IDENTICAL LOCI AND RELATIONSHIP

G. MALÉCOT
UNIVERSITY OF LYON

1. Introduction

I shall call identical loci, two loci bearing genes identical by descent; that is, going back to the same locus of one common ancestor.

If we consider one diploid individual called K, his two homologous loci may be identical if he is "inbred"; that is, if his two parents I and J have some common ancestors A_i (figure 1).

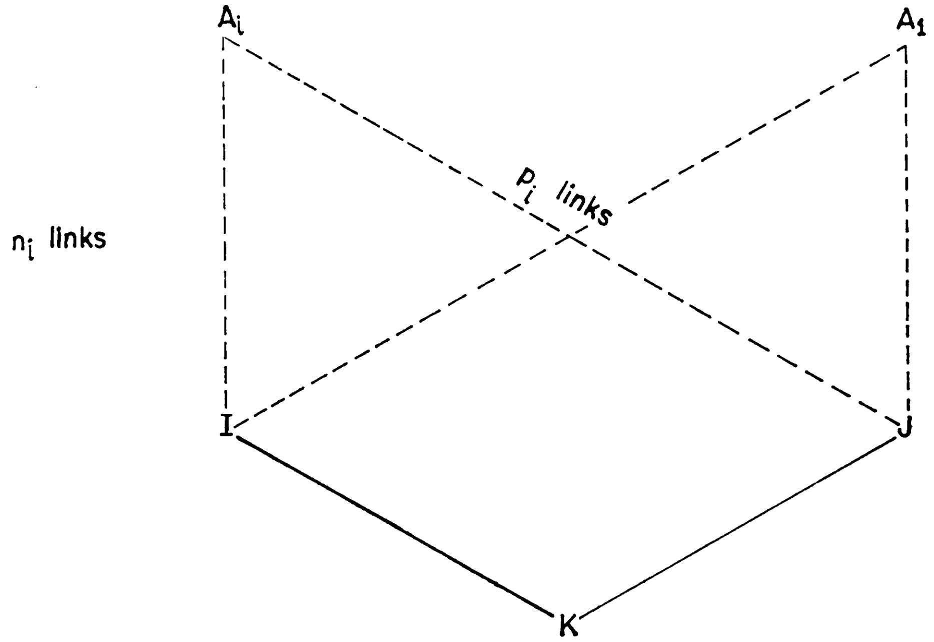

FIGURE 1

Example of identical loci due to common ancestors.

The probability of identity of his two homologous loci is named his *coefficient of inbreeding* f_K,

$$f_K = \sum_i \left(\frac{1}{2}\right)^{n_i+p_i} \frac{1}{2} (1 + f_{A_i}), \tag{1.1}$$

where A_i are unrelated ancestors. It is remembered that $1 - f_K$ is the probability of nonidentity; that is, the probability that the two homologous loci trace back, as far as we go back, to different ancestors. They are then "independent" in the probabilistic sense; this means that they may or may not possess the same gene (a for instance), but knowing one of the two genes *does not imply* anything about the other. We call q the probability that a random locus bears the gene a with $1 - q = p$ being the overall probability of the other alleles grouped under the symbol A. The independence of the two loci gives for the three genotypes aa, aA, AA, the probabilities q^2, $2pq$, and p^2; whereas the identity of the two loci gives the probabilities q, 0, p. So the *a priori* probabilities that a random individual K bears either of the three genotypes are

$$P = (1 - f_K)q^2 + f_K q, \qquad 2Q = 2(1 - f_K)pq, \qquad R = (1 - f_K)p^2 + f_K p. \tag{1.2}$$

The probabilities of the homozygotes are, in case of inbreeding, greater than in cases of random mating. The probabilities (and then, in general, the frequencies) of heterozygotes, are decreased by random mating. The variance of a quantitative character increases by inbreeding when the gene effects are additive. Now we may define more generally the *coefficient of kinship* of the two individuals I and J. We shall define it in such a manner as to obtain, if they are parents of an individual K, the coefficient of inbreeding of K. The coefficient of kinship of I and J is the probability that one locus chosen at random among the two loci of I, and one locus chosen at random among the two homologous loci of J, are identical. So, the coefficient of *inbreeding* of K is *equal* to the coefficient of *kinship* of his parents I and J

$$f_K = f_{IJ}. \tag{1.3}$$

If we account for *mutation*, which can modify a locus transmitted from parent to child with probability u, we have

$$f_K = (1 - u)^2 f_{IJ} \tag{1.4}$$

and

$$f_{ij} = \sum_i \left[\frac{1}{2}(1 - u)\right]^{n_i + p_i} \frac{1}{2}(1 + f_{A_i}). \tag{1.5}$$

2. Coefficient of kinship

In a natural population, the probability of two individuals I and J bearing, in randomly chosen homologous chromosomes, two identical loci decreases when their distance increases because the probability of common ancestors decreases. It is the phenomenon of isolation by distance, going up to racial unlikeness if the distance is so large that there are very few identical loci between I and J. It is possible to calculate, as a function of the distance between I and J, their coefficient of kinship; that is, the probability for two randomly chosen loci in each to be identical. This function is dependent on the *migration law;* that is, the probability law ruling the distance between the birthplace of the child and the birthplace of his parents, called the *parental distance,* whereas the distance between the birthplace of the two *parents* of each individual will be called *marital distance*. We shall call x the parental distance from child to parent. This x is an oriented vector in the two dimensional plane, but the arrows will be omitted for typographical simplicity. The coordinates of x will be called x_1 and x_2; points of the plane, locating positions of individuals will be called b, and so forth, with coordinates b_1, b_2, and so forth. The elementary area of location of an individual will be called $db_1\, db_2 = db$ or if this is the birthplace of parent, $dx_1\, dx_2 = dx$. The migration law, the probability law of parental distance, will be characterized by the probability density $\ell(x_1, x_2)$, the elementary probability $\ell(x_1, x_2)\, dx_1\, dx_2 = \ell(x)\, dx$ for shortness, or the distribution $L(x_1, x_2)$, with $dL = \ell\, dx$ in the continuous case.

With the assumption of separate generations, let us consider two individuals I and J, belonging to the same generation F_n, and separated by vectorial distance y, with coordinates y_1 and y_2. Their coefficient of kinship will be called $\varphi_n(y)$; it is obviously related to the coefficient of kinship of previous generations. If we were allowed to suppose that the parents of I and J are *always* distinct, we would have

$$\varphi_n(y) = (1 - u)^2 \int \varphi_{n-1}(y + z - x)\ell(z)\ell(x)\, dz\, dx, \tag{2.1}$$

where u is the mutation rate.

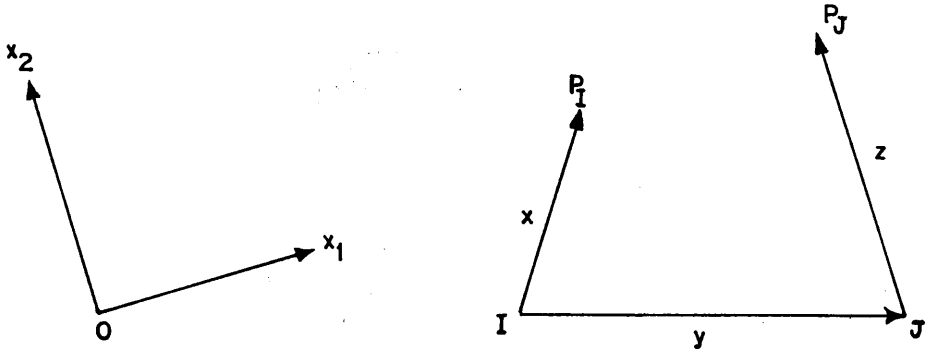

FIGURE 2

Distance between two individuals I and J, and the parents P_I and P_J supplying the two loci.

But formula (2.1) does not account for the possibility that P_I and P_J, the parents who supplied the two loci chosen in I and J, are the same individual, which is possible since the adults are in limited number. Indeed, P_I and P_J may be the same if they are born in the same elementary area dx, occupied by $\delta\, dx$ individuals, for which the corresponding probability is

$$\ell(x)\, dx\, \ell(-y + x)\, dx, \tag{2.2}$$

and if also they are the same, the probability of which is $1/\delta\, dx$. So, the probability of P_I and P_J being the same individual located at distance x from I is

$$\ell(x)\ell(x-y)\,dx/\delta. \tag{2.3}$$

REMARK 1. The reasoning may also be used in the discontinuous case with partition into separate groups of N individuals each, the probability of migration from a group at vectorial distance d being $\ell(d)$, with d having discrete, perhaps entire, coordinates. The probability that the parents come from the same group is then $\ell(d)\ell(d-y)$, and the conditional probability that they are the same is 1 divided by N. The probability of being the same individual located at distance d from I is then

$$\ell(d)\ell(d-y)/N. \tag{2.4}$$

When we tend to the case of a continuous distribution, we obtain formula (2.3).

REMARK 2. In the previous reasoning, sex was not yet taken into account. If we now suppose the densities of males and females are δ_1 and δ_2 per unit area, the parents P_I and P_J who have the randomly chosen loci of I and J may both be males (with probability 1/4) and identical with conditional probability of location

$$\ell(x)\ell(x-y)\,dx/\delta_1, \tag{2.5}$$

or they may both be females with probability 1/4 and identical with conditional probability of location

$$\ell(x)\ell(x-y)\,dx/\delta_2. \tag{2.6}$$

So formula (2.3) is valid if we put $1/4\delta_1 + 1/4\delta_2 = 1/\delta$. If δ_1 and δ_2 are equal, then $\delta = 2\delta_1$, and is then total density; but in the general case δ is only double the harmonic mean between δ_1 and δ_2.

2.1. *Refinement of equation* (2.1) *to account for finiteness of population.* We have to replace some infinitesimal terms of (2.1) corresponding to the same location of parents; that is, to the relation $y + z = x$ and $dz = dx$. In cases of probability $1/\delta\, dx$, the coefficient of kinship $\varphi_{n-1}(0)$ of neighboring but distinct individuals is to be replaced by the probability that two loci given to I and J by the same parent P_I were the same. This probability is $1/2 + f_0/2$, where f_0 is the coefficient of inbreeding of P_I. So we have to add to formula (2.1) the corrective term

$$(1-u)^2 \int \left[\frac{1}{2}(1+f_0) - \varphi_{n-1}(0)\right] [\ell(x)\ell(x-y)\,dx]/\delta. \tag{2.7}$$

We shall now put

$$\left[\frac{1}{2}(1+f_0) - \varphi_{n-1}(0)\right] \Big/ \delta = s_{n-1} > 0 \tag{2.8}$$

as a correction for the same parent. We have now to calculate the function $\varphi_n(x)$ and the constant s_{n-1} giving the coefficient of inbreeding f_0, by solving equations of which one is the finite difference integral equation

$$\begin{aligned}\varphi_n(y) = {} & (1-u)^2 \int \varphi_{n-1}(y+z-x)\ell(z)\ell(x)\,dz\,dx \\ & + (1-u)^2 s_{n-1} \int \ell(x)\ell(x-y)\,dx.\end{aligned} \tag{2.9}$$

It is easy to see that, because the factor $(1-u)^2$ is less than one, the function $\varphi_n(y)$ tends to a limit $\varphi(y)$ as n tends to infinity. If we call s the limit of s_{n-1}; that is, if we put

$$\frac{1+f_0-2\varphi(0)}{2\delta}=s \tag{2.10}$$

and if we put, for simplicity,

$$(1-u)\ell(x)=g(x) \tag{2.11}$$

giving us

$$\int g(x)\,dx=(1-u) \tag{2.12}$$

(instead of $\int \ell(x)\,dx=1$), we now have to solve

$$\varphi(y)=\int \varphi(y+z-x)g(z)g(x)\,dz\,dx+s\int g(x)g(x-y)\,dx. \tag{2.13}$$

The convolutions in (2.13) will be replaced by algebraic multiplications if we introduce the following bidimensional Fourier transforms (where v is an arbitrary vector of coordinates v_1 and v_2):

$$\begin{aligned} G(v) &= \int e^{ivx}g(x)\,dx && \text{(vectorial form)} \\ &= \int e^{i(v_1x_1+v_2x_2)}g(x_1,x_2)\,dx_1\,dx_2, && \text{(scalar form)} \end{aligned} \tag{2.14}$$

$$K(v)=\int e^{ivy}\varphi(y)\,dy=\int e^{i(v_1y_1+v_2y_2)}\varphi(y_1,y_2)\,dy_1\,dy_2. \tag{2.15}$$

By multiplying each member of (2.13) by

$$e^{ivy}=e^{iv(y+z-x)}e^{-ivz}e^{ivx}=e^{ivx}e^{-iv(x-y)} \tag{2.16}$$

and integrating with respect to y, we obtain the formula

$$K(v)=K(v)G(-v)G(v)+sG(v)G(-v) \tag{2.17}$$

which gives

$$K(v)=s\frac{G(v)G(-v)}{1-G(v)G(-v)}. \tag{2.18}$$

Then the inversion of Fourier transform (2.15) gives

$$\begin{aligned} \varphi(y) &= \frac{1}{4\pi^2}\int e^{-ivy}K(v)\,dv && \text{(vectorial form),} \\ &= \frac{1}{4\pi^2}\int e^{-i(v_1y_1+v_2y_2)}K(v_1,v_2)\,dv_1\,dv_2 && \text{(scalar form).} \end{aligned} \tag{2.19}$$

These formulas may be simplified in the *isotropic case* where the migration law, defined by $\ell(x_1,x_2)$, or by $g(x_1,x_2)=(1-u)\ell(x_1,x_2)$, is only dependent on the absolute value, or "norm" of vector x; that is, $(x_1^2+x_2^2)^{1/2}=r$.

We then have, putting $x_1=r\cos\theta$ and $x_2=r\sin\theta$,

$$\ell(x_1,x_2)\,dx_1\,dx_2=\ell(r)r\,dr\,d\theta. \tag{2.20}$$

$G(v)$ given by (2.14) is then dependent only on the norm $|v| = (v_1^2 + v_2^2)^{1/2}$, and is then known to be the "Hankel transform"

$$G(|v|) = 2\pi(1 - u) \int_0^{+\infty} J_0(r|v|)r\ell(r)\, dr. \tag{2.21}$$

We use the notation $G(|v|)$ to indicate a function depending only on the norm $|v|$ of vector v: and we shall henceforth omit the sign $|\ |$ and write only $G(v)$.

The same can be said about equation (2.15), which gives the Hankel transform of the coefficient of kinship $\varphi(y_1, y_2)$. But it is better to use the inverse transform (2.19) thus putting $(y_1^2 + y_2^2)^{1/2} = a$, $\varphi(y) = \varphi(a)$, and $K(v_1, v_2) = K(|v|) = K(v)$ the value of $K(v)$ being given by (2.18).

Then we may write

$$\varphi(a) = \frac{1}{2\pi} \int_0^{+\infty} vJ_0(a|v|)K(v)\, dv. \tag{2.22}$$

Combining formulas (2.18), (2.21) and (2.22), we have

$$G(v) = (1 - u) \int_0^{+\infty} J_0(rv)2\pi r\ell(r)\, dr \tag{2.23}$$

$$\varphi(a) = \frac{s}{2\pi} \int_0^{+\infty} vJ_0(av) \frac{G^2(v)}{1 - G^2(v)}\, dv, \tag{2.24}$$

formulas which express the coefficient of kinship of I and J as a function of their distance a, in the stationary case, provided we know the migration law defined by the probability density $2\pi r\ell(r)$ of the *parental* distance r (that is, the distance between parent and child).

REMARK. Another interpretation is that, in the present isotropic case $G^2(|v|) = G(\overrightarrow{v})G(-\overrightarrow{v})$ is, apart from the factor $(1 - u)^2$, the Fourier transform of the difference of two independent vectors $\overrightarrow{x}$ and $\overrightarrow{z}$, each having as probability law the migration law; so, *in the case of complete random mating*, $G^2(v)$ is the

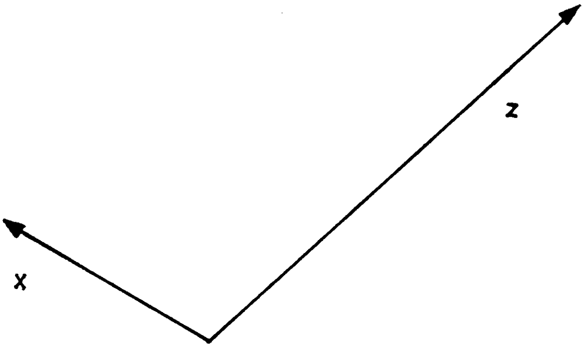

FIGURE 3

Interpretation as the difference of two independent vectors.

bidimensional Fourier transform of the vectorial *marital* distance, or the Hankel transform of the absolute *marital* distance; then the distribution of *marital* distance may be introduced directly in (2.24). We put

$$M(v) = (1 - u)^2 \int_0^{+\infty} J_0(rv)m(r)\, dr, \tag{2.25}$$

where $m(r)$ is the probability density of the marital distance, and $M(v)$ is, when v increases from zero, a decreasing function with the initial value $M(0) = (1 - u)^2$. Then we have

$$\varphi(a) = \frac{s}{2\pi} \int_0^{+\infty} \frac{M(v)}{1 - M(v)} J_0(av) v\, dv. \tag{2.26}$$

3. Application to a normal migration law

If, for the marital distance, $m(r)$ is normal (the distribution of the parental distance is then itself normal, in case of complete random mating), we have $M(v) = (1 - u)^2 \exp(-\sigma^2 v^2/4)$, where σ^2 is the second moment of the marital distance and

$$\varphi(a) = \frac{s}{2\pi} \int_0^{+\infty} \frac{(1 - u)^2 \exp(-\sigma^2 v^2/4)}{1 - (1 - u)^2 \exp(-\sigma^2 v^2/4)} J_0(av) v\, dv, \tag{3.1}$$

which gives by way of an easy integration

$$\varphi(0) = -\frac{s}{\pi\sigma^2} \log[1 - (1 - u)^2] \sim \frac{s}{\pi\sigma^2} \log\left(\frac{1}{2u}\right). \tag{3.2}$$

Recall that s is related to $\varphi(0)$; see the last paragraphs. In case of "complete random mating" where $f_0 = \varphi(0)$, we get ([1], p. 59)

$$\varphi(0) = f_0 = \left[1 + 2\pi\sigma^2\delta/\log\left(\frac{1}{2u}\right)\right]^{-1}. \tag{3.3}$$

To obtain an asymptotic expression of $\varphi(a)$ when a is large, we shall use, taking account of the smallness of the mutation rate, the asymptotic properties of the Fourier-Hankel transform. From (2.26) we see that $\varphi(a)$ is (apart from the factor $s/2\pi$) the transform of $M(v)/[1 - M(v)]$ which, in the vicinity of $v = 0$, is very large, equivalent to $1/2u$; and then decreases quickly from $1/2u$ to zero (when v increases); the principal part of $\varphi(a)$ for a large is given by the first derivatives (for $v = 0$) of the function $M(v)/[1 - M(v)]$, which are also the first derivatives of

$$\frac{(1 - u)^2(1 - \sigma^2 v^2/4)}{1 - (1 - u)^2(1 - \sigma^2 v^2/4)} \sim \frac{1}{2u + \sigma^2 v^2/4}. \tag{3.4}$$

The Hankel transform of order zero of $(m^2 + v^2)^{-1}$ is $K_0(am)$ ([2], p. 282), and thus we obtain

$$\varphi(a) \sim \frac{4s}{2\pi\sigma^2} K_0\left[\frac{a}{\sigma}(8u)^{1/2}\right] \tag{3.5}$$

whence, using a well known equivalence

$$\varphi(a) \sim \frac{2s}{\pi\sigma^2}\left[\frac{\pi\sigma}{2a(8u)^{1/2}}\right]^{1/2} \exp\left[-\frac{a}{\sigma}(8u)^{1/2}\right]. \tag{3.6}$$

This asymptotic form was given by me in 1959 ([3], p. 191) with the notation σ^2 instead of $\sigma^2/4$ (variance of parental distance along an axis, instead of variance of marital distance in the plane) using another method of approximation from ([4], p. 210) replacing the difference equation of p. 188, extended to two dimensions, by the partial differential equation

$$2u\varphi(x, y) - 2m\left(\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2}\right) = \frac{1 - \varphi(0)}{2N} \text{ times a Dirac distribution.} \tag{3.7}$$

The isotropic solution $\varphi(x, y) = \varphi(a)$ of the homogeneous equation is given by

$$\frac{\partial^2\varphi}{\partial a^2} + \frac{1}{a}\frac{\partial\varphi}{\partial a} - \frac{u}{m}\varphi = 0. \tag{3.8}$$

The general solution ([5], p. 114) is a linear combination of $I_0[a(u/m)^{1/2}]$ and $K_0[a(u/m)^{1/2}]$; but φ must be bounded, and is then proportional to $K_0[a(u/m)^{1/2}]$; that is, to

$$\left[\frac{\pi}{2a(u/m)^{1/2}}\right]^{1/2} e^{-a(u/m)^{1/2}}. \tag{3.9}$$

The variance of parental distance along each axis is $\sigma^2 = 2m$ in [3], in which I gave

$$\varphi(x) = K\frac{1}{x^{1/2}} e^{-(2u)^{1/2}x/\sigma} \tag{3.10}$$

on p. 191. Weiss and Kimura ([6], p. 142) have shown that the same formula numbered (4.40) is valid in the two dimensional discontinuous case, thus correcting my first approximation of this case ([7], p. 55). In this 1950 paper, I had already extended the formula for the discontinuous unidimensional case to a general migration law, with rate ma_j between places separated by distance j. Supposing m small and u small with respect to m, I obtained (pp. 51–52)

$$\varphi(|j - i|) = f_{ij} \sim K\alpha_1^{|i-j|} \sim K \exp\left[-\left(\frac{2u}{2m\sum\limits_n n^2a_n}\right)^{1/2} |i - j|\right], \tag{3.11}$$

which is still the same formula (putting $\sum_n 2mn^2a_n = \sigma^2$) as in the continuous unidimensional case ([8], p. 52). The analogous two dimensional comparison may be made between my "continuous normal case" bidimensional result dealt with in ([1], p. 59), recalled in ([8], p. 53, formula 5), and completed in the present paper and the discontinuous bidimensional case of Weiss and Kimura ([6], p. 139 and p. 146).

The asymptotic reasoning in the normal bidimensional case may be extended to every law giving a quickly decreasing Hankel transform $M(v)$ of the form

$$M(v) = (1 - u)^2 - kv^2 + o(v^2) \tag{3.12}$$

(implying the existence of a second moment, equal to $4k$, of the marital distribution). This will be verified, in the next paragraph, for the interesting case of a "K distribution." A further step would be to investigate the case of a Beta distribution without second moment.

4. Application to a "*K* distribution"

By "K distribution" we shall mean the distribution of elementary probability, where K_β is the modified Bessel function of the second kind,

$$\frac{1}{2^\beta \Gamma(\beta + 1)} K_\beta(rh)(hr)^{\beta+1}\, d(hr), \qquad \beta > -1. \tag{4.1}$$

Its Hankel transform (or characteristic function) is

$$\left(\frac{h^2}{h^2 + v^2}\right)^{\beta+1}. \tag{4.2}$$

When $\beta = -1/2$, we have the exponential distribution, the limiting case between a "continuous" distribution (when $\beta > -1/2$), and a distribution infinite at $x = 0$ (when $\beta < -1/2$).

The usefulness of the K distribution as a two dimensional isotropic distribution is that every convolution of K distributions with the same parameters h is itself a K distribution, since the powers of the characteristic functions are of the same form and the parameters $\beta + 1$ add up. (This is the same property in the two dimensional case as that of Gamma distributions in the one dimensional case.) So all convolution formulas will be easy to deal with; in particular, it is equivalent to suppose a K distribution for the parental distance or for the marital distance (in the case of complete random mating). Let us suppose that the marital distance obeys the K law with parameter $\beta + 1 = b$ for the sake of simplicity. Now b is >0 and we have

$$M(v) = G^2(v) = (1 - u)^2(1 + v^2/h^2)^{-b} \tag{4.3}$$

and then, from (2.24) (with the approximation $(1 - u)^2 - 1 \sim -2u$),

$$\varphi(a) = \frac{s}{2\pi} \int_0^{+\infty} \frac{J_0(av)v\, dv}{(1 + v^2/h^2)^b - 1 + 2u}. \tag{4.4}$$

This integral is uniformly convergent if b is assumed >1, where a is $\geqq 0$.

If b is an integer, let us call w_k the b distinct roots of the equation $w^b = 1 - 2u$, given by the formula $w_k = (1 - 2u)^{1/b} \exp\,(2i\pi k/b)$; we then may write

$$\frac{1}{(1 + v^2/h^2)^b - w^b} = \sum_{k=1}^{b} \frac{A_k}{1 + v^2/h^2 - w_k}, \tag{4.5}$$

with $A_k = w_k/(1 - 2u)b$, so that

$$\varphi(a) = \frac{sh^2}{2\pi} \sum_{k=1}^{b} \frac{w_k}{(1 - 2u)b} \int_0^{+\infty} \frac{J_0(av)v\, dv}{h^2(1 - w_k) + v^2}. \tag{4.6}$$

Each integral is easily calculated by the formula

$$\int_0^{+\infty}\frac{J_0(av)v\,dv}{m^2+v^2}=K_0(am); \tag{4.7}$$

see Guelfand and Chilov ([2], p. 282) which gives

$$\varphi(a)=\frac{sh^2}{2\pi}\sum_{k=1}^{b}\frac{w_k}{(1-2u)b}K_0[ah(1-w_k)^{1/2}]. \tag{4.8}$$

The expression $(1-w)^{1/2}$ may be chosen with an argument between $-\pi/4$ and $+\pi/4$.

If a is large, this sum $\sum_{k=1}^{b}$ contains a term much larger than the others; indeed, the root $w_b=(1-2u)^{1/b}\sim 1-2u/b$ is (if b is not large), much nearer 1 than the others, making the corresponding value of $|K_0|$ much larger than the others (the real part and the modulus of $ah(1-w_b)^{1/2}$ being much less than those of $ah(1-w_k)^{1/2}$ if $k\neq b$). Therefore,

$$\varphi(a)\sim\frac{sh^2}{2\pi b}K_0[ah(2u/b)^{1/2}]=\frac{2s}{\pi M_2}K_0[a(8u/M_2)^{1/2}]. \tag{4.9}$$

(The second moment M_2 is $4b/h^2$.)

The analyticity of (2.24) with respect to $b>1$ ensures validity even when $b>1$ is not an integer. Then we may apply the approximation

$$\begin{aligned}\varphi(a)&\sim\frac{sh^2}{2\pi b}\left[\frac{\pi}{2ah(2u/b)^{1/2}}\right]^{1/2}e^{-ah(2u/b)^{1/2}}\\&=\frac{2s}{\pi M_2}\left[\frac{\pi M_2^{1/2}}{2a(8u)^{1/2}}\right]^{1/2}e^{-a(8u/M_2)^{1/2}}\end{aligned} \tag{4.10}$$

which is the same as in the case of the normal law.

For $a=0$, formula (2.24) gives directly

$$\begin{aligned}\varphi(0)&=\frac{sh^2}{4\pi}\int_0^{+\infty}\frac{d(1+v^2/h^2)}{(1+v^2/h^2)^b-1+2u}\\&=\frac{sh^2}{4}\int_1^{+\infty}\frac{dz}{z^b-1+2u}\end{aligned} \tag{4.11}$$

which is convergent only if $b>1$. In the vicinity of $z=1$, we now write

$$\begin{gathered}z^b=1+w,\qquad bz^{b-1}\,dz=dw,\\ \varphi(0)=\frac{sh^2}{4\pi b}\int_0^{+\infty}\frac{dw}{(2u+w)(1+w)^{(b-1)/b}}.\end{gathered} \tag{4.12}$$

Because u is small this incomplete B integral is equivalent to

$$\int_0^1\frac{dw}{2u+w}\sim-\log(2u)=\log\frac{1}{2u}. \tag{4.13}$$

When $\varphi(0)=f_0$, which is the case of completely random mating, we have, by replacing s with $[1-\varphi(0)]/2\delta$, and $4b/h^2$ with the second moment M_2, the relation

$$\varphi(0) = f_0 = \left[1 + 8\pi b\delta \Big/ \left(h^2 \log \frac{1}{2u}\right)\right]^{-1} = \left[1 + 2\pi M_2\, \delta \Big/ \left(\log \frac{1}{2u}\right)\right]^{-1}. \tag{4.14}$$

Again, the formula is the same as for the normal law. (See [1], p. 59.)

5. Calculation of f_0

5.1. *Complete random mating.* If there is *complete* random mating within the local gametic pool (not excluding the possibility of the two uniting gametes coming from the same individual; that is, allowing self-fertilization), there is then equality between f_0, which is the probability of identity of the two loci of the same individual, and $\varphi(0)$, which is the probability of identity of two random loci of two individuals born at distance 0. Then s is equal to $[1 - \varphi(0)]/2\delta$, as was supposed in my previous works ([1], p. 57, compare with previous formula; [4], p. 209).

A more general case, leading to nearly the same results, will be dealt with hereafter.

5.2. *Exclusion of some matings of high consanguinity.* In human genetics, the inbreeding coefficient f_0 of an individual K is equal to the coefficient of kinship of his two parents I and J, who are *not* independently chosen random individuals; given their vectorial distance y (at their birth), they are chosen so as to exclude the nearest type of matings. Thus, their coefficient of kinship is not $\varphi(y)$ (valid only for random individuals of distance y), but a new function $\psi(y)$, which we shall calculate by assuming that the first possible common ancestor of the mates I and J is of order p ($p = 0$ corresponding to allowance of self-fertilization, and $p = 1$ to allowance of brother-sister mating); the probability distribution of location of each ancestor of order p will be called $\ell^{(p)}$, defined by the convolution

$$\ell^{(p)}(x) = \int \ell^{(p-1)}(x')\ell(x - x')\, dx', \tag{5.1}$$

with $\ell^{(1)} = \ell$. Putting $g^{(p)}(x) = (1 - u)^p \ell^{(p)} x$, we have

$$\int e^{ivx} g^{(p)}(x)\, dx = G^p(v). \tag{5.2}$$

If we suppose that the two ancestors, A_I and A_J, of order p who may have given the chosen loci of I and J have been randomly chosen (we will call this "p removed random mating"), their joint probability law of location in places z and x is $\ell^{(p)}(z)\ell^{(p)}(x)\, dz\, dx$; their probability of being the same individual (in place x) is

$$[\ell^{(p)}(x)\ell^{(p)}(x - y)/\delta]\, dx. \tag{5.3}$$

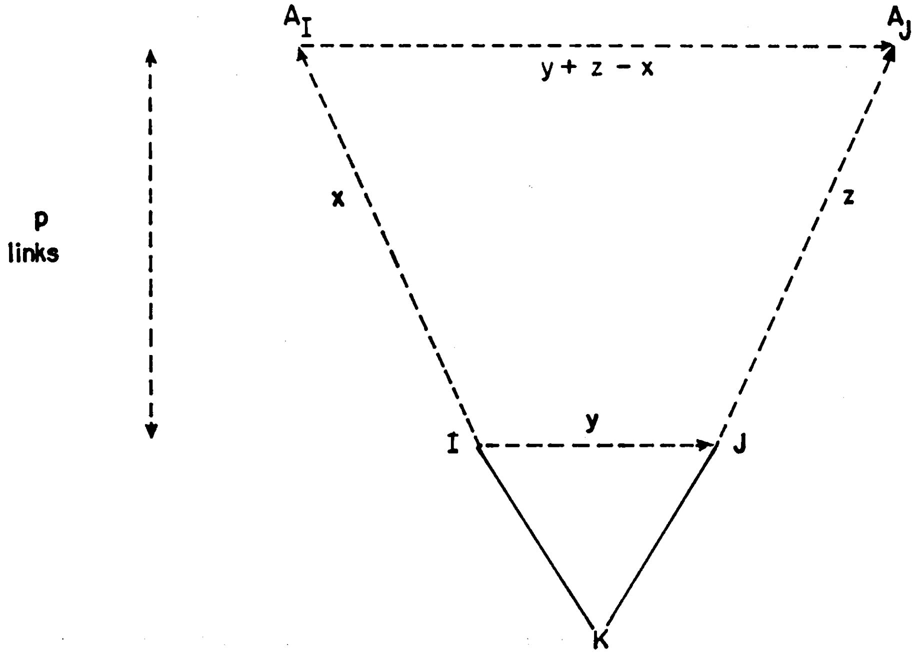

FIGURE 4

Determination of joint probability law of location of two ancestors, and of the probability of their being the same individual.

So $\psi(y)$ is given by

$$\psi(y) = \int \varphi(y + z - x) g^{(p)}(z) g^{(p)}(x)\, dz\, dx + s \int g^{(p)}(x) g^{(p)}(x - y)\, dx \tag{5.4}$$

with, as before,

$$s = \left[\frac{1 + f_0}{2} - \varphi(0)\right] \Big/ \delta. \tag{5.5}$$

Compare (5.4) to (2.13).

The Fourier transform gives

$$\begin{aligned} H(v) &= \iint e^{ivy}\psi(y)\, dy = K(v)[G(v)]^{2p} + s[G(v)]^{2p} \\ &= \left[\frac{sG^2(v)}{1 - G^2(v)} + s\right] G^{2p}(v) = s\,\frac{G^{2p}}{1 - G^2}, \end{aligned} \tag{5.6}$$

using (2.18).

Hence,

$$\psi(y) = \frac{1}{4\pi^2} \iint e^{-ivy} H(v)\, dv = \frac{1}{2\pi} \int_0^{+\infty} J_0(|y||v|) H(|v|) |v|\, d|v|. \tag{5.7}$$

It is now possible to calculate f_0. If we suppose we know the bidimensional probability density $\overline{m}(y)$ of "vectorial marital distance" between I and J and the corresponding transform

$$(1 - u)^2 \int e^{ivy}\overline{m}(y)\,dy = M(v) \tag{5.8}$$

($M(v)$ is now not equal to $G^2(v)$, owing to the fact that the two mates I and J are not independently chosen), we have

$$f_0 = \int (1 - u)^2\overline{m}(y)\psi(y)\,dy \tag{5.9}$$

which may be calculated by noting that $\psi(y) = \psi(-y)$, and introducing (with arbitrary z) the convolution

$$(1 - u)^2 \int \overline{m}(y)\psi(z - y)\,dy = \theta(z), \tag{5.10}$$

of which the Fourier transform is

$$\int e^{ivz}\theta(z)\,dz = M(v)H(v). \tag{5.11}$$

The value f_0, being equal to $\theta(0)$, requires the reciprocal formula

$$\theta(z) = \frac{1}{4\pi^2} \int e^{-ivz}M(v)H(v)\,dv, \tag{5.12}$$

whence,

$$f_0 = \frac{1}{4\pi^2} \int M(v)H(v)\,dv \qquad \text{(two dimensional integral).} \tag{5.13}$$

In the same manner, formulas (2.18) and (2.19) give

$$\varphi(0) = \frac{1}{4\pi^2} \int K(v)\,dv = \frac{s}{4\pi^2} \int \frac{G^2(v)}{1 - G^2(v)}\,dv \quad \text{(two dimensional integral).} \tag{5.14}$$

Comparing (5.13) (where $H(v)$ is given by (5.6)) with (5.14) and with the definition of s (2.10), we are now able to calculate the three constants s, $\varphi(0)$, and f_0.

$$2\delta s = 1 + \frac{s}{4\pi^2} \int M(v) \frac{G^{2p}(v)}{1 - G^2(v)}\,dv - \frac{2s}{4\pi^2} \int \frac{G^2(v)}{1 - G^2(v)}\,dv, \tag{5.15}$$

or in polar coordinates, with unidimensional integrals

$$2\delta s = 1 + \frac{s}{2\pi} \int_0^{+\infty} M(|v|) \frac{G^{2p}(|v|)}{1 - G^2(|v|)}\,|v|\,d|v| - \frac{2s}{2\pi} \int_0^{+\infty} \frac{G^2(|v|)}{1 - G^2(|v|)}\,|v||dv|. \tag{5.16}$$

This equation, where the sign $||$ will now be omitted, gives the numerical value of s, from which using (2.24), (5.6) and (5.7) and putting $|y| = a$ gives

$$\varphi(a) = \frac{s}{2\pi} \int_0^{+\infty} J_0(av) \frac{G^2(v)}{1 - G^2(v)}\,v\,dv \tag{5.17}$$

and

$$\psi(a) = \frac{s}{2\pi} \int_0^{+\infty} J_0(av) \frac{G^{2p}(v)}{1 - G^2(v)}\,v\,dv. \tag{5.18}$$

5.3. *Numerical results.* We have

$$s = \left[2\delta + \frac{1}{2\pi}\int_0^{+\infty} \frac{2G^2(v) - M(v)G^{2p}(v)}{1 - G^2(v)}\, v\, dv\right]^{-1}, \tag{5.19}$$

$$\varphi(0) = \frac{s}{2\pi}\int_0^{+\infty} \frac{G^2}{1 - G^2}\, v\, dv, \tag{5.20}$$

$$\psi(0) = \frac{s}{2\pi}\int_0^{+\infty} \frac{G^{2p}}{1 - G^2}\, v\, dv, \tag{5.21}$$

and, by (5.13),

$$f_0 = \frac{s}{2\pi}\int_0^{+\infty} M(v)\, \frac{G^{2p}}{1 - G^2}\, v\, dv. \tag{5.22}$$

5.4. *Particular case: K distributions.* All the formulas are easy to work out if the parental distance and the marital distance are "K distributed," the Hankel transform for law of *parental* distance being

$$G(v) = (1 - u)\left(1 + \frac{v^2}{h^2}\right)^{-b/2}. \tag{5.23}$$

For simplicity let us put $G^2(v) = M(v)$. (Accounting for nonrandomness of mating could be done by taking for $m(r)$ a K distribution with a parameter different from b, but this correction has a small effect on formulas (5.19) and (5.22)). We then have

$$G^2 = (1 - u)^2\left(1 + \frac{v^2}{h^2}\right)^{-b}. \tag{5.24}$$

The calculation of the integral in (5.21) is easy and gives the integrals in (5.20) and (5.22) (with $M(v) = G^2(v)$). All integrals are convergent if b is supposed >1. Let us write

$$\begin{aligned} I_p &= \int_0^{+\infty} \frac{G^{2p}}{1 - G^2}\, v\, dv \\ &= (1 - u)^{2p}\int_0^{+\infty} \frac{(1 + v^2/h^2)^{-bp}}{1 - (1 - u)^2(1 + v^2/h^2)^{-b}}\, d\left(\frac{v^2}{2}\right), \end{aligned} \tag{5.25}$$

which, putting $z = 1 + v^2/h^2$, is

$$(1 - u)^{2p}\frac{h^2}{2}\int_1^{+\infty} \frac{z^{-bp}}{1 - (1 - u)^2 z^{-b}}\, dz, \tag{5.26}$$

or, again putting

$$z^b = 1 + wetbz^{b-1}\, dz = dw.$$

$$I_p = (1 - u)^{2p}\frac{h^2}{2b}\int_0^{+\infty} \frac{(1 + w)^{-p+1/b}}{1 + w - (1 - u)^2}\, dw. \tag{5.27}$$

The two equations (5.20) and (5.22) then give, jointly with the relation $2s\delta = 1 + f_0 - 2\varphi(0)$,

$$(5.28) \qquad 2s\delta = 1 + \frac{s}{2\pi}(I_{p+1} - 2I_1)$$

$$= 1 + (1-u)^{2p}\frac{sh^2}{2\pi b}\int_0^{+\infty}\frac{(1+w)^{-1-p+1/b} - 2(1+w)^{-1+1/b}}{1-(1-u)^2+w}\,dw.$$

The rapid increase of the denominator from the very small initial value $1-(1-u)^2 \sim 2u$ allows us to replace the integral by the approximation

$$(5.29) \qquad \int_0^1 \frac{-1}{2u+w}\,dw = -\int_0^1 \log(2u+w) \sim \log 2u.$$

Replacing $(1-u)^{2p}$ by 1 (since p is not very large) and $4b/h^2$ by M_2, we obtain

$$(5.30) \qquad 2s\delta - 1 \sim \frac{s}{\pi M_2}\log 2u, \quad s \sim \left[2\delta + \frac{\log(1/2u)}{\pi M_2}\right]^{-1},$$

which is a particular case of formula (5.19). Then

$$(5.31) \qquad f_0 = \frac{2s}{4\pi} I_{p+1} \quad \text{and} \quad \varphi(0) = \frac{2s}{4\pi} I_1;$$

or I_{p+1} and I_1 are, by the same reasoning, equivalent to

$$(5.32) \qquad \frac{h^2}{2b}\int_0^1 \frac{dw}{2u+w} \sim \frac{2}{M_2}\log\left(\frac{1}{2u}\right).$$

We then have

$$(5.33) \qquad f_0 \sim \varphi(0) \sim \frac{s}{\pi M_2}\log\left(\frac{1}{2u}\right) \sim \left[1 + 2\pi M_2\delta/\log\left(\frac{1}{2u}\right)\right]^{-1}.$$

This differs very little from the case of complete random mating.

In the case of a normal law, the result would be the same (and also in the case of other laws of finite second moment M_2). It remains to investigate the case of migration laws without a second moment.

I am grateful to the Berkeley Symposium and to the Universities of Hawaii, Stanford, and Michigan State for the help they offered to me by their kind invitation to discuss this paper in their lectures or seminars.

REFERENCES

[1] G. Malécot, *Les Mathématiques de l'Hérédité*, Vol. 1, Paris, Masson, 1948.

[2] I. M. Guelfand and G. E. Chilov, *Les Distributions*, Vol. 1, Paris, Dunod, 1962 (French translation of Russian *Generalized Functions* by I. M. Gel'fand and G. E. Shilov, Moscow, Fiz. Mat. Giz., 1959).

[3] G. Malécot, "Les modèles stochastiques en génétique de population," *Publ. Inst. Statist. Univ. Paris*, Vol. 8 (1959), pp. 173–210.

[4] ———, "Migration et parenté génétique moyenne," *Entretiens de Monaco en sciences humaines*, 1962, pp. 205–212.

[5] I. N. Sneddon, *Special Functions of Mathematical Physics and Chemistry*, Edinburgh, Oliver and Boyd; New York, Interscience, 1956.

[6] G. H. Weiss and M. Kimura, "A mathematical analysis of the stepping stone model of genetic correlation," *J. Appl. Probability*, Vol. 2 (1965), pp. 129–149.

[7] G. Malécot, "Quelques schémas probabilités sur la variabilité des populations naturelles" *Ann. Univ. Lyon, Sciences*, Section A, Vol. 13 (1950), pp. 37–60.

[8] ———, "The decrease of relationship with distance," *Cold Spring Harbor Symp. Quant. Biol.*, Vol. 20 (1955), pp. 52–53.