\sigma) := \prod_{x\in V} P(\sigma x, x)$.*

The chain X can be represented as a $\overline{\mu}$-random walk (Y, N), so that the processes $X \overset{d}{=} Y$ have the same law. The idea behind the process (Y, N) is to choose independently a successor for each point in V according to P, and only then to look at the value of Y to apply the right transition.

What's more, this law $\overline{\mu}$ has a maximal support: for any random colour μ inducing the kernel P, necessarily $\mu(\sigma) \leq P(\sigma x, x)$ so $\overline{\mu}(\sigma) \geq \mu(\sigma)^{|V|}$. Thus the maximality of the support, supp$(\mu) \subset$ supp$(\overline{\mu})$.

Note that, just like a graph G can have several colourations, a transition law P can be induced by several laws μ. For example, in Figure 1, a uniform choice between red and blue induces a Chain $(X_n)_{n\in I} \overset{\text{iid}}{\sim} \mathcal{U}(\{L, R\})$ in both cases, which in turn induces the random colour law $\overline{\mu} = \mathcal{U}(\Sigma)$. Now, by inclusion of the supports, a transition law P can be induced by some random colour law μ with synchronising support if and only if the law $\overline{\mu}$ induced by P has a synchronising support itself.

1.3 The Random Road Colouring Theorem

In all this subsection, we consider a random colour μ that induces a strongly-connected aperiodic graph G, so that the results of Trahtman [7] hold and, by Perron-Frobenius theorem, the induced transition law P has a unique invariant law λ.

We also assume from now on that the time interval is $\mathbb{Z}$ (or eventually $-\mathbb{N}$), the important point being that $\inf I = -\infty$, so that each variable X_n has an infinite past and displays an asymptotic behaviour.

The proof of the lemma below can be found on p. 268 in [2]. As it does not make use of any specificity of the finite case, the results of the lemma are still true in the countable case, and they may be used in the following sections.

Lemma 3 (Uniqueness of the μ-Random Walk)**.** *Under the previous hypotheses, there exists a unique μ-random walk, up to identity in law.*

By uniqueness, the process is necessarily stationary and the common law of $(X_n)_{n\in\mathbb{Z}}$ is the P-invariant law λ.

While the uniqueness is a mere consequence of Perron-Frobenius theorem, the existence part of the lemma uses Kolmogorov extension theorem, which basically states that if a family of laws does not hold any internal contradiction then it can be represented by a family of random variables on a probability space. More details on this fundamental result can be found in these online notes [8].

Definition 2 (Strong μ-Random Walk)**.** A μ-random walk $(X_n, N_n)_{n\in\mathbb{Z}}$ is called *strong* if, for any $n \in \mathbb{Z}$, we have $X_n \in \sigma(N_j\,,\ j \leq n)$; in other words, there is a function $f_n : \Sigma^{\mathbb{N}} \to V$ so that $X_n = f_n(N_n, N_{n-1} \dots)$.

The random road colouring problem is, in a general framework, the research of implications or equivalences between the strongness and other properties of μ. In particular, in the finite case and under this subsection's hypotheses, Kouji Yano proved the following theorem which is stated on p. 262 in [2]:

Theorem 1 (Random Road Colouring Theorem)**.** *Under the previous hypotheses, for any μ-random walk (X, N), the following affirmations are equivalent:*

1. *$supp(\mu)$ is synchronising;*
2. *$\forall n \in \mathbb{Z}$, $(N_n \circ N_{n-1} \circ \dots \circ N_{n-k})_{k\geq 0}$ almost surely converges in Σ when $k \to \infty$;*
3. *(X, N) is strong.*

The forward implications are quite forward actually. $(1 \Rightarrow 2)$ is a mere consequence of the second Borel-Cantelli lemma: we are certain to encounter a synchronising sequence of colours in finite time and everything happening before has no effect. $(2 \Rightarrow 3)$ is a bit trickier out of the blue, but with the uniqueness in Lemma 3 in mind, it is enough to remark that the process (Y, N) is strong, where $Y_n := \lim N_n \circ N_{n-1} \circ \dots \circ N_{n-k}(x_0)$ (for some chosen x_0) is almost surely well defined.

However, the last implication $(3 \Rightarrow 1)$ (its contrapositive, actually) is much harder to demonstrate: the whole proof takes a dozen of pages, so only the main ideas will be introduced in Appendix A. The whole proof can be found in [2], mostly in section 4.

2 Generalisation to the Countable Case

In this second section, my objective is to see how the previously studied road colouring problem may be generalised. In order to do so, we will first briefly introduce the general framework for Markov chains (on uncountable spaces). Then, we will see how the random colours can relate to the notion of coupling. Finally, we will introduce a notion of synchronising behaviour for couplings and see how it compares to the previously introduced synchronising behaviours for μ-random walks.

2.1 Generalised Markov Chains

Nowadays, the typical space of a Markov chain is not a finite nor countable set, but a Polish space E. This new framework allows us to study Brownian motions and Lévy processes on $\mathbb{R}^d$ for example.

Definition 3 (Polish Space)**.** A topological space $(E, \mathcal{T})$ is said *Polish* when:

- the space E is separable; and
- the topology $\mathcal{T}$ is induced by some complete metric d over E.

Note that when a space E has the discrete topology (induced by the complete metric $d(x,y) = 1-\delta_{x,y}$), it is separable (thus Polish) if and only if it is countable; thus the notions of countable set and discrete topology are equivalent for Polish spaces. Note also that a countable space $(E, \mathcal{T})$ is Polish if and only if $\mathcal{T}$ is the discrete topology.

The next step is to generalise the notion of transition law. In the discrete case, transition laws over E are exactly families of probability laws in $\mathcal{P}(E)$ indexed on E. However, in the more general case, we have to assume some more structure on the Markov kernel Π:

Definition 4 (Markov Kernel)**.** A *Markov kernel* Π (over E), which we will simply call a *kernel*, is an application $\Pi : E \times \mathcal{B}(E) \to \mathbb{R}^+$ such that:

- for any state $x \in E$, $\Pi_x \in \mathcal{P}(E)$ is a probability law over E; and
- for any measurable $A \in \mathcal{B}(E)$, the application $x \mapsto \Pi_x(A)$ is measurable.

When the following integral is defined, we denote $\Pi_x(f) := \int_E f(y) d\Pi_x(y)$. We can also define the product of kernels as follows:

$$\forall A \in \mathcal{B}(E)\,, \ \left(\Pi^{(2)}\Pi^{(1)}\right)_x (A) = \int\limits_E \Pi^{(2)}_y(A) d\Pi^{(1)}_x(y)\,.$$

To put it bluntly, the previous assumptions on the space E are the bare minimum so that the Kolmogorov extension theorem can always guarantee the existence of a Π-Markov chain (X_n), defined as follows:

Definition 5 (Π-Markov Chain)**.** A random process $(X_n)_{n\in\mathbb{N}}$ is called a Π-*Markov chain* if for any $n \in \mathbb{N}$ and any measurable $f : E \to \mathbb{R}^+$ we have

$$\mathbb{E}[f(X_{n+1}) | X_0, \ldots, X_n] = \Pi_{X_n}(f)\,. \tag{1}$$

If E is a countable space, the notions of kernel Π and transition law P are equivalent with the equality $P(y,x) = \Pi_x(\{y\})$, as these two notions induce the same random processes X.

These definitions are everything we need to state problems for general Markov chains. Much more details on this general case can be found in [9] *Markov chains and stochastic stability*, which I found more accessible for newcomers than the much more technical classic [10] *Markov chains* (Revuz).

Definition 6 (Coupling of Kernels)**.** Let Π be a kernel over E. A *coupling* of Π is a kernel Q over $E \times E$ so that the marginal transition laws always correspond to Π:

$$\forall (x,y) \in E^2\,,\ Q_{x,y}(\bullet \times E) = \Pi_x(\bullet) \text{ and } Q_{x,y}(E \times \bullet) = \Pi_y(\bullet)\,. \tag{2}$$

Remark 1 (Underlying Idea Behind Coupling)**.** The general idea behind a coupling is to assume that if you consider two Markov chains on E happening simultaneously, the way they will interact and influence each other can be predicted as well.

It is quite clear that if (X,Y) is a Q-Markov chain, then X and Y are Π-Markov chains. Two identical ($X = Y$) or independent processes ($X \perp\!\!\!\perp Y$) naturally induce a coupled kernel. However, given two Π-Markov chains X and Y, the couple (X,Y) is not necessarily a time-homogeneous Markov chain itself, and thus does not necessarily induce a coupled kernel.

For example, consider $(\xi_i)_{i>0} \overset{\text{iid}}{\sim} \mathcal{U}(\{0,1\})$, $X_n = X_0 + \sum_{i=1}^n \xi_i \bmod 2$ and $Y_n = Y_0 + \sum_{i=1}^n \xi_i + \lfloor \frac{n}{2} \rfloor \bmod 2$. At even times ($2k \to 2k+1$ transitions), X and Y take opposite decisions, only one of them moves, while at odd times, X and Y make the same choice of moving or not. If we look back on Figure 1, it would mean to alternate between the left and the right colourations, choosing one of the two available colours uniformly at each step. Such a process cannot be represented by a time-homogeneous Q-Markov chain.

2.2 Couplings and Random Colours

In this subsection, we will go back to the finite case. We have seen that any Markov chain can be represented as a μ-random walk, and that a law μ induces one (and only one) kernel Π.

Definition 7 (Coupling Induced by a Random Colour)**.** Let μ be a law over colours and Π the kernel on E induced by μ. A coupling Q of Π is said to be *induced by μ* if

$$Q_{x,y}(u,v) = \mu\left(\{\sigma \in \Sigma\,,\ \sigma x = u \text{ and } \sigma y = v\}\right)\,. \tag{3}$$

Note that, if (X_n) and (Y_n) are two μ-random walks – defined on the same probability space – sharing the same random colours (N_n), then (X,Y) is a coupled Q-Markov chain.

However, the converse property is not obvious. In fact, is false in general that a coupling can be induced by some random colour. A quick look at a coupling Q induced by μ shows that it must at the very least be *sticky* and *symmetrical*, two notions I naturally came to introduce while studying this problem.

Definition 8 (Sticky Coupling)**.** A coupling Q of Π is called *sticky* if for any initial state $x \in E$, starting from the couple (x,x), we have the following transition law:

$$Q_{x,x}(u,v) = \begin{cases} \Pi_x(u) & \text{if} \quad u = v\,, \\ 0 & \text{if} \quad u \neq v\,. \end{cases}$$

In the general situation of a Polish space, the previous relation may be generalised as $Q_{x,x}(A \times B) = \Pi_x(A \cap B)$.

Definition 9 (Symmetrical Coupling)**.** A coupling Q of Π is called *symmetrical* if for any initial pair $(x,y) \in E^2$, we have $Q_{x,y}(u,v) = Q_{y,x}(v,u)$.

In the general situation of a Polish space, the previous relation may be generalised as $Q_{x,y}(A \times B) = Q_{y,x}(B \times A)$.

I conjectured at first that this necessary condition was also sufficient, ie any symmetrical sticky coupling could be induced by a random colour. However, while studying this subject, I noticed that the main obstruction to this result was the existence of a colour compatible with the coupling, and figured that such an obstruction should arise as soon as $|E| > 2$, which gave birth to the following counterexample:

Example 1 (Coupling Incompatible With Mappings)**.** Consider the fully uniform process on 3 elements: $(X_n)_{n\in\mathbb{N}} \overset{\text{iid}}{\sim} \mathcal{U}(\mathbb{F}_3)$. This process is a Π-Markov chain where, for any $i,j \in \mathbb{F}_3$, we have $\Pi_i(j) := \frac{1}{3}$.

We want to define a symmetrical sticky coupling, thus we only need to set the coupled laws for the initial pairs $(1,2)$, $(1,3)$ and $(2,3)$. We define the coupling as follows, with the each row being a choice of x and each column a choice of y in $Q_{i,j}(x,y)$.

$Q_{1,2}$	1	2	3
1	0	1/3	0
2	1/3	0	0
3	0	0	1/3

$Q_{1,3}$	1	2	3
1	0	0	1/3
2	0	1/3	0
3	1/3	0	0

$Q_{2,3}$	1	2	3
1	0	0	1/3
2	1/3	0	0
3	0	1/3	0

Figure 2: Definition of the Sticky Symmetrical Coupling Q

The interest of this representation of the coupling Q is that the sum over a row x (respectively a column y) in $Q_{i,j}$ must be equal to $\Pi_i(x)$ (resp. $\Pi_j(y)$) – hence a much more readable table than a "linear" enumeration of $Q_{i,j}$ over the couples in $\mathbb{F}_3 \times \mathbb{F}_3$.

The main tip behind this coupling is to maximize the number of impossible transitions, in order to force any value of a hypothetical random colour to have null probability, thus an absurdity. In order to do so, the most direct way was to have exactly one positive value on each column and each row, which in turn imposed the choice of the uniform kernel Π.

The problem is now a purely combinatorial one. We know that in all generality, if Q is generated by a random colour law μ, then for any $(x, y) \in E^2$ we have $\mu(\sigma) \leq Q_{x,y}(\sigma x, \sigma y)$.

$\sigma(1,2,3)$	$Q_{x,y}(\sigma x, \sigma y) = 0$
$(1,1,*)$	$Q_{1,2}(1,1)$
$(1,2,1)$	$Q_{1,3}(1,1)$
$(1,2,2)$	$Q_{1,3}(1,2)$
$(1,2,3)$	$Q_{2,3}(2,3)$
$(1,3,*)$	$Q_{1,2}(1,3)$
$(2,1,1)$	$Q_{1,3}(2,1)$
$(2,1,2)$	$Q_{2,3}(1,2)$
$(2,1,3)$	$Q_{1,3}(2,3)$

$\sigma(1,2,3)$	$Q_{x,y}(\sigma x, \sigma y) = 0$
$(2,2,*)$	$Q_{1,2}(2,2)$
$(2,3,*)$	$Q_{1,2}(2,3)$
$(3,1,*)$	$Q_{1,2}(3,1)$
$(3,2,*)$	$Q_{1,2}(3,2)$
$(3,3,1)$	$Q_{2,3}(3,1)$
$(3,3,2)$	$Q_{1,3}(3,2)$
$(3,3,3)$	$Q_{1,3}(3,3)$

Figure 3: Incompatibilities Between Q and the $\sigma \in \Sigma$

Now, this figure quickly shows that for any mapping $\sigma : E \to E$, there is (at least) one coupled transition with null probability in Q. Thus, the coupling Q is entirely incompatible with any random colour μ.

Even in a simple finite case, the notion of (symmetrical sticky) coupling is a strict generalisation of the one of random colours. A problem stated on couplings will therefore be *a priori* more general than its counterpart on colours, though they may be equivalent, as the existence of a coupling satisfying some property may be equivalent to the existence of such a coupling induced by a random colour.

One could wonder why we should study couplings instead of random mappings. One of the reasons is that the set of mappings on E has the cardinality $2^{|E|}$. This means in particular that we lose the main interest of the countable case by introducing a continuous law. Therefore, in what follows, we assume that a random colour law μ *must* have a countable support.

2.3 Successful Couplings and a Generalised Road Colouring Theorem

We will now introduce the notion of successful coupling, a kind of synchronising behaviour for couplings, and then compare it to the notion of strong μ-random walk.

Definition 10 (Successful Coupling)**.** Let E be a countable Polish space. A coupling Q is *successful* if for any $(x, y) \in E^2$, the Π-Markov chain (X, Y) has the following property:

$$\mathbb{P}_{x,y}(\exists n_0 \, , \; \forall n \geq n_0 \, , \; X_n = Y_n) = 1 \, . \tag{4}$$

In the general case of a Polish space (E, d), the notion of successful coupling is too demanding, and quite unrealistic as the diagonal $\{(x, x), x \in E\} \subset E \times E$

typically has a null measure. Thus, in the uncountable case, one may introduce a property like $\mathbb{P}_{x,y}(d(X_n, Y_n) \to 0) = 1$ instead.

Lemma 4 (Stickiness of a Successful Coupling)**.** *Consider some coupling Q, which is not necessarily sticky. We define the induced sticky coupling $\overline{Q}$ by*

$$\overline{Q}_{x,y}(u,v) = \begin{cases} Q_{x,y}(u,v) & \text{if} \quad x \neq y\,, \\ \Pi_x(u) & \text{if} \quad x = y \text{ and } u = v\,, \\ 0 & \text{if} \quad x = y \text{ and } u \neq v\,. \end{cases}$$

If Q is successful, then so is $\overline{Q}$.

It follows from this lemma that when studying synchronising behaviours, we can restrict ourselves to sticky couplings without loss of generality. However, while we will mostly consider symmetrical couplings, it is still unclear to me now whether the existence of a successful coupling is equivalent to the existence of its symmetrical counterpart or not, so these considerations will be left open for now.

Remark 2 (Exact Sampling as an Application of Couplings)**.** While the model of μ-random walks is quite recent and the results on it are sparce, the notion of successful couplings has been studied since the 1970s and is useful to add a notion of correlation into statistical models. One of the most known application of Markov chains coupling is exact sampling.

When an irreducible kernel is positive-recurrent, you can start from any point and let the walk run long enough to reach a probability distribution close to the invariant law. However, in such situations, the probability distribution is only an approximation of your invariant law and you do not always have a good control over the rate of convergence.

In some situations, using for example monotonicity properties of your transitions, you can find some successful coupling so that the probability of meeting on a given point is equal to the invariant law: this method is called exact sampling. More details and examples of applications can be found in [11].

Note that in the case of a colour law μ with synchronising support, you can similarly generate a random sequence of mappings according to μ and stop when you observe a synchronising word: the image of the obtained constant application follows the desired law.

We have introduced the basics of our notion of synchronising behaviour for couplings. Until the end of this section, we will see how this notion compares to the notion of strong μ-random walk.

Lemma 5 (Equivalence Between the Notions of Synchronisation in the Finite Case)**.** *We are here in the finite case, with $E = [\![1, n]\!]$. Consider some random colour law μ and Q the induced coupling. Then supp(μ) is synchronising if and only if the induced coupling Q is successful.*

Proof.

We have already seen that when the coupling Q is induced by a law μ, we can represent a Q-Markov chain (X,Y) as two μ-random walks (X,N) and (Y,N) driven by the same random mappings N. As a synchronising word appears in the random sequence N with probability 1, for any initial couple $(x,y) \in E$ we have $\mathbb{P}_{x,y}(\exists n\,,\ X_n = Y_n) = 1$, and by stickiness the coupling is thus successful.

The converse implication is a bit more convoluted, as it only gives pairwise synchronising sequences. In other words, because Q is successful, then for any starting couple (x,y), there is (at least) one word in $\mathrm{supp}(\mu)^+$ so that $\langle s\rangle\, x = \langle s\rangle\, y$. Consider s_2 so that $\langle s_2\rangle\, 1 = \langle s_2\rangle\, 2$. By induction, we pick s_{i+1} so that $\langle s_{i+1} \dots s_2\rangle\, i = \langle s_{i+1} \dots s_2\rangle\, (i+1)$. Then, $t = s_n \dots s_2 \in \mathrm{supp}(\mu)^+$ is synchronising word for E, thus $\mathrm{supp}(\mu)$ is synchronising. $\square$

We previously remarked that if some μ has synchronising support, then by maximality the law $\overline{\mu}$ we built also has synchronising support. Now, note that in the previous proof, each letter of the word t is a possible transition on the graph of Π, so it is in $\mathrm{supp}(\overline{\mu})$; in fact, even when Q is not induced by a random colour μ, we can similarly build such a word t with letters in $\mathrm{supp}(\overline{\mu})$. Therefore, the existence of a synchronising colouration and a successful coupling are equivalent in the finite case.

In the countable case, things get a bit more tedious. While the notions of successful coupling and of strong μ-random walk can be unambiguously generalised, how they relate to each other becomes quite uncertain. I proved the following result:

Theorem 2 (Generalisation of the Road Colouring Theorem)**.** *Let E be a countable space and μ a random colour law, ie a discrete probability measure (with countable support) over $\{\sigma : E \to E\}$, inducing an aperiodic irreducible positive recurrent kernel Π, and a coupling Q. Consider the following properties:*

1. *Q is successful;*
2. *any finite $E_0 \subset E$ has a synchronising word $s \in supp(\mu)^+$ such that $|\langle s\rangle\, E_0| = 1$; and*
3. *the μ-random walk (X,N) (unique up to identity in law) is strong.*

We have the direct implications $(1 \Rightarrow 2)$ *and* $(2 \Rightarrow 3)$.

Proof.
The implication $(1 \Rightarrow 2)$ comes from a direct induction: because Q is successful, for any initial couple (x,y), the coupled process (X,Y) which can be seen as two μ-random walk with the same transitions N is merged with probability 1. Thus, there is a synchronising word s for the pair $\{x,y\}$. Then, we can find t so that t synchronises $\{\langle s\rangle\, x, \langle s\rangle\, z\}$. Thus, ts synchronises $\{x,y,z\}$. By iterating this process, we obtain synchronising words for any finite subset of E.

This second property can be seen as the generalisation of a synchronising support $\mathrm{supp}(\mu)$ from the finite to the countable case. However, unlike the finite case, we cannot simply wait a (random) finite amount of time for N to merge E into one point.

Let us show $(2 \Rightarrow 3)$ now. Consider the stationary μ-random walk $(X_n, N_n)_{n\in\mathbb{Z}}$ and λ the common law of X. We define an enumeration $(x_n)_{n\in\mathbb{N}}$ of E and $E_i = \{x_n\,,\ n \le i\}$. Clearly $(\lambda(E_i))_{i\in\mathbb{N}}$ is strictly increasing to 1.

Now, for a given $i \in \mathbb{N}$, E_i has a synchronising word s_i. We define the stopping time τ_i so that s_i is a suffix of $(N_0, \dots, N_{\tau_i})$, without any other occurrence (by Borel-Cantelli lemma, this random time is well-defined in $-\mathbb{N}$).

We denote $\sigma_i = N_0 \circ \cdots \circ N_{\tau_i}$. We have $X_0 = \sigma_i(X_{\tau_i-1})$. Because the event $\{\omega \in \Omega\,,\ \tau_i(\omega) = k\}$ is measurable in $\sigma(N_j\,,\ j \ge k)$, by making a disjunction over the values of τ_i, one can easily prove the independence of the variables $\sigma_i \perp\!\!\!\perp X_{\tau_i-1}$ and that $X_{\tau_i-1} \overset{d}{=} \lambda$. Now we define $Y_i := \sigma_i(x_0) \in \sigma(N_n\,,\ n \le 0)$.

$$
\begin{array}{rcll}
\mathbb{P}(X_0 \ne Y_i) & = & \mathbb{P}(X_0 \ne \sigma_i(0)\,,\ X_{\tau_i-1} \in E_i) & +\ \ \mathbb{P}(X_0 \ne Y_i\,,\ X_{\tau_i-1} \in E_i^c) \\
& \le & \mathbb{P}(X_0 \ne \sigma_i(X_{\tau_i-1})\,,\ X_{\tau_i-1} \in E_i) & +\ \ \mathbb{P}(X_{\tau_i-1} \in E_i^c) \\
& \le & \mathbb{P}(X_0 \ne \sigma_i(X_{\tau_i-1})) & +\ \ \lambda(E_i^c) \\
& \le & 0 & +\ \ \underset{i\to\infty}{o}(1)
\end{array}
$$

Because of this, we have the convergence in probability $Y_i \xrightarrow{\mathbb{P}} X_0$. Up to some extraction ϕ, we have consequently $Y_{\phi i} \xrightarrow{\text{a.s.}} X_0$, hence $X_0 = \lim Y_{\phi i}$ is measurable in $\sigma(N_n\,,\ n \le 0)$, and the μ-random walk (X, N) is strong. $\square$

Remark 3 (Extraction of a Convergent Sequence)**.** In the previous theorem, we used the fact that if $Y_n \xrightarrow{\mathbb{P}} X_0$, then there exists a subsequence of $(Y_n)_{n\in\mathbb{N}}$ almost-surely converging to X_0, which is generally true on Polish spaces (see for example lemma 3.2 p. 40 in [12] *Foundations of modern probability*). However, this gives us no information at all on which subsequence is actually converging.

Assume that $\sum_{i\in\mathbb{N}} \lambda(E_i^c) = \sum_{i\in\mathbb{N}} i \times \lambda(x_i) < \infty$. Then, by the first Borel-Cantelli lemma, only a finite number of events $\{\omega\,,\ X_{\tau_i-1} \notin E_i\}$ will be realised, thus a random index $J(\omega)$ such that $\mathbb{P}(\forall i \ge J\,,\ X_{\tau_i-1} \in E_i) = 1$, which implies in turn that $\mathbb{P}(\forall i \ge J\,,\ X_0 = Y_i) = 1$. In other words, the sequence Y is almost surely stationary, thus $Y_i \xrightarrow{\text{a.s.}} X_0$. More generally, if your extraction ϕ is such that $\sum_{i\in\mathbb{N}} \lambda(E_{\phi i}^c) < \infty$, then $Y_{\phi i} \xrightarrow{\text{a.s.}} X_0$.

This result can also be adapted to any sequence of finite subsets $(\tilde{E}_i)_{i\in\mathbb{N}}$ such that $(\lambda(\tilde{E}_i))_{i\in\mathbb{N}}$ is strictly increasing to one.

We have seen that the synchronising behaviour of the coupling Q implies the synchronising behaviour of the support $\mathrm{supp}(\mu)$, which in turn implies the strongness of the μ-random walk in the positive-recurrent case. This justifies that in order to study μ-random walks, a good comprehension of couplings may be helpful.

2.4 Sufficient Condition on Random Colours for Successful Couplings

We will begin with an example of random colour on a countable set, to illustrate the results of the previous subsection:

Example 2 (Example of Successful Coupling)**.** Consider the case $E = \mathbb{Z}$. We define three mappings f, g and h on the figure below.

	$x < 0$	$x = 0$	$x > 0$
f	$x + 1$	x	$x - 1$
g	$x - 1$	$x + 1$	$x + 1$
h	$x - 1$	$x - 1$	$x + 1$

Figure 4: Support of the Random Colour Law

Let $0 < p < q < 1$, $p + q = 1$ and define the colour law $\mu = q\delta_f + \frac{p}{2}(\delta_g + \delta_h)$. The Markov kernel Π induced by μ is irreducible, positive-recurrent (because $p < q$) and aperiodic (because $f(0) = 0$ and $\mu(f) > 0$), thus the generalised road colouring theorem holds. Because $f^i([\![-i, i]\!]) = \{0\}$, the property (2) of the theorem is true and therefore the μ-random walk is strong.

Consider the invariant law λ such that $\lambda = \Pi\lambda$. When $x > 1$, this means $\lambda(x) = \lambda(x+1)\Pi_{x+1}(x) + \lambda(x-1)\Pi_{x-1}(x) = q\lambda(x+1) + p\lambda(x-1)$. Now, as $qX^2 - X + p = (X-1)(qX - p)$, the roots of the polynomial are 1 and $\frac{p}{q}$. On the right side, for $n \geq 1$, we have the behaviour $\lambda(n) = a \times 1 + b \times \left(\frac{p}{q}\right)^n$. As the total mass of $\mathbb{N}$ must be finite, we have $a = 0$. Now, because Π has a symmetrical effect on $\pm\mathbb{N}$, for $n \neq 0$ we have $\lambda(n) = \frac{\lambda(0)}{2} \times \left(\frac{p}{q}\right)^{|n|}$. Finally, $\lambda(0) = \frac{q-p}{q}$.

If we consider the sets $\tilde{E}_i = [\![-i, i]\!]$,

$$\sum_{i=0}^{\infty} \lambda(\tilde{E}_i^c) = \lambda(0) \times \frac{p}{q} \times \sum_{i=0}^{\infty} i \left(\frac{p}{q}\right)^{i-1} = \frac{p}{q-p} < \infty \,, \tag{5}$$

thus the previous remark holds.

Therefore, we do not only have the implication ($2 \Rightarrow 3$) here, but also obtain naturally an almost sure convergence without any intricate extraction. Now, if the conjecture ($3 \Rightarrow 1$) is true, the coupling Q induced by μ should be successful. That's what we will demonstrate now.

To prove it, we consider a coupled process $(X_n, Y_n)_{n\in\mathbb{N}}$ with the initial state $(x, y) \in \mathbb{Z}^2$. We introduce the random distance $d_n = |X_n - Y_n|$ which is adapted to the filtration $\mathcal{F}_n = \sigma(X_i\,,\ Y_i\,,\ 0 \leq i \leq n)$. By studying each possible configuration of the starting couple (x, y), we obtain

$$\begin{array}{rcl} \mathbb{E}[d_{n+1}|\mathcal{F}_n] & = & \mathbb{E}[d_{n+1}|X_n, Y_n]\,, \\ \mathbb{E}[d_{n+1}|\mathcal{F}_n] & \leq & d_n - (q-p)\chi(X_n \neq Y_n)\chi(X_n \leq 0 \leq Y_n \text{ or } X_n \geq 0 \geq Y_n)\,, \\ \mathbb{E}[d_{n+1}|\mathcal{F}_n] & \leq & d_n\,, \end{array}$$

hence $(d_n)_{n\in\mathbb{N}}$ is a supermartingale.

Now, by recurrence of the kernel Π, the stopping times $\tau_{-1} = -1$ and $\tau_{i+1} = \inf\{t > \tau_i\,,\ X_t = 0\}$ are such that, for any $i \in \mathbb{N}$, $X_{\tau_i} = 0$ almost surely, and $\{\omega\,,\ \tau_i = k\} \in \mathcal{F}_k$.

This implies that $\mathbb{E}[d_{\tau_i+1}] \leq \mathbb{E}[d_{\tau_i}] - (q-p)\mathbb{P}(Y_{\tau_i} \neq 0)$. Because (d_n) is a supermartingale, by induction, $0 \leq \mathbb{E}[d_{\tau_i}] \leq d_0 - (q-p) \times \sum_{j\in[\![0,i-1]\!]} \mathbb{P}(Y_{\tau_j} \neq 0)$. Therefore, for any $i \in \mathbb{N}$, we have

$$0 \leq (q-p) \times \sum_{j=0}^{i-1} \mathbb{P}(Y_{\tau_j} \neq 0) \leq d_0\,, \tag{6}$$

and thus $\sum_{j\in\mathbb{N}} \mathbb{P}(Y_{\tau_j} \neq 0) \leq \frac{d_0}{q-p} < \infty$.

By the first Borel-Cantelli lemma, only a finite amount of such events happens simultaneously, thus some random index J such that $\mathbb{P}(Y_{\tau_J} = 0) = 1$. This implies in particular that, if we denote $T = \tau_J$, then $\mathbb{P}_{x,y}(X_T = Y_T) = 1$ and so Q is a successful coupling.

This example shows that the study of couplings is not necessarily easier than the study of random colours, depending on the situation. The reasoning in this example can be generalised as a sufficient condition on μ for Q to be successful:

Proposition 1 (Sufficient Condition on μ for a Successful Coupling)**.** *Let (E, d_E) be a countable Polish space and μ a random colour law like in Theorem 2 above. We define the following properties:*

1. $\forall x, y \in E\,,\ \mathbb{E}_{x,y}[d_E(X_1, Y_1)] = \int_\Sigma d_E(\sigma x, \sigma y) d\mu(\sigma) \leq d_E(x, y)\,,$

2. $\exists x_0 \in E\,,\ \inf_{y \neq x_0} (d_E(x_0, y) - \mathbb{E}_{x_0,y}[d_E(X_1, Y_1)]) = \theta > 0\,.$

If these properties on μ are true, then the induced coupling Q is successful.

Proof.
Because of the first property, $d_n = d_E(X_n, Y_n)$ is a supermartingale for the filtration $\mathcal{F}_n$ like in the previous example. If we replace 0 by x_0 (given by the second property) in the example, we can still define the stopping times τ_i; now, if we replace $(q-p)$ by θ in the calculations, we obtain a proof of this proposition. $\square$

Note that we do not use the positive-recurrence of the kernel in this result but only its recurrence, so this sufficient condition may be used in a recurrent framework, where μ-random walks with infinite past do not necessarily exist.

The next logical step would be to prove the remaining implication ($3 \Rightarrow 1$); considering how long the proof is in the finite case, this is out of reach for now, assuming that ($3 \Rightarrow 1$) holds. This implication would also give us ($2 \Rightarrow 1$), ie one would only need to study the support of the law μ to conclude on the (non-)successful nature of the coupling Q, which is quite convenient. However, considering how long the proof is in the finite case, this seems beyond my current abilities, and will stay an open conjecture for now.

3 Successful Couplings for Countable Irreducible Kernels

In this last section, we will see some simple examples of couplings to get a better idea of how successful behaviours emerge in the countable case. A study of the general case of non-irreducible kernels allows dynamics too convoluted to be studied, so we will only see some typical behaviours of an irreducible kernel.

3.1 Successful Coupling for Positive-Recurrent Kernels

The following theorem echoes in fact the construction of $\overline{\mu}$ in the first section. In both situations, the basic idea is to consider independently a successor for each starting point. As a coupling Q considers only two starting points x and y (eventually equal to each other), this idea works pretty well. However, the probability law $\overline{\mu}$ has to consider a successor for *each* element of E: $\mathrm{supp}(\overline{\mu})$ is countable if and only if there is only a finite amount of states with more than one successor, exhibiting a non-deterministic transition under the kernel Π, which is quite a demanding.

Theorem 3. *We define the* independent sticky *coupling Q induced by Π by*

$$Q_{x,y}(u,v) = \begin{cases} \Pi_x(u) \times \Pi_y(v) & \textit{if} \quad x \neq y\,, \\ \Pi_x(u) & \textit{if} \quad x = y \textit{ and } u = v\,, \\ 0 & \textit{if} \quad x = y \textit{ and } u \neq v\,. \end{cases}$$

Now, if the transition kernel Π is irreducible positive-recurrent, then the independent sticky coupling Q is successful.

Proof.
Consider some starting couple (x,y), and two independent Π-Markov chains X and Z starting on x and y. We introduce the (eventually infinite) sticking time $T = \inf\{n \in \mathbb{N}\,,\ X_n = Z_n\} \in \overline{\mathbb{N}}$ and define the process Y by

$$Y_n := \begin{cases} Z_n & \text{if} \quad T \geq n\,, \\ X_n & \text{if} \quad T \leq n\,. \end{cases}$$

Clearly, (X,Y) is the (unique up to identity in law) Q-Markov chain starting on (x,y). What's more, $\mathbb{P}(\exists n\,,\ X_n = Y_n) = \mathbb{P}(T < \infty)$.

Now, thanks to Proposition 3 in Appendix B, the process (X, Z) is recurrent, which implies $\mathbb{P}(T < \infty) = 1$, thus Q is a successful coupling. $\square$

We have seen that the independent sticky coupling can always be induced by the probability law $\overline{\mu}$ induced by Π, but this probability law does not have a countable support in most of the cases. However, it is still possible to obtain a satisfying random colour, like in the last example of the previous section which used three colours only.

While the positive-recurrence of the kernel Π can be useful, it is not a requirement to have a successful coupling Q, as we will see in the next examples. However, without this assumption, we cannot guarantee the existence nor uniqueness of a μ-random walks with an infinite past; thus, while μ-random walks indexed on $\mathbb{N}$ may be studied in the following frameworks, the notion of strong process with an infinite past is not really adapted anymore.

3.2 Random Colour in a Drifting Transient Case

In the transient case, things get worse for random colours and successful couplings. We will now see an example of kernel which can be easily represented by a random colour law (with finite support) but which cannot be associated to any successful coupling. This example is the well-know unbalanced random walk on $\mathbb{Z}$:

Example 3 (Mirrored Random Walk on $\mathbb{Z}$ Drifting to $\pm\infty$)**.** Consider $p, q, r > 0$ such that $p + q + r = 1$. The Figure 5 clearly represents a transition law.

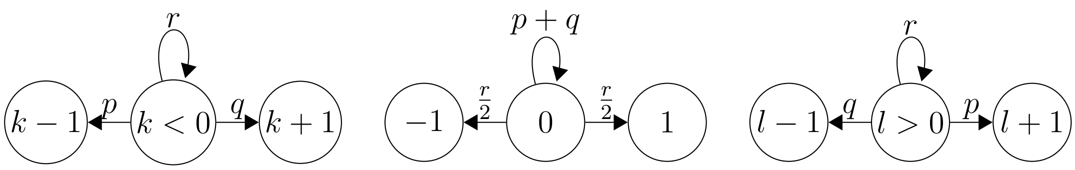

Figure 5: Symmetrical Transition Law

With the assumption $q > p$, we would obtain a positive-recurrent kernel, similar to the one in Example 2. Now we consider the case $p > q$, so that the kernel Π is transient instead of positive-recurrent. Note that the choice of Π_0, or even of a finite amount of transitions, is not important, as what matters is the asymptotic behaviour in $\pm\infty$.

Consider some coupling Q, a state $x > 0$ and (X, Y) the Q-Markov chain starting on the initial couple $(x, -x)$. Because both X and Y values cannot increase or decrease of more than one at each step, to have $X_j = Y_j$ at some time j, we must have $X_i = 0$ or $Y_i = 0$ at some prior time $i \leq j$. Hence, by symmetry of the transition laws:

$$\begin{aligned} \mathbb{P}(\exists n\,,\; X_n = Y_n) &\leq \mathbb{P}(\exists n\,,\; X_n = 0 \text{ and } Y_n = 0)\,, \\ &\leq 2 \times \mathbb{P}_x(\exists n\,,\; X_n = 0)\,. \end{aligned}$$

Now, by transience of the transition law in Figure 5 we have the convergence $\mathbb{P}_x(\exists n\,,\ X_n = 0) \underset{x\to\infty}{\longrightarrow} 0$. Thus there is some big enough x so that, for any coupling Q, we have $\mathbb{P}_{x,-x}(\exists n\,,\ X_n = Y_n) < 1$. Therefore, this mirrored kernel Π has no successful coupling.

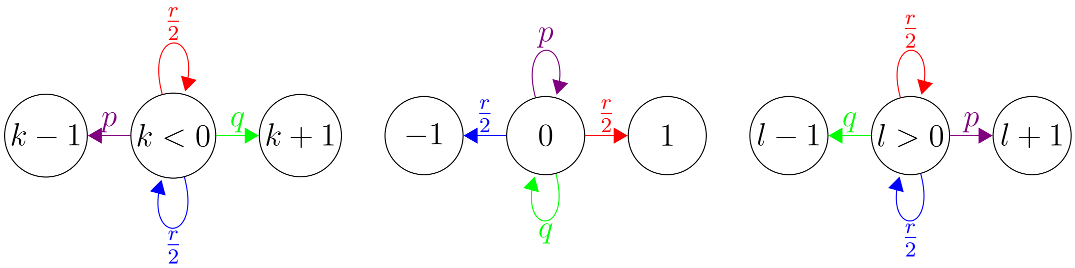

Figure 6: 4-Colouration of the Transition Law

Even in such a transient situation, we can define a random colour μ as in Figure 6, with a finite support which is minimal most of the time; it is however meaningless to study a μ-random walk with infinite past here.

The main issue in order to obtain a successful coupling here is that the model has two "limbs" ($\mathbb{Z}^+$ and Z^-) drifting in two clearly incompatible directions. However, if your system drifts in one direction only, there is still hope. For example, in the case of a random walk on $(\mathbb{Z}, +)$ driven by transitions of law $(p\delta_{-1} + q\delta_1) \in \mathcal{P}(\mathbb{Z})$ $(0 < p < q\,,\ p + q = 1)$, the independent sticky coupling is successful, because the distance between the two processes can be seen as a recurrent Markov chain; however, as already said, such a coupling doesn't induce a probability law on Σ with countable support.

3.3 Multidimensional Decomposition of a Coupling

The last scenario we will illustrate in this report is the case where a walk is transient but does not exhibit a drifting behaviour; instead, it wanders around in a space too arborescent for recurrence to appear. Note that this wandering behaviour is not incompatible with drifting in the general case, like in the uniform walk on the free group with two generators.

Example 4 (Successful Coupling on $\mathbb{Z}^3$)**.** It is known that the usual isotropic random walk on $\mathbb{Z}^3$ is transient. However, the projection on each axis has a recurrent behaviour. Thus, instead of a mere independent sticky coupling, we consider the symmetrical sticky coupling which chooses the same coordinate for both walks and then behaves like the independent sticky coupling over $\mathbb{Z}$ – which we already studied – on this coordinate.

For example, consider some vectors $\vec{x}$ and $\vec{y}$ so that $x_1 \neq y_1$ but $x_2 = y_2$ and $x_3 = y_3$. The coupling $Q_{\vec{x},\vec{y}}$ is, under these assumptions, equal to the figure below.

$Q_{\vec{x},\vec{y}}$	$\vec{y}+e_1$	$\vec{x}-e_1$	$\vec{x}+e_2$	$\vec{x}-e_2$	$\vec{x}+e_3$	$\vec{x}-e_3$
$\vec{x}+e_1$	1/12	1/12	0	0	0	0
$\vec{x}-e_1$	1/12	1/12	0	0	0	0
$\vec{x}+e_2$	0	0	1/6	0	0	0
$\vec{x}-e_2$	0	0	0	1/6	0	0
$\vec{x}+e_3$	0	0	0	0	1/6	0
$\vec{x}-e_3$	0	0	0	0	0	1/6

Figure 7: Coupling Q When $x_1 \neq y_1$, $\hat{x}_1 = \hat{y}_1$

With this coupling Q, it is quite clear that if $\left(\vec{X}_0 - \vec{Y}_0\right) \in (2\mathbb{Z})^3$, then $\mathbb{P}\left(\exists n\,,\ \vec{X}_n = \vec{Y}_n\right) = 1$.

The key idea of this example is to remark that while we are following a random walk on $\mathbb{Z}^3 = \mathbb{Z}_1 \times \mathbb{Z}_2 \times \mathbb{Z}_3$, at each step we only really perform a random walk on one of the $\mathbb{Z}_i$. Thus, more generally, it may be good to seek a good multidimensional decomposition of your initial space to use the following theorem:

Theorem 4 (Multidimensional Coupling). *Consider some countable space E and a kernel $\Pi : E \to \mathcal{P}(E)$ which admits a multidimensional decomposition: there is a law $\lambda \in \mathcal{P}(\llbracket 1, r \rrbracket)$ and kernels $\Pi^{(i)} : E_i \to \mathcal{P}(E_i)$ such that $E = E_1 \times \cdots \times E_r$ and $\Pi_{\vec{x}}(\vec{y}) = \sum_{j=1}^{r} \delta_{\hat{x}_j, \hat{y}_j} \times \lambda(j) \times \Pi^{(j)}_{x_j}(y_j)$.*

In other words, we first choose a coordinate i according to the law λ and then apply a transition driven by the kernel $\Pi^{(i)}$ on this coordinate of $\vec{x}$.

If each $\Pi^{(i)}$ has a successful coupling $Q^{(i)}$, then the coupling Q obtained by choosing at random i following λ and then applying $Q^{(i)}$ to the i-th coordinate is also successful. What's more, if each $Q^{(i)}$ is induced by a random colour $\mu^{(i)}$ then Q is also induced by a colour law μ with countable support.

Note that, while the reasoning in the previous theorem can be generalised to a countable product of spaces without struggle, the whole space E becomes uncountable then.

This kind of multidimensional decomposition reminds me of some game theory models, where each E_i represents a game and where the players play a game at random at each turn. Up until now we mostly considered symmetrical couplings; however, with this framework in mind, we could see a coupling Q as the strategy of a second player, following the same sequence of games as the first one, with the intent of eventually reaching the same result.

At the beginning of the second section, we wondered whether the existence of a successful coupling is equivalent to the existence of its symmetrical counterpart or not. This kind of methods may allow us to reach an answer. However, game theory results are generally quite abstract and have strong hypothesis, so this study clearly exceeds my current reach.

A Sketch of the Proof of the Road Colouring Theorem

While not explicitly used in the proof, the following result (p. 263 in [2]) is the most essential of this section: all the proof does is to recreate a similar setting starting from the hypothesis of a non-synchronising support $\mathrm{supp}(\mu)$, which is more general than using solely permutations:

Proposition 2 (Weakness of a Permutation Walk). *Consider some random colour law μ so that $\mathrm{supp}(\mu)$ contains exclusively permutations over V – the support of μ is as far from the synchronising case as possible.*

As $(\Sigma, \circ)$ is a semigroup, we can define the convolution of probability laws $\mu^{(1)}, \mu^{(2)} \in \mathcal{P}(\Sigma)$ by

$$\mu^{(1)} * \mu^{(2)}(\sigma) := \sum_{f,g\in\Sigma\,,\; f\circ g=\sigma} \mu^{(1)}(f) \times \mu^{(2)}(g)\,.$$

With this notion in mind, we do not make any assumptions on the induced graph G itself, but we assume that for any transition from x to $y \in V$, we have

$$\mu^{*n}\left(\{\sigma\,,\ \sigma(x) = y\}\right) \underset{n\to\infty}{\longrightarrow} \frac{1}{|V|}\,.$$

Under these assumptions, the common law of X is the uniform probability law $\lambda = \mathcal{U}(V)$ and at any given time $n \in \mathbb{Z}$ we have the independence $X_n \perp\!\!\!\perp \sigma(N_j, j \in \mathbb{Z})$, thus a clearly non-strong random walk.

For any word $s = (\sigma_1 \ldots \sigma_r) \in \mathrm{supp}(\mu)^+$ we denote $\langle s\rangle = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_r$ the composition of the mappings.

Now, in the general framework of the theorem, assume that $\mathrm{supp}(\mu)$ is non-synchronising. This assumption implies that there is some minimal cardinality $\hat{m} = \min_{s\in\mathrm{supp}(\mu)^+} |\langle s\rangle\, V| \geq 2$; note that in Proposition 2, we are in the maximal case where $\hat{m} = |V|$.

The key is then to introduce some carefully chosen s so that $|\langle s\rangle\, V| = \hat{m}$. For such a word, there is some enumeration $\langle s\rangle\, V = \{x_i\,,\ i \in \hat{V}\}$ with $\hat{V} = [\![1, \hat{m}]\!]$. Now, the trick is to introduce stopping times at each occurrence of s in the random sequence N. This induces a process $(\hat{X}, \hat{N})$ where $\hat{X}$ is $\hat{V}$-valued, and by minimality of $\hat{m}$, the random colours $\hat{N}$ over $\hat{V}$ are in fact permutations. This induced process is in fact a non-strong $\hat{\mu}$-random walk as in proposition 2.

This final result is stated on p. 279 in [2] and roughly states that the weakness of $(\hat{X}, \hat{N})$ can be propagated to (X, N):

Theorem 5 (Weakness of the Main Process). *Let $k \in \mathbb{Z}$. Consider $K(k)$ the random time of the most recent occurrence of s in the sequence N before time k. Then $X_k \in \sigma(\hat{X}_{K(k)}\,,\ N_j\,,\ j \in \mathbb{Z})$ has a genuine dependence on $\hat{X}_{K(k)}$.*

Because $\hat{X}_{K(k)} \perp\!\!\!\perp \sigma(N_j\,,\ j \in \mathbb{Z})$ and $\hat{X}_{K(k)}$ is a non-constant variable, we have $X_k \notin \sigma(N_j\,,\ j \leq k)$. Hence a μ-random walk (X, N) which is non-strong.

B Positive-Recurrence of Independent Couplings

The next result is well-known; as I proved it implicitly while working on Theorem 3, I will include here my proof for completeness of this report.

Proposition 3. *Consider two (countable) positive-recurrent irreducible aperiodic kernels $\Pi^{(1)}$ over E and $\Pi^{(2)}$ over F, of invariant laws $\lambda \in \mathcal{P}(E)$ and $\mu \in \mathcal{P}(F)$. Consider their independent coupling $Q_{x,y}(u,v) := \Pi^{(1)}_x(u) \times \Pi^{(2)}_y(v)$. Then Q is also a positive-recurrent aperiodic kernel, of invariant law $\lambda \otimes \mu$.*

Proof.
Consider a Q-Markov chain (X,Y) starting from (x,y), or in other words X a $\Pi^{(1)}$-Markov chain and Y a $\Pi^{(2)}$-Markov so that $X \perp\!\!\!\perp Y$. Clearly, $\lambda \otimes \mu$ is an invariant law for Q, and the coupled chain is irreducible aperiodic, so we only need to show that, almost surely, (X,Y) takes the value $(x,y) \in E \times F$ again at some positive time.

We will instead study the probability of never going back to this state. For a sequence of positive integers $(n_i)_{i>0}$ and an initial couple $(u,v) \in E \times F$ we define

$$\phi_{u,v}(n_1, n_2, \dots) := \mathbb{P}_{u,v}\left((X_{n_1}, Y_{n_1}) \neq (x,y), (X_{n_1+n_2}, Y_{n_1+n_2}) \neq (x,y), \dots\right) \leq 1\,.$$

We have the simple majoration $\mathbb{P}_{x,y}(\forall n > 0, (X_n, Y_n) \neq (x,y)) \leq \inf_{(n_i)} \phi_{x,y}(n_1 \dots)$. Now, we can use Markov property on ϕ.

$$\begin{aligned}
\phi_{i,j}(n_1 \dots) &= \sum_{(u,v)\neq(x,y)} \mathbb{P}_i(X_{n_1} = u) \times \mathbb{P}_j(Y_{n_1} = v) \times \phi_{u,v}(n_2 \dots) \\
&\leq \sum_{(u,v)\neq(x,y)} \lambda(u) \times \mathbb{P}_j(Y_{n_1} = v) \times \phi_{u,v}(n_2 \dots) + |\mathbb{P}_i(X_{n_1} = u) - \lambda(u)| \\
&\leq \left(\sum_{(u,v)\in E\times F} \lambda(u) \times \mu(v) \times \phi_{u,v}(n_2 \dots) \right) \\
&\quad + 2d_{TV}(X_{n_1}|_{X_0=i}, \lambda) + 2d_{TV}(Y_{n_1}|_{Y_0=j}, \mu)
\end{aligned}$$

Let $\epsilon > 0$. By σ-additivity, we can find some finite subset $A_\epsilon \subset E$ (resp. $B_\epsilon \subset F$) so that $\lambda(A_\epsilon) \geq 1 - \epsilon$ (resp. $\mu(B_\epsilon) \geq 1 - \epsilon$). By positive-recurrence of $\Pi^{(1)}$ and $\Pi^{(2)}$ we can also find some (deterministic) time $n(i,j)$ so that, for any n_1 bigger than n, we have $d_{TV}(X_n|_{X_0=i}, \lambda) \leq \epsilon$ and $d_{TV}(Y_n|_{Y_0=j}, \mu) \leq \epsilon$, hence

$$\phi_{i,j}(n_1 \dots) \leq \left(\sum_{(u,v)\in A_\epsilon \times B_\epsilon} \lambda(u) \times \mu(v) \times \phi_{u,v}(n_2 \dots) \right) + 6\epsilon\,.$$

Note that λ and μ are not trivial probabilities, thus $\langle \lambda, \mu \rangle_{L^2} \leq ||\lambda||_\infty < 1$. Finally, by replacing ϵ by $\frac{\epsilon}{6}$, we can obtain some finite set $C \subset E \times F$ so that, for any $(i,j) \in E \times F$, we have some time n so that for any (n_i) such that $n_1 \geq n$, we have the inequality

$$\phi_{i,j}(n_1 \dots) \leq ||\lambda||_\infty \times \sup_{(i,j)\in C} \phi_{i,j}(n_2 \dots) + \epsilon \leq ||\lambda||_\infty + \epsilon \,.$$

Because C is finite, we have $n_2 := \max_{(i,j)\in C} n(i,j) < \infty$. Consider the constant sequence $s = (n_2)_{i>0}$. By taking the supremum over C on the left term above, it appears that

$$\begin{aligned} \sup_{(i,j)\in C} \phi_{i,j}(s) &\leq ||\lambda||_\infty \times \sup_{(i,j)\in C} \phi_{i,j}(s) + \epsilon \,, \\ \sup_{(i,j)\in C} \phi_{i,j}(s) &\leq \tfrac{\epsilon}{1-||\lambda||_\infty} \,. \end{aligned}$$

Thus, $\phi_{x,y}\left(n(x,y), s\right) \leq ||\lambda||_\infty \sup_C \phi_{i,j}(s) + \epsilon \leq \left(\frac{||\lambda||_\infty}{1-||\lambda||_\infty} + 1\right)\epsilon$. Finally, as this kind of inequality holds for any $\epsilon > 0$, we have $\mathbb{P}_{x,y}(\forall n > 0, (X_n, Y_n) \neq (x,y)) = 0$. $\square$

Note that consequently, if $\Pi^{(1)}$ is i-periodic and $\Pi^{(2)}$ is j-periodic (instead of 1-periodic), then Q is positive-recurrent $\mathrm{lcm}(i,j)$-periodic.

C List of Notations

I	interval of integers in $\mathbb{Z}$
$\mathbb{N}$	set of non-negative integers
$[\![a, b]\!]$	interval of integers $\{n \in \mathbb{Z}\,,\ a \le n \le b\}$
$\mathbb{R}^+$	set of non-negative real numbers
V	finite set
G	directed graph allowing superposed edges
Σ	finite set of transformations over V
σ	mapping in Σ, also called a colour
$\mathcal{P}(S)$	set of probabilities over a measurable space S
μ	random colour law in $\mathcal{P}(\Sigma)$ with countable support
P	transition law over V
λ	invariant law of a transition law P
X, Y	Markov chains
N	random colour of law μ
S^+	set of the words of finite length on the alphabet S
(E, d)	Polish space
$\mathcal{B}(S)$	Borel algebra of a topological space S
Π	Markov kernel over E
Q	coupling of a kernel Π
$\vec{x}$	vector in $E_1 \times \cdots \times E_r$
$\hat{x}_j$	vector $(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots x_r)$ induced by $\vec{x}$

References

[1] Kouji Yano and Kenji Yasutomi. Realization of an ergodic Markov chain as a random walk subject to a synchronizing road coloring. *J. Appl. Probab.*, 48(3):766–777, 2011.

[2] Kouji Yano. Random walk in a finite directed graph subject to a road coloring. *J. Theoret. Probab.*, 26(1):259–283, 2013.

[3] Kouji Yano and Kenji Yasutomi. Random walk in a finite directed graph subject to a synchronizing road coloring. *ArXiv e-prints*, May 2011.

[4] Marc Yor. Tsirel'son's equation in discrete time. *Probab. Theory Related Fields*, 91(2):135–152, 1992.

[5] Jean Brossard and Christophe Leuridan. Chaînes de Markov indexées par $-\mathbf{N}$: existence et comportement. *Ann. Probab.*, 29(3):1033–1046, 2001.

[6] Ramon van Handel. On the exchange of intersection and supremum of σ-fields in filtering theory. *Israel J. Math.*, 192(2):763–784, 2012.

[7] A. N. Trahtman. The road coloring problem. *Israel Journal of Mathematics*, 172(1):51–60, Jul 2009.

[8] Jordan Bell. The Kolmogorov extension theorem. http://individual.utoronto.ca/jordanbell/notes/kolmogorov.pdf.

[9] Sean Meyn and Richard L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. With a prologue by Peter W. Glynn.

[10] D. Revuz. *Markov chains*, volume 11 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1984.

[11] James Gary Propp and David Bruce Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996.

[12] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.