

# On Aldous's Question about the Metropolis Chain on Cayley Graphs: Non-Monotonicity of the Relaxation Time

Alper Ferudun\*

October 9, 2026

## Abstract

The list of open problems of D. Aldous contains the following question, dated March 2009. On a finite Cayley graph let  $\mu_p$ ,  $0 < p < 1$ , be the law of  $X(T_p - 1)$ , where  $X$  is the random walk started at the identity and  $T_p$  is a geometric time with parameter  $p$ ; let  $\mu_0$  be the uniform law; and let  $\tau(p)$  be the relaxation time of the Metropolis chain based on the random walk with stationary law  $\mu_p$ . Can anything be proved about  $\tau(p)$  in this generality, for instance (i) that  $\tau(p)$  is monotone decreasing in  $p$ , or (ii) that  $\tau(p) \leq C\tau(\infty)$  for a universal constant  $C$ ? The symbol  $\tau(\infty)$  is not defined in the source. We give a partial answer: items (i) and (ii), as literally posed, have negative answers, for the relaxation time  $1/(1 - \lambda_2)$  and for both readings of  $\tau(\infty)$ , namely  $\tau(0)$  and  $\lim_{p \rightarrow 1} \tau(p)$ . We prove, on every finite Cayley graph of degree  $d$  with identity  $e$ : (1)  $\tau(p) \rightarrow d$  as  $p \rightarrow 1$ ; (2) for every  $p$ ,  $\tau(p) \geq 2 \text{Var}(\text{dist}(e, X(T_p - 1)))$  and  $\tau(p) \geq (1 - p)m(1 - m)/(m - p)$  with  $m = \mu_p(e)$ ; (3)  $1/\tau(p) = 1/\tau(0) - s^*p + O(p^2)$  as  $p \rightarrow 0$ , with an explicit constant  $s^* > 0$ , so that  $\tau(p) > \tau(0)$  for all small  $p > 0$ . By (3), item (i) fails on every finite Cayley graph. On the hypercube of dimension  $d$  we have  $\tau(0) = d/2$ ,  $\lim_{p \rightarrow 1} \tau(p) = d$  and  $\tau(1/d) \geq d(d - 1)/15$ ; and on Cayley graphs of bounded degree whose walk eigenvalues other than  $\pm 1$  are bounded away from  $\pm 1$ ,  $\sup_p \tau(p) \geq c(\log n)^2$ , where  $n$  is the number of vertices. Hence (ii) fails under both readings. On the complete graph  $K_n$  one has  $\tau(p) = (n - 1)(1 + (n - 2)p)/(n - p)$ , which is strictly increasing; this is a corollary of a known formula (Liu; Diaconis and Saloff-Coste; Aldous and Fill), and since it already contradicts item (i) and, under the first reading, item (ii), the question that was intended may differ from the literal one. The methods are standard. The third item of the source, which asks for a decreasing bound on  $\tau(p)$ , and its open-ended opening question are not answered. This is an unrefereed note.

## 1 Introduction

### 1.1 The problem

The list of open problems maintained by D. Aldous contains a page entitled “Metropolis on Cayley graphs” [Ald09]. Its content is the following. On a finite Cayley graph let  $X(0), X(1), \dots$  be the random walk started at the identity, and let  $T_p$  have the geometric distribution with parameter  $p$ . For  $0 < p < 1$  let  $\mu(p)$  be the distribution of  $X(T_p - 1)$ , and take  $\mu(0)$  to be the uniform distribution. Consider the Metropolis chain, based on the random walk, with stationary distribution  $\mu(p)$ , and let  $\tau(p)$  be its relaxation time. The page asks whether one can prove anything about  $\tau(p)$  in this general setting, and it names three instances:

- (i) whether “ $\tau(p)$  is monotone decreasing in  $p$ ”;

---

\*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: <https://github.com/AlperTheKing>.  
Unrefereed preprint. CC BY 4.0.

- (ii) whether  $\tau(p) \leq C \tau(\infty)$  for some universal constant  $C$ ;
- (iii) whether there is some decreasing bound on  $\tau(p)$  in terms of  $\tau(\infty)$ , of  $p$ , of the number of states and of familiar parameters of the graph.

On the page the third item carries the label (ii) a second time; we call it (iii). According to its history line the page was first written in March 2009. It gives no reference and no further definition. In the index of the list it is entry “(1.9)” among the open research problems. The problem is record AMR-096-0008 of the UnsolvedMath collection [ula26], where it is listed as open.

Before the question can be read, the following has to be settled.

*The symbol  $\tau(\infty)$ .* The parameter  $p$  ranges over  $[0, 1)$ , and the value  $\infty$  does not occur. We treat both possible readings:

- (a)  $\tau(\infty) = \tau(0)$ , the relaxation time of the random walk itself;  $p = 0$  is the end at which the geometric time is infinite;
- (b)  $\tau(\infty) = \tau(1^-) := \lim_{p \uparrow 1} \tau(p)$ , the end at which  $\mu(p)$  concentrates at the identity; this is the end “ $\infty$ ” if the family is parametrised by a rate (Remark 2.2(c)).

Under reading (a) the three items form a chain: (i) would give (ii) with  $C = 1$ , and (iii) asks for a quantitative form of (ii).

*The orientation of  $p$ .* Since one symbol of the page does not fit its parametrisation, we also record what happens to (i) when the orientation of  $p$  is reversed, that is, for the statement that  $\tau$  is non-decreasing in  $p$ . This is not what the page says.

The generating set, the geometric time and the relaxation time are fixed in the next subsection, in the way that is compatible with the page and with the monograph [AF02] of Aldous and Fill.

## 1.2 Conventions

Throughout,  $G$  is a finite group with  $n \geq 2$  elements and identity  $e$ , and  $S \subset G \setminus \{e\}$  is a generating set with  $S = S^{-1}$ ; we put  $d = |S|$ . The Cayley graph has the edges  $\{x, xs\}$ ,  $x \in G$ ,  $s \in S$ ; it is a connected  $d$ -regular graph without loops and multiple edges. The random walk has the kernel

$$K(x, y) = \frac{1}{d} \text{ if } x^{-1}y \in S, \quad K(x, y) = 0 \text{ otherwise.}$$

$K$  is symmetric and irreducible, and  $K(ax, ay) = K(x, y)$  for all  $a \in G$ . Its eigenvalues are  $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_n \geq -1$ . We write  $|x|$  for the word length of  $x \in G$  with respect to  $S$ , which is the graph distance from  $e$  to  $x$ . The page does not say that  $S$  is symmetric or that  $e \notin S$ ; this is the usual meaning of a random walk on a Cayley graph, and a lazy walk or a generating set containing  $e$  only reparametrises the family (Remark 2.2).

*The geometric time.* The expression  $X(T_p - 1)$  needs  $T_p \geq 1$ , and the convention that  $\mu(0)$  is uniform needs  $T_p \rightarrow \infty$  as  $p \rightarrow 0$ . Hence  $\mathbb{P}(T_p = k) = p(1 - p)^{k-1}$  for  $k = 1, 2, \dots$ , and,  $T_p$  being independent of  $X$ ,

$$\mu_p(x) := \mathbb{P}(X(T_p - 1) = x) = p \sum_{k \geq 0} (1 - p)^k K^k(e, x) \quad (0 < p < 1), \quad \mu_0(x) := \frac{1}{n}. \quad (1)$$

By Lemma 2.1(c),  $\mu_p \rightarrow \mu_0$  as  $p \downarrow 0$  on every Cayley graph, bipartite or not, so that  $\mu_0$  is the continuous extension; and  $\mu_p \rightarrow \delta_e$  as  $p \uparrow 1$ . We write  $\mu_p$  for the measure  $\mu(p)$  of the page.

The Metropolis chain [MRR<sup>+</sup>53, Has70] based on the symmetric kernel  $K$  with stationary distribution  $\mu_p$  is, as in [AF02, Section 11.2.1],

$$P_p(x, y) = K(x, y) \min\left\{1, \frac{\mu_p(y)}{\mu_p(x)}\right\} \quad (y \neq x), \quad P_p(x, x) = 1 - \sum_{y \neq x} P_p(x, y). \quad (2)$$

It is reversible with respect to  $\mu_p$  and irreducible (Section 2.2), so that its eigenvalues are real,  $1 = \beta_1(p) > \beta_2(p) \geq \dots \geq \beta_n(p) \geq -1$ .

The relaxation time is, as in [AF02, (3.38)–(3.39)],

$$\tau(p) = \frac{1}{1 - \beta_2(p)}, \quad (3)$$

which is also the relaxation time of the continuous-time chain with generator  $P_p - I$ . In particular  $P_0 = K$  and  $\tau(0) = 1/(1 - \lambda_2)$ . All statements of this note are for the definition (3). The variant with the absolute spectral gap is discussed in Remark 7.2. The function  $\tau$  is continuous on  $[0, 1)$  (Corollary 2.5).

### 1.3 Results

The first statement is known; it is placed here because it already answers the literal items.

**Proposition 1.1** (complete graph; known). *Let  $S = G \setminus \{e\}$ , so that the Cayley graph is the complete graph  $K_n$ ,  $n \geq 2$ . Then  $\mu_p(e) = \frac{1+(n-2)p}{n-p}$ ,  $\mu_p(x) = \frac{1-p}{n-p}$  for  $x \neq e$ , and*

$$\tau(p) = (n-1) \mu_p(e) = \frac{(n-1)(1+(n-2)p)}{n-p} = \frac{n-1}{n} + \frac{(n-1)^3}{n} \cdot \frac{p}{n-p} \quad (0 \leq p < 1).$$

Hence  $\tau$  is strictly increasing on  $[0, 1)$ ,  $\tau(0) = \frac{n-1}{n}$ ,  $\tau(1^-) = n-1$ , and  $\sup_p \tau(p)/\tau(0) = n$ .

On  $K_n$  the Metropolis chain based on the walk with *any* target  $\pi$  has relaxation time  $(n-1) \max_x \pi(x)$  (Proposition 3.1). This is the complete-graph form of the eigenvalue analysis of Metropolized independent sampling by Liu [Liu96], quoted as Theorem 2.2 in the survey [DSC98] of Diaconis and Saloff-Coste, and of [AF02, Section 11.4.2, Proposition 11.2(b)]. Only the evaluation of  $\mu_p(e)$ , a geometric series, is specific to the problem.

**Theorem 1.2** (the limit  $p \rightarrow 1$ ). *For every finite Cayley graph of degree  $d$ ,  $\lim_{p \uparrow 1} \tau(p) = d$ .*

**Theorem 1.3** (two lower bounds). *For every finite Cayley graph and every  $0 < p < 1$ , with  $m = \mu_p(e)$ :*

$$(a) \quad \tau(p) \geq \frac{(1-p)m(1-m)}{m-p}, \text{ with equality on the complete graphs;}$$

$$(b) \quad \tau(p) \geq 2 \operatorname{Var}(|X(T_p - 1)|), \text{ twice the variance of the distance from } e \text{ under } \mu_p.$$

**Corollary 1.4** (hypercube). *Let  $G = (\mathbb{Z}/2)^d$  with the  $d$  unit vectors as generators,  $d \geq 2$ , and put  $q = 1 - p$ . Then  $\tau(0) = d/2$ ,  $\tau(1^-) = d$ , and for  $0 < p < 1$*

$$\tau(p) \geq H_d(p) := d \cdot \frac{pd + 2q}{pd + 4q} \cdot \left(1 + \frac{pq d^2}{(pd + 2q)^2}\right). \quad (4)$$

In particular:

$$(a) \quad \tau(1/d) \geq d(d-1)/15;$$

(b)  $\tau(p) \geq (1-p)d/(3p)$  whenever  $pd \geq 4$ ;

(c)  $H_d(p) > d$  if and only if  $p > 4/(d^2 - 2d + 4)$ ; hence, if  $d \geq 3$ , then  $\tau(p) > d = \tau(1^-)$  for all  $p$  with  $4/(d^2 - 2d + 4) < p < 1$ .

Consequently  $\sup_p \tau(p)/\tau(0) \geq 2(d-1)/15$  and  $\sup_p \tau(p)/\tau(1^-) \geq (d-1)/15$ .

**Corollary 1.5** (bounded degree, two-sided spectral gap). *Let  $d \geq 2$  and  $n \geq 32d^6$ , let  $\rho$  be the largest absolute value of an eigenvalue of  $K$  different from  $\pm 1$ , and put  $\kappa = \min\{1, \log(1/\rho)/\log d\}$ . Then*

$$\sup_{0 < p < 1} \tau(p) \geq \frac{\kappa}{40} (\log_d n - 6)^2, \quad \text{while} \quad \tau(0) \leq \frac{1}{1-\rho} \quad \text{and} \quad \tau(1^-) = d.$$

Consequently, along every sequence of Cayley graphs with fixed degree  $d$ , with  $\rho \leq \rho_0 < 1$  and with  $n \rightarrow \infty$ , both  $\sup_p \tau(p)/\tau(0)$  and  $\sup_p \tau(p)/\tau(1^-)$  tend to infinity, at least like  $(\log n)^2$ .

The expander property used is thus the following: all eigenvalues of the walk other than 1 and (in the bipartite case)  $-1$  have absolute value at most  $\rho_0 < 1$ . Sequences of Cayley graphs of fixed degree with this property exist; for instance the Ramanujan graphs of Lubotzky, Phillips and Sarnak [LPS88] are Cayley graphs of the groups  $\text{PSL}_2$  or  $\text{PGL}_2$  over prime fields, of degree  $\ell + 1$  and with  $\rho \leq 2\sqrt{\ell}/(\ell + 1)$ . We quote this existence statement; it was not checked for this note.

**Theorem 1.6** (the increase at  $p = 0$ ). *For every finite Cayley graph there is a constant  $s^* > 0$  such that*

$$\frac{1}{\tau(p)} = \frac{1}{\tau(0)} - s^*p + O(p^2) \quad (p \downarrow 0). \quad (5)$$

In particular  $\tau$  has the right derivative  $s^*\tau(0)^2 > 0$  at  $p = 0$ , and  $\tau(p) > \tau(0)$  for all sufficiently small  $p > 0$ . Explicitly, let  $V$  be the eigenspace of  $K$  for  $\lambda_2$ , let  $m_2 = \dim V$ , and let  $g: G \rightarrow \mathbb{R}$  be the solution of  $(I - K)g = \delta_e - \frac{1}{n}$  with  $\sum_x g(x) = 0$ . Then

$$s^* = \max \left\{ Q(f) : f \in V, \sum_x f(x)^2 = 1 \right\}, \quad (6)$$

$$Q(f) = \frac{nf(e)^2 - 1}{2} + \frac{n}{4} \sum_{x,y} K(x,y) |g(x) - g(y)| (f(x) - f(y))^2,$$

and  $s^* \geq Q(f_0) > \frac{m_2-1}{2} \geq 0$ , where  $f_0$  is the normalised orthogonal projection of  $\delta_e$  to  $V$ .

Theorems 1.2, 1.3 and 1.6 and Corollary 1.5 hold, with the same proofs, for the simple random walk on every finite connected vertex-transitive graph (Remark 2.6).

*Consequences for the question.* They are collected in Table 1 in Section 7. For the relaxation time (3), item (i) as literally posed fails on every finite Cayley graph (Theorem 1.6). With the orientation of  $p$  reversed it fails on the cycles  $C_n$ ,  $n \geq 7$  (Corollary 4.1), and on the hypercubes of dimension  $d \geq 3$  (Corollary 1.4(c); for  $d \geq 17$  already by Corollary 1.4(a), since then  $d(d-1)/15 > d$ ). Item (ii) fails under both readings on hypercubes and on Cayley graphs of bounded degree with a two-sided spectral gap (Corollaries 1.4 and 1.5). Item (iii) and the opening question are open-ended, and they are not answered here.

## 1.4 What is known, what is new, and what stays open

*Known.* Proposition 1.1 is a corollary of the results of [Liu96, DSC98, AF02] quoted after it. Two families of Metropolis chains on Cayley graphs have been analysed explicitly in print since 1992: Diaconis and Hanlon [DH92] treat the targets proportional to  $\theta^{|x|}$ ,  $0 < \theta \leq 1$ , on the hypercube

and on the symmetric group generated by all transpositions, and they give the eigenvalues of the chains to which the Metropolis chain can be lumped, the chain of the Hamming weight and the chain of the conjugacy class. On the hypercube the relaxation time is  $d/(1 + \theta)$ . In both families the relaxation time of the lumped chain increases as the target concentrates ( $\theta \downarrow 0$ ), from the relaxation time of the walk up to the degree of the graph. These are prior instances of the phenomenon behind Theorems 1.2 and 1.6; the formulas, and what they give for the Metropolis chains themselves, are in Remark 4.2. The survey [DSC98] points to them. We read [DH92] in its technical-report version. The targets  $\mu_p$  are in general not of the form  $\theta^{|x|}$ , and neither [DH92] nor [DSC98] contains the family  $\mu_p$  or Theorems 1.2, 1.3 and 1.6. For another dynamics, the Glauber dynamics of the mean-field Ising model, Kargin [Kar11] proves that the relaxation time decreases monotonically in the temperature; in the family  $\mu_p$  the relaxation time likewise increases at the uniform end  $p = 0$  (Theorem 1.6), but it is in general not monotone on  $[0, 1)$  (Corollary 4.1(b)).

*Methods.* They are standard. Theorem 1.2 is a zero-temperature limit, in which the limit matrix is block triangular; low-temperature asymptotics of relaxation times of Metropolis algorithms is a classical subject, see for instance [Mic02]. Theorem 1.3 is the variational characterisation of the relaxation time with the test functions  $\mathbf{1}_{\{e\}}$  and  $|x|$ . Theorem 1.6 is first-order perturbation theory of a multiple eigenvalue of a symmetric matrix [Rel37, Kat95].

*New, as far as we found.* We found no treatment of the family  $\mu_p$  in print and no citation of the problem page; the searches are described in the paragraph “Scope and priority” at the end. A search that finds nothing is not a proof of novelty, and no priority is claimed.

*A caveat on the literal reading.* Proposition 11.2(b) of [AF02] yields Proposition 1.1, and with it negative answers on the complete graphs to item (i) and to item (ii) under reading (a). These two items, as literally posed, are thus already contradicted by a consequence of a result in the monograph [AF02]. The question that was intended on the page may therefore differ from the literal one, for instance by a restriction on the graphs, by a normalisation, or by the orientation of the parameter. We cannot decide this from the sources. For this reason the conclusions of this note are formulated for the items as literally posed, under the readings listed above, and we do not say that the problem is solved.

*Open.* Item (iii) is not answered: we have no upper bound for  $\tau(p)$  other than the comparison bound of Remark 7.1, and the theorems only give constraints on such a bound. The opening question is open-ended; Theorems 1.2, 1.3 and 1.6 are statements in the general setting, but we do not regard them as an answer to it. The shape of  $\tau$  on  $(0, 1)$  is not understood in general: computations show three different shapes (Remark 7.3).

## 2 Preliminaries

### 2.1 The target

We write  $q = 1 - p$  throughout. Let  $\Pi_\lambda$  be the orthogonal projection of  $\mathbb{R}^G$  onto the eigenspace of  $K$  for the eigenvalue  $\lambda$ ; thus  $K = \sum_\lambda \lambda \Pi_\lambda$ , the sum being over the distinct eigenvalues, and  $\Pi_1 f$  is the constant function with value  $\frac{1}{n} \sum_x f(x)$ . For  $x \in G$  let  $N(x) \geq 1$  be the number of words  $(s_1, \dots, s_k) \in S^k$  of length  $k = |x|$  with  $s_1 \cdots s_k = x$ , that is, the number of geodesics from  $e$  to  $x$ . By (1),  $\mu_p = p(I - qK)^{-1} \delta_e$ : up to the factor  $p$ , the target is the resolvent (the Green function) of the walk at the point  $e$ .

**Lemma 2.1.** *Let  $0 < p < 1$ .*

- (a)  $\mu_p(x) = p \delta_e(x) + \frac{q}{d} \sum_{s \in S} \mu_p(xs)$  for all  $x \in G$ , and  $\mu_p$  is the only solution of this system.
- (b)  $\mu_p(x) > 0$  for all  $x$ , and  $\mu_p(e) > \mu_p(x)$  for all  $x \neq e$ .

(c)  $\mu_p = \frac{1}{n} + \sum_{\lambda \neq 1} \frac{p}{(1-\lambda)+p\lambda} \Pi_\lambda \delta_e$ . Hence  $\mu_p(x) \rightarrow \frac{1}{n}$  as  $p \downarrow 0$ , for every  $x$ .

(d)  $\mu_p(x) = p q^{|x|} (N(x) d^{-|x|} + \varepsilon_x(p))$  with  $0 \leq \varepsilon_x(p) \leq q/p$ .

*Proof.* (a) By (1), as row vectors,  $q\mu_p K = p \sum_{k \geq 1} q^k \delta_e K^k = \mu_p - p\delta_e$ . Since  $S$  is symmetric,  $K(y, x) = \frac{1}{d}$  exactly when  $y = xs$  with  $s \in S$ , so  $(\mu_p K)(x) = \frac{1}{d} \sum_s \mu_p(xs)$ . The matrix  $I - qK$  is invertible, because the eigenvalues of  $qK$  have absolute value at most  $q < 1$ .

(b)  $K$  is irreducible, so  $K^k(e, x) > 0$  for some  $k$ . The matrix  $R = (I - qK)^{-1} = \sum_k q^k K^k$  is symmetric with the eigenvalues  $1/(1 - q\lambda_j) > 0$ , hence positive definite, and  $R(ax, ay) = R(x, y)$  for  $a \in G$ , because  $K$  has this property. In particular  $R(x, x) = R(e, e)$ . For  $x \neq e$  and  $v = \delta_e - \delta_x$  we get  $0 < \langle v, Rv \rangle = R(e, e) + R(x, x) - 2R(e, x) = 2(R(e, e) - R(e, x))$ , and  $\mu_p(x) = pR(e, x)$ .

(c) Apply  $(I - qK)^{-1} = \sum_\lambda (1 - q\lambda)^{-1} \Pi_\lambda$  to  $p\delta_e$ , and note that  $1 - q\lambda = (1 - \lambda) + p\lambda$ , which is  $p$  for  $\lambda = 1$ , and that  $\Pi_1 \delta_e \equiv \frac{1}{n}$ . For  $\lambda \neq 1$  the coefficient  $p/((1 - \lambda) + p\lambda)$  tends to 0 as  $p \downarrow 0$ ; this includes  $\lambda = -1$ , where it is  $p/(2 - p)$ .

(d)  $K^k(e, x) = 0$  for  $k < |x|$ , and  $K^{|x|}(e, x) = N(x) d^{-|x|}$ . So (1) gives the formula with  $\varepsilon_x(p) = \sum_{j \geq 1} q^j K^{|x|+j}(e, x) \leq \sum_{j \geq 1} q^j = q/p$ .  $\square$

*Remark 2.2* (conventions). In (a) and (b) below, a change of convention replaces  $p$  by an increasing bijection of  $[0, 1)$  and multiplies  $\tau$  by a constant; the monotonicity properties of  $\tau$  and all ratios of its values are not affected.

(a) *Lazy walk*  $K' = \frac{1}{2}(I + K)$ . Since  $I - qK' = (1 - \frac{q}{2})(I - q'K)$  with  $q' = q/(2 - q)$ , the target built from  $K'$  is  $\mu_{p'}$  with  $p' = 1 - q' = 2p/(1 + p)$ . The Metropolis chain based on  $K'$  is  $\frac{1}{2}(I + P_{p'})$ , and its relaxation time is  $2\tau(p')$ .

(b) *A generating set containing  $e$* , that is, one loop at each vertex and the walk  $K'' = (I + dK)/(d + 1)$ . In the same way the target is  $\mu_{p''}$  with  $p'' = p(d + 1)/(d + p)$ , the Metropolis chain is  $(I + dP_{p''})/(d + 1)$ , and its relaxation time is  $\frac{d+1}{d}\tau(p'')$ .

(c) *Continuous time*. Let  $(Y_t)_{t \geq 0}$  be the continuous-time walk which jumps at rate 1 according to  $K$ , and let  $\xi$  be an independent exponential time of rate  $\theta > 0$ . The number of jumps before  $\xi$  equals  $k$  with probability  $\frac{\theta}{1+\theta}(\frac{1}{1+\theta})^k$ , so  $Y_\xi$  has the law  $\mu_p$  with  $p = \theta/(1 + \theta)$ . In the parametrisation by  $\theta \in (0, \infty)$  the uniform law is the end  $\theta = 0$ , and the point mass at  $e$  is the end  $\theta = \infty$ .

(d) *Bipartite graphs*. The walk then has the eigenvalue  $-1$ , but  $\mu_p \rightarrow \mu_0$  by Lemma 2.1(c), and  $\tau(0) = 1/(1 - \lambda_2)$  is finite. For  $p > 0$  the Metropolis chain is aperiodic on every Cayley graph, since  $P_p(e, e) = 1 - \frac{1}{d} \sum_{s \in S} \mu_p(s)/\mu_p(e) > 0$  by Lemma 2.1(b).

*Remark 2.3* (the target and the distance). By Lemma 2.1(b) the target is largest at  $e$ . If  $x$  and  $y$  are adjacent and  $|y| = |x| + 1$ , then  $\mu_p(y)/\mu_p(x) \leq q(N(y)d^{-|y|} + q/p)/(N(x)d^{-|x|}) \rightarrow 0$  as  $p \uparrow 1$ , by Lemma 2.1(d); so for  $p$  close to 1 the target decreases along every edge which leads away from  $e$ . On the hypercube  $\mu_p$  is a decreasing function of  $|x|$  for every  $p$  (Proposition 5.1(a)). In general, however,  $\mu_p$  is not monotone in the distance: on  $\mathbb{Z}/16$  with  $S = \{\pm 1, \pm 3\}$  and  $p = \frac{1}{100}$  the vertices 6 and 5 are adjacent,  $|6| = 2$ ,  $|5| = 3$ , and  $\mu_p(5) > \mu_p(6)$  (an exact computation, Section 8).

## 2.2 The chain, the Dirichlet form and the spectrum

For  $0 \leq p < 1$  and  $x \neq y$  put

$$c_p(x, y) := \mu_p(x)P_p(x, y) = K(x, y) \min\{\mu_p(x), \mu_p(y)\}.$$

This is symmetric in  $x$  and  $y$ , so  $P_p$  is reversible with respect to  $\mu_p$ ; and  $P_p$  is irreducible, because  $K$  is irreducible and  $\mu_p > 0$ . For  $f: G \rightarrow \mathbb{R}$  let

$$\mathcal{E}_p(f, f) = \frac{1}{2} \sum_{x \neq y} c_p(x, y) (f(x) - f(y))^2, \quad \text{Var}_p(f) = \sum_x \mu_p(x) f(x)^2 - \left( \sum_x \mu_p(x) f(x) \right)^2.$$

The variational characterisation of the relaxation time [AF02, Theorem 3.25] reads

$$\frac{1}{\tau(p)} = 1 - \beta_2(p) = \min \left\{ \frac{\mathcal{E}_p(f, f)}{\text{Var}_p(f)} : f: G \rightarrow \mathbb{R} \text{ not constant} \right\}. \quad (7)$$

We also use the symmetric form. Let  $D_p$  be the diagonal matrix with the entries  $\mu_p(x)$ . The matrix

$$S(p) := D_p^{1/2}(I - P_p)D_p^{-1/2}, \quad S(p)(x, y) = \begin{cases} -K(x, y) \frac{\min\{\mu_p(x), \mu_p(y)\}}{\sqrt{\mu_p(x)\mu_p(y)}}, & x \neq y, \\ \sum_z K(x, z) \min\left\{1, \frac{\mu_p(z)}{\mu_p(x)}\right\}, & x = y, \end{cases} \quad (8)$$

is symmetric and has the eigenvalues  $1 - \beta_j(p)$ . For a real symmetric matrix  $A$  we denote its eigenvalues, with multiplicities, by  $\sigma_1(A) \leq \sigma_2(A) \leq \dots$ . Then

$$\frac{1}{\tau(p)} = \sigma_2(S(p)), \quad S(0) = I - K. \quad (9)$$

**Lemma 2.4** (continuity of ordered real spectra). *Let  $A_m \rightarrow A$  be real  $n \times n$  matrices, and suppose that every  $A_m$  has only real eigenvalues  $b_1^{(m)} \geq \dots \geq b_n^{(m)}$ . Then  $A$  has only real eigenvalues  $b_1 \geq \dots \geq b_n$ , and  $b_j^{(m)} \rightarrow b_j$  for each  $j$ .*

*Proof.* The numbers  $b_j^{(m)}$  are bounded, since they are bounded by the maximal absolute row sums of the  $A_m$ , which converge. Take a subsequence along which every  $b_j^{(m)}$  converges, say to  $c_j$ ; then  $c_1 \geq \dots \geq c_n$ . Along it  $\det(zI - A_m) = \prod_j (z - b_j^{(m)}) \rightarrow \prod_j (z - c_j)$  for every  $z$ , and also  $\det(zI - A_m) \rightarrow \det(zI - A)$ . So the  $c_j$  are the eigenvalues of  $A$ , in decreasing order. The limit does not depend on the subsequence, hence the whole sequence converges.  $\square$

**Corollary 2.5.** *The function  $\tau$  is continuous on  $[0, 1)$ . In particular  $\tau(p) \rightarrow \tau(0)$  as  $p \downarrow 0$ .*

*Proof.* By Lemma 2.1(a) and Cramer's rule every  $\mu_p(x)$  is a rational function of  $p$  without poles in  $(0, 1)$ , and it is positive there. So  $P_p$  depends continuously on  $p \in (0, 1)$ . As  $p \downarrow 0$  we have  $\mu_p(y)/\mu_p(x) \rightarrow 1$  by Lemma 2.1(c), hence  $P_p \rightarrow K = P_0$ . All  $P_p$  have only real eigenvalues, so  $\beta_2$  is continuous on  $[0, 1)$  by Lemma 2.4, and  $\beta_2(p) < 1$ .  $\square$

*Remark 2.6* (vertex-transitive graphs). Apart from the special graphs of Proposition 1.1 and Corollary 1.4, the only properties of a Cayley graph which are used in this note are that  $K$  is the walk on a connected  $d$ -regular graph without loops and multiple edges, and that  $K(\sigma x, \sigma y) = K(x, y)$  for a transitive group of permutations  $\sigma$  of the vertices. The second property is used three times, in Lemma 2.1(b), in Step 1 of the proof of Corollary 1.5 and in the proof of Theorem 1.6, each time to see that the diagonal entries of a function of  $K$  are all equal. Hence Theorems 1.2, 1.3 and 1.6 and Corollary 1.5 hold, with the same proofs, for the simple random walk on a finite connected vertex-transitive graph with any base vertex  $e$  in place of the identity.

### 3 The complete graph

Nothing in this section is new.

**Proposition 3.1** (Liu; Diaconis and Saloff-Coste; Aldous and Fill). *Let  $\pi$  be a positive probability vector on a set of  $n \geq 2$  points, let  $\pi_{\max} = \max_x \pi(x)$ , and let  $P(x, y) = \frac{1}{n-1} \min\{1, \pi(y)/\pi(x)\}$  for  $y \neq x$  and  $P(x, x) = 1 - \sum_{y \neq x} P(x, y)$ ; this is the Metropolis chain based on the walk on  $K_n$  with stationary distribution  $\pi$ . Then the second largest eigenvalue of  $P$  is  $1 - \frac{1}{(n-1)\pi_{\max}}$ , that is, the relaxation time is  $(n-1)\pi_{\max}$ .*

In [AF02, Proposition 11.2(b)] the relaxation time of the Metropolis chain over independent proposals is found to be  $1/(1 - \rho)$ , where  $\rho$  is the probability that a proposal is rejected at a state of largest weight. For the uniform proposal on the  $n$  points, which may propose the current state,  $\rho = \frac{1}{n} \sum_y (1 - \pi(y)/\pi_{\max}) = 1 - 1/(n\pi_{\max})$ , so that the relaxation time is  $n\pi_{\max}$ . The walk on  $K_n$  proposes one of the other  $n - 1$  points; its Metropolis chain  $P$  and the chain  $P^{\text{ind}}$  with uniform independent proposals satisfy  $P - I = \frac{n}{n-1}(P^{\text{ind}} - I)$ , and the relaxation time becomes  $(n - 1)\pi_{\max}$ . Liu's theorem [Liu96], quoted as [DSC98, Theorem 2.2], gives all eigenvalues of  $P^{\text{ind}}$ . For completeness we include a short proof; the eigenvector in it is the one used in [AF02].

*Proof.* Let  $\pi(x_1) = \pi_{\max}$  and  $\beta = 1 - \frac{1}{(n-1)\pi_{\max}}$ . From  $x_1$  a proposal  $y$  is accepted with probability  $\pi(y)/\pi_{\max}$ , hence

$$\delta_{x_1}P = \frac{\pi}{(n-1)\pi_{\max}} + \beta \delta_{x_1}.$$

As  $\pi P = \pi$ , this gives  $(\delta_{x_1} - \pi)P = \beta(\delta_{x_1} - \pi)$ , and  $\delta_{x_1} \neq \pi$ . So  $\beta$  is an eigenvalue of  $P$ , and  $\beta < 1$ ; since the eigenvalue 1 is simple,  $\beta$  is at most the second largest eigenvalue  $\beta_2$  of  $P$ . Conversely,  $\min\{\pi(x), \pi(y)\} \geq \pi(x)\pi(y)/\pi_{\max}$ , so for every  $f$  the Dirichlet form of  $P$  satisfies

$$\begin{aligned} \mathcal{E}(f, f) &= \frac{1}{2(n-1)} \sum_{x \neq y} \min\{\pi(x), \pi(y)\} (f(x) - f(y))^2 \\ &\geq \frac{1}{(n-1)\pi_{\max}} \cdot \frac{1}{2} \sum_{x, y} \pi(x)\pi(y) (f(x) - f(y))^2 = \frac{\text{Var}_{\pi}(f)}{(n-1)\pi_{\max}}. \end{aligned}$$

By the variational characterisation,  $1 - \beta_2 \geq \frac{1}{(n-1)\pi_{\max}} = 1 - \beta$ .  $\square$

*Proof of Proposition 1.1.* Here  $K(x, y) = \frac{1}{n-1}$  for  $x \neq y$ . Let  $u_k = K^k(e, e)$ . Then  $u_0 = 1$  and  $u_{k+1} = (1 - u_k)/(n - 1)$ , so  $u_k = \frac{1}{n} + (1 - \frac{1}{n})(-\frac{1}{n-1})^k$ , and by (1)

$$\mu_p(e) = p \sum_{k \geq 0} q^k u_k = \frac{1}{n} + \left(1 - \frac{1}{n}\right) \frac{p}{1 + \frac{q}{n-1}} = \frac{1}{n} + \frac{(n-1)^2}{n} \cdot \frac{p}{n-p} = \frac{1 + (n-2)p}{n-p}.$$

Every permutation of  $G \setminus \{e\}$  is an automorphism of  $K_n$  fixing  $e$ , so  $\mu_p$  is constant on  $G \setminus \{e\}$ , with the value  $(1 - \mu_p(e))/(n - 1) = (1 - p)/(n - p)$ . For  $p = 0$  these are the values of the uniform law. Since  $\mu_p(e) \geq \mu_p(x)$ , Proposition 3.1 gives  $\tau(p) = (n - 1)\mu_p(e)$ . This is the first formula for  $\tau(p)$ , and the second one follows from  $n(1 + (n - 2)p) = (n - p) + (n - 1)^2 p$ . Finally  $\frac{d\tau}{dp} = (n - 1)^3/(n - p)^2 > 0$ ,  $\tau(0) = \frac{n-1}{n}$ , and  $\tau(p) \rightarrow n - 1$  as  $p \uparrow 1$ .  $\square$

## 4 The limit $p \rightarrow 1$

*Proof of Theorem 1.2.* Let  $y = xs$  with  $s \in S$ ; then  $|y| - |x| \in \{-1, 0, 1\}$ . By Lemma 2.1(d),

$$\frac{\mu_p(y)}{\mu_p(x)} = q^{|y|-|x|} \cdot \frac{N(y) d^{-|y|} + \varepsilon_y(p)}{N(x) d^{-|x|} + \varepsilon_x(p)}, \quad \varepsilon_x(p), \varepsilon_y(p) \rightarrow 0 \quad (p \uparrow 1).$$

So, as  $p \uparrow 1$ , the ratio tends to 0 if  $|y| = |x| + 1$ , to  $N(y)/N(x)$  if  $|y| = |x|$ , and to  $\infty$  if  $|y| = |x| - 1$ . Hence  $P_p \rightarrow P_1$  entrywise, where for  $y \neq x$

$$P_1(x, y) = K(x, y) \alpha(x, y), \quad \alpha(x, y) = \begin{cases} 0, & |y| = |x| + 1, \\ \min\{1, N(y)/N(x)\}, & |y| = |x|, \\ 1, & |y| = |x| - 1, \end{cases} \quad (10)$$

and  $P_1(x, x) = 1 - \sum_{y \neq x} P_1(x, y) \geq 0$ .

The matrix  $P_1$  is stochastic, but it is neither irreducible nor reversible: the state  $e$  is absorbing. This does not matter, since only Lemma 2.4 will be applied to it. Order  $G$  by word length, and let  $D$  be the largest word length. Since  $P_1(x, y) = 0$  for  $|y| > |x|$ , the matrix  $P_1$  is block lower triangular with the diagonal blocks  $B_k = (P_1(x, y))_{|x|=|y|=k}$ ,  $k = 0, \dots, D$ , and the spectrum of  $P_1$  is the union, with multiplicities, of the spectra of the  $B_k$ . We have  $B_0 = (1)$ , because all neighbours of  $e$  have length 1.

Let  $k \geq 1$  and  $|x| = k$ , and let  $d^-(x)$  be the number of neighbours of  $x$  of length  $k - 1$ . Then  $d^-(x) \geq 1$ , since a geodesic from  $e$  to  $x$  passes through such a neighbour. The row of  $P_1$  at  $x$  has the sum 1, it has the mass  $d^-(x)/d$  on the vertices of length  $k - 1$ , and no mass on those of length  $k + 1$ . Hence the row sum of  $B_k$  at  $x$  is  $1 - d^-(x)/d \leq 1 - \frac{1}{d}$ . As  $B_k$  has non-negative entries, all its eigenvalues have absolute value at most  $1 - \frac{1}{d}$ . For  $x \neq y$  of the same length,  $N(x)P_1(x, y) = K(x, y) \min\{N(x), N(y)\}$  is symmetric in  $x$  and  $y$ ; so  $B_k$  is self-adjoint for the inner product with the weights  $N(x)$ , and its eigenvalues are real. For  $k = 1$  the vertices are the elements of  $S$ , and  $e$  is the only vertex of length 0; so  $d^-(x) = 1$  for every  $x \in S$ , all row sums of  $B_1$  are equal to  $1 - \frac{1}{d}$ , and  $1 - \frac{1}{d}$  is an eigenvalue of  $B_1$ .

Consequently all eigenvalues of  $P_1$  are real, the largest one is 1 and is simple, the second largest is exactly  $1 - \frac{1}{d}$ , and all eigenvalues other than 1 have absolute value at most  $1 - \frac{1}{d}$ . By Lemma 2.4,  $\beta_2(p) \rightarrow 1 - \frac{1}{d}$  as  $p \uparrow 1$ , that is,  $\tau(p) \rightarrow d$ .  $\square$

**Corollary 4.1** (the two ends). (a)  $\tau(1^-)/\tau(0) = d(1 - \lambda_2)$ . In particular  $\tau(0) < \tau(1^-)$  if and only if  $\lambda_2 < 1 - \frac{1}{d}$ ; this holds for the complete graphs, for the hypercubes, and for the cycles  $C_3, C_4, C_5$ .

(b) If  $\lambda_2 > 1 - \frac{1}{d}$ , then  $\tau(0) > \tau(1^-)$ , and  $\tau$  is neither non-increasing nor non-decreasing on  $[0, 1)$ . This holds for the cycles  $C_n$  with  $n \geq 7$ , where  $\tau(0) = 1/(1 - \cos \frac{2\pi}{n})$  and  $\tau(1^-) = 2$ ; along the cycles  $\tau(0)/\tau(1^-) \rightarrow \infty$ .

*Proof.* (a) follows from Theorem 1.2 and  $\tau(0) = 1/(1 - \lambda_2)$ . For  $K_n$  we have  $\lambda_2 = -\frac{1}{n-1}$ , for the hypercube  $\lambda_2 = 1 - \frac{2}{d}$  (Section 5.2), and for the cycle  $C_n$  we have  $d = 2$  and  $\lambda_2 = \cos \frac{2\pi}{n}$ , which is smaller than  $\frac{1}{2}$  for  $n \leq 5$ , equal to  $\frac{1}{2}$  for  $n = 6$ , and larger for  $n \geq 7$ . (b) By Theorem 1.6, which is proved in Section 6 independently of this section,  $\tau(p) > \tau(0)$  for small  $p > 0$ ; and  $\tau(p) \rightarrow \tau(1^-) < \tau(0)$  as  $p \uparrow 1$ .  $\square$

*Remark 4.2* (the targets  $\theta^{|x|}$ ). The proof of Theorem 1.2 uses the form of  $\mu_p$  only through the limits of the ratios  $\mu_p(y)/\mu_p(x)$ . It applies word for word to the targets  $\pi_\theta(x) \propto \theta^{|x|}$ ,  $0 < \theta \leq 1$ , on any Cayley graph: here  $\pi_\theta(y)/\pi_\theta(x) = \theta^{|y|-|x|}$ , the limit matrix for  $\theta \downarrow 0$  is given by (10) with  $\min\{1, N(y)/N(x)\}$  replaced by 1, and the relaxation time of the Metropolis chain with target  $\pi_\theta$  tends to  $d$  as  $\theta \downarrow 0$ . The two families of Diaconis and Hanlon are instances. On the hypercube the chain of the Hamming weight, to which the Metropolis chain for  $\pi_\theta$  can be lumped, has the eigenvalues  $1 - \frac{j}{d}(1 + \theta)$ ,  $j = 0, \dots, d$  [DH92, Section 3, Theorem 1]; the Metropolis chain itself, a product chain in which a coordinate chosen uniformly is set to 0 if it is 1 and to 1 with probability  $\theta$  if it is 0, has the same eigenvalues, with the multiplicities  $\binom{d}{j}$ . So its relaxation time  $d/(1 + \theta)$  increases from  $\frac{d}{2}$  to  $d$  as  $\theta$  decreases from 1 to 0. On the symmetric group  $S_N$  with all transpositions, the chain lumped to the conjugacy classes has the second largest eigenvalue  $1 - \frac{2\theta}{N} - \frac{2}{N(N-1)}$  [DH92, Section 4]; its relaxation time  $N(N-1)/(2(\theta(N-1) + 1))$  increases from  $\frac{N-1}{2}$ , the relaxation time of the walk, to  $\binom{N}{2} = d$ . The relaxation time of the Metropolis chain on  $S_N$  is at least this number, since the eigenvalues of the lumped chain are among its eigenvalues; in the computations of Section 8 the two are equal for  $N \leq 7$ , although the Metropolis chain has further eigenvalues. The section and theorem numbers are those of the technical report. The general statement about the limit  $\theta \downarrow 0$  is not made in [DH92].

## 5 Two lower bounds

### 5.1 Proof of Theorem 1.3

(a) Take  $f = \mathbf{1}_{\{e\}}$  in (7). Then  $\text{Var}_p(f) = m(1 - m)$ . By Lemma 2.1(b),  $\min\{\mu_p(e), \mu_p(s)\} = \mu_p(s)$  for  $s \in S$ , so  $\mathcal{E}_p(f, f) = \sum_{s \in S} c_p(e, s) = \frac{1}{d} \sum_{s \in S} \mu_p(s)$ , and by Lemma 2.1(a) at  $x = e$  this equals  $(m - p)/q$ . In particular  $m > p$ , and (7) gives  $1/\tau(p) \leq (m - p)/(q m(1 - m))$ . On  $K_n$  we have  $m - p = q^2/(n - p)$  and  $1 - m = (n - 1)q/(n - p)$  by Proposition 1.1, so the bound is  $(n - 1)m = \tau(p)$ .

(b) Take  $f(x) = |x|$ . If  $c_p(x, y) > 0$ , then  $x$  and  $y$  are adjacent and  $|f(x) - f(y)| \leq 1$ . So  $\mathcal{E}_p(f, f) \leq \frac{1}{2} \sum_x \sum_{y \neq x} c_p(x, y) = \frac{1}{2} \sum_x \mu_p(x)(1 - P_p(x, x)) \leq \frac{1}{2}$ , and  $\text{Var}_p(f) = \text{Var}(|X(T_p - 1)|) > 0$ .  $\square$

### 5.2 The hypercube

Let  $G = (\mathbb{Z}/2)^d$  and  $S = \{e_1, \dots, e_d\}$ , the unit vectors, with  $d \geq 2$ ; then  $|x|$  is the Hamming weight of  $x$ . For  $A \subset \{1, \dots, d\}$  the character  $\chi_A(x) = (-1)^{\sum_{i \in A} x_i}$  satisfies  $K\chi_A = (1 - \frac{2|A|}{d})\chi_A$ . Hence  $\lambda_2 = 1 - \frac{2}{d}$  and  $\tau(0) = \frac{d}{2}$ . We write  $\mathbb{E}_p$  for the expectation with respect to  $\mu_p$ .

**Proposition 5.1.** *Let  $0 < p < 1$  and  $q = 1 - p$ .*

(a)  $\mu_p(x) = w(|x|)$  with  $w(k) = \int_{(0,1)} t^k \nu(dt)$ ,  $k = 0, \dots, d$ , for a finite positive measure  $\nu$  on  $(0, 1)$ . Consequently  $w$  is strictly decreasing, and  $w(k + 1)^2 \leq w(k)w(k + 2)$ .

(b)  $\mathbb{E}_p|x| = \frac{qd}{pd + 2q}$  and  $\text{Var}_p|x| = \frac{qd}{pd + 4q} \left(1 + \frac{pq d^2}{(pd + 2q)^2}\right)$ .

(c) For  $f(x) = |x|$  we have  $\mathcal{E}_p(f, f) = \frac{1}{d} \mathbb{E}_p|x|$ .

*Proof.* (a) For  $|A| = j$ , (1) and  $K^k \chi_A = (1 - \frac{2j}{d})^k \chi_A$  give  $\sum_x \mu_p(x) \chi_A(x) = p \sum_k q^k (1 - \frac{2j}{d})^k = \frac{pd}{pd + 2jq} =: a_j$ . By Fourier inversion and  $a_j = \int_0^\infty pd e^{-(pd + 2jq)s} ds$ ,

$$\begin{aligned} \mu_p(x) &= 2^{-d} \sum_A a_{|A|} \chi_A(x) = 2^{-d} \int_0^\infty pd e^{-pds} \prod_{i=1}^d (1 + (-1)^{x_i} e^{-2qs}) ds \\ &= \int_0^\infty \tanh(qs)^{|x|} 2^{-d} pd e^{-pds} (1 + e^{-2qs})^d ds, \end{aligned}$$

because  $(1 - e^{-2qs})/(1 + e^{-2qs}) = \tanh(qs)$ . The image of the measure  $2^{-d} pd e^{-pds} (1 + e^{-2qs})^d ds$  under  $s \mapsto \tanh(qs)$  is the measure  $\nu$ . Strict monotonicity of  $w$  is clear, and the inequality is the Cauchy–Schwarz inequality for  $\int t^{k/2} t^{(k+2)/2} \nu(dt)$ .

(b) We have  $|x| = \frac{1}{2}(d - \sum_i \chi_{\{i\}}(x))$  and  $\chi_{\{i\}} \chi_{\{j\}} = \chi_{\{i,j\}}$  for  $i \neq j$ . Hence  $\mathbb{E}_p|x| = \frac{d}{2}(1 - a_1) = \frac{qd}{pd + 2q}$  and

$$\text{Var}_p|x| = \frac{1}{4}(d + d(d - 1)a_2 - d^2 a_1^2) = \frac{1}{4}(d(1 - a_2) + d^2(a_2 - a_1^2)),$$

and  $1 - a_2 = \frac{4q}{pd + 4q}$ ,  $a_2 - a_1^2 = \frac{4pq^2 d}{(pd + 4q)(pd + 2q)^2}$ .

(c) Every edge joins a vertex  $x$  to a vertex  $y$  with  $|y| = |x| - 1$ , and then  $c_p(x, y) = \frac{1}{d} \mu_p(x)$  by (a). A vertex  $x$  has  $|x|$  such neighbours. So  $\mathcal{E}_p(f, f) = \sum_x \mu_p(x)|x|/d$ .  $\square$

*Proof of Corollary 1.4.* We have seen that  $\tau(0) = \frac{d}{2}$ , and  $\tau(1^-) = d$  by Theorem 1.2. By (7) with  $f(x) = |x|$  and by Proposition 5.1,  $\tau(p) \geq \text{Var}_p|x|/\mathcal{E}_p(f, f) = d \text{Var}_p|x|/\mathbb{E}_p|x| = H_d(p)$ .

(a) For  $p = \frac{1}{d}$  we get  $H_d(\frac{1}{d}) = d \cdot \frac{1+2q}{1+4q} \left(1 + \frac{qd}{(1+2q)^2}\right) \geq \frac{qd^2}{(1+4q)(1+2q)} \geq \frac{qd^2}{15} = \frac{d(d-1)}{15}$ .

(b) If  $pd \geq 4$ , then  $4q \leq pd$ , hence  $(pd + 4q)(pd + 2q) \leq 3p^2 d^2$  and  $H_d(p) \geq \frac{pq d^3}{(pd + 4q)(pd + 2q)} \geq \frac{qd}{3p}$ .

(c) A direct computation gives

$$H_d(p) - d = \frac{dq(p(d^2 - 2d + 4) - 4)}{(pd + 4q)(pd + 2q)}, \quad (11)$$

which is positive if and only if  $p > 4/(d^2 - 2d + 4)$ ; and  $4/(d^2 - 2d + 4) < 1$  if and only if  $d \geq 3$ .

The two ratios follow from (a),  $\tau(0) = \frac{d}{2}$  and  $\tau(1^-) = d$ .  $\square$

The bound (4) is sharp at both ends:  $H_d(p) \rightarrow \frac{d}{2} = \tau(0)$  as  $p \rightarrow 0$  and  $H_d(p) \rightarrow d = \tau(1^-)$  as  $p \rightarrow 1$ . By Corollary 1.4(c) and  $\tau(0) = \frac{d}{2} < d$ , on the hypercubes of dimension  $d \geq 3$  the function  $\tau$  is neither non-increasing nor non-decreasing.

The next proposition is used only for the computations of Section 8.

**Proposition 5.2** (level chain). *Let  $0 < p < 1$ . On the hypercube the eigenvalue  $\beta_2(p)$  of  $P_p$  has an eigenfunction which depends only on  $|x|$ . Consequently  $\tau(p)$  is the relaxation time of the birth-and-death chain  $L_p$  on  $\{0, \dots, d\}$  which moves from  $k$  to  $k - 1$  with probability  $\frac{k}{d}$  and from  $k$  to  $k + 1$  with probability  $\frac{d-k}{d} \theta_k$ , where  $\theta_k = w(k + 1)/w(k)$ .*

*Proof.* By Proposition 5.1(a), a proposal  $x \rightarrow x - e_i$  (with  $x_i = 1$ ) is always accepted, a proposal  $x \rightarrow x + e_i$  (with  $x_i = 0$ ) is accepted with probability  $\theta_{|x|} \in (0, 1)$ , and  $\theta_k$  is non-decreasing in  $k$ . Consider the continuous-time chain with generator  $P_p - I$  and its semigroup  $H_t = e^{t(P_p - I)}$ . It has the following graphical construction. Each coordinate  $i$  has two independent Poisson clocks of rate  $\frac{1}{d}$ . When the first clock of  $i$  rings, coordinate  $i$  is set to 0. When the second one rings, an independent uniform variable  $U$  on  $(0, 1)$  is drawn, and coordinate  $i$  is set to 1 if it is 0 and  $U < \theta_{|x|}$ , where  $x$  is the current state. Run two copies from states  $x \leq y$  (coordinatewise) with the same clocks and the same variables  $U$ . A ring of a first clock preserves the order. At a ring of the second clock of  $i$ : if  $x_i = y_i = 0$  and the lower copy flips, then  $U < \theta_{|x|} \leq \theta_{|y|}$  and the upper copy flips too; if  $x_i = 0$  and  $y_i = 1$ , the upper copy does not change; if  $x_i = 1$ , then  $y_i = 1$  and nothing happens. So the order is preserved, and  $H_t$  maps non-decreasing functions (for the coordinatewise order) to non-decreasing functions.

Let  $E$  be the eigenspace of  $P_p$  for  $\beta_2(p)$ , let  $\Pi_E$  be the projection onto  $E$  which is orthogonal in  $\ell^2(\mu_p)$ , and let  $\gamma = 1 - \beta_2(p)$ . For every  $h$  we have  $e^{\gamma t}(H_t h - \sum_x \mu_p(x) h(x)) \rightarrow \Pi_E h$  as  $t \rightarrow \infty$ ; so  $\Pi_E h$  is non-decreasing whenever  $h$  is. The indicator functions of the sets  $\{y : y \geq x\}$ ,  $x \in G$ , are non-decreasing and form a basis of  $\mathbb{R}^G$ ; hence  $\varphi := \Pi_E h \neq 0$  for one of them. The function  $\varphi$  is non-decreasing and not constant, since it is orthogonal to the constants; therefore  $\varphi(1, \dots, 1) > \varphi(0, \dots, 0)$ . The permutations of the coordinates preserve  $\mu_p$  and commute with  $P_p$ . The average  $\bar{\varphi}$  of  $\varphi$  over these permutations lies in  $E$ , depends only on  $|x|$ , and has the same values as  $\varphi$  at the two points  $(0, \dots, 0)$  and  $(1, \dots, 1)$ ; so it is not zero.

Finally,  $P_p$  maps a function  $\psi(|x|)$  to the function  $(L_p \psi)(|x|)$ . So every eigenvalue of  $L_p$  is an eigenvalue of  $P_p$ , the eigenvalue 1 of  $L_p$  is simple, and by the first part  $\beta_2(p)$  is an eigenvalue of  $L_p$ . Hence  $\beta_2(p)$  is the second largest eigenvalue of  $L_p$ .  $\square$

### 5.3 Bounded degree: proof of Corollary 1.5

The eigenvalue 1 of  $K$  is simple, the eigenvalue  $-1$  has multiplicity at most one, and all other eigenvalues have absolute value less than 1; so  $\rho < 1$ . Since  $K^2(x, x) = \frac{1}{d}$  for all  $x$ , we have  $\frac{n}{d} = \sum_j \lambda_j^2 \leq 2 + (n - 2)\rho^2$ , hence  $\rho > 0$  as  $n > 2d$ . So  $0 < \kappa \leq 1$  and  $\rho \leq d^{-\kappa}$ . Let  $r \geq 6$  be an integer with  $32d^r \leq n$ . We show that  $\tau(p) \geq \kappa r^2/40$  for some  $p$ .

*Step 1:*  $K^j(e, x) \leq \frac{2}{n} + \rho^{j-1}$  for  $j \geq 1$  and  $x \in G$ . For  $a + b = j$ , the Cauchy–Schwarz inequality and the symmetry of  $K$  give  $K^j(e, x) = \sum_z K^a(e, z) K^b(z, x) \leq (K^{2a}(e, e) K^{2b}(x, x))^{1/2}$ . Moreover

$K^{2a}(x, x) = K^{2a}(e, e) = \frac{1}{n} \sum_i \lambda_i^{2a} \leq \frac{2}{n} + \rho^{2a}$ . Take  $a = \lfloor j/2 \rfloor$  and  $b = \lceil j/2 \rceil$ ; then  $2b \geq 2a \geq j - 1$ , and the claim follows because  $\rho < 1$ .

*Step 2.* The ball  $\{x : |x| \leq r\}$  has at most  $1 + d + \dots + d^r \leq 2d^r$  elements. So  $\mathbb{P}(|X(j)| \leq r) \leq 2d^r \left(\frac{2}{n} + \rho^{j-1}\right)$ .

*Step 3.* Let  $t = 1 + \lceil (r+4)/\kappa \rceil$ . For  $j \geq t$  we have  $\rho^{j-1} \leq d^{-\kappa(j-1)} \leq d^{-(r+4)}$ , and therefore  $\mathbb{P}(|X(j)| \leq r) \leq \frac{4d^r}{n} + 2d^{-4} \leq \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$ .

*Step 4.* Let  $p = \frac{1}{t}$ ,  $J = T_p - 1$  and  $Y = |X(J)|$ ; thus  $\mathbb{P}(J \geq j) = (1-p)^j$ , and  $t \geq r+5 \geq 11$ . Consider the events

$$A = \{J \leq \lfloor r/2 \rfloor\} \quad \text{and} \quad B = \{J \geq t, Y \geq r+1\}.$$

On  $A$  we have  $Y \leq J \leq \frac{r}{2}$ , so  $A$  and  $B$  are disjoint. Since  $t \leq 2 + \frac{r+4}{\kappa} \leq \frac{r+6}{\kappa}$  and  $r \geq 6$ , we have  $\frac{r}{2t} \geq \frac{\kappa r}{2(r+6)} \geq \frac{\kappa}{4}$ , and so  $\mathbb{P}(A) = 1 - (1 - \frac{1}{t})^{\lfloor r/2 \rfloor + 1} \geq 1 - e^{-r/(2t)} \geq 1 - e^{-\kappa/4} \geq \frac{\kappa}{4} \left(1 - \frac{\kappa}{8}\right) \geq \frac{7\kappa}{32} \geq \frac{\kappa}{5}$ . By Step 3,  $\mathbb{P}(B) = \sum_{j \geq t} \mathbb{P}(J = j) \mathbb{P}(|X(j)| \geq r+1) \geq \frac{3}{4} (1 - \frac{1}{t})^t \geq \frac{3}{4} \left(\frac{10}{11}\right)^{11} \geq \frac{1}{4}$ . For an independent copy  $Y'$  of  $Y$ , with the corresponding events  $A'$  and  $B'$ ,

$$\text{Var}(Y) = \frac{1}{2} \mathbb{E}(Y - Y')^2 \geq \frac{1}{2} (\mathbb{P}(A \cap B') + \mathbb{P}(B \cap A')) \left(\frac{r}{2} + 1\right)^2 = \mathbb{P}(A) \mathbb{P}(B) \left(\frac{r}{2} + 1\right)^2 \geq \frac{\kappa r^2}{80}.$$

By Theorem 1.3(b),  $\tau(\frac{1}{t}) \geq 2 \text{Var}(Y) \geq \kappa r^2/40$ .

*Conclusion.* Let  $r = \lfloor \log_d(n/32) \rfloor$ . Then  $32d^r \leq n$ ,  $r \geq 6$  because  $n \geq 32d^6$ , and  $r \geq \log_d n - \log_d 32 - 1 \geq \log_d n - 6 \geq 0$ . So  $\sup_p \tau(p) \geq \frac{\kappa}{40} (\log_d n - 6)^2$ . Finally,  $\lambda_2 \neq \pm 1$  because  $n > 2$ , so  $\lambda_2 \leq \rho$  and  $\tau(0) \leq 1/(1-\rho)$ ; and  $\tau(1^-) = d$  by Theorem 1.2. If  $d$  is fixed and  $\rho \leq \rho_0 < 1$ , then  $\kappa \geq \min\{1, \log(1/\rho_0)/\log d\} > 0$ , and the two ratios are at least a positive constant, which depends on  $d$  and  $\rho_0$ , times  $(\log n)^2$ .  $\square$

## 6 The increase at $p = 0$

By (9),  $1/\tau(p)$  is the second smallest eigenvalue of the symmetric matrix  $S(p)$ . We expand  $S(p)$  to first order at  $p = 0$  and apply the first-order perturbation theory of eigenvalues of symmetric matrices, which goes back to Rellich [Rel37]; see [Kat95, Chapter II]. Since  $S(p)$  is defined through minima and is given for  $p \geq 0$  only, we state and prove the elementary one-sided form that is needed. We use the Euclidean norm  $|v|$  on  $\mathbb{R}^N$ , the corresponding inner product  $\langle \cdot, \cdot \rangle$  and the operator norm  $\|A\|$ , and the min-max principle [HJ12, Section 4.2]: for a real symmetric  $N \times N$  matrix  $A$ ,

$$\sigma_k(A) = \min_{\dim W = k} \max_{v \in W, |v|=1} \langle v, Av \rangle = \max_{\dim W = N-k+1} \min_{v \in W, |v|=1} \langle v, Av \rangle. \quad (12)$$

**Lemma 6.1** (first-order perturbation of a multiple eigenvalue). *Let  $A_0, A_1$  be real symmetric  $N \times N$  matrices, and for  $0 \leq p \leq p_1$  let  $A(p)$  be real symmetric matrices with  $\|A(p) - A_0 - pA_1\| \leq C_0 p^2$ . Let  $k \geq 1$  and  $\sigma := \sigma_k(A_0)$ , and suppose that  $\sigma_{k-1}(A_0) < \sigma$  if  $k \geq 2$ . Let  $E$  be the eigenspace of  $A_0$  for  $\sigma$ , and*

$$\nu := \min\{\langle f, A_1 f \rangle : f \in E, |f| = 1\}.$$

*Let  $\gamma > 0$  be such that every eigenvalue of  $A_0$  other than  $\sigma$  has distance at least  $\gamma$  from  $\sigma$ , and let  $a = \|A_1\|$ . Then*

$$|\sigma_k(A(p)) - \sigma - p\nu| \leq \left(\frac{2a^2}{\gamma} + C_0\right)p^2 \quad \text{for } 0 \leq p \leq \min\{p_1, \gamma/(4a)\},$$

*where  $\gamma/(4a)$  is read as  $\infty$  if  $a = 0$ .*

*Proof.* Let  $F_0$  and  $F_1$  be the spans of the eigenvectors of  $A_0$  for the eigenvalues smaller than  $\sigma$  and larger than  $\sigma$ , respectively. Then  $\mathbb{R}^N = F_0 \oplus E \oplus F_1$  orthogonally,  $\dim F_0 = k-1$ ,  $\langle b, A_0 b \rangle \leq (\sigma - \gamma)|b|^2$  for  $b \in F_0$ , and  $\langle b, A_0 b \rangle \geq (\sigma + \gamma)|b|^2$  for  $b \in F_1$ . Note that  $|\nu| \leq a$ . We may assume  $a > 0$ ; for  $a = 0$  the claim follows directly from (12). Let  $0 \leq p \leq \min\{p_1, \gamma/(4a)\}$ ; then  $\gamma - 2ap \geq \gamma/2$ , and

$$(\gamma - 2ap)s^2 - 2ap s \geq \frac{\gamma}{2}s^2 - 2ap s \geq -\frac{2a^2 p^2}{\gamma} \quad \text{for all } s \geq 0. \quad (13)$$

*Lower bound.* By (12) with  $W = F_0^\perp = E \oplus F_1$ ,  $\sigma_k(A(p)) \geq \min\{\langle v, A(p)v \rangle : v \in E \oplus F_1, |v| = 1\}$ . Let  $v = f + b$  with  $f \in E$ ,  $b \in F_1$  and  $|f|^2 + |b|^2 = 1$ . Then  $\langle v, A_0 v \rangle \geq \sigma + \gamma|b|^2$  and

$$\langle v, A_1 v \rangle = \langle f, A_1 f \rangle + 2\langle f, A_1 b \rangle + \langle b, A_1 b \rangle \geq \nu|f|^2 - 2a|b| - a|b|^2 \geq \nu - 2a|b| - 2a|b|^2,$$

because  $\nu|f|^2 = \nu - \nu|b|^2 \geq \nu - a|b|^2$ . Hence, by (13),  $\langle v, A(p)v \rangle \geq \sigma + p\nu + (\gamma - 2ap)|b|^2 - 2ap|b| - C_0 p^2 \geq \sigma + p\nu - (\frac{2a^2}{\gamma} + C_0)p^2$ .

*Upper bound.* Choose  $f_* \in E$  with  $|f_*| = 1$  and  $\langle f_*, A_1 f_* \rangle = \nu$ , and let  $W = F_0 \oplus \mathbb{R}f_*$ , which has dimension  $k$ . By (12),  $\sigma_k(A(p)) \leq \max\{\langle v, A(p)v \rangle : v \in W, |v| = 1\}$ . Let  $v = b + sf_*$  with  $b \in F_0$  and  $|b|^2 + s^2 = 1$ . Then  $\langle v, A_0 v \rangle \leq \sigma - \gamma|b|^2$  and  $\langle v, A_1 v \rangle = \nu s^2 + 2s\langle b, A_1 f_* \rangle + \langle b, A_1 b \rangle \leq \nu + 2a|b| + 2a|b|^2$ . Hence, by (13),  $\langle v, A(p)v \rangle \leq \sigma + p\nu - (\gamma - 2ap)|b|^2 + 2ap|b| + C_0 p^2 \leq \sigma + p\nu + (\frac{2a^2}{\gamma} + C_0)p^2$ .  $\square$

**Lemma 6.2** (expansion at  $p = 0$ ). *Let  $g := \sum_{\lambda \neq 1} (1 - \lambda)^{-1} \Pi_\lambda \delta_e$ .*

- (a)  $(I - K)g = \delta_e - \frac{1}{n}$  and  $\sum_x g(x) = 0$ , and  $g$  is determined by these two properties. Moreover  $g(e) > g(x)$  for all  $x \neq e$ .
- (b)  $|\mu_p(x) - \frac{1}{n} - pg(x)| \leq C_1 p^2$  for  $0 \leq p \leq \frac{1}{2}$  and all  $x$ , with  $C_1 = 2n/(1 - \lambda_2)^2$ .
- (c) Let  $S_1$  be the symmetric matrix with the entries

$$S_1(x, y) = \frac{n}{2} K(x, y) |g(x) - g(y)| \quad (x \neq y), \quad S_1(x, x) = -n \sum_y K(x, y) (g(x) - g(y))^+.$$

Then there are constants  $C_2$  and  $p_1 > 0$ , depending only on  $(G, S)$ , such that  $\|S(p) - (I - K) - pS_1\| \leq C_2 p^2$  for  $0 \leq p \leq p_1$ .

- (d) For every  $f: G \rightarrow \mathbb{R}$ , with  $\Gamma(f) := \sum_{x,y} K(x, y) |g(x) - g(y)| (f(x) - f(y))^2$ ,

$$\langle f, S_1 f \rangle = -\frac{1}{2} \left( n f(e)^2 - \sum_x f(x)^2 \right) - \frac{n}{4} \Gamma(f).$$

*Proof.* (a)  $(I - K)g = \sum_{\lambda \neq 1} \Pi_\lambda \delta_e = \delta_e - \Pi_1 \delta_e = \delta_e - \frac{1}{n}$ , and  $\sum_x g(x) = 0$  because  $\Pi_\lambda \delta_e$  is orthogonal to the constants for  $\lambda \neq 1$ . The kernel of  $I - K$  consists of the constants, so  $g$  is determined. Let  $u = g(e) - g$ . Then  $u(x) - (Ku)(x) = \frac{1}{n} > 0$  for  $x \neq e$ . If  $u$  attained its minimum at some  $x_0 \neq e$ , we would get  $u(x_0) = (Ku)(x_0) + \frac{1}{n} \geq \min u + \frac{1}{n} > u(x_0)$ . So the minimum of  $u$  is attained only at  $e$ , where  $u(e) = 0$ .

(b) By Lemma 2.1(c),  $\mu_p - \frac{1}{n} - pg = \sum_{\lambda \neq 1} \left( \frac{p}{(1-\lambda)+p\lambda} - \frac{p}{1-\lambda} \right) \Pi_\lambda \delta_e$ , and the coefficient is  $-p^2 \lambda / ((1-\lambda)((1-\lambda)+p\lambda))$ . For  $0 \leq p \leq \frac{1}{2}$  we have  $(1-\lambda) + p\lambda \geq \frac{1}{2}(1-\lambda)$ : if  $\lambda \geq 0$  this is clear, and if  $\lambda < 0$  the left side is  $1 + |\lambda|(1-p) \geq \frac{1}{2}(1+|\lambda|)$ . So the coefficient has absolute value at most  $2p^2/(1-\lambda_2)^2$ . There are at most  $n-1$  terms, and  $|\langle \Pi_\lambda \delta_e, x \rangle| \leq 1$ .

(c) By (b) there is  $p_1 > 0$  such that  $\ell_x(p) := \log(n\mu_p(x)) = np g(x) + O(p^2)$  for  $0 \leq p \leq p_1$ ; here and below  $O(p^2)$  stands for a quantity bounded by a constant times  $p^2$ , the constant depending only on  $(G, S)$ . For adjacent  $x, y$ , by (8),

$$S(p)(x, y) = -\frac{1}{d} \exp(-\frac{1}{2}|\ell_x(p) - \ell_y(p)|) = -\frac{1}{d} + \frac{np}{2d} |g(x) - g(y)| + O(p^2),$$

and  $\min\{1, \mu_p(y)/\mu_p(x)\} = \exp(-(\ell_x(p) - \ell_y(p))^+) = 1 - np(g(x) - g(y))^+ + O(p^2)$ , so that  $S(p)(x, x) = 1 - np \sum_y K(x, y)(g(x) - g(y))^+ + O(p^2)$ . Thus all entries of  $S(p) - (I - K) - pS_1$  are  $O(p^2)$ .

(d) Since  $K(x, x) = 0$ ,

$$\langle f, S_1 f \rangle = -n \sum_x f(x)^2 \sum_y K(x, y)(g(x) - g(y))^+ + \frac{n}{2} \sum_{x, y} K(x, y) |g(x) - g(y)| f(x) f(y).$$

In the second sum write  $f(x)f(y) = \frac{1}{2}(f(x)^2 + f(y)^2) - \frac{1}{2}(f(x) - f(y))^2$ ; by the symmetry of  $K(x, y)|g(x) - g(y)|$  it becomes  $\frac{n}{2} \sum_x f(x)^2 \sum_y K(x, y)|g(x) - g(y)| - \frac{n}{4} \Gamma(f)$ . Since  $|s| - 2s^+ = -s$  for real  $s$ , we obtain

$$\langle f, S_1 f \rangle = -\frac{n}{2} \sum_x f(x)^2 (g(x) - (Kg)(x)) - \frac{n}{4} \Gamma(f) = -\frac{n}{2} \left( f(e)^2 - \frac{1}{n} \sum_x f(x)^2 \right) - \frac{n}{4} \Gamma(f),$$

by (a). □

*Proof of Theorem 1.6.* We apply Lemma 6.1 to  $A(p) = S(p)$ ,  $A_0 = I - K$  and  $A_1 = S_1$ , with  $N = n$  and  $k = 2$ ; the hypothesis on  $A(p)$  is Lemma 6.2(c). The eigenvalues of  $I - K$  are  $1 - \lambda_j$ , so  $\sigma_1(A_0) = 0 < \sigma_2(A_0) = 1 - \lambda_2 = 1/\tau(0)$ , and the eigenspace of  $A_0$  for  $\sigma_2(A_0)$  is  $V$ . By Lemma 6.2(d), for  $f \in V$  with  $\sum_x f(x)^2 = 1$  we have  $\langle f, S_1 f \rangle = -Q(f)$  with  $Q$  as in (6). So  $\nu = -s^*$ , and Lemma 6.1 and (9) give (5). The statements on  $\tau$  follow from (5), once we know that  $s^* > 0$ .

Let  $\Pi_V$  be the orthogonal projection onto  $V$  and  $m_2 = \dim V$ . Since  $\Pi_V$  is a polynomial in  $K$ , we have  $\Pi_V(ax, ay) = \Pi_V(x, y)$  for  $a \in G$ , hence  $\Pi_V(x, x) = \Pi_V(e, e) = \frac{1}{n} \text{tr } \Pi_V = \frac{m_2}{n}$ . Let  $f_0 = \Pi_V \delta_e / |\Pi_V \delta_e|$ . Then  $f_0 \in V$ ,  $\sum_x f_0(x)^2 = 1$ , and  $f_0(e) = \Pi_V(e, e) / \Pi_V(e, e)^{1/2} = (m_2/n)^{1/2}$ , so that  $n f_0(e)^2 = m_2$  and

$$s^* \geq Q(f_0) = \frac{m_2 - 1}{2} + \frac{n}{4} \Gamma(f_0).$$

It remains to show that  $\Gamma(f_0) > 0$ . We have  $f_0(e) > 0$  and  $(Kf_0)(e) = \lambda_2 f_0(e) < f_0(e)$ ; so the mean of  $f_0$  over the neighbours of  $e$  is smaller than  $f_0(e)$ , and  $f_0(s) \neq f_0(e)$  for some  $s \in S$ . Moreover  $g(s) < g(e)$  by Lemma 6.2(a). Hence  $\Gamma(f_0) \geq 2K(e, s) |g(e) - g(s)| (f_0(e) - f_0(s))^2 > 0$ . □

By Corollary 2.5,  $\tau$  is then not non-increasing on the open interval  $(0, 1)$  either: if it were, we would have  $\tau(p) \geq \tau(p') > \tau(0)$  for all  $0 < p < p'$  and a fixed small  $p'$ , which contradicts  $\tau(p) \rightarrow \tau(0)$ .

*Remark 6.3.* (a) *Rellich's theorem.* Each difference  $\mu_p(x) - \mu_p(y)$  is a rational function of  $p$ , so it has a constant sign on some interval  $(0, p_3)$ . There  $S(p)$  coincides with a family of symmetric matrices which is real-analytic in a neighbourhood of  $p = 0$ , and by Rellich's theorem [Rel37], [Kat95, Chapter II] the eigenvalues of such a family are given by analytic branches; the first-order terms of the branches through  $1 - \lambda_2$  are the eigenvalues of the compression of  $S_1$  to  $V$ . This is the way in which (5) was first derived; Lemma 6.1 replaces it by an elementary argument.

(b) *Examples.* On  $K_n$ , Proposition 1.1 gives  $s^* = n - 1$ . On the hypercube, Corollary 1.4 gives  $1/\tau(p) \leq 1/H_d(p) = \frac{2}{d} - \frac{d+1}{2} p + O(p^2)$ , hence  $s^* \geq \frac{d+1}{2}$ ; in the computations of Section 8,  $s^* = Q(f_0) = \frac{d+1}{2}$  for  $d \leq 8$ . In general  $s^*$  can be larger than  $Q(f_0)$ : for  $S_4$  with the three adjacent transpositions the computed values are  $Q(f_0) = 1.62763$  and  $s^* = 1.62894$ .

| Item                                  | Reading                                               | Answer       | Witness (proved)                                                                                                                                                                                                               |
|---------------------------------------|-------------------------------------------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (i) $\tau$ is non-increasing in $p$   | as posed                                              | no           | every finite Cayley graph: $\tau(p) > \tau(0)$ for all small $p > 0$ (Theorem 1.6); on $K_n$ , $\tau$ is strictly increasing (Proposition 1.1, known)                                                                          |
| (i)                                   | orientation of $p$ reversed: $\tau$ is non-decreasing | no           | cycles $C_n$ , $n \geq 7$ : $\tau(0) > 2 = \tau(1^-)$ (Corollary 4.1); hypercubes $Q_d$ , $d \geq 3$ : $\tau(p) > d = \tau(1^-)$ for $p$ close to 1 (Corollary 1.4(c))                                                         |
| (ii) $\tau(p) \leq C\tau(\infty)$     | (a) $\tau(\infty) = \tau(0)$                          | no           | $K_n$ : $\sup_p \tau(p)/\tau(0) = n$ (known); $Q_d$ : $\sup_p \tau(p)/\tau(0) \geq \frac{2(d-1)}{15}$ (Corollary 1.4); bounded degree with a two-sided spectral gap: $\sup_p \tau(p)/\tau(0) \geq c(\log n)^2$ (Corollary 1.5) |
| (ii)                                  | (b) $\tau(\infty) = \tau(1^-) = d$                    | no           | cycles: $\tau(0)/\tau(1^-) \rightarrow \infty$ (Corollary 4.1); $Q_d$ : $\sup_p \tau(p)/\tau(1^-) \geq \frac{d-1}{15}$ ; bounded degree with a two-sided spectral gap: $\sup_p \tau(p)/\tau(1^-) \geq c(\log n)^2$             |
| (iii) a decreasing bound on $\tau(p)$ | —                                                     | not answered | necessary conditions only (Remark 7.1)                                                                                                                                                                                         |
| opening question                      | —                                                     | not answered | statements valid on every finite Cayley graph: Theorems 1.2, 1.3 and 1.6                                                                                                                                                       |

Table 1: The items of the problem page [Ald09], as literally posed, for the relaxation time  $\tau = 1/(1 - \beta_2)$ . “No” means that the statement fails on the graphs named. The constants  $c > 0$  depend on the degree and on the spectral gap.

(c) *The size of the interval.* The constant in Lemma 6.1 contains  $1/\gamma$ , and here  $\gamma \leq \lambda_2 - \lambda'$  for the next smaller eigenvalue  $\lambda'$  of  $K$ . When this difference is small, the first-order term dominates only on a small interval: in the computations there are Cayley graphs on which  $(1/\tau(0) - 1/\tau(p))/p$  at  $p = 10^{-5}$  still differs from  $s^*$  by ten per cent.

## 7 Conclusions and remarks

Table 1 collects what the results say about the items of the problem page, for the relaxation time (3) and under the readings of the introduction. It has to be read together with the caveat of Section 1.4: the question that was intended may differ from the literal one.

*Remark 7.1* (item (iii)). We have no answer to item (iii). The theorems give necessary conditions. If  $B(p)$  is an upper bound for  $\tau(p)$  on a given Cayley graph, then  $B(p) \geq \max\{2 \operatorname{Var}_p|x|, qm(1 - m)/(m - p)\}$  for every  $p$ ,  $\liminf_{p \rightarrow 1} B(p) \geq d$ , and  $B(p) > \tau(0)$  for all small  $p > 0$ . If  $B$  is non-increasing in  $p$ , then  $B(p) \geq \sup_{p' \geq p} \tau(p')$ ; on the hypercube this forces  $B(p) \geq d(d - 1)/15$  for  $p \leq \frac{1}{d}$ , although  $\tau(0) = \frac{d}{2}$  and  $\tau(1^-) = d$ . So a non-increasing bound cannot be a universal constant times  $\tau(\infty)$ , under either reading. The only general upper bound we know is the comparison bound

$$\tau(p) \leq \tau(0) \cdot \frac{\max_x \mu_p(x)}{\min_x \mu_p(x)}, \quad (14)$$

which follows from (7), because  $c_p(x, y) \geq n \min_z \mu_p(z) \cdot c_0(x, y)$  and  $\operatorname{Var}_p(f) \leq n \max_z \mu_p(z) \cdot \operatorname{Var}_0(f)$ . It is an instance of the comparison of Dirichlet forms in [DH92, Section 2, Lemma 2], where it is also remarked that such bounds can be far off for Metropolis chains. The bound (14) is not decreasing in  $p$ , and it tends to infinity as  $p \rightarrow 1$ , in contrast to Theorem 1.2.

*Remark 7.2* (the absolute spectral gap). Some authors define the relaxation time by the absolute gap,  $\tau_{\text{abs}}(p) = 1/(1 - \max\{\beta_2(p), |\beta_n(p)|\})$ . The results change as follows.

(a) Theorems 1.2 and 1.3 hold for  $\tau_{\text{abs}}$ :  $\tau_{\text{abs}} \geq \tau$ , and by the proof of Theorem 1.2 all eigenvalues of  $P_1$  other than 1 have absolute value at most  $1 - \frac{1}{d}$ , so that  $\tau_{\text{abs}}(p) \rightarrow d$  as well.

(b) On  $K_n$  with  $n \geq 3$  the eigenvalues of  $P_p$  are 1,  $\beta_2(p) = 1 - 1/\tau(p)$ , and  $-\frac{1}{n-1}$  with multiplicity  $n-2$ : from every  $y \neq e$  all proposals are accepted, so  $P_p f = -\frac{1}{n-1}f$  for the functions  $f$  with  $f(e) = 0$  and  $\sum_y f(y) = 0$ . Hence  $\tau_{\text{abs}}(p) = \max\{\tau(p), \frac{n-1}{n-2}\}$ . For  $n \geq 4$  this is non-decreasing and not constant. For  $K_3$  it is constant, equal to 2. For  $K_2$  the eigenvalues are 1 and  $-(1-p)$ , so  $\tau_{\text{abs}}(p) = 1/p$ , which is strictly decreasing, while  $\tau(p) = 1/(2-p)$ . Thus, for  $\tau_{\text{abs}}$ , item (i) holds on  $K_2$  and, in the weak sense, on  $K_3$ ; the first row of Table 1 is a statement about  $\tau$  and is not true for  $\tau_{\text{abs}}$ .

(c) Theorem 1.6 holds for  $\tau_{\text{abs}}$  on the Cayley graphs with  $|\lambda_n| < \lambda_2$ , since then  $\tau_{\text{abs}} = \tau$  near  $p = 0$  by Lemma 2.4. Otherwise it can fail. On bipartite graphs, such as even cycles and hypercubes,  $\tau_{\text{abs}}(0) = \infty$ . On the odd cycles  $|\lambda_n| = \cos \frac{\pi}{n} > \lambda_2$ , and in the computations for  $C_5, \dots, C_{13}$  one finds  $\tau_{\text{abs}}(p) < \tau_{\text{abs}}(0)$  for small  $p > 0$ . We have not proved this.

*Remark 7.3* (the shape of  $\tau$ : computations). What is proved about the shape of  $\tau$  on  $[0, 1]$  is this:  $\tau$  is continuous, it increases at  $p = 0$  with the slope  $s^* \tau(0)^2$ , and it tends to  $d$ ; it is increasing on the complete graphs; and it is not monotone on the cycles  $C_n$ ,  $n \geq 7$ , and on the hypercubes  $Q_d$ ,  $d \geq 3$ . Everything else in this remark is computation (Section 8), on finite grids of values of  $p$ , and proves nothing.

Three shapes occur (Figure 1). Among the 1204 Cayley graphs of the 23 groups of order at most 12 (all symmetric generating sets), on a grid of 23 values of  $p$  from 0.01 to 0.995,  $\tau$  is increasing on 776 graphs and goes up and then down on 428. These two numbers depend on the grid: on a grid of 209 values which reaches  $p = 1 - 10^{-6}$  they are 764 and 440, because on twelve Cayley graphs of the dihedral group of order 12, of degree 5 and with the same spectrum,  $\tau$  has its maximum near  $p = 0.99$  and exceeds  $d$  there by less than  $10^{-3}$ . Among 2000 random Cayley graphs of 29 groups of orders 15 to 120, 21 show neither of these two shapes on the grid of 23 values. Two of them were recomputed on the grid  $j/400$  and go up, down and up again:  $\mathbb{Z}/16$  with  $S = \{\pm 3, \pm 5, 8\}$  (Figure 1(c)), and  $\mathbb{Z}/21$  with  $S = \{\pm 2, \pm 7, \pm 9\}$  (local maximum 5.923 near  $p = 0.41$ , local minimum 5.854 near  $p = 0.725$ ,  $d = 6$ ). So  $\tau$  is not always increasing, and not always unimodal.

On the cycles  $C_n$  (16 values of  $n$  from 5 to 256)  $\tau$  goes up and then down, and  $\max_p \tau(p)/\tau(0)$  decreases from 1.69 for  $n = 5$  to 1.05 for  $n = 256$ ; so on cycles item (ii) under reading (a) may well hold, and for large  $n$  item (i) fails by a few per cent only. On the hypercubes (by Proposition 5.2; ten values of  $d$  from 24 to 512) the maximum of  $\tau(p)/d^2$  over the grids lies between 0.15 and 0.19, and  $\tau(1/d) \approx 0.16 d^2$ ; so the order  $d^2$  in Corollary 1.4(a) is the right one in these computations, and the constant  $\frac{1}{15}$  is not. The convergence in Theorem 1.2 can be slow: on the cycles  $C_n$ ,  $5 \leq n \leq 14$ ,  $\tau(1 - \varepsilon) - d$  is of the order  $\sqrt{\varepsilon}$ .

## 8 Computations and reproducibility

All theorems, propositions, corollaries and lemmas of this note are proved without computation. The computations test the statements and intermediate steps of the proofs on explicit graphs, in exact rational arithmetic or with high precision, and they produce Figure 1 and the observations of Remark 7.3. The reproducibility package has four parts.

*The exact program of this note* (`writing_stage/`). In the module `metro.py` the target (1) is the exact rational solution of the system in Lemma 2.1(a); the kernel (2), the limit matrix (10), the function  $g$  and the level chain of Proposition 5.2 are exact as well; only eigenvalues are computed in floating point, with 30 digits. The program `check_note.py` tests the statements as printed on 34 graphs:  $K_n$  ( $n \leq 6$ ),  $C_n$  ( $n \leq 12$ ),  $Q_d$  ( $d \leq 5$ ),  $S_3$  and  $S_4$  with all transpositions,  $S_4$  with adjacent transpositions, two dihedral groups, seven further Cayley graphs of abelian groups,  $A_4$ , and

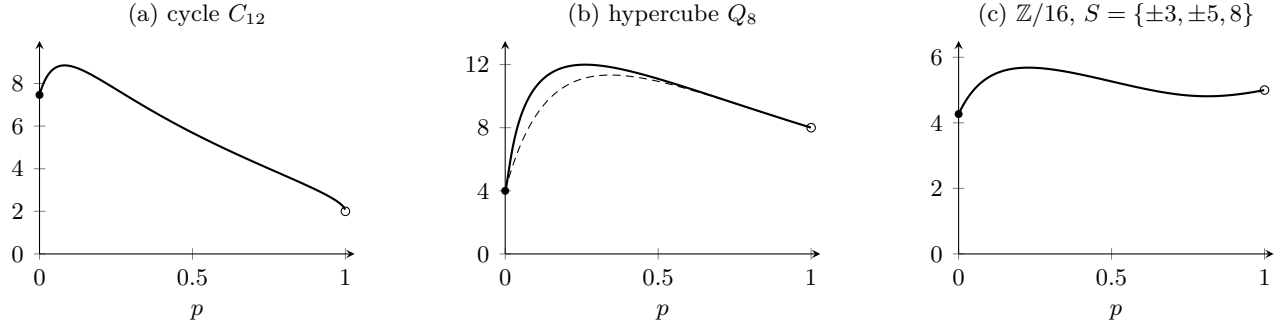


Figure 1: The function  $\tau$  on three Cayley graphs (computed on a grid of 162 values of  $p$ ). The dot is  $\tau(0)$ , and the circle at  $p = 1$  is the limit  $d$  of Theorem 1.2. (a)  $C_{12}$ :  $\tau(0) = 7.46$ , maximum 8.85 near  $p = 0.08$ ,  $d = 2$ . (b)  $Q_8$ :  $\tau(0) = 4$ , maximum 11.99 near  $p = 0.26$ ,  $d = 8$ ; the dashed curve is the lower bound  $H_8$  of Corollary 1.4. (c)  $\mathbb{Z}/16$  with  $S = \{\pm 3, \pm 5, 8\}$ :  $\tau(0) = 4.27$ , local maximum 5.68 near  $p = 0.23$ , local minimum 4.81 near  $p = 0.81$ ,  $d = 5$ .

the Petersen graph, which is vertex-transitive but not a Cayley graph (Remark 2.6). Its parts test: (A) Lemma 2.1, Remark 2.2 and the example of Remark 2.3, exactly; (B) Proposition 3.1 on 42 random rational targets and Proposition 1.1 for  $n \leq 12$ , exactly; (C) the structure of  $P_1$ , exactly, the limit in Theorem 1.2, and the formulas of [DH92] quoted in Remark 4.2; (D) Theorem 1.3 at 25 values of  $p$  on every graph; (E) Propositions 5.1 and 5.2 for  $d \leq 5$ , and Corollary 1.4(a)–(c) with (11) in rational arithmetic for  $d$  up to 1000 and symbolically; (F) the steps of the proof of Corollary 1.5; (G) Lemma 6.2, the formula (6) against difference quotients of  $1/\tau$ , and Lemma 6.1 with its constant on 300 random symmetric matrices; (H) (14), Remark 7.2 and the shapes of Figure 1. All checks pass; the run takes less than one minute. The program `make_figures.py` produced the data and the code of Figure 1.

*The programs with which the results were first obtained (original/)*: a symbolic check of Proposition 1.1 for  $n \leq 8$ ; the exact limit matrix  $P_1$  on 38 graphs; the bounds of Theorem 1.3 on 40 graphs and the slope at  $p = 0$  on 41 graphs, in double precision; both limits on six graphs with 60 digits; the hypercube, with the level chain up to  $d = 1024$ ; the steps of Corollary 1.5 on eleven graphs; and the families of Remark 7.3.

*The programs of the first independent verification run (independent\_run/)*, written without reading or running those of `original/`, with exact rational targets and eigenvalues with 40 to 60 digits: Proposition 1.1 symbolically for  $n \leq 8$  and in 280 exact rational cases; Theorems 1.2, 1.3 and 1.6 on 46 graphs; all 1204 Cayley graphs of the groups of order at most 12 and 2000 random Cayley graphs of groups of orders 15 to 120, in double precision, without a violation of any statement; the hypercube up to  $d = 400$ ; the steps of the proof of Corollary 1.5 on  $\text{SL}_2(\mathbb{Z}/53)$ , which has 148,824 elements and satisfies  $n \geq 32d^6$ ; the formulas of [DH92] with 50 digits; the absolute gap; and the shapes.

*The programs of the second independent verification run (independent\_run\_2/)*, written from the text of this note before the other programs were opened; exact rational targets, and eigenvalues with 25 to 50 digits where the size permits. They test the statements of Sections 2 to 6 on 40 graphs, and among other things: Proposition 3.1 on 56 random rational targets; the blocks of  $P_1$  by exact counts of real roots; all steps of the proof of Corollary 1.5 on  $\text{PGL}_2(\mathbb{Z}/31)$  with three generators (29,760 elements,  $n \geq 32d^6$ ); Lemma 6.1 with its constant in 7200 random cases, also with multiple eigenvalues, and on the matrices  $S(p)$ ; the 1204 Cayley graphs of the groups of order at most 12, built anew, on three grids; the two families of Remark 4.2, lumped and not lumped ( $S_N$  for  $N \leq 7$ ,  $Q_d$  for  $d \leq 8$ ); and the numbers of Figure 1 and Remark 7.3. All 106 checks pass.

All programs of the first three parts with a recorded output were run again when this note was

written, one process at a time, and once more in the second verification run, from an extracted copy of the archive. All recorded outputs were reproduced, up to fields which record running times; see the file `README.md` of the package.

**Verification.** The results were obtained in three stages, all AI-assisted. In the first stage the formula of Proposition 1.1 was found by a direct computation. The second stage derived it again and found Theorems 1.2 and 1.3, Corollary 1.4(a), (b), Corollary 1.5, and Theorem 1.6 in the form  $1/\tau(p) \leq 1/\tau(0) - Q(f_0)p + O(p^2)$ , with proofs and with the programs in `original/`. A first independent verification run then examined the written results of the second stage. It re-derived every proof, wrote its own programs, and tried to break each statement, among other things by other conventions for  $p$ ,  $T_p$ , the walk and the relaxation time, by bipartite graphs, by multiple and nearly multiple eigenvalues  $\lambda_2$ , and by a search for a counterexample among the 3204 Cayley graphs mentioned above. It found no gap and no counterexample. It required six corrections, which are all made in this note: the attribution of Proposition 1.1 to [Liu96, DSC98, AF02]; the reading of [DSC98, DH92], which had not been accessible in the second stage, and the presentation of the two families of [DH92] as prior instances; the formulation “as literally posed” with the caveat of Section 1.4; the restriction of all statements to the relaxation time (3), with Remark 7.2; a remark on the shapes which does not suggest that there are only two; and proofs for the equality in Proposition 5.1(c), for Proposition 5.2 and for the exact derivative (5), which the second stage had only tested. The representation in Proposition 5.1(a), the proof of Proposition 5.2, the derivative (5) (through Rellich’s theorem) and the examples with three monotone pieces also come from this run.

When this note was written, the following was added or changed after that run. Section 6 is organised around the matrices  $S(p)$  and the elementary Lemma 6.1, which takes the place of Rellich’s theorem; Lemma 6.2(c) and (d) are the corresponding reformulation of the expansion. Corollary 1.4(c) with (11) is new; the run had the reversed orientation on hypercubes for  $d \geq 17$  only, from Corollary 1.4(a). Remark 2.3 with its example, the bound  $s^* \geq \frac{d+1}{2}$  of Remark 6.3(b), and the observation that  $\rho > 0$  in the proof of Corollary 1.5 were added.

A second independent verification run, also AI-assisted, examined the final text, and these parts first. It re-derived Lemmas 6.1 and 6.2 and the proof of Theorem 1.6 line by line, the identity (11), the example of Remark 2.3 and the constants in the proof of Corollary 1.5, and then all other proofs as they are printed here. It read the problem page and the index again and compared them with the introduction; compared the citations of [AF02], [DSC98] and [DH92] with the texts named below; wrote its own programs (Section 8); ran the package again from the archive; and repeated the literature search. It found no wrong statement and no gap. It asked for ten corrections of the text and of the package, which were made. Those which concern the content are that Section 1.4 and Remark 4.2 now distinguish the lumped chains, whose eigenvalues are given in [DH92], from the Metropolis chains themselves; that the bound (14) is credited; that Remark 7.3 gives the numbers of shapes also on a finer grid, on which twelve graphs change their class; and that the caveat of Section 1.4 names the two items which the complete graphs contradict. No proof depends on a computation.

**Scope and priority.** This note gives a partial answer to the question of [Ald09]: negative answers to its items (i) and (ii) as literally posed, for the relaxation time (3) and under the readings of the introduction. Item (iii) and the opening question are not answered, and we do not know what was intended (Section 1.4). The problem is due to Aldous [Ald09], and the definitions are those of Aldous and Fill [AF02]. Proposition 1.1 is a corollary of results of Liu [Liu96], of Diaconis and Saloff-Coste [DSC98] and of Aldous and Fill [AF02]. That the relaxation time of a Metropolis chain on a Cayley graph increases up to the degree when the target concentrates at the identity can be seen in the formulas of Diaconis and Hanlon [DH92] for two families (Remark 4.2), and the comparison bound

(14) is an instance of a lemma of [DH92]. The Metropolis algorithm goes back to [MRR<sup>+</sup>53, Has70], the perturbation theory to Rellich [Rel37] (see [Kat95]), and the expander families quoted after Corollary 1.5 to Lubotzky, Phillips and Sarnak [LPS88]. To our knowledge Theorems 1.2, 1.3 and 1.6 and Corollaries 1.4 and 1.5 are not in print; the methods are standard.

What was read. The problem page and the index of the list were read on 9 October 2026, the last time in the second verification run (a copy of 3 December 2015 in the Wayback Machine has the same text, except that the range  $0 < p < 1$  in the first sentence is missing there and one word is misspelt; a copy of the index of 18 July 2010 already lists the entry; the list of 2003 which is linked from the index does not contain the problem). Of [AF02] we read, in the HTML edition of 2014, Section 3.4 around (3.36)–(3.41), Sections 3.6.1 and 3.6.3, Section 11.2 and Section 11.4; the rest was not read, and a text search of the PDF edition found no statement on Metropolis chains on Cayley graphs. The first independent verification run read all 17 pages of [DSC98], in the copy on the web page of its second author, and all pages of [DH92] in the technical-report version (Technical Report No. 392, March 1992; 29 scanned pages); the second run read Sections 2.1 and 2.2 of [DSC98] and Sections 1–4 of the report again. The published version of [DH92] was not read and may differ; the site of its publisher refused automated access in both runs. In the scan, the general eigenvalue formula in Section 4 of the report shows a plus sign where the example  $N = 3$  of the report and its formula for the second eigenvalue require a minus sign; this note uses only the second eigenvalue, which the computations confirm for  $N \leq 7$ . The paper [Liu96] was not read, and the second run found no open copy of it; what we say about it is taken from [AF02] and [DSC98]. The paper [LPS88] was not read for this note, and of [Mic02] and [Kar11] only the abstracts were read. The works [MRR<sup>+</sup>53, Has70, Rel37, Kat95, HJ12] were not consulted for this note; from them only the min–max principle (12) is used in a proof. The bibliographic data of all references with a DOI were checked with Crossref.

The searches were made on 9 October 2026, in the second stage and in the two independent verification runs: 13, 34 and 38 completed queries to the arXiv API (Metropolis combined with Cayley graphs, relaxation time, geometric time, Green function, finite group, hypercube, word length, monotonicity, temperature and similar terms; the queries are listed in the package); 21 and 17 queries to OpenAlex in the second stage and in the first run, among them the lists of the works citing [DSC98] (154, and 31 for its conference version) and [DH92] (41), of which the titles were read; queries to Crossref and zbMATH (14 and 10 of them in the second run); and five web searches. In the second run the search queries to OpenAlex (17) and to Semantic Scholar (6) were refused because of the limits for requests without a key; they were not repeated. No treatment of the family  $\mu_p$ , no answer to the question, and no citation of the problem page was found. The nearest titles concern the targets  $\theta^{|x|}$  on the symmetric group and related models, the effect of the energy gap on the mixing time of the Metropolis algorithm, and Metropolis chains with uniform stationary distribution on graphs with symmetries; these papers were not read. The arXiv API returned only 415 records with Metropolis in the title, so its index is incomplete for this purpose and this part of the search is weak evidence. A search that finds nothing is not a proof of novelty, and no priority is claimed.

AI-assisted tools were used to search the problem corpus and the literature, to find the arguments, to write and run the verification programs, and to help prepare this manuscript. The author has reviewed the proofs, the computations and the sources and is responsible for all claims. The note has not been peer reviewed.

## References

- [AF02] David Aldous and James Allen Fill. Reversible Markov chains and random walks on graphs. Unfinished monograph, recompiled 2014, <https://www.stat.berkeley.edu/~aldous/RWG/book.html>, 2002.

- [Ald09] David Aldous. Metropolis on Cayley graphs. Open problem page, <https://www.stat.berkeley.edu/~aldous/Research/OP/cayley.html>, 2009. First written March 2009; entry (1.9) of the list of open problems <https://www.stat.berkeley.edu/~aldous/Research/OP/index.html>; read on 9 October 2026.
- [DH92] Persi Diaconis and Phil Hanlon. Eigen-analysis for some examples of the Metropolis algorithm. In Donald St. P. Richards, editor, *Hypergeometric Functions on Domains of Positivity, Jack Polynomials, and Applications (Tampa, FL, 1991)*, volume 138 of *Contemporary Mathematics*, pages 99–117. American Mathematical Society, Providence, RI, 1992. Technical-report version: Technical Report No. 392, Department of Statistics, Stanford University, March 1992, <https://purl.stanford.edu/wr236pp7937>. doi:10.1090/conm/138/1199122.
- [DSC98] Persi Diaconis and Laurent Saloff-Coste. What do we know about the Metropolis algorithm? *Journal of Computer and System Sciences*, 57(1):20–36, 1998. doi:10.1006/jcss.1998.1576.
- [Has70] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. doi:10.1093/biomet/57.1.97.
- [HJ12] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, second edition, 2012. doi:10.1017/CB09781139020411.
- [Kar11] Vladislav Kargin. Relaxation time is monotone in temperature in the mean-field Ising model. *Statistics & Probability Letters*, 81(8):1094–1097, 2011. doi:10.1016/j.spl.2011.03.001.
- [Kat95] Tosio Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Springer, Berlin, Heidelberg, 1995. Reprint of the 1980 edition. doi:10.1007/978-3-642-66282-9.
- [Liu96] Jun S. Liu. Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Statistics and Computing*, 6(2):113–119, 1996. doi:10.1007/BF00162521.
- [LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988. doi:10.1007/BF02126799.
- [Mic02] Laurent Miclo. About relaxation time of finite generalized Metropolis algorithms. *The Annals of Applied Probability*, 12(4):1492–1515, 2002. doi:10.1214/aoap/1037125871.
- [MRR<sup>+</sup>53] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953. doi:10.1063/1.1699114.
- [Rel37] Franz Rellich. Störungstheorie der Spektralzerlegung. I. Mitteilung. Analytische Störung der isolierten Punkteigenwerte eines beschränkten Operators. *Mathematische Annalen*, 113(1):600–619, 1937. doi:10.1007/BF01571652.
- [ula26] ulamai. UnsolvedMath. Hugging Face dataset, <https://huggingface.co/datasets/ulamai/UnsolvedMath>, 2026. Version 1.6.0; record AMR-096-0008.