

A note on adiabatic theorem for Markov chains

Yevgeniy Kovchegov*

Abstract

We state and prove a version of an adiabatic theorem for Markov chains using well known facts about mixing times. We extend the result to the case of continuous time Markov chains with bounded generators.

1 Introduction

The adiabatic theorem was first stated by Max Born and Vladimir Fock in 1928. It asserts that “a physical system remains in its instantaneous eigenstate if a given perturbation is acting on it slowly enough and if there is a gap between the eigenvalue and the rest of the Hamiltonian’s spectrum” (see Wikipedia page on adiabatic theorem).

In this work, we state and prove the corresponding theorem for Markov chains using some of the machinery of mixing times and relaxation times of Markov chains that was developed so successfully in the last thirty years (see Aldous (1983), Aldous and Fill, Burton and Kovchegov (2009), Levin et al (2008) and references therein). Stated in terms of Markov chains, the adiabatic theorem is intuitively simple and accessible. It is important to point out that despite clear similarities and relation between the two, the quantum adiabatic theorem and the adiabatic theorem of this paper are different.

First we state a version of the adiabatic theorem. Given two Hamiltonians, $\mathcal{H}_{initial}$ and \mathcal{H}_{final} , acting on a quantum system. Let $\mathcal{H}(s) = (1-s)\mathcal{H}_{initial} + s\mathcal{H}_{final}$. Suppose the system evolves according to $\mathcal{H}(t/T)$ from time $t = 0$ to time T (the so called *adiabatic evolution*). The adiabatic theorem of quantum mechanics states that for T large enough, the final state of the system will be close to the ground state of \mathcal{H}_{final} . They are ϵ close in ℓ^2 norm whenever $T \geq \frac{C}{\epsilon\beta^3}$, where β is the least spectral gap of $\mathcal{H}(s)$ over all $s \in [0, 1]$, and C depends linearly on a square of the distance between $\mathcal{H}_{initial}$ and \mathcal{H}_{final} . See Messiah (1958). There are many versions of the adiabatic theorem.

Now, let us list the main concepts involved in the theory of mixing and relaxation times that we will need in the following sections. We refer the readers to Levin et al (2008) and Aldous and Fill for details.

*Department of Mathematics, Oregon State University, Corvallis, OR 97331-4605, USA
kovchegy@math.oregonstate.edu

Definition 1. If μ and ν are two probability distributions over Ω , then the *total variation distance* is

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = \sup_{A \subset \Omega} |\mu(A) - \nu(A)|$$

Observe that the total variation distance measures the coincidence between the distributions on a scale from zero to one.

Definition 2. Suppose P is an irreducible and aperiodic finite Markov chain with stationary distribution π , i.e. $\pi P = \pi$. Given an $\epsilon > 0$, the mixing time $t_{mix}(\epsilon)$ is defined as

$$t_{mix}(\epsilon) = \inf \{t : \|\nu P^t - \pi\|_{TV} \leq \epsilon, \text{ for all probability distributions } \nu\}$$

Now, suppose $P = \left(p(x, y)\right)_{x, y}$ is *reversible*, i.e. P satisfies the detailed balance condition

$$\pi(x)p(x, y) = \pi(y)p(y, x) \quad \text{for all } x, y \text{ in the sample space } \Omega$$

It is easy to check that if P is reversible, then P is self-adjoint with respect to an inner product induced by π , and as such will have all real eigenvalues. The difference $\beta = 1 - |\lambda_{(2)}|$ between the largest eigenvalue (i.e. $\lambda_1 = 1$) and the second largest (in absolute value) eigenvalue $\lambda_{(2)}$ is called the *spectral gap* of P . The relaxation time is defined as

$$\tau_{rlx} = \frac{1}{\beta}$$

One can show the following relationship between mixing and relaxation times.

Theorem 1. *Suppose P is a reversible, irreducible and aperiodic Markov chain with state space Ω and stationary distribution π . Then*

$$(\tau_{rlx} - 1) \log(2\epsilon)^{-1} \leq t_{mix}(\epsilon) \leq \tau_{rlx} \log(\epsilon \min_{x \in \Omega} \pi(x))^{-1}$$

2 Adiabatic theorem for Markov chains

Given two transition probability operators, $P_{initial}$ and P_{final} , with a finite state space Ω . Suppose P_{final} is irreducible and aperiodic and π_f is the unique stationary distribution for P_{final} . Let

$$P_s = (1 - s)P_{initial} + sP_{final}$$

Definition 3. Given $\epsilon > 0$, a time T_ϵ is called the *adiabatic time* if it is the least T such that

$$\max_{\nu} \|\nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{T-1}{T}} P_1 - \pi_f\|_{TV} \leq \epsilon,$$

where the maximum is taken over all probability distributions ν over Ω .

Theorem 2. Let t_{mix} denote the mixing time for P_{final} . Then the adiabatic time

$$T_\epsilon = O\left(\frac{t_{mix}(\epsilon/2)^2}{\epsilon}\right)$$

Proof. Observe that

$$\nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{T-1}{T}} P_1 = \frac{T!}{N! T^{T-N}} \nu_N P_{final}^{T-N} + \mathcal{E},$$

where $\nu_N = \nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{N}{T}}$ and \mathcal{E} is the rest of the terms. Hence, by the triangle inequality,

$$\max_\nu \|\nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{T-1}{T}} P_1 - \pi_f\|_{TV} \leq \max_\nu \|\nu P_{final}^{T-N} - \pi_f\|_{TV} \cdot \frac{T!}{N! T^{T-N}} + S_N,$$

where $0 \leq S_N \leq 1 - \frac{T!}{N! T^{T-N}}$.

Let $T = K t_{mix}(\epsilon/2)$ and $N = (K-1) t_{mix}(\epsilon/2)$, so that $\max_\nu \|\nu P_{final}^{T-N} - \pi_f\|_{TV} \leq \epsilon/2$. Observe that

$$e^{\int_N^T \log x dx - (T-N) \log T} \leq \frac{T!}{N! T^{T-N}} = e^{\sum_{j=N+1}^T \log j - (T-N) \log T} \leq 1$$

and therefore, expressing T and N via $t_{mix}(\epsilon/2)$, and simplifying, we obtain

$$\left(\frac{\left(1 + \frac{1}{K-1}\right)^{K-1}}{e}\right)^{t_{mix}(\epsilon/2)} = e^{N \log \frac{T}{N} - (T-N)} \leq \frac{T!}{N! T^{T-N}} \leq 1$$

So

$$0 \leq S_N \leq 1 - \frac{T!}{N! T^{T-N}} \leq 1 - \left(\frac{\left(1 + \frac{1}{K-1}\right)^{K-1}}{e}\right)^{t_{mix}(\epsilon/2)}$$

and we need to find the least K such that

$$1 - \left(\frac{\left(1 + \frac{1}{K-1}\right)^{K-1}}{e}\right)^{t_{mix}(\epsilon/2)} \leq \epsilon/2$$

Now, since $\log(1+x) = x - \frac{x^2}{2} + O(x^3)$, the least such K is approximated as follows

$$K \approx \frac{t_{mix}(\epsilon/2)}{-2 \log(1 - \epsilon/2)} \approx \frac{t_{mix}(\epsilon/2)}{\epsilon}$$

Thus for $T = K t_{mix}(\epsilon/2) \approx \frac{t_{mix}(\epsilon/2)^2}{\epsilon}$,

$$\max_\nu \|\nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{T-1}{T}} P_1 - \pi_f\|_{TV} \leq \epsilon$$

□

Observe that Theorem 2 is independent of the distance. We can use ℓ^2 norm in the definitions of adiabatic and mixing times, and arrive to the same result.

A Markov chain P_{final} over finite sample space Ω , if it is reversible, will have a spectral gap. Here we apply Theorem 1 to the result in Theorem 2.

Corollary. *Suppose $P_{initial}$ and P_{final} are Markov chains with state space Ω . If P_{final} is reversible, irreducible and aperiodic with its spectral gap $\beta > 0$, then*

$$T_\epsilon = O\left(\frac{\log^2\left[\frac{2}{\epsilon \min_{x \in \Omega} \pi_f(x)}\right]}{\epsilon \beta^2}\right)$$

3 Continuous time Markov processes

In the case of continuous time Markov chains, an equivalent result is produced via the method of uniformization and order statistics. Suppose Q is a bounded Markov generator for a continuous time Markov chain $P(t)$, and $\lambda \geq \max_{i \in \Omega} \sum_{j: j \neq i} q(i, j)$ is the upper bound on the departure rates over all states. The method of *uniformization* provides an expression for the transition probabilities $P(t)$ as follows:

$$P(t) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} P_\lambda^n, \quad \text{where } P_\lambda = I + \frac{1}{\lambda} Q$$

The expression is obtained via conditioning on the number of arrivals in a Poisson process with rate λ .

The definition of a Mixing time is similar in the case of continuous time processes.

Definition 4. Suppose $P(t)$ is an irreducible and finite continuous time Markov chain with stationary distribution π . Given an $\epsilon > 0$, the mixing time $t_{mix}(\epsilon)$ is defined as

$$t_{mix}(\epsilon) = \inf \{t : \|\nu P(t) - \pi\|_{TV} \leq \epsilon, \quad \text{for all probability distributions } \nu\}$$

Suppose $Q_{initial}$ and Q_{final} are two bounded generators for continuous time Markov processes over a finite state space Ω , and π_f is the only stationary distribution for Q_{final} . Let

$$Q[s] = (1 - s)Q_{initial} + sQ_{final}$$

be a time non-homogeneous generator. Given $T > 0$, let $P_T(t_1, t_2)$ ($0 \leq t_1 \leq t_2 \leq T$) denote a matrix of transition probabilities of a Markov process generated by $Q\left[\frac{t}{T}\right]$ over $[t_1, t_2]$ time interval.

Observe that the continuous time Markov adiabatic evolution is governed by

$$\frac{d\nu_t}{dt} = \nu_t Q\left[\frac{t}{T}\right], \quad t \in [0, T]$$

while the quantum adiabatic evolution is described via the corresponding Schrödinger's equation $\frac{d\nu_t}{dt} = -i\nu_t \mathcal{H}\left(\frac{t}{T}\right)$.

Definition 5. Given $\epsilon > 0$, a time T_ϵ is called the *adiabatic time* if it is the least T such that

$$\max_{\nu} \|\nu P_T(0, T) - \pi_f\|_{TV} \leq \epsilon,$$

where the maximum is taken over all probability distributions ν over Ω .

We will state and prove the following adiabatic theorem.

Theorem 3. Let t_{mix} denote the mixing time for Q_{final} . Take λ such that

$$\lambda \geq \max_{i \in \Omega} \sum_{j: j \neq i} q_{initial}(i, j) \quad \text{and} \quad \lambda \geq \max_{i \in \Omega} \sum_{j: j \neq i} q_{final}(i, j),$$

where $q_{initial}(i, j)$ and $q_{final}(i, j)$ are the rates in $Q_{initial}$ and Q_{final} respectively. Then the adiabatic time

$$T_\epsilon \leq \frac{\lambda t_{mix}(\epsilon/2)^2}{\epsilon} + \theta,$$

where $\theta = t_{mix}(\epsilon/2) + \epsilon/(4\lambda)$.

Proof. Observe that $\lambda \geq \max_{i \in \Omega} \sum_{j: j \neq i} q_t(i, j)$, where $q_t(i, j)$ are the rates in $Q[\frac{t}{T}]$ ($0 \leq t \leq T$).

Take

$$T = K \left(1 - \frac{1}{2K}\right)^{-1} t_{mix}(\epsilon/2), \quad N = (K-1) \left(1 - \frac{1}{2K}\right)^{-1} t_{mix}(\epsilon/2),$$

and let $P_{final}(t) = e^{tQ_{final}}$ denote the transition probability matrix associated with the generator Q_{final} .

Now, we let $P_0 = I + \frac{1}{\lambda}Q_{initial}$ and $P_1 = I + \frac{1}{\lambda}Q_{final}$. Then P_0 and P_1 are discrete Markov chains and

$$\nu P_T(0, T) = \nu_N P_T(N, T) = \nu_N \left(\sum_{n=0}^{\infty} \frac{(\lambda(T-N))^n}{n!} e^{-\lambda(T-N)} I_n \right),$$

where $\nu_N = \nu P_T(0, N)$ and

$$I_n = \frac{n!}{(T-N)^n} \int \dots \int_{N < x_1 < x_2 < \dots < x_n < T} \left[\left(1 - \frac{x_1}{T}\right) P_0 + \frac{x_1}{T} P_1 \right] \dots \left[\left(1 - \frac{x_n}{T}\right) P_0 + \frac{x_n}{T} P_1 \right] dx_1 \dots dx_n$$

i.e. the order statistics of n arrivals within the $[N, T]$ time interval. We used the fact that, when conditioned on the number of arrivals, the arrival times of a Poisson process are distributed as an order statistics of uniform random variables. Hence

$$\begin{aligned} \nu P_T(0, T) &= \nu_N \left(\sum_{n=0}^{\infty} (\lambda(T-N))^n \cdot \frac{e^{-\lambda(T-N)}}{(T-N)^n T^n n!} P_1^n \int_N^T \dots \int_N^T x_1 \dots x_n dx_1 \dots dx_n \right) + \mathcal{E} \\ &= e^{-\lambda(1-\frac{1}{2K})^{-1} t_{mix}(\epsilon/2)} \nu_N \left(\sum_{n=0}^{\infty} \frac{\lambda^n (t_{mix}(\epsilon/2))^n}{n!} P_1^n \right) + \mathcal{E} \\ &= e^{-\frac{\lambda t_{mix}(\epsilon/2)}{2K-1}} \nu_N P_{final}(t_{mix}(\epsilon/2)) + \mathcal{E}, \end{aligned}$$

where \mathcal{E} is the rest of the terms. Thus, the total variation distance,

$$\max_{\nu} \|\nu P_T(0, T) - \pi_f\|_{TV} \leq e^{-\frac{\lambda t_{mix}(\epsilon/2)}{2K-1}} \epsilon/2 + S_N$$

Taking $K \geq \frac{\lambda t_{mix}(\epsilon/2)}{\epsilon} + 1/2$, we bound the error term

$$S_N = \|\mathcal{E} - \pi_f\|_{TV} \leq 1 - e^{-\lambda(1-\frac{1}{2K})^{-1} t_{mix}(\epsilon/2)} \sum_{n=0}^{\infty} \frac{\lambda^n (t_{mix}(\epsilon/2))^n}{n!} = 1 - e^{-\frac{\lambda t_{mix}(\epsilon/2)}{2K-1}} \leq \epsilon/2$$

as $\epsilon < -2 \log(1 - \frac{\epsilon}{2})$. Therefore $T_\epsilon \leq \left(\frac{\lambda t_{mix}(\epsilon/2)}{\epsilon} + 1/2\right) (t_{mix}(\epsilon/2) + \epsilon/(2\lambda))$. \square

In the end, we would like to mention a possible application. The Glauber dynamics of a finite Ising model is a continuous time Markov process. Its mixing time can be estimated using path coupling. The adiabatic transformation of the Markov generator corresponds to an adiabatic transformation of the Hamiltonian. The above result can be applied to obtain the adiabatic time for the transformation. The result can be adjusted when the adiabatic evolution of the generator is nonlinear.

3.1 Application: adiabatic Glauber dynamics of a one dimensional Ising model

In the original, non-adiabatic case, the Glauber dynamics is used to generate distribution

$$\pi(x) = \frac{1}{Z(\beta)} e^{-\mathcal{H}(x)}$$

over all spin configurations $x \in \{-1, +1\}^S$, where S denotes all the sites of a graph, $\mathcal{H}(x) = -\beta \sum_{u \sim v} x(u)x(v)$, β is the reciprocal of the temperature, and $Z(\beta)$ is the normalization constant.

Consider a non-linear adiabatic Glauber dynamics of an Ising model on $S = \mathbb{Z}/n\mathbb{Z}$. There, for time $s \in [0, 1]$, the Hamiltonian

$$\mathcal{H}_s = (1 - \psi(s))\mathcal{H}_{initial} + \varphi(s)\mathcal{H}_{final},$$

where $\psi(s)$ and $\varphi(s)$ are continuous functions on $[0, 1]$ such that $\psi(0) = \varphi(0) = 0$ and $\psi(1) = \varphi(1) = 1$. Each site $j \in \mathbb{Z}/n\mathbb{Z}$ has an independent exponential clock with parameter one associated with it. When it rings, the spin $x(j)$ is reselected using the following probability

$$P(x(j) = +1) = \frac{e^{-\mathcal{H}_s(x_+)}}{e^{-\mathcal{H}_s(x_-)} + e^{-\mathcal{H}_s(x_+)}}$$

where $x_+(i) = x_-(i) = x(i)$ for $i \neq j$, $x_+(j) = +1$ and $x_-(j) = -1$. Here \mathcal{H}_s is the Hamiltonian at the transition time s . The time non-homogeneous generator associated with the above dynamics is denoted by $Q[s]$.

We will compare the above adiabatic Glauber dynamics with the ordinary Glauber dynamics used for generating Ising models, where at each site $j \in S$, its spin is reset after waiting for an exponential interarrival time with parameter one, and the probabilities are determined by the local Hamiltonians. Let $Q_{initial}$ and Q_{final} be the time homogenous generators of ordinary Glauber dynamics associated with the respective Hamiltonians, $\mathcal{H}_{initial}$ and \mathcal{H}_{final} .

Let us consider the case when $\mathcal{H}_{initial}(x) = 0$, $\mathcal{H}_{final}(x) = -\beta \sum_{j \in \mathbb{Z}/n\mathbb{Z}} x(j)x(j+1)$ and

$$\varphi(s) = \frac{1}{4\beta} \log \frac{p(s)}{1-p(s)},$$

where $p(s) = (1-s)\frac{1}{2} + s\frac{e^{2\beta}}{e^{-2\beta} + e^{2\beta}}$. Then $\frac{e^{2\beta\varphi(s)}}{e^{-2\beta\varphi(s)} + e^{2\beta\varphi(s)}} = p(s)$, and therefore

$$Q[s] = (1-s)Q_{initial} + sQ_{final}$$

So the question of how slow the nonlinear adiabatic transition

$$\mathcal{H}_{t/T} = (1 - \psi(t/T))\mathcal{H}_{initial} + \varphi(t/T)\mathcal{H}_{final}$$

has to happen in order to obtain the distribution that is ϵ close to that of Ising model with Hamiltonian \mathcal{H}_{final} at the end of the adiabatic evolution is addressed in Theorem 3. In particular, for $\beta < \frac{1}{2}$, the adiabatic time

$$T_\epsilon \leq O\left(\frac{n}{\epsilon} \cdot \left[\frac{\log n + \log(2/\epsilon)}{1 - \tanh(2\beta)}\right]^2\right)$$

as $t_{mix}(\epsilon/2) \leq \frac{\log n + \log(2/\epsilon)}{1 - \tanh(2\beta)}$ by Theorem 15.1 in Levin et al (2008).

The above example connects an adiabatic evolution of a physical system to that of Markov operators, and invites to further study the latter.

Acknowledgment

The author would like to thank the anonymous referee for pointing at the errors in an earlier draft, and suggesting an improvement.

References

Aldous, D., 1983. Random Walks on Finite Groups and Rapidly Mixing Markov Chains. *Seminaire de Probabilites XVII, Lecture Notes in Math.*, **986**, 243–297.

Aldous, D., Fill, J.A. Reversible Markov Chains and Random Walks on Graphs. On the web: <http://www.stat.berkeley.edu/users/aldous/>

Burton, R.M., Kovchegov, Y., 2009. Mixing times via super-fast coupling. In press.

Levin, D.A., Peres, Y., Wilmer, E.L., 2008. Markov Chains and Mixing Times. Amer. Math. Soc., Providence, RI.

Messiah, A., 1958. Quantum mechanics. John Wiley and Sons, NY