Journal of The Faculty of Environmental Science and Technology, Okayama University
Vol.2, No.1, pp.1-8, January 1997

A Critical Phenomenon Appearing in the Process of Particle Diffusion in Classical Statistical Mechanics

Akihito HORA *

(Received October 29, 1996)

Abstract

An aspect of the *cut-off phenomenon* is reviewed. It is a sort of critical phenomenon observed in a wide variety of diffusion. We introduce a framework in which such a phenomenon can be understood and analyzed in a rigorous manner. We illustrate the mechanism with some concrete models.

Keywords: cut-off phenomenon, critical time, diffusion, Markov chain on a graph, symmetry

1991 Mathematics Subject Classification: 60J, 60B, 82C

1 Introduction

1.1 Particle diffusion

Diffusion is a quite common phenomenon which we see in our daily life. One can imagine smoke in the air, dripped milk in a cup of coffee, spreading rumor in a community, repeated shuffle of a deck of cards, etc. In contrast with its familiar appearance, the mechanism is complicated. What are typical features of these phenomena? Essential is the fact that a large — actually, huge — number of elements, or particles, are involved. Every element behaves in a certain manner or rahter at random, and anyway obeys a *microscopic* law. Such small movements getting together, the interrelated totality shows a *macroscopic* effect, in our context, diffusion. Thus we adopt an approach of (classical) statistical mechanics in this article. It means that we intend to combine the behavior of each particle with the macroscopically observed diffusion through the structure of the system considerd instead of analyzing diffusion phenomenalogically by using, say, differential equations in macroscopic variables.

One knows at least empirically that diffusion of particles will reach a stationary state, namely, a well mixed or almost uniformly spread situation after a long time period. We are concerned with careful persuing this process of diffusion and then observing a remarkable critical phenomenon. Before explaining the critical phenomenon in the following subsection, let us mention some concrete models which are analyzed later.

The Bernoulli-Laplace diffusion model Let us imagine a box divided into two parts by an imaginary partition, the left part containing d red balls and the right one containing $v-d$ white balls. Every unit time, one ball is picked up at random from each part of the box and the two balls are interchanged. Then, as time goes on, two kinds of balls are mixed up. This simple model, introduced by D. Bernoulli and Laplace, imitates diffusion of two initially separated incompressible liquids or sparse gases.

The Ehrenfests' urn model Let us imagine two urns and d balls contained in them. At each stage, one specifies a ball among d at random and moves it into the other urn. This is also a simple model describing

*Department of Environmental and Mathematical Sciences, Faculty of Environmental Science and Technology, Okayama University, Okayama, 700 Japan.

certain diffusion of gases. Equivalently, one can formulate this model as a random change of binary sequences consisting of, say, + and −.

Card shuffling Let us imagine a deck of 52 cards. A familiar way of shuffling the cards is so-called riffle shuffle: one divides a deck into two and riffles them from both hands sides. Another shuffling rule causes a different model. Card shuffling is an attractive and effective material which vividly shows us the essence of the phenomenon dealt with in this article.

1.2 Critical phenomenon

In general, we say a critical phenomenon occurs if the state of a system changes drastically at a specific value of a parameter contained in the system. The specific value is called a critical value. Let us recall, for example, the process in which water is heated and turn into vapor. One knows the remarkable shape of the time-temperature curve associated with this process. The specific temperature, $100°C$, separates the two states, where such macroscopic variables as volumes or pressures take quite different values in the two states.

In particle diffusion, as is noted in the previous subsection, we see that an initial orderly state will lose its order as time goes on and eventually settle down to a disorderly one. Is this process of diffusion flat? Or does it have a special feature like boiling water? One cannot stir her/his coffee million times. To start a game, one cannot help stopping shuffling the cards at some moment. Therefore, we further want to know how many times shuffle is "necessary and sufficient" to reach a well mixed state, even if it is assured after a long time period.

Recent quantitative analysis of Markov chains enables us to attack such a problem. It can be shown that we observe a critical phenomenon described as above in a wide variety of concrete models of diffusion. In other words, we can separate an orderly state and a disorderly one in the system at a specific time, called critical time, in an appropriate sense. Diaconis is a main developer in this field. In collaboration with Aldous, Graham, Shahshahani and others, he has studied many interesting models including those mentioned in the previous subsection and constructed a substantial theory. The term *cut-off phenomenon* has been used to denote this critical phenomenon since [1]. We refer to [3] and [4] as comprehensive treatises on the cut-off phenomenon with bibliographical information.

In Sect. 3, we introduce a precise definition of the cut-off phenomenon and then study on individual models. Sect. 2 is devoted to preliminary discussions on mathematical modeling of our object.

2 Mathematical Modeling

In this section, we prepare some necessary notions to describe and analyze the cut-off phenomenon in Sect. 3. We consider *a Markov chain on a graph* as mathematical modeling of particle diffusion.

2.1 Markov chains

Markov chains are often used as a mathematical device to describe random affairs. Let X be a finite set, which becomes the state space of the chain. We consider a random motion on X by assigning a real value $P_{x,y}$ for $x, y \in X$, which is regarded as the probability in which the chain makes a transition from current state x to new state y in one unit time (say, a second). Interpretation of $P_{x,y}$ as a probability requires that $P_{x,y} \geq 0$ and $\sum_{y\in X} P_{x,y} = 1$. Square matrix $P = [P_{x,y}]_{x,y\in X}$ of degree $|X|$ is called the transition matrix of the Markov chain. Here and after, $|X|$ denotes the cardinality of finite set X. For each $x \in X$, sequence $\{P_{x,y}|y \in X\}$, or probability measure $P_{x,\cdot}$ on X, gives the distribution of the chain starting at x after one unit time. The transition probability from x to y in n unit times is expressed by products of the transition matrix:

$$(P^n)_{x,y} = \sum_{x_1,\cdots,x_{n-1}\in X} P_{x,x_1}\cdots P_{x_i,x_{i+1}}\cdots P_{x_{n-1},y} \ .$$

Then, the distribution after n unit times is given by $\{(P^n)_{x,y}|y \in X\}$ when the chain started at x. Probability measure π on X (i.e. $\pi(x) \geq 0$, $\sum_{x\in X}\pi(x) = 1$) is called an invariant measure of the Markov chain if it satisfies

$$\sum_{x\in X} \pi(x)P_{x,y} = \pi(y) \qquad (y \in X) \ .$$

This implies that distribution π is stationary under time evolution, therefore describes stationary (or equilibrium) state of the chain. Under some mild assumptions, it can be proved that the chain tends to its stationary state, in other words, gets well mixed up, as time goes on. The most typical invariant measure is the uniform one: $\pi(x) = 1/|X|$ for every $x \in X$.

2.2 Graphs

A graph consists of vertices and edges. For the sake of simplicity, we here deal only with finite undirected simple graphs. Let X be a finite set and E a subset of $X \times X$. E is assumed to be symmetric, i.e. $(x,y) \in E \iff (y,x) \in E$ and not to contain such a diagonal pair as (x,x). A *graph* is, by definition, pair $\Gamma = (X,E)$. X and E are called the vertices and the edges of Γ respectively. Two vertices x and y are said to be adjacent if $(x,y) \in E$. To visualize a graph, it is convenient to assign a point to each vertex and to join two points if the two vertices are adjacent. The number of adjacent vertices of vertex x is called the degree of x and denoted by d_x : $d_x = |\{y \in X|(x,y) \in E\}|$. If every vertex x has a common degree $\kappa = d_x$, the graph is said to be κ-regular. Graphs can be further classified according to certain symmetry which they enjoy.

2.3 Examples of concrete models

In Subsect. 1.1, we mentioned three models as examples of particle diffusion. Now we formulate them precisely as a Markov chain on a graph by using the notions in the preceding subsections.

Example A The Bernoulli-Laplace diffusion model Every state in this model is specified by d members in the left part. Hence, letting S be a finite set of cardinality v, we set state space $X = \{x \subset S||x| = d\}$. Further set $E = \{(x,y) \in X \times X||x \cap y| = d-1\}$. $(x,y) \in E$ means that x differs from y by exactly one member. The degree of each vertex x is $d(v-d)$. Thus $\Gamma = (X,E)$ becomes a $d(v-d)$-regular graph. Γ is called a Johnson graph and denoted by $J(v,d)$. Every unit time, one member of each part is interchanged, hence the chain should move from vertex x to one of the adjacent vertices in equal probability. The transition matrix of this Markov chain is given by

$$P_{x,y} = \begin{cases} 1/d(v-d) & \text{if} \quad (x,y) \in E \\ 0 & \text{if} \quad (x,y) \notin E \ . \end{cases}$$

Example B The Ehrenfests' urn model and its extension A state is specified by determining the urn, say $+$ or $-$, in which each ball is contained. Set $F = \{+,-\}$, $X = F^d$ (totality of d-sequences of $+$ and $-$), and $E = \{(x,y) \in X \times X|\,|\{i|x_i \neq y_i\}| = 1$ where $x = (x_i)_{i=1}^d$, $y = (y_i)_{i=1}^d\}$. Then, $\Gamma = (X,E)$ is a d-regular graph. Every unit time, exactly one coordinate of vertex $x = (x_i)_{i=1}^d$ is flipped, which gives rise to a Markov chain on X with transition matrix

$$P_{x,y} = \begin{cases} 1/d & \text{if} \quad (x,y) \in E \\ 0 & \text{if} \quad (x,y) \notin E \ . \end{cases}$$

Unfortunately, this chain does not tend to stationary state because of the parity problem. It is then natural to modify the above transition matrix into

$$P_{x,y} = \begin{cases} 1/d+1 & \text{if} \quad x = y \\ 1/d+1 & \text{if} \quad (x,y) \in E \\ 0 & \text{otherwise} . \end{cases}$$

To give an extension of the Ehrenfests' urn model, we replace 2-set $\{+,-\}$ by an n-set. Then, $\Gamma = (X,E)$ becomes an $(n-1)d$-regular graph. Γ is called a Hamming graph and denoted by $H(d,n)$. The corresponding Markov chain is given by transition matrix

$$P_{x,y} = \begin{cases} 1/(n-1)d & \text{if} \quad (x,y) \in E \\ 0 & \text{if} \quad (x,y) \notin E . \end{cases}$$

Example C Card shuffling Let us imagine n cards. One shuffle corresponds to a permutation of $\{1,2,\cdots,n\}$. Totality of the permutations of n letters is the symmetric group of degree n and denoted by $\mathcal{S}_n$. Thus our Markov chain has $\mathcal{S}_n$ as its state space X. A shuffling rule determines a graph structure in which $X = \mathcal{S}_n$ is the vertices. Let us begin with a simple shuffling rule consisting of transpositions. At each stage, one chooses two cards (which possibly coincide) among n at random and transposes them. The transition matrix of the Markov chain describing this shuffle is

$$P_{x,y} = \begin{cases} 1/n & \text{if} \quad x = y \\ 2/n^2 & \text{if} \quad \sigma x = y \text{ for some transposition } \sigma \\ 0 & \text{otherwise} . \end{cases}$$

Let Ω be the set of transpositions and set $E = \{(x,y) \in X \times X | \sigma x = y \text{ for some } \sigma \in \Omega\}$. Note that Ω generates the whole $\mathcal{S}_n$. Then, $\Gamma = (X,E)$ is an $\binom{n}{2}$-regular graph. Γ is a Cayley graph of $\mathcal{S}_n$ generated by Ω.

Next we consider the riffle shuffle. We follow the Gilbert-Shannon-Reeds model. See [2], [12], and [4]. At each stage, one divides a deck of n cards into two piles according to the binomial distribution (so that k and $n-k$ piles are realized in probability $\binom{n}{k}/2^n$) and then riffles the two piles together, where, if the left hand pile has p cards and the right hand pile has q cards at some moment, then the next card is dropped either from the left pile in probability $p/(p+q)$ or from the right pile in probability $q/(p+q)$. Starting at the unit element e of $\mathcal{S}_n$, the chain can move to vertex x ($\in X = \mathcal{S}_n$) which contains at most two rising sequences. Let R be the set of n-permutations with two rising sequences. R generates $\mathcal{S}_n$ since R contains every transposition of two consecutive numbers. Set $E = \{(x,y) \in X \times X | \sigma x = y \text{ for some } \sigma \in R\}$. Since E is not symmetric, $\Gamma = (X,E)$ is a directed graph. The transition matrix of the Gilbert-Shannon-Reeds model is

$$P_{x,y} = \begin{cases} (n+1)/2^n & \text{if} \quad x = y \\ 1/2^n & \text{if} \quad \sigma x = y \text{ for some } \sigma \in R \\ 0 & \text{otherwise} . \end{cases}$$

2.4 Convergence to a stationary state

In every model mentioned in the preceding subsection, the stationary state of the Markov chain is unique and uniformly distributed: $\pi(x) = 1/|X|$. Hence,

$$(P^n)_{x,y} \longrightarrow \frac{1}{|X|} \qquad \text{as} \quad n \to \infty \qquad \text{for} \quad x,y \in X . \tag{1}$$

We are not content with (1) but concerned with the manner of convergence in (1). For this purpose, let us define

$$\|(P^n)_{x,\cdot} - \pi\| = \frac{1}{2} \sum_{y \in X} \left| (P^n)_{x,y} - \frac{1}{|X|} \right| \tag{2}$$

as a quantity to measure closeness to the stationary state of a chain starting at x after n unit times. If the chain considered enjoys certain symmetry (which is assumed in our models discussed), (2) does not depend on starting vertex x and hence coincides with

$$D(n) = \frac{1}{2|X|} \sum_{x\in X} \sum_{y\in X} \left| (P^n)_{x,y} - \frac{1}{|X|} \right| . \tag{3}$$

Our task is to evaluate (3) carefully. We will obtain asymptotic properties of (3) depending on both time n and the size of a graph.

3 The Cut-Off Phenomenon

3.1 Formulation

We formulate the cut-off phenomenon at two different levels. The point is that we do not fix a single Markov chain but consider an infinite family of Markov chains. The thermodynamical limit in statistical mechanics motivates our definition. Let Λ be a directed set and $\{\Gamma_\lambda = (X_\lambda, E_\lambda) | \lambda \in \Lambda\}$ a family of graphs. Γ_λ is assumed to get larger (i.e. $|X_\lambda| \to \infty$) as $\lambda \to \infty$. For each $\lambda \in \Lambda$, we consider a Markov chain with transition matrix P_λ and quantity $D_\lambda(n)$ defined in (3).

Definition 1 ([9], [10]) Assume that we can take $n_\lambda \in \mathbf{N}$ for each $\lambda \in \Lambda$ satisfying the following conditions:
(i) $n_\lambda \longrightarrow \infty \quad$ as $\quad \lambda \to \infty$
(ii) $\forall\epsilon > 0$, $\exists\lambda_\epsilon \in \Lambda$ and $\exists h_{\epsilon,\lambda} > 0$ such that, $h_{\epsilon,\lambda}/n_\lambda \longrightarrow 0$ as $\lambda \to \infty$ and, if $\lambda > \lambda_\epsilon$,

$$\begin{aligned} 0 \le n \le n_\lambda - h_{\epsilon,\lambda} \quad &\Longrightarrow \quad D_\lambda(n) \ge 1 - \epsilon \\ n \ge n_\lambda + h_{\epsilon,\lambda} \quad &\Longrightarrow \quad D_\lambda(n) \le \epsilon . \end{aligned}$$

Then we say that the *cut-off phenomenon at level 1* occurs for (the family of) the Markov chains and call n_λ the *critical time*.

Definition 1 shows a remarkable shape of the n–$D(n)$ curve. As $\lambda \to \infty$ (namely, as the size of the system grows), it looks "cut-off" at critical time n_λ. Let us regard n_λ as a time of macroscopic order. Then, $|X_\lambda|$ is huge (almost infinite), while $h_{\epsilon,\lambda}$ is so small as to be negligible. Before $n_\lambda - h_{\epsilon,\lambda}$, $D_\lambda(n)$ stays almost 1, which implies that the system is still orderly. After $n_\lambda + h_{\epsilon,\lambda}$, $D_\lambda(n)$ becomes almost 0, which implies in turn that diffusion has been almost completed and the system is disorderly. Definition 1 thus shows that transition from orderly state to disorderly one is suddenly made near the critical time.

Information on the behavior near a critical value is crucial in studying a physical system. Deepening Definition 1, we introduce the following one.

Definition 2 ([4], see also [11]) Assume that we can take $n_\lambda \in \mathbf{N}$ and $h_\lambda > 0$ for each $\lambda \in \Lambda$ satisfying the following conditions:
(i) $n_\lambda \longrightarrow \infty$, $\quad h_\lambda/n_\lambda \longrightarrow 0 \quad$ as $\quad \lambda \to \infty$
(ii) $D_\lambda(\lfloor n_\lambda + \theta h_\lambda \rfloor) \longrightarrow c(\theta) \quad$ as $\quad \lambda \to \infty \quad$ for $\quad \forall\theta \in \mathbf{R}$
where function $c(\theta)$ satisfies $0 \le c(\theta) \le 1$, $c(-\infty) = 1$ and $c(\infty) = 0$.
Then we say that the *cut-off phenomenon at level 2* occurs for the Markov chains.

Compared with in Definition 1, time scale near the critical time is extended by n_λ/h_λ times in Definition 2. Explicit calculation of function $c(\theta)$ in Definition 2 is an important and challenging problem in concrete models.

3.2 Results

Theorem 1 In every model mentioned in Subsect. 2.3, the cut-off phenomenon at level 1 occurs. (i) The critical time n_λ and (ii) the constraint for λ are as follows.

(A) The Bernoulli-Laplace diffusion model ([7], [3], [8], [9]). Set $\lambda = (v, d)$.

$$\text{(i)} \quad n_\lambda = \frac{d}{2}\Big(1 - \frac{d}{v}\Big)\log v$$

$$\text{(ii)} \quad d \to \infty\ , \quad 2d \le v\ , \quad \limsup_{d\to\infty} \frac{\log v}{2\log d} \le 1\ .$$

(B1) The Ehrenfests' urn model: $H(d,2)$ case ([3], [5], [13], [4]). Set $\lambda = d$.

$$\text{(i)}\ n_\lambda = \frac{d}{4}\log d \qquad\qquad \text{(ii)}\ d \to \infty\ .$$

(B2) $H(d,n)$ case ([9]). Set $\lambda = (d,n)$.

$$\text{(i)} \quad n_\lambda = \frac{d}{2}\Big(1 - \frac{1}{n}\Big)\log(n-1)d$$

$$\text{(ii)} \quad d \to \infty\ , \quad n \ge 3\ , \quad \limsup_{(d,n)\to\infty} \frac{\log n}{\log d} \le 1\ .$$

(C1) Card shuffling by random transpositions ([6], [3]). Set $\lambda = n$.

$$\text{(i)}\ n_\lambda = \frac{n}{2}\log n \qquad\qquad \text{(ii)}\ n \to \infty\ .$$

(C2) Riffle shuffle ([2], [12], [4]). Set $\lambda = n$.

$$\text{(i)}\ n_\lambda = \frac{3}{2}\log_2 n \qquad\qquad \text{(ii)}\ n \to \infty\ .$$

For other models in which the cut-off phenomenon at level 1 is observed, see some of the references in [4].

Theorem 2 The cut-off phenomenon at level 2 is proved in the following models with (iii) h_λ and (iv) $c(\theta)$ below.

(B1) The Ehrenfests' urn model: $H(d,2)$ case ([5], [13], [4]). Set $\lambda = d$.

$$\text{(iii)}\ h_\lambda = \frac{d}{4} \qquad\qquad \text{(iv)}\ c(\theta) = Erf\Big(\frac{e^{-\theta/2}}{2\sqrt{2}}\Big)$$

where

$$Erf(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt$$

is the error function.

(B2) $H(d,n)$ case ([11]). Set $\lambda = (d,n)$. Let $n \ge 3$.

$$\text{(iii)} \quad h_\lambda = \frac{d}{2}\Big(1 - \frac{1}{n}\Big)$$

$$\text{(iv.1)} \quad c(\theta) = Erf\Big(\frac{e^{-\theta/2}}{2\sqrt{2}}\Big) \quad \text{if} \quad d \to \infty \text{ and } n/d \to 0$$

$$\text{(iv.2)} \quad c(\theta) = F_{1/\tau}(\theta^*) - F_{(1/\tau)+(e^{-\theta/2}/\sqrt{\tau})}(\theta^*) \quad \text{if} \quad d \to \infty \text{ and } n/d \to \tau > 0$$

where $1/\theta^* = e^{\theta/2}\sqrt{\tau}\log(1 + e^{-\theta/2}\sqrt{\tau})$ and F_α denotes the Poisson distribution function with intensity α.

(C2) Riffle shuffle ([2], [4]). Set $\lambda = n$.

$$\text{(iii)}\ h_\lambda = 1 \qquad\qquad \text{(iv)}\ c(\theta) = Erf\Big(\frac{2^{-\theta}}{4\sqrt{6}}\Big)\ .$$

3.3 Methods

Harmonic analysis provides us a successful method to evaluate (3). Symmetry of a graph, or state space X, is inherited by transition matrix P of a Markov chain on X. This enables us to calculate P^n explicitly by using 'nice functions' induced by the symmetry. Examples include permutaion groups, matrix groups over a finite field and Q-polynomial distance-regular graphs. In [9], we proposed some general criteria for the cut-off phenomenon at level 1 in Q-polynomial distance-regular graphs. Details of the proofs of Theorem 1 and 2 are left to individual articles cited in the preceding subsection.

Symmetry of a state space not only helps our calculation but also is an essential cause of the cut-off phenomenon. When (3) is estimated in terms of characters or spherial functions by the method of harmonic analysis, the dominant term of the expression is controlled by the second eigenvalue and its multiplicity of the transition matrix. Generally speaking, a large symmetry induces high multiplicity (or degeneration) of th eigenvalues and hence tends to cause the cut-off phenomenon. See [4] for more details.

4 Discussion

The cut-off phenomenon at level 1 is now proved in a lot of models. However, we do not have many models yet in which the cut-off phenomenon at level 2 is established. It is hence worth-while to try to work out fine evaluation of (3) in various concrete models. On that occasion, numerical experiments will serve us in finding out a possible value of the critical time.

Symmetry of a state space surely plays an essential role in the cut-off phenomenon as is stated in Subsect. 3.3. However, character of the problem suggests that *approximate symmetry* may suffice in a sense. As is mentioned in [4], it is an important trial to break symmetry and to perturbate the transition matrix. In the context of particle diffusion, it may imply putting a bit of impurities. Such a discussion will lead to structural stability (or instability) of the cut-off phenomenon.

References

[1] D.Aldous and P.Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly **93** (1986), 333–348.

[2] D.Bayer and P.Diaconis, *Trailing the dovetail shuffle to its lair*, Ann. Appl. Probab. **2** (1992), 294–313.

[3] P.Diaconis, Group Representations in Probability and Statistics, IMS, Hayward, California, 1988.

[4] P.Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Nat. Acad. Sci. USA **93** (1996), 1659–1664.

[5] P.Diaconis, R.L.Graham and J.A.Morrison, *Asymptotic analysis of a random walk on a hypercube with many dimensions*, Random Struct. Algor. **1** (1990), 51–72.

[6] P.Diaconis and M.Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahr. verw. Geb. **57** (1981), 159–179.

[7] P.Diaconis and M.Shahshahani, *Time to reach stationarity in the Bernoulli-Laplace diffusion model*, SIAM J. Math. Anal. **18** (1987), 208–218.

[8] P.Donnelly, P.Lloyd and A.Sudbury, *Approach to stationarity of the Bernoulli-Laplace diffusion model*, Adv. Appl. Probab. **26** (1994), 715–727.

[9] A.Hora, *Critical phenomena for random walks on P- and Q-polynomial association schemes*, Preprint (1995).

[10] A.Hora, *Towards critical phenomena for random walks on various algebraic structures*, In H.Heyer and T.Hirai (eds.), Transactions of a German-Japanese Symposium "Infinite Dimensional Harmonic Analysis" in Tübingen, D.+M. Gräbner, Bamberg, 1996, 113–127.

[11] A.Hora, *Evaluation of the cut-off phenomenon for a random walk on a Hamming graph with varying large size*, Preprint (1996).

[12] B.Mann, *How many times should you shuffle a deck of cards?*, In J.L.Snell (ed.), Topics in Contemporary Probability and Its Applications, CRC Press, Boca Raton, Florida, 1995, 261–289.

[13] M.Voit, *Asymptotic distributions for the Ehrenfest urn and related random walks*, J. Appl. Probab. **33** (1996), to appear.