C. R. Math. Rep. Acad. Sci. Canada Vol. **32** (1) 2010, pp. 13–18

THE AVALANCHE PRINCIPLE AND SOME DEVIATION PROBABILITIES

D. GOLDSTEIN AND I. GOLDSTEIN

Presented by George Elliott, FRSC

Abstract. We prove certain probabilistic inequalities for long matrix products generated by a pair of 2×2 matrices. Our main tool is the so-called Goldstein-Schlag avalanche principle.

Résumé. Nous prouvons certaines inégalités probabilistes concernant les longs produits de matrices générés par une paire de matrices 2×2. Notre objectif principal concerne le principe d'avalanche, énoncé par Goldstein et Schlag.

1. **Introduction.** In this article we continue the study of a long matrix product generated by a pair of 2×2 matrices based on the *avalanche principle.* This principle has been established by M. Goldstein and W. Schlag [G-S]. In our previous work (see [G-G]), we have described a quick algorithm which computes an *averaged joint spectral radius* for a pair of 2×2 matrices satisfying an avalanche condition, (see Theorem 1.1). The next step of our research relates to estimations of some classical deviation probabilities associated with the long matrix products.

Theorem 1.1 (Avalanche Principle [G-S, Proposition 2.2]). *Let $A_1, \dots, A_n$ be a sequence of unimodular 2×2 matrices. Suppose that*

(i) $\min_{1\le j\le n} \|A_j\| \ge \mu > n$,

(ii) $\max_{1\le j\le n-1}[\log \|A_{j+1}\| + \log \|A_j\| - \log \|A_{j+1}A_j\|] < \frac{1}{2}\log \mu$.

Then

$$\Big| \log \|A_n \cdots A_1\| + \sum_{j=2}^{n-1} \log \|A_j\| - \sum_{j=1}^{n-1} \log \|A_{j+1}A_j\| \Big| < c\frac{n}{\mu}.$$

2. **Deviation probabilities.** Now we begin to study some probabilistic aspects of the avalanche principle. We define the four following quantities:

$$l_1 = \log \|A^2\| - \log \|A\|, \qquad l_2 = \log \|AB\| - \log \|B\|,$$

$$l_3 = \log \|B^2\| - \log \|B\|, \qquad l_4 = \log \|BA\| - \log \|A\|,$$

Received by the editors on September 9, 2009.
AMS Subject Classification: Primary: 60F10; secondary: 37H15.
Keywords: avalanche principle, deviation probabilities.

and their average by

$$\gamma = \frac{l_1 + l_2 + l_3 + l_4}{4}.$$

For any word $\Pi_n = C_n \cdots C_1$, $C_i \in \{A, B\}$ it follows by the avalanche principle that $n^{-1} \log \|\Pi_n\|$ can be approximated by $n^{-1}(n_1 l_1 + n_2 l_2 + n_3 l_3 + n_4 l_4)$, where n_1, n_2, n_3, n_4 denote the appearance frequencies of the combinations

$$A^2, AB, B^2, BA$$

in the word $\Pi_n = C_n \cdots C_1$. Indeed, the avalanche principle implies that

$$\Big|\log \|C_n \dots C_1\| + \sum_{j=2}^{n-1} \log \|C_j\| - \sum_{j=1}^{n-1} \log \|C_{j+1} C_j\|\Big| < c\frac{n}{\mu_0}.$$

On the other hand, one can write

$$n_1 l_1 + n_2 l_2 + n_3 l_3 + n_4 l_4 + \chi(C_n) = \sum_{j=1}^{n-1} \log \|C_{j+1} C_j\| - \sum_{j=2}^{n-1} \log \|C_j\|,$$

where

$$\chi(C_n) = \begin{cases} \log \|A\| & \text{if } C_n = A, \\ \log \|B\| & \text{if } C_n = B. \end{cases}$$

Therefore,

$$\big|n^{-1} \log \|\Pi_n\| - n^{-1}(n_1 l_1 + n_2 l_2 + n_3 l_3 + n_4 l_4)\big| < \frac{c}{\mu_0} + \frac{\chi(C_n)}{n}.$$

Our next goal is to estimate the following classical deviation probabilities:

$$P\Big(\Big|n_i - \frac{n}{4}\Big| > M\sqrt{n}\Big), \quad i = 1, 4.$$

Let us observe that by symmetry arguments it will be sufficient to compute only two of these probabilities

$$P\Big(\Big|n_1 - \frac{n}{4}\Big| > M\sqrt{n}\Big) \quad \text{and} \quad P\Big(\Big|n_2 - \frac{n}{4}\Big| > M\sqrt{n}\Big),$$

say.

Lemma 2.1. *For $k \in N$ we have*

$$P(n_1 = k) = 2^{-n} \sum_{m \geq k+1}^{\lfloor \frac{n+k+1}{2} \rfloor} \binom{m-1}{k} \binom{n-m+1}{m-k}$$

Proof. Let A occur $m \geq k+1$ times in the word $\Pi_n = C_n \cdots C_1$. It is clear that we have to count the number of outcomes in which A^2 (as a combination) occurs in the word exactly $\Pi_n k$ times. All such events can be described by the following schemes: $B^{d_0}AB^{d_1}\cdots AB^{d_{m-1}}AB^{d_m}$, where k members of $d_1, \dots, d_{m-1}$ are zeros (which means the appearance of k couples $AA = A^2$), and the rest of them, including d_0 and d_m, are positive integer numbers. Obviously $d_0, d_1, \dots, d_{m-1}, d_m$ satisfy the following Diophantine equation $d_0 + \cdots + d_m = n - m$. We need to find the number of integer solutions of this equation under the above-mentioned constraints. Choose k members from $d_1, \dots, d_{m-1}$ to be zeros. By using a standard change of variables, one can reduce the equation $d_0 + \cdots + d_m = n - m$ under the following (above-mentioned) constraints:

$$d_0, d_m \geq 0, \quad d'_1 = d'_2 = \cdots = d'_k = 0, \quad d'_{k+1}, \dots, d'_{m-1} \geq 1,$$

(here $d'_i \in \{d_1, \dots, d_{m-1}\}$ for all $i = 1, \dots, m-1$) to a new standard equation

$$x_1 + \cdots + x_{m-k+1} = (n-m) - (m-k-1) = n - 2m + k + 1,$$

where $x_1, \dots, x_{m-k+1} \geq 0$. (Indeed, put $x_1 = d_0$, $x_2 = d'_{k+1} - 1, \dots, x_{m-k} = d'_{m-1}$, $x_{m-k+1} = d_m$.) Summarizing: the number of integer non-negative solutions of the last equation is

$$\binom{m-1}{k}\binom{n-m+1}{m-k}.$$

Thus,

$$P(n_1 = k \mid A \text{ occurs } m \text{ times }) = 2^{-n}\binom{m-1}{k}\binom{n-m+1}{m-k}.$$

Finally, summing these probabilities over all $m = k+1, \dots, \lfloor \frac{n+k+1}{2} \rfloor$ we obtain

$$P(n_1 = k) = 2^{-n} \sum_{m \geq ; k+1}^{\lfloor \frac{n+k+1}{2} \rfloor} \binom{m-1}{k}\binom{n-m+1}{m-k},$$

as claimed. □

Lemma 2.2. *For $k \in N$ we have*

$$P(n_2 = k) = 2^{-n}\binom{n+1}{2k+1}.$$

PROOF. Arguing in just the same way as for Lemma 2.1, we count the number of outcomes in which AB occurs in the word Π_n exactly k times. The corresponding scheme is

$$B^{d_0}A^{q_0}B^{d_1}A^{q_1}\cdots B^{d_k}A^{q_k},$$

where all of $d_1,\dots,d_{k-1}$ and $q_0,\dots,q_{k-1}$ are strictly positive integer numbers (which implies the appearance of k couples AB), d_0 and q_k are non-negative integer numbers. Note that $d_0,\dots,d_k,q_0,\dots,q_{k-1}$ satisfy the following Diophantine equation:

$$\sum_{i=0}^{k} d_i + \sum_{j=0}^{k} q_j = n.$$

Redefine the variables $s_0 = d_0$, $s_i = d_i - 1$, $0 \le i \le k-1$, $s_{k+j+1} = q_j - 1$ for $0 \le j \le k$ and $s_{2k+1} = q_k$. Now the previous Diophantine equation can be reduced to $\sum_{i=0}^{2k+1} s_i = n - 2k$, where $s_i \ge 0$. Hence, the number of non-negative integer solutions of this equation is $\binom{n+1}{2k+1}$. This yields the statement of Lemma 2.2. □

Now we prove that the deviation probabilities computed above are exponentially small in M.

LEMMA 2.3.

$$P\Big(\Big|n_i - \frac{n}{4}\Big| > M\sqrt{n}\Big) < Ke^{-M^2}, \quad i = \overline{1,4},$$

where K is an absolute constant.

PROOF. The probabilities $(|n_i - \frac{n}{4}| > M\sqrt{n})$ are well known from the random allocation methods and may be estimated by using various classical inequalities. It seems that the simplest way to estimate $P(|n_1 - \frac{n}{4}| > M\sqrt{n})$ is the so-called *Chernoff Bounds* (see [C] or [A-S])

$$P(X \le \beta pn) \le e^{-(1-\beta^2)np/2},$$

where $X \sim B(n,p)$, $\beta < 1$. In particular, for $p = 0.5$ we have

$$P(X \le \beta n/2) \le e^{-(1-\beta^2)n/4}.$$

Thus,

$$\begin{aligned}
P\Big(\Big|n_2 - \frac{n}{4}\Big| > M\sqrt{n}\Big) &= P\Big(\Big|n_4 - \frac{n}{4}\Big| > M\sqrt{n}\Big) \\
&= \sum_{k<\frac{n}{4}-M\sqrt{n}} 2^{-n}\binom{n+1}{2k+1} + \sum_{k>\frac{n}{4}+M\sqrt{n}} 2^{-n}\binom{n+1}{2k+1} \\
&= \sum_{m<\frac{n+1}{4}-2M\sqrt{n}} 2^{-n}\binom{n+1}{m} + \sum_{m>\frac{n+1}{4}+2M\sqrt{n}} 2^{-n}\binom{n+1}{m}
\end{aligned}$$

(by the binomial symmetry)

$$= \sum m < \frac{n+1}{4} + 2M\sqrt{n} 2^{-n+1} \binom{n+1}{m}$$
$$= P\Big(X \leq \frac{n+1}{4} + 2M\sqrt{n}\Big) \leq Ke^{-M^2\sqrt{n+1}}.$$

To estimate the equal probabilities

$$P\Big(\Big|n_1 - \frac{n}{4}\Big| > M\sqrt{n}\Big) \quad \text{and} \quad P\Big(\Big|n_3 - \frac{n}{4}\Big| > M\sqrt{n}\Big)$$

we note that

$$P\Big(\Big|n_1 - \frac{n}{4}\Big| > M\sqrt{n}\Big) = 2^{-n} \sum_{k > \frac{n}{4} + M\sqrt{n}} \sum_{m \geq k+1}^{\lfloor \frac{n+k+1}{2} \rfloor} \binom{m-1}{k} \binom{n-m+1}{m-k}$$
$$+ 2^{-n} \sum_{k < \frac{n}{4} - M\sqrt{n}} \sum_{m \geq k+1}^{\lfloor \frac{n+k+1}{2} \rfloor} \binom{m-1}{k} \binom{n-m+1}{m-k}.$$

Straightforward calculations using Stirling's approximation lead to the claimed estimation $P(|n_1 - \frac{n}{4}| > M\sqrt{n}) < Ke^{-M^2}$. □

Now we can formulate our main result.

THEOREM 2.4. *Let $\Pi_n = C_n C_{n-1} \cdots C_1$ be a random word such that $C_i \in \{A, B\}$, $i = \overline{1, n}$. Then*

$$P\Big(\Big|\frac{1}{n} \log \|\Pi_n\| - \gamma\Big| > \frac{M}{\sqrt{n}}\Big) < Ke^{-(M/L)^2},$$

where K is an absolute constant and $L = \max_{1 \leq i \leq 4} l_i$.

PROOF. For a suitable absolute constant $\alpha > 0$ we have

$$P\Big(\Big|\frac{1}{n} \log \|\Pi_n\| - \gamma\Big| > \frac{M}{\sqrt{n}}\Big) \leq P\Big(\Big|\frac{n_1 l_1 + n_2 l_2 + n_3 l_3 + n_4 l_4}{n} - \gamma\Big| > \alpha \frac{M}{\sqrt{n}}\Big)$$
$$= P(|n_1 l_1 + n_2 l_2 + n_3 l_3 + n_4 l_4 - n\gamma| > \alpha M \sqrt{n})$$
$$= P\Big(\Big|\Big(n_1 - \frac{n}{4}\Big) l_1 + \Big(n_2 - \frac{n}{4}\Big) l_2 + \Big(n_3 - \frac{n}{4}\Big) l_3$$
$$+ \Big(n_4 - \frac{n}{4}\Big) l_4\Big| > \alpha M \sqrt{n}\Big)$$
$$\leq P\Big(4 \max_{i = \overline{1,4}} \Big|n_i - \frac{n}{4}\Big| \cdot L > c_1 \sqrt{n}\Big) < Ke^{-(\alpha M/L)^2}.$$

□

References

[**A-S**] N. Alon and J. Spencer, *The Probabilistic Method.* Second edition. Wiley-Interscience, New York, 2000.

[**C**] H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations.* Annals of Mathematical Statistics **23** (1952), 493–507.

[**G-G**] D. Goldstein and I. Goldstein, *The avalanche principle from joint to averaged joint spectral radius.* C. R. Math. Rep. Acad. Sci. Canada **28** (2006), no. 4, 97–104.

[**G-S**] M. Goldstein and W. Schlag, *Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions.* Ann. of Math. **154** (2001), no. 1, 155–203.

Holon Institute of Technology,
Holon 58102, Israel
email: dmitryg@hit.ac.il

Department of Mathematics,
Ben-Gurion University,
Beer-Sheva 84105, Israel
email: ilyago@bgu.ac.il