ON THE THEORY OF MARKOFF CHAINS

By Elliott W. Montroll

University of Pittsburgh

1. Summary. Although there exists voluminous literature on the theory of probability of independent events, and powerful techniques have been developed for the analysis of most of the interesting problems in this field, the theory of probability of dependent events has been rather neglected. The first detailed investigations in this subject were published by A. Markoff [1]. S. Bernstein [2] has extended the fundamental limit theorems to chains of dependent events. The most extensive exposition of this field has been made by M. Fréchet [3].

In the present paper we shall develop methods of averaging functions over chains of dependent variables and find the probability distribution of these functions. It will be shown that for certain types of chains these averages and distribution functions can be expressed in terms of the characteristic values and vectors of a certain operator equation. Many of the methods discussed here have been applied to problems in statistical mechanics [4, 5, 6, 7, 8]. The most important application has been made by L. Onsager [8] who proved rigorously (on the basis of a simplified model) that Boltzmann's energy distribution in a solid with cooperative elements leads to a phase transition. The first explicit application of linear operator theory (through matrices and integral equations) to probability chains has apparently been made by Hostinsky [9].

2. Introductory Remarks. Suppose there exists a chain of events each of which might lead to one of ν possible results, and which are correlated in such a manner that the probability of n successive events leading to a chain of results

$$\alpha_1, \alpha_2, \cdots, \alpha_n$$

is proportional to

$$P_n(\alpha_1, \alpha_2, \cdots, \alpha_n).$$

The probability of a given function $F(\alpha_1, \alpha_2, \cdots, \alpha_n)$ having a value corresponding to the sequence of α's would be proportional to

$$F(\alpha_1, \alpha_2, \cdots, \alpha_n)\ P_n(\alpha_1, \cdots, \alpha_n)$$

and its average value over all configurations of the chain would be

$$\bar{F} = F_1/F_0 = \sum_{\{\alpha_j\}} F(\alpha_1, \alpha_2, \cdots, \alpha_n)P_n(\alpha_1, \alpha_2, \cdots, \alpha_n) / \sum_{\{\alpha_j\}} P_n(\alpha_1, \cdots, \alpha_n) \tag{1}$$

where

$$F_m = \sum_{\{\alpha_j\}} [F(\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_n)]^m\ P_n(\alpha_1, \cdots, \alpha_n) \tag{1a}$$

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

and the summation extends over all values of

$$\{\alpha_j\} = (\alpha_1, \alpha_2, \cdots, \alpha_n).$$

The probability of a result α_1 of the first event leading to a result α_n of the nth event is

$$P_n(\alpha_1, \alpha_n) = (1/F_0) \sum_{\alpha_2, \cdots, \alpha_{n-1}} P_n(\alpha_1, \alpha_2, \cdots, \alpha_n). \tag{2}$$

In order to find the probability of a given function $F(\alpha_1, \cdots, \alpha_n)$ having a value between ξ and $\xi + h$ it is useful to know the moments and Thiele semi-invariants of $F(\alpha_1, \cdots, \alpha_n)$. Both of these functions of F can be calculated from

$$Z_n(x) = \sum_{\{\alpha_j\}} P_n(\alpha_1, \cdots, \alpha_n) \exp \{xF(\alpha_1, \alpha_2, \cdots, \alpha_n)\}. \tag{3}$$

Obviously

$$F_m = \lim_{x \to 0} \partial^m Z_n(x)/\partial x^m. \tag{4}$$

It is known [10] that the mth Thiele semi-invariant is given by

$$\Lambda_m = \lim_{x \to 0} \partial^m \log Z_n(x)/\partial x^m. \tag{5}$$

In the notation of Cramér $Z_n(i\omega)/Z_n(0) = f(\omega)$, the characteristic function of F.

If $G(z)$ is defined so that $G(\xi + h) - G(\xi)$ is the probability that the function $F(\alpha_1, \cdots, \alpha_n)$ has a value between $\xi \leq F(\alpha_1, \cdots, \alpha_n) < \xi + h$, then it is well known that [5] if $G(z)$ is continuous at $x = \xi$ and $x = \xi + h$

$$G(\xi + h) - G(\xi) = \frac{1}{2\pi i} \lim_{T \to \infty} \int_{-T}^{T} \frac{(1 - e^{-i\omega h})e^{-i\omega\xi}}{\omega} \exp [\log f(\omega)] \, d\omega \tag{6}$$

where

$$\log f(\omega) = \sum_{m=1}^{\infty} \frac{\Lambda_m (i\omega)^m}{m!} = \sum_{m=1}^{k} \Lambda_m (i\omega)^m/m! + o(\omega^k). \tag{6a}$$

When the derivative of $G(\xi)$ with respect to ξ exists, the probability of

$$F(\alpha_1, \cdots, \alpha_n)$$

having a value between ξ and $\xi + d\xi$ is

$$\varphi(\xi) \, d\xi = (\partial G/\partial \xi) \, d\xi = \frac{d\xi}{2\pi} \lim_{T \to \infty} \int_{-T}^{T} \exp \left\{ \sum_{m=1}^{\infty} \Lambda_m (i\omega)^m/m! \right\} e^{-i\omega\xi} \, d\omega. \tag{6b}$$

From (4)

$$\sum_{m=1}^{\infty} \Lambda_m (i\omega)^m/m! = -\log Z_n(0) + \lim_{x \to 0} e^{-i\omega \partial/\partial x} \log Z_n(x). \tag{7}$$

Since, for a constant c independent of x,

$$e^{c\partial/\partial x} f(x) = f(x + c)$$

we have

$$\sum_{m=1}^{\infty} \Lambda_m (i\omega)^m/m! = \log \{Z_n(i\omega)/Z_n(0)\}, \tag{8}$$

and from (6)

$$G(\xi + h) - G(\xi) = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-i\omega\xi}(1 - e^{-i\omega h})Z_n(i\omega)\, d\omega}{\omega Z_n(0)}. \tag{9}$$

Equations (3), (4), (5) and (9) indicate that much information concerning a chain of correlated events can be obtained from a knowledge of $Z_n(x)$. We shall now introduce procedures for the determination of $Z_n(x)$ for several general forms of $P(\alpha_1, \cdots, \alpha_n)$.

When α is a continuous variable, the results of this section and those to follow are easily generalized by replacing the summations operations over all values of the α's by integrals, and by replacing the matrix equations of the next section by integral equations.

3. Simple Chains, $P_n(\alpha_1, \cdots, \alpha_n) = \prod_{j=1}^{n-1} p(\alpha_j, \alpha_{j+1})$.

a. General theory. By a simple chain we shall mean a sequence of events, each of which leads to one of ν possible results and which occur in such a manner that if the result of the kth event is α_k, the probability of the $(k + 1)$st one yielding a result α_{k+1} to proportional to $p(\alpha_k, \alpha_{k+1})$. This implies that the probability of the occurrence of the sequence of results

$$\alpha_1, \alpha_2, \cdots, \alpha_n$$

is

$$\prod_{i=1}^{n-1} p(\alpha_i, \alpha_{i+1}) \Big/ \sum_{\{\alpha_j\}} \prod_{i=1}^{n-1} p(\alpha_i, \alpha_{i+1}), \tag{10}$$

and the probability of a first result α_1, leading to an nth result α_n is

$$P_n(\alpha_1, \alpha_n) = \sum_{\alpha_2, \cdots, \alpha_{n-1}} \prod_{j=1}^{n} p(\alpha_j, \alpha_{j+1}) \Big/ \sum_{\{\alpha_j\}} \prod_{j=1}^{n} p(\alpha_j, \alpha_{j+1}). \tag{11}$$

The summations are to be extended over all ν possible values of each α_i indicated on the summation indices. Chains of this type are sometimes called simple Markoff chains after the first author who studied them systematically.

From (1), the average value of a function $F(\alpha_1, \cdots, \alpha_n)$ is

$$F_1/F_0 = \frac{\sum_{\alpha_1} \cdots \sum_{\alpha_n} F(\alpha_1, \cdots, \alpha_n) \prod_{j=0}^{n-1} p(\alpha_j, \alpha_{j+1})}{\sum_{\alpha_1} \cdots \sum_{\alpha_n} \prod_{j=1}^{n-1} p(\alpha_j, \alpha_{j+1})}. \tag{12}$$

Many chain functions $F(\alpha_1, \cdots, \alpha_n)$ of interest are either additive or multiplicative and of one of the forms

(13a) a) $F_1(\alpha_1, \cdots, \alpha_n) = h(\alpha_1, \alpha_2) + h(\alpha_2, \alpha_3) + \cdots + h(\alpha_{n-1}, \alpha_n)$

(13b) b) $F_2(\alpha_1, \cdots, \alpha_n) = g(\alpha_1, \alpha_2)\, g(\alpha_2, \alpha_3) \cdots g(\alpha_{n-1}, \alpha_n)$.

In case (b) it is convenient to define a new function $h(\alpha_i, \alpha_j)$ by

(14) $$g(\alpha_i, \alpha_j) = \exp[xh(\alpha_i, \alpha_j)]$$

and in both cases to consider a function of the form

(15) $$Z_n(x) = \sum_{\{\alpha_i\}} \prod_{j=1}^{n-1} p(\alpha_j, \alpha_{j+1}) \exp[xh(\alpha_j, \alpha_{j+1})],$$

for then the values of F_1 and F_2 averaged over the entire chain are given by

(16a) $$<F_1>_{av.} = \lim_{x\to 0} \partial \log Z_n(x)/\partial x$$

and

(16b) $$<F_2>_{av.} = Z_n(1)/Z_n(0).$$

When n is large, the direct evaluation of (15) may become quite difficult because of the large number of variables involved. As an alternative we shall now introduce a procedure that is based on the observation that $Z_n(x)$ is the sum of the elements of the nth power of the matrix

(17) $$\boldsymbol{P}_x = \begin{pmatrix} p_x(1, 1) & p_x(1, 2) & \cdots & p_x(1, \nu) \\ p_x(2, 1) & p_x(2, 2) & \cdots & p_x(2, \nu) \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ p_x(\nu, 1) & p_x(\nu, 2) & \cdots & p_x(\nu, \nu) \end{pmatrix}$$

where the elements $p_x(\alpha, \beta)$ are defined as

(18) $$p_x(\alpha, \beta) = p(\alpha, \beta) \exp[xh(\alpha, \beta)].$$

α and β range over the same set of values as one of the "result" parameters α_j; and each of the ν possible results is represented by a unique integer of the set $1, 2, \cdots, \nu$. Thus $Z_n(x)$ = sum of elements of $\boldsymbol{P}_x^{n-1}$. To employ this observation to advantage, let us consider the characteristic values and vectors of the matrix $\boldsymbol{P}_x$. It is well known that if the characteristic values are simple the characteristic vectors form a biorthogonal set; that is, if

(19a) $$\Phi_{i,x} = \{\varphi_{i,x}(1), \varphi_{i,x}(2), \cdots, \varphi_{i,x}(\nu)\}, \qquad (i = 1, 2, \cdots, \nu),$$

and

(19b) $$\Psi_{i,x} = \begin{bmatrix} \psi_{i,x}(1) \\ \psi_{i,x}(2) \\ \psi_{i,x}(\nu) \end{bmatrix},$$

satisfy the operator equations

(20a) $$\Phi_{i,x} \cdot \boldsymbol{P}_x = \lambda_{i,x}\Phi_{i,x}$$

(20b) $$\boldsymbol{P}_x \cdot \Psi_{i,x} = \lambda_{i,x}\Psi_{i,x}$$

where $\lambda_{i,x}$ is the ith characteristic value of (17), then

$$\Phi_{i,x}\cdot\Psi_{j,x} = \sum_{\alpha=1}^{\nu} \varphi_{i,x}(\alpha)\psi_{j,x}(\alpha) = 0 \quad \text{when} \quad i \neq j.$$

We shall for convenience always assume that the φ's and ψ's are normalized:

$$\Phi_{i,x} \cdot \Psi_{i,x} = 1$$

so that in general:

(21) $$\Phi_{i,x}\cdot\Psi_{j,x} = \delta_{ij} = \begin{matrix} 0 & \text{when} & i \neq j \\ 1 & \text{when} & i = j. \end{matrix}$$

It is well known from matrix theory that one can expand a matrix element as

(22) $$p_x(\alpha, \beta) = \sum_{i=1}^{\nu} \lambda_{i,x}\varphi_{i,x}(\beta)\psi_{i,x}(\alpha)$$

and that

(23) $$\lambda_{i,x} = \Phi_{i,x} \cdot \boldsymbol{P}_x \cdot \Psi_{i,x}\,.$$

By substituting (22) into the expression for $Z_n(x)$ in terms of $\boldsymbol{P}_x^{n-1}$, one can show that

(24) $$\begin{aligned} Z_n(x) &= \sum_{i=1}^{\nu} \{\lambda_{i,x}\}^{n-1}\left\{\sum_{\beta=1}^{\nu}\varphi_{i,x}(\beta)\right\}\left\{\sum_{\alpha=1}^{\nu}\psi_{i,x}(\alpha)\right\} \\ &= \sum_{i=1}^{\nu}\lambda_{i,x}^{n-1}(\Phi_{i,x}\cdot\mathbf{1})(\mathbf{1}\cdot\Psi_{i,x}). \end{aligned}$$

Therefore $Z_n(x)$ can be determined from a knowledge of the characteristic vectors and values of the matrix $\boldsymbol{P}_x$.

If there exists a largest characteristic root $\lambda_{i,x}$ such that

(25) $$\lambda_{L,x} > |\lambda_{i,x}| \qquad \text{if } i \neq L,$$

one can obtain some interesting results. Before deriving these, we shall give a sufficient condition (which is satisfied in many chains) for the existance of this inequality. Frobenius [11] has shown that if all the elements of a finite matrix are > 0, then the characteristic value of largest absolute value of the matrix is real, positive, and simple (nondegenerate). Thus, as long as ν is finite and $p_x(\alpha, \beta) > 0$ for all α and β, (25) is valid.

We shall now prove that

(25a) $$\lim_{n\to\infty}\left\{\frac{Z_n(x)}{\lambda_{L,x}^{n-1}(\Phi_{L,x}\cdot\mathbf{1})(\mathbf{1}\cdot\Psi_{L,x})} - 1\right\} = 0$$

that is,

$$Z_n(x) \sim \lambda_{L,x}^{n-1}(\Phi_{L,x} \cdot 1)(1 \cdot \Psi_{L,x}). \tag{25b}$$

First let us consider the case in which P_x is a symmetrical matrix. Then $\varphi_{j,x}(\alpha) = \psi_{j,x}(\alpha)$, all the characteristic values are real, and

$$Z_n(x) = \lambda_{L,x}^{n-1}(\Phi_{L,x}\cdot 1)^2 + \sum_{i\neq L} \lambda_{i,x}^{n-1}(\Phi_{L,x}\cdot 1)^2.$$

From Cauchy's inequality and (21)

$$|\Phi_{i,x}\cdot 1|^2 = \left|\sum_{\alpha=1}^{\nu} \varphi_{i,x}(\alpha)\right|^2 \leq \left[\sum_{\alpha=1}^{\nu} \varphi_{i,x}^2(\alpha)\right]\left[\sum_{\alpha=1}^{\nu} 1\right] = \nu.$$

Therefore,

$$\left|\sum_{i\neq L} \lambda_{i,x}^{n-1}(\varphi_{i,x}\cdot 1)^2\right| \leq \nu \left|\sum_{i\neq L} \lambda_{i,x}^{n-1}\right| \leq \nu(\nu - 1)\,|\lambda_{s,x}^{n-1}|$$

where $\lambda_{s,x}$ is the characteristic value of P_x second largest in absolute value. This inequality yields

$$\left|\frac{Z_n(x)}{\lambda_{L,x}^{n-1}(\Phi_{L,x}\cdot 1)^2} - 1\right| \leq \frac{\nu(\nu - 1)}{(\Phi_{L,x}\cdot 1)^2}\left|\left(\frac{\lambda_{s,x}}{\lambda_{L,x}}\right)^{n-1}\right| \tag{25c}$$

and (25a) (since $\lambda_{s,x}/\lambda_{L,x} < 1$) follows. When P_x is not symmetrical, one can easily derive the analogous expression

$$\left|\frac{Z_n(x)}{\lambda_{L,x}^{n-1}(\Phi_{L,x}\cdot 1)(1\cdot\Psi_{L,x})} - 1\right| \leq \frac{A(\nu - 1)\,|\lambda_{s,\alpha}/\lambda_{L,x}|^{n-1}}{(\varphi_{L,x}\cdot 1)(1\cdot\psi_{L,x})}$$

where

$$A = [\max\{|(\Phi_{i,x} \cdot 1)|\}][\max\{|(1 \cdot \Psi_{i,x})|\}]$$

For brevity, when $x = 0$, we write $\lambda_{i,x}$ as λ_i, $\Psi_{i,x}$ as Ψ_i and $\Phi_{i,x}$ as Φ_i. By summing (10) over all α's except α_1, α_k and α_n we obtain the probability of an intermediate event leading to a result α_k if the results of the first and last events are known to have been α_1 and α_n. With the aid of (21) and (22) it is easy to show that this probability is exactly:

$$\frac{\sum_{i,j=1}^{\nu} \lambda_j^{n-k}\lambda_i^{k-1}\psi_i(\alpha_1)\varphi_i(\alpha_k)\psi_j(\alpha_k)\varphi_j(\alpha_n)}{\sum_{i=1}^{\nu} \lambda_i^{n-1} \sum_{\alpha_1,\alpha_n} \psi_i(\alpha_1)\varphi_i(\alpha_n)}. \tag{26}$$

When n is very large, and when we have simultaneously $n >> k >> 1$, we can rewrite this equation to include λ_L, and neglect all terms containing other i's and j's. This leads to the results

a) If the number of events, n, in a simple chain is very large, the probability $P_n(\alpha_k)$ of a kth event far removed from the first and the last, yielding a result α_k when α_1, and α_n are unspecified is

$$P_n(\alpha_k) \sim \psi_L(\alpha_k)\;\varphi_L(\alpha_k)\;/\;(\Phi_L \cdot 1)(1 \cdot \Psi_L). \tag{27}$$

b) When $k = n$, the probability of the result $\alpha_1 \cdot$ of the first event leading to the result α_n of the nth event is

(28a)
$$P_n(\alpha_1, \alpha_n) = \frac{\sum_{i=1}^{\nu} \lambda_i^{n-1} \psi_i(\alpha_1)\varphi_i(\alpha_n)}{\sum_{i=1}^{\nu} \lambda_i^{n-1} \sum_{\alpha_1,\alpha} \psi_i(\alpha_1)\varphi_i(\alpha_n)}.$$

So, as $n \to \infty$

(28b)
$$P_n(\alpha_1, \alpha_n) \sim \frac{\psi_L(\alpha_1)\phi_L(\alpha_n)}{(\Phi_L \cdot 1)(1 \cdot \Psi_L)}.$$

c) When there exists no knowledge concerning the result of the first event, the probability of the nth event yielding the result α_n is

(29)
$$P_n(\alpha_n) = \sum_{\alpha_1} P_n(\alpha_1, \alpha_n) \sim \Phi_L(\alpha_n)/(\mathbf{1} \cdot \Phi_L).$$

In chains of sufficient length for (25) to be valid, the probability of

$$F(\alpha_1, \cdots, \alpha_n)$$

having a value between ξ and $\xi + h$ has an especially simple asymptotic form. From (6) this probability is (if for a given n we let $T = an^{\frac{1}{2}}$)

(30)
$$G(\xi + h) - G(\xi) = \frac{1}{2\pi i} \lim_{a\to\infty} \int_{-an^{1/2}}^{an^{1/2}} \left(\frac{d\omega}{\omega}\right) e^{-i\omega(\xi - \Lambda_1)} (1 - e^{-i\omega h}) \exp\left\{-\tfrac{1}{2}\omega^2 \Lambda_2 - \frac{i\Lambda_3 \omega^3}{3!} + \cdots\right\}$$

and from (25) and (5)

(31)
$$\Lambda_m \sim n \lim_{x\to 0} \partial^m \log \lambda_{L,x}/\partial x^m = nL_m$$

if

(32)
$$L_m \equiv \lim_{x\to 0} \partial^m \log \lambda_{L,x}/\partial x^m.$$

Letting $y = \omega n^{\frac{1}{2}}$, (30) becomes

(33)
$$G(\xi + h) - G(\xi) \sim \frac{1}{2\pi i} \lim_{a\to\infty} \int_{-a}^{a} \frac{dy}{y} (e^{-iy\mu_1} - e^{-iy\mu_2}) e^{-\frac{1}{2}y^2 L_2} \left\{1 - \frac{L_3 y^3 i}{6n^{\frac{1}{2}}} + \cdots\right\}$$

where

(34a)
$$\mu_1 = (\xi - \Lambda_1)/n^{\frac{1}{2}}$$
$$\mu_2 = (\xi + h - \Lambda_1)/n^{\frac{1}{2}}$$

(34b)
$$\Lambda_1 = \text{average value of } F(\alpha_1, \cdots, \alpha_n) = \bar{F}.$$

Integrating (33)

$$G(\xi + h) - G(\xi) \sim \frac{1}{(2\pi L^2)^{\frac{1}{2}}} \int_{\mu_1}^{\mu_2} e^{-\mu^2/2L_2}[1 + O(1/n)]\, d\mu. \tag{35}$$

As $n \to \infty$ and $h \to 0$

$$G(\xi + h) - G(\xi) \sim \frac{h}{(2\pi n L_2)^{\frac{1}{2}}} \exp\, (-\tfrac{1}{2})[\xi - \bar{F}]/nL_2), \tag{35a}$$

and the probability that $\xi \leq F < \xi + h$ becomes Gaussian.

b. Examples of a simple chain. As an example of a simple Markoff chain let us consider an event which can lead to either of two possible results, say "-1" or "1". Further, let us suppose that the probability of a given result being followed by an identical one is p and by one of another type is $(1 - p)$; that is,

$$p(-1, -1) = p(1, \ \ 1) = p$$
$$p(-1, \ \ 1) = p(1, -1) = 1 - p.$$

This chain would be encountered in an analysis of a sequence of tosses of a coin with a "memory" so that the probability of two successive tosses showing the same face of the coin would be p and that of showing opposite faces $(1 - p)$.

A question one might ask concerning such a chain is—What is the probability of the occurrence of a given number of transitions from one kind of result to another? In the chain of results

$$-1, -1, -1, 1, 1, -1, 1, -1, -1, -1$$

there would be four transitions, one corresponding to each -1 followed by a 1 and to each 1 followed by a -1. The function giving the number of transitions in a sequence of n events is

$$F(\alpha_1, \cdots, \alpha_n) = \sum_{i=1}^{n-1} h(\alpha_i, \alpha_{i+1}) \tag{36}$$

where

$$h(-1, -1) = h(1, \ \ 1) = 0$$
$$h(-1, \ \ 1) = h(1, -1) = 1.$$

Even though the α's are dependent, in this special case, $h(\alpha_i, \alpha_{i+1})$ and $h(\alpha_{i+1}, \alpha_{i+2})$ are independent so that (40) could have been obtained on this basis.

To apply the methods described in the beginning of this section we must find the characteristic values and vectors of the matrix

$$P_x = \begin{pmatrix} p & (1 - p)e^x \\ (1 - p)e^x & p \end{pmatrix} \tag{37}$$

(the configuration index α has the value either -1 or 1 in this case instead of

"1" and "2" as given in (17)). The characteristic values are the roots of the equation

$$\begin{vmatrix} p - \lambda & (1 - p)e^x \\ (1 - p)e^x & p - \lambda \end{vmatrix} = 0$$

that is,

$$\begin{aligned} \lambda_{1,x} &= p + (1 - p)e^x \\ |\lambda_{2,x}| &= |p - (1 - p)e^x| < \lambda_{1,x} \end{aligned} \tag{38}$$

and the characteristic vectors are

$$\psi_{1,x} = 2^{-\frac{1}{2}}\begin{pmatrix}1\\1\end{pmatrix} \quad \text{and} \quad \psi_{2,x} = 2^{-\frac{1}{2}}\begin{pmatrix}1\\-1\end{pmatrix}.$$

The ψ and φ vectors have the same components in this case because of the sym metry of the P_x matrix. Clearly

$$\lambda_L = \lambda_1 = \lambda_{1,0} = 1; \qquad \lambda_2 = \lambda_{2,0} = 2p - 1$$

$$\psi_1(\alpha) = 2^{-\frac{1}{2}} \quad \text{and} \quad \psi_2(\alpha) = -\alpha \cdot 2^{-\frac{1}{2}}.$$

From (26) we see that if the result of the first event in the chain is α_1, and that of the nth event is α_n, the probability of the kth event yielding the result α_k is

$$\frac{[(2p - 1)^{k-1}\alpha_1\alpha_k + 1][1 + (2p - 1)^{n-k}\alpha_k\alpha_n]}{2[1 + (2p - 1)^{n-1}\alpha_1\alpha_n]}.$$

As k, n_1 and $(n - k)$ simultaneously get very large, $P_n(\alpha_k) \sim \frac{1}{2}$, independently of α_k.

The probability of an initial result α_1 leading to a final result α_n is (from 28a)

$$P_n(\alpha_1, \alpha_n) = (\tfrac{1}{4})\{1 + (2p - 1)^{n-1}\alpha_1\alpha_n\}$$

so that

$$\begin{aligned} P_n(\ 1, 1) &= P_n(-1, -1) = (\tfrac{1}{4})\{1 + (2p - 1)^{n-1}\} \\ P_n(-1, 1) &= P_n(\ 1, -1) = (\tfrac{1}{4})\{1 - (2p - 1)^{n-1}\}. \end{aligned}$$

Now, to answer our original question regarding the probability distribution of the transition function (36)

$$F(\alpha_1, \cdots, \alpha_n) = \sum_{i=1}^{n-1} h(\alpha_i, \alpha_{i+1}), \tag{39}$$

we use the expression for $Z_n(x)$ determined from (24)

$$Z_n(x) = 2[p + (1 - p)e^x]^{n-1} \tag{39}$$

From (9) the probability of there being between ξ and $\xi + h$ transitions in a sequence of $n + 1$ events is

$$G(\xi + h) - G(\xi) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} e^{i\omega\xi}(1 - e^{i\omega\xi})\{p + (1 - p)e^{i\omega}\}^n \, d\omega/\omega \tag{40}$$

$$= \frac{1}{2\pi i} \int_{-\infty}^{\infty} (e^{-i\omega\xi} - e^{-i\omega(\xi - h)}) \sum_{k=0}^{n} \frac{n!(1 - p)^k p^{n-k}}{(n - k)!k!}.$$

Letting $x = \omega h/2$ and rearranging

$$G(\xi + h) - G(\xi) = \frac{1}{\pi} \sum_{k=0}^{n} \frac{n!(1 - p)^k p^{n-k}}{(n - k)!k!} D\left(1 + \frac{2}{h}(\xi + h)\right),$$

where $D(\lambda)$ is the Dirichlet integral

$$D(\lambda) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin x \cos \lambda x}{x} \, dx = \begin{cases} 0 & \text{if } |\lambda| > 1 \\ \frac{1}{2} & |\lambda| = 1 \\ 1 & |\lambda| < 1. \end{cases}$$

We therefore have, when $[\xi + h] \leq n$

$$G(\xi + h) - G(\xi) = \sum_{k=[\xi+1]}^{[\xi+h]} \frac{n!(1 - p)^k p^{n-k}}{(n - k)!k!}. \tag{41}$$

Here $[x]$ denotes the greatest integer not exceeding x. The sum is zero if $[\xi + h] < [\xi + 1]$. When $[\xi + h] > n$

$$G(\xi + h) - G(\xi) = \sum_{k=[\xi+1]}^{n} \frac{n!(1 - p)^k p^{n-k}}{k!(n - k)!}. \tag{42}$$

When n is large it is difficult to get a clear picture of the function $G(\xi)$ from (41) and (42), so we shall develop asymptotic results for large n by using (6) instead of (9).

By employing (5), we see that (this section will be developed on the basis of $n + 1$ trials instead of n)

$$\Lambda_1 = \bar{F} = n(1 - p)$$

$$\Lambda_2 = np(1 - p)$$

$$\Lambda_3 = np(1 - p)(2p - 1) \text{ etc.}$$

Therefore, from (6)

$$\Delta G = G(\xi + h) - G(\xi) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{-i\omega(\xi - \Lambda_1)}(1 - e^{-i\omega h})}{\omega}$$

$$\exp\left[-\tfrac{1}{2}np(1 - p)\omega^2 - inp(1 - p)(2p - 1)\omega^3/6 - \cdots\right] d\omega.$$

Letting $u = \omega n^{\frac{1}{2}}$, we have

$$\Delta G = \frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{du}{u}[e^{-iu(\xi-\Lambda_1)/n^{\frac{1}{2}}} - e^{-iu(\xi+h-\Lambda_1)/n^{\frac{1}{2}}}]$$

$$\left[1 - \frac{ip(1-p)(2p-1)u^3}{6n^{\frac{1}{2}}} + O\left(\frac{u^4}{n}\right)\right]e^{-\frac{1}{2}u^2p(1-p)}$$

$$= \frac{1}{2\pi}\int_{\mu_1}^{\mu_2} d\lambda \int_{-\infty}^{\infty} e^{-iu\lambda}\left[1 - \frac{ip(1-p)(2p-1)u^3}{6n^{\frac{1}{2}}} + O\left(\frac{u^4}{n}\right)\right]e^{-\frac{1}{2}u^2p(1-p)}\,du$$

where

$$\mu_1 = (\xi + h - \Lambda_1)/n^{\frac{1}{2}}$$

$$\mu_2 = (\xi - \Lambda_1)/n^{\frac{1}{2}}.$$

Since

$$\int_{-\infty}^{\infty} e^{-au^2}e^{-i\lambda u}\,du = (\pi/a)^{\frac{1}{2}}\exp(-\lambda^2/4a)$$

$$i\int_{-\infty}^{\infty} u^3 e^{-au^2}e^{-i\lambda u}\,du = \frac{3\lambda\pi^{\frac{1}{2}}}{4a^{5/2}}\left(1 - \frac{\lambda^2}{6a}\right)e^{-\lambda^2/4a},$$

we have for large n

(43a)
$$\Delta G \sim \frac{1}{[2\pi p(1-p)]^{\frac{1}{2}}}\int_{-\mu_1}^{\mu_2} e^{-\lambda^2/2p(1-p)} \left\{1 - \frac{(2p-1)\lambda}{2p(1-p)n^{\frac{1}{2}}}\left(1 - \frac{\lambda^2}{3p(1-p)}\right) + O\left(\frac{1}{n}\right)\right\}d\lambda.$$

As $n \to \infty$ and $h \to 0$, this becomes

(43b)
$$G(\xi + h) - G(\xi) \sim \frac{h\exp\{-[\xi - \bar{F}]^2/2p(1-p)n\}}{[2\pi np(1-p)]^{\frac{1}{2}}} \left\{1 - \frac{(2p-1)(\xi - \bar{F})}{2p(1-p)n} + O\left(\frac{1}{n^2}\right)\right\}.$$

A similar problem which occurs in statistics of high polymers can be stated abstractly as follows. Suppose there exists a sequence of events each of which leads to a translation of length a of a point either to the right or to the left, and that the probability of a translation continuing in the same direction as its predecessor is p while that of changing its direction is $(1 - p)$. After n translations what is the probability of a point being displaced a distance ξ from its origin.

If "-1" represents a translation to the left and "$+1$" a translation to the right,

$$p(-1, -1) = p(1, 1) = p$$

$$p(-1, 1) = p(1, -1) = (1 - p)$$

The function giving the distance of the point from its origin after n displacements is (when $\alpha = \pm 1$)

$$F(\alpha_1, \cdots, \alpha_n) = a \sum_{j=1}^{n} \alpha_j = \tfrac{1}{2}a\alpha_1 + h(\alpha_1, \alpha_2) + \cdots + h(\alpha_{n-1}, \alpha_n) + \tfrac{1}{2}a\alpha_n$$

where

$$h(1, \;\; 1) = a, \quad h(-1, -1) = -a$$
$$h(1, -1) = h(-1, 1) = 0.$$

Neglecting the terms $a\alpha_1/2$ and $a\alpha_n/2$ in $F(\alpha_1, \cdots, \alpha_n)$, one can answer questions concerning this problem by evaluating $Z_n(x)$ as defined by (15). In this case P_x has the form

$$P_x = \begin{pmatrix} pe^{ax} & 1-p \\ 1-p & pe^{-ax} \end{pmatrix}.$$

Its characteristic roots are

$$\lambda_{1,x} \;=\; p \cosh ax + [p^2 \cosh^2 ax + (1 - 2p)]^{\frac{1}{2}} = \lambda_{L,x}$$
$$|\lambda_{2,x}| = |p \cosh ax - [p^2 \cosh^2 ax + (1 - 2p)]^{\frac{1}{2}}| < \lambda_{1,x}.$$

and its characteristic vectors:

$$\psi_{1,x} = [(p-1)^2 + (pe^{ax} - \lambda_1)^2]^{-\frac{1}{2}} \begin{pmatrix} p-1 \\ pe^{ax} - \lambda_1 \end{pmatrix}$$

$$\psi_{2,x} = [(p-1)^2 + (pe^{ax} - \lambda_2)^2]^{-\frac{1}{2}} \begin{pmatrix} p-1 \\ pe^{ax} - \lambda_2 \end{pmatrix}.$$

Since

$$\bar{F} = \Lambda_1 = \lim_{x \to 0} \partial \log Z_n(x)/\partial x,$$

one can show in the present problem that $\bar{F} = 0$. Therefore, the probability of the translated point being a distance between ξ and $\xi + h$ from the origin after $(n + 1)$ translations, is, as $n \to \infty$ and $h \to 0$

$$F(\xi + h) - F(\xi) \sim h(2\pi n L_2)^{-\frac{1}{2}} e^{-\xi^2/2nL_2}$$

where L_2 is by (32):

$$L_2 = \lim_{x \to 0} \partial^2 \log \lambda_{L,x}/\partial x = a^2 p/(1-p).$$

Thus,

$$F(\xi + h) - F(\xi) \sim h[a^2 2\pi n p/(1-p)]^{-\frac{1}{2}} e^{-\xi^2(1-p)/2na^2p}.$$

When $p = 2/3$ this problem is equivalent to the determination of the probability distribution of the components in an arbitrary direction of the distance between the ends of a linear polymer. In this case

$$F(\xi + h) - F(\xi) \sim h(4a^2\pi n)^{-\frac{1}{2}} \exp(-\xi^2/4na^2)$$

a result obtained by Tobolsky [12] after a lengthy and complicated combinatory calculation.

Another type of simple chain is encountered in the determination of the "life span" of a particle which is displaced a unit distance to the right or left per unit time along a straight line until it collides with an absorbing boundary either $-(q + 1)$ or $(p + 1)$ units from the starting point. This problem has been analyzed by M. Kac using the methods discussed in the present paper. We shall generalize his results to include the effect of an attraction of the particle toward one end of the line so that displacements toward that end are more probable than those in the other direction.

Following the notation of Kac [13] we let X_j represent the jth displacement, m_j its length, and $\delta(m)$ the probability of a given displacement having the length m. Then,

$$\delta(m) = \begin{cases} s & \text{if } m = 1 \\ 1 - s & \text{if } m = -1 \\ 0 & \text{otherwise.} \end{cases}$$

If N represents the life span of a particle, the probability of its exceeding n is

$$\begin{aligned} \text{Prob}\,\{N > n\} &= \text{Prob}\,\{-q \le X_1 \le p, -q \le X_1 + X_2 \le p, \cdots, \\ &\quad -q \le X_1 + X_2 + \cdots + X_n \le p\} = \Sigma\, \delta(m_1)\delta(m_2) \cdots \delta(m_n) \end{aligned}$$

where the summation extends over all integers $m_1, m_2, \cdots, m_n$ such that $-q \le m_1 \le p, -q \le m_1 + m_2 \le p, \cdots, -q \le m_1 + m_2 + \cdots + m_n \le p$.

Defining the new set of variables

$$\alpha_j = q + m_1 + m_2 + \cdots + m_j \qquad (j = 1, 2, \cdots n)$$

we see that

$$\text{Prob}\,\{N > n\} = \sum_{\alpha_1, \cdots, \alpha_n = 0}^{p+q} \delta(\alpha_1 - q)\delta(\alpha_2 - \alpha_1) \cdots \delta(\alpha_n - \alpha_{n-1}).$$

As before, if we introduce the P matrix (of $p + q + 1$ rows and columns)

$$P = (\delta(\alpha - \beta)) = \begin{Bmatrix} 0 & 1 - s & 0 & 0 & \cdots \\ s & 0 & 1 - s & 0 & \cdots \\ 0 & s & 0 & 1 - s & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{Bmatrix}$$

we obtain after applying the equivalent of (22)

$$\text{Prob }\{N > n\} = \sum_{j=1}^{p+q+1} \lambda_j^n \varphi_j(q) \sum_{\alpha_n=0}^{p+q} \psi_j(\alpha_n).$$

Where λ_j is the jth characteristic value of P, and ψ_j and φ_j are its associated characteristic vectors as defined by (19) and (20) (here the range of α starts from 0 instead of 1 as in (17) and (19)).

It is easy to show that the characteristic values of P are

$$\lambda_j = 2[\mathrm{s}(1 - \mathrm{s})]^{\frac{1}{2}}\cos \zeta_j \ (j = 1, 2, \cdots, p + q + 1)$$

where

$$\zeta_j = \pi j/(p + q + 2)$$

and that the components of the characteristic vectors are

$$\psi_j(\alpha) = [2/(p + q + 2)]^{\frac{1}{2}}[\mathrm{s}/(1 - s)]^{\frac{1}{2}\alpha} \sin (\alpha + 1)\zeta_j \qquad (\alpha = 0, 1, \cdots, p + q)$$

and

$$\varphi_j(\alpha) = [2/(p + q + 2)]^{\frac{1}{2}}[(1 - \mathrm{s})/s]^{\frac{1}{2}\alpha} \sin (\alpha + 1)\zeta_j .$$

Since

$$\sum_{\alpha_n=0}^{p+q} \psi_j(\alpha_n) = \frac{\sqrt{2}\,(1 - s)}{\sqrt{p + q + 2}} \, \frac{\{1 - 1(-1)^j[s/1 - s]^{\frac{1}{2}(p+q+2)}\} \sin \zeta_j}{1 - 2[s(r - s)]^{\frac{1}{2}} \cos \zeta_j}$$

we finally have

$$\text{Prob }\{N > n\} = \frac{(1 - s)^{\frac{1}{2}(n+q+2)} 2^{n+1} s^{\frac{1}{2}(n-q)}}{p + q + 2} \sum_{j=1}^{p+q+1} \frac{\{1 - (-1)^j(s/1 - s)^{\frac{1}{2}(p+q+2)}\} \cos^n \zeta_j \sin \zeta_j \sin (q + 1)\zeta_j}{1 - 2\sqrt{s(1 - s)} \cos \zeta_j}$$

When $s = \frac{1}{2}$ this reduces to the result of Kac: (* means summation is only over even j's

$$\text{Prob }\{N > n\} = \frac{2}{p + q + 2} \sum_{j=1}^{p+q+1} {}^{*} \cos^n \zeta_j \sin (q + 1)\zeta_j \cot \tfrac{1}{2}\zeta_j .$$

4. Simple Chains with Restrictions. Often when studying chains of dependent events, certain functions averaged over the entire chains are known to be restricted between definite limits. That is, there might exist k functions $g_j(\alpha_1, \alpha_2, \cdots, \alpha_n)$ such that

$$(44) \qquad -\Delta G_j < G_j - g_j(\alpha_1, \cdots, \alpha_n) < \Delta G_j, \qquad (j = 1, 2, \cdots k),$$

where the G_j's and ΔG_j's are preassigned constants. To calculate averages of other functions (1) is no longer valid, for it is an unrestricted sum over all sets of α's, including those incompatible with (44). All unrestricted sums in this formula (and other similar ones) must be replaced by sums over only those sets of α's compatible with (44). Since it is sometimes more difficult to evaluate restricted sums than unrestricted ones, we shall apply an idea of Markoff [1] to the reduction of the former to the latter type.

Let us seek an explicit expression for a function $P_n^*(\alpha_1, \alpha_2, \cdots, \alpha_n)$ which has the property:

$$P_n^*(\alpha_1, \cdots, \alpha_n) = \begin{cases} P_n(\alpha_1, \cdots, \alpha_n) & \text{when } \alpha\text{'s are chosen so that (44) is satisfied of all } j. \\ 0 & \text{otherwise.} \end{cases}$$

Since the Dirichlet integrals

$$\delta_j = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin(\rho_j \Delta G_j)}{\rho_j} \exp(i\rho_j \gamma_j)\, d\rho_j$$

have the property

$$\delta_j = \begin{cases} 1 \text{ when } -\Delta G_j < \gamma_j < \Delta G_j \\ 0 \text{ otherwise,} \end{cases}$$

$$P_n^*(\alpha_1, \cdots, \alpha_n) = \delta_1\delta_2 \cdots \delta_k P_n(\alpha_1, \cdots, \alpha_n)$$

has the required character provided

$$\gamma_j = G_j - g_j(\alpha_1, \cdots, \alpha_n).$$

The average value of a function $F(\alpha_1, \cdots, \alpha_n)$ can be written in terms of the unrestricted sum

$$\bar{F} = \sum_{\{\alpha_e\}} F(\alpha_1, \cdots, \alpha_n) P_n^*(\alpha_1, \cdots, \alpha_n) \Big/ \sum_{\{\alpha_e\}} P_n^*(\alpha_1, \cdots, \alpha_n),$$

where the summation extends over the complete set of $\{\alpha_e\}$'s

$$\{\alpha_e\} = (\alpha_1, \alpha_2, \cdots, \alpha_n).$$

As in the case of chains without auxiliary restrictions, a useful function is

$$\begin{aligned} Z_n(x) &= \sum_{\{\alpha_e\}} P_n^*(\alpha_1, \cdots, \alpha_n) \exp\{xF(\alpha_1, \cdots, \alpha_n)\} \\ &= \frac{1}{\pi^k} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} S_n(x, \rho_1, \cdots, \rho_k) \prod_{m=1}^{k} \left\{ \frac{\sin(\rho_m \Delta G_m)}{\rho_m} e^{i\rho_m G_m} d\rho_m \right\} \end{aligned} \tag{45}$$

where

$$S_n(x, \rho_1, \cdots, \rho_k) = \sum_{\{\alpha_e\}} P_n(\alpha_1, \cdots, \alpha_n) \exp\left\{ xF(\alpha_1, \cdots, \alpha_n) - i \sum_{j=1}^{k} \rho_j g_j(\alpha_1, \cdots, \alpha_n) \right\}.$$

When $F(\alpha_1, \cdots, \alpha_n)$ and $\{g_j(\alpha_1, \cdots, \alpha_n)\}$ are all additive or multiplicative functions of the form (13a) and (13b), say

$$F(\alpha_1, \cdots, \alpha_n) = \sum_{k=1}^{n-1} h(\alpha_k, \alpha_{k+1})$$

$$g_j(\alpha_1, \cdots, \alpha_n) = \sum_{k=1}^{n-1} g_j(\alpha_k, \alpha_{k+1})$$

and the probability chain is a simple one, $Z_n(x)$ reduces to a simple form.

Suppose

$$P_n(\alpha_1, \cdots, \alpha_n) = \sum_{j=1}^{n-1} p(\alpha_j, \alpha_{j+1})$$

then following the derivation of (24), we have

$$S_n(x, \rho_1, \cdots, \rho_k) = \sum_{l=1}^{\nu} \{\lambda_{l,x,\rho}\}^{n-1}(\Phi_{l,x,\rho}\cdot\mathbf{1})(\mathbf{1}\cdot\Psi_{l,x,\rho}) \tag{46}$$

where $\lambda_{l,x,\rho}$, $\Phi_{l,x,\rho}$ and $\Psi_{l,x,\rho}$ are characteristic values and vectors of the matrix

$$P_{x,\rho} = \begin{pmatrix} p_{x,\rho}(1, 1) & \cdots & p_{x,\rho}(1, \nu) \\ \cdots & \cdots & \cdots \\ p_{x,\rho}(\nu, 1) & \cdots & p_{x,\rho}(\nu, \nu) \end{pmatrix}$$

and

$$p_{x,\rho}(\alpha, \beta) = p(\alpha, \beta) \exp \{xh(\alpha, \beta) - i \sum_j \rho_j g_j(\alpha, \beta)\}.$$

Substitution of (46) into (45) allows one to calculate $Z_n(x)$.

5. More Complicated Chains. In a chain of N events in which the result of each event depends on those of its n predecessors ($n << N$), the calculation of $Z_n(x)$ proceeds in essentially the same manner as in the case of a simple chain. Let the N events be divided into N/n sets of "grand events" of n simple events each (for simplicity we assume N is divisible by n, this can easily be avoided). Thus, if each simple event could lead to any one of ν possible results, a grand event could lead to any one of ν^n possible results and a complicated chain becomes a simple chain of grand events with the result of each grand event depending on the preceeding grand event. Quantitative calculations thus proceed formally in the same manner as in a simple chain.

6. Continuous Case. In this section we generalize, by studying an example, to the case in which each event in a simple chain may lead to any one of a continuum of results. The example is a problem arising in statistical mechanics of molecular chains.

Consider a linear chain of n identical molecules whose centers of mass remain at a set of fixed regularly spaced positions, but which may rotate about their centers of mass in a plane. Suppose, that the potential energy of interaction between neighboring pairs of molecules is a function of the angles a specified axis of the molecules makes with the line connection the centers of mass of the molecules; that is, the potential energy of interaction between pairs of adjacent molecules can be written as $V(\theta_j, \theta_{j+1})$. Assuming that forces are sufficiently short ranged for interaction between more distant neighbors can be neglected, Boltzmann's theorem states that the probability of the axis of the first molecule making an angle between θ_1 and $\theta_1 + a\theta_1$ with the line of centers of the chain, the second between θ_2 and $\theta_2 + \alpha\theta_2$ and the nth between θ_n and $\theta_n + d\theta_n$ is proportional to

$$\exp\,[-kT\,\{V(\theta_1, \theta_2) + V(\theta_2, \theta_3) + \cdots + V(\theta_{n-1}, \theta_n)\}]\, d\theta_1 \cdots d\theta_n$$

where k is Boltzmann's constant and T is the absolute temperature. The contribution of the interaction to the thermodynamic properties of the chain can be derived from the partition function

$$Z_n = \int_0^{2\pi}\int_0^{2\pi}\cdots\int_0^{2\pi} \exp\left\{-\frac{1}{kT}\,[V(\theta_1, \theta_2) + \cdots + V(\theta_{n-1}, \theta_n)]\right\} d\theta_1 \cdots d\theta_n. \tag{47}$$

For example, the internal energy is

$$\bar{E} = \partial \log Z_n/\partial(-1/kT)$$

and the specific heat is $c = \partial E/\partial T$.

It is to be noted that Z_n is exactly the integral of the iterated kernel of the integral equation

$$\lambda\psi(\theta_1) = \int_0^{2\pi} \psi(\theta_2) \exp\left\{-\frac{1}{kT} V(\theta_1, \theta_2)\right\} d\theta_2. \tag{48}$$

If $V(\theta_1, \theta_2)$ is symmetrical in θ_1 and θ_2, this linear homogeneous integral equation has a set of orthonormal characteristic functions $\{\psi_j(\theta)\}$ such that

$$\int_0^{2\pi} \psi_j(\theta)\psi_k(\theta)\, d\theta = \delta_{jk}. \tag{49}$$

To each of these characteristic functions there corresponds a characteristic value λ_j. Now it is well known that the kernel of (48) can be expanded as a series in its characteristic functions

$$\exp\left\{-\frac{1}{kT} V(\theta_1, \theta_2)\right\} = \sum_j \lambda_j \psi_j(\theta_1)\psi_j(\theta_2).$$

Introduction of this expression into (47) and applying the orthogonality conditions (49) one obtains

$$Z_n = \sum_j \lambda_j^{n-1}\left\{\int_0^{2\pi} \psi_j(\theta)\, d\theta\right\}^2. \tag{47a}$$

Probably the most interesting example of a molecular chain of the type described above is a chain of magnetic dipoles which are restricted to rotate only in a plane. In that case

$$V(\theta_j, \theta_{j+1}) = \frac{\mu^2}{r^3}[\cos(\theta_j - \theta_{j+1}) - 3\cos\theta_j \cos\theta_{j+1}].$$

Where μ is the magnetic moment of each dipole and r is the distance between a pair of adjacent centers of mass. This potential function leads to the integral equation

$$\lambda\psi(\theta_1) = \int_0^{2\pi} \psi(\theta_2) \exp\left\{-\frac{\mu^2}{r^3kT}[\cos(\theta_1 - \theta_2) - 3\cos\theta_1\cos\theta_2]\right\} d\theta_2 .$$

Since this equation is rather complicated to solve, we shall devote the rest of the section to a potential function of less physical interest, but which leads to a less formidable integral equation.

In studying hindered rotation of molecules, one sometimes uses potential functions of the form:

$$V(\theta_j, \theta_{j+1}) = -\beta\cos(\theta_j - \theta_{j+1})$$

where β is a constant. With this potential function (48) becomes

$$\lambda\psi(\theta_1) = \int_0^{2\pi} \psi(\theta_2) \exp\{J\cos(\theta_1 - \theta_2)\}\, d\theta_2 \tag{50}$$

where

$$J = \beta/kT.$$

The characteristic functions and characteristic values of (50) are easily found with the aid of the Fourier Series for exp (J cos θ):

$$\exp(J\cos\theta) = I_0(J) + 2\sum_{m=1}^{\infty} I_m(J)\cos m\,\theta \tag{51}$$

where $I_m(J)$ is the mth Bessel function of imaginary argument:

$$I_m(J) = \sum_{m=0}^{\infty} \frac{(\frac{1}{2}J)^{2k+m}}{(m+k)!k!} .$$

From (51)

$$\exp[J\cos(\theta_1 - \theta_2)] = I_0(J) + 2\sum_{m=1}^{\infty} I_m(J)(\cos m\theta_1 \cos m\theta_2 + \sin m\theta_1 \sin m\theta_2).$$

Substituting this expression into (50) we have

$$\lambda\psi(\theta_1) = \int_0^{2\pi} \psi(\theta_2)\left\{I_0(J) + 2\sum_{m=1}^{\infty} I_m(J)(\cos m\theta_1\cos m\theta_2 + \sin m\theta_1 \sin m\theta_2)\right\} d\theta_2 .$$

Because of the orthogonality of the trigonometric functions, one can verify by direct substitution that the characteristic functions are

$$\psi_0(\theta) = 1/(2\pi)^{\frac{1}{2}}$$

$$\psi_m^{(1)}(\theta) = \pi^{-\frac{1}{2}} \sin m\theta; \qquad \psi_m^{(2)} = \pi^{-\frac{1}{2}} \cos m\theta, \ (m = 1, 2, \cdots)$$

and the corresponding characteristic values are

$$\lambda_0 = 2\pi I_0(J)$$

$$\lambda_m^{(1)} = \lambda_m^{(2)} = 2\pi I_m(J) \qquad m > 0.$$

Introduction of these characteristic functions and values into (47a) we obtain the simple formula for the partition function:

$$Z_n = 2\pi\{2\pi I_0(J)\}^{n-1}.$$

The internal energy of the molecular chain is therefore

$$\begin{aligned}\bar{E} &= \partial \log Z_n/\partial(-1/kT) \\ &= -\beta(n - 1)\, I_1(J)/I_0(J),\end{aligned}$$

and the specific heat is:

$$C = \partial\bar{E}/\partial T = \tfrac{1}{2}k(n - 1)J^2\left\{1 + \frac{I_2(J)}{I_0(J)} - 2\left[\frac{I_1(J)}{I_0(J)}\right]^2\right\}.$$

REFERENCES

[1] A. Markoff, *Wahrscheinlichkeitsrechnung*, Leipzig, 1912.

[2] S. Bernstein, "Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dependantes," *Math. Ann.*, Vol. 97 (1927), p. 1.

[3] M. Frechét, *Recherces Theoretiques Moderns sur La Theorie des Probabilites*, Vol. 2, Paris (1937).

[4] H. Kramers and G. Wannier, "Statistics of the two-dimensional ferromagnet: Part I," *Phys. Rev.*, Vol. 60, (1941), p. 252.

[5] E. Montroll, "Statistical mechanics of nearest neighbor systems," *Jour. Chem. Phys.*, Vol. 9 (1941), p. 708; Vol. 10 (1942), p. 61.

[6] E. Lassettre and J. Howe, "Thermodynamic properties of binary solid solutions on the basis of the nearest neighbor approximation," *Jour. Chem. Phys.*, Vol. 9 (1941), p. 747.

[7] J. Ashkin and W. E. Lamb, "The propagation of order in crystal lattices," *Phys. Rev.*, Vol. 64 (1943), p. 159.

[8] L. Onsager, "Crystal statistics I. A two-dimensional model with an order-disorder transition," *Phys. Rev.*, Vol. 65, (1944), p. 117.

[9] M. Hostinsky, *Methodes generales du Calculu des Probabilites*, Paris, 1931.

[10] H. Cramér, *Random Variables and Probability Distributions*, Cambridge Univ. Press, 1937, Chap. 4.

[11] G. Frobenius, "Über Matrizen aus positiven Elementen. II.," *Preuss. Acad. Wiss. Sitz.*, (1909), p. 514.

[12] A. Tobolsky, Powell and H. Eyring, an article in *Chemistry of Large Molecules*, Interscience Publishers, 1943, pp. 156, 182.

[13] M. Kac, "Random walk in the presence of absorbing barriers," *Annals of Math. Stat.* Vol. 14, (1945), p. 62.