CHAPTER 4: THEORY OF STRUCTURAL NUMBERS

4.0 Introduction

In theoretical chemistry, we can roughly distinguish two broad areas of wor the theory of atomic structure which tells which are the possible elements and hc they may relate to one another; and structure determination, in which one studie exactly which structure(s) exists in a particular case. In some sense, this enti manuscript is about the first problem: which structures are possible, and what would be the properties of the various possible "pure elements"

In their "purest form" elements are often found in crystals with definite regular structure, and this chapter lays the foundations for the study of the "c stal structure" of mating systems. From this point of view, the last chapter la the ground work for description of any "compound" structure, whether it is a "si crystal", a "chain molecule", or any mixture. The following chapters begin the cussion of what are the statistical properties of systems composed of the "pure materials" but not necessarily in their crystalline state. Thus, in the present chapter, we have as results, generally nice simple diagrams, graphs, and numbers whose appearance is deceptively simple. In the following chapters, we learn how *these simple numbers, which are properties of marriage rules and not of populati can predict statistics which are properties of populations following the rules*.

This chapter is the actual "first" so far as theoretical developments are cerned: We develop the theory of structural numbers in detail. The reader may to see introductory sections of chapter 6.4 for a very short summary of the pres developments or for a list of further examples of the rules discussed here.

4.1 Definitions

DEFINITION 4.1: A *generation* *G* is a set of elements p, q, ... called *individual* A *configuration* *C* is a triple (*B*,*M*,*S*) such that:

- i) *B* is a partition of *G*;
- ii) *M* is a collection of two element subsets of *G*;
- iii) There are subsets of *G* labeled m and f such that $|m \cap f| = 0$ and $m \cup f = \underline{G}$ and *S* is the partition of *G* imposed by these labelings.
- iv) $|m \cap M| = 1$ and $|f \cap M| = 1$ for all $M \in \underline{M}$.

In anthropological applications, *G*; is interpreted as a population of peo of a particular "generation", whose parents (by Definition 4.2 below) are membe the previous generation and whose descendants are of the following. Thus, generations are "discrete". The partition $\underline{B}$ sorts individuals into sets of siblings who are descendants of the same parents. $\underline{M}$ is a monogamous "marriage" or "mating" relation between individuals of different sexes. Sequential marriages, etc. are not encompassed by these definitions.

DEFINITION 4.2: A descent relation d from a generation $\underline{G}$ with a configuration $\underline{C}$ to a generation $\underline{G}'$ with configuration $\underline{C}'$ is a map d: $\underline{B}' \rightarrow \underline{M}$, which is 1-1 into $\underline{M}$ and *such that the inverse map is 1-1, from a subset of* $\underline{M}$ *onto* $\underline{B}'$

An infinite sequence of generations $\underline{G}^{(n)}$ is a sequence . . . $\underline{G}^{(n)}, \underline{G}^{(n+1)}$, . . . of generations together with configurations . . . $\underline{G}^{(n)}, \underline{C}^{(n+1)}$, . . . and together with descent maps $d^{(n)}:\underline{B}^{(n+1)} \xrightarrow{} \underline{M}^{(a)}$.

A finite sequence of generations is a sequence $\underline{G}^{(n)}, \underline{G}^{(n+1)}, \ldots \underline{G}^{(n+k)}$ of generations together with configurations $\underline{C}^{(n)}, \underline{C}^{(n+1)}, \ldots, \underline{C}^{(n+k)}$ and together with descent maps $d^{(i)}:\underline{B}^{(i+1)} \xrightarrow{} \underline{M}^{(i)}$, $1 \leq i \leq k$.

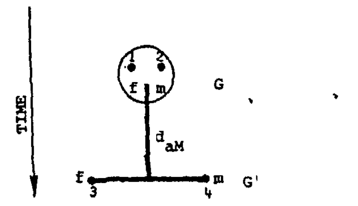
Notice that while a of $\underline{B}'$ is mapped to some distince element M of $\underline{M}$, there may be elements of $\underline{M}$ which have no members of $\underline{B}'$ mapped to them.

Notational Convenience:

"Let $\underline{G}^{(n)}$ be a generation in a sequence of generations" will be used to mean: "Let . . . , $\underline{G}^{(n)}, \underline{G}^{(n+1)}$, . . . be an (infinite) sequence of generations with configurations . . . ,$\underline{C}^{(n)}, \underline{C}^{(n+1)}$, . . . , and maps . . . , $d^{(n)}:\underline{B}^{(n+1)} \xrightarrow{} \underline{M}^{(n)}$, . . ., for all n and let $\underline{G}^{(n)}$ be a generation in this sequence."

Note that under certain circumstances, the existence of a finite sequence permits construction of an infinite sequence of generations. For this reason, "sequence of generations" will mean "infinite sequence of generations" unless otherwise stated. A theorem on cyclic sequence construction is presented later.

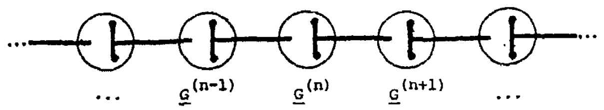
Notation is developed for displaying these pairs ($\underline{G}$,$\underline{G}'$) in a graphic manner. Use a dote "." to represent individuals $\epsilon\underline{G}$ and use a labeled dot ".m" or ".f" to show that the individual p represented by the dot is either pϵf or pϵm, respectively. A line between dots (labeled or unlabeled dots) will show equivalence mod $\underline{B}$ for individuals in $\underline{G}$, and mod $\underline{B}'$ for individuals in $\underline{G}'$. A circle around two dots shows equivalence mod $\underline{M}$ or mod $\underline{M}'$ respectively. Lines and circles may be labeled. On occasion, a line will be allowed to branch, to include more than two individuals mod $\underline{B}$, etc. A line with no dot at the end, from inside a circle to a line with dots, or to a single dot, shows a particular descent map element.



In this diagram, the pictorial representation depicts the following sets: $\underline{G} = \{1,2\}$; $\underline{B} = \{(1), (2)\}$; $\underline{G} = \{1,2\}$; $\underline{B} = \{(1),(2)\}$; $\underline{M} = \{N\} = \{\{1,2\}\}$; $\underline{S} = m\ f = \{\{2\}\ \{1\}\}$; $\underline{G}' = \{3,4\}$; $\underline{B}' = \{a\} = \{\{3,4\}\}$; $\underline{M}' = \emptyset$; $\underline{S}' = \{m \cup f\} = \{\{4\} \cup \{3\}\}$; $d = \{d_{aM}\}$. Thus, $\underline{G}'$ is a generation with configuration $\underline{C}' = \{\underline{M}',\underline{B}',\underline{S}'\}$. The elements of $\underline{M}'$, $\underline{B}'$ and $\underline{S}'$ are enumerated above. Individuals 1 ar 2 of $\underline{G}'$ are the ancestors of 3 and 4 in $\underline{G}$, which has configuration $C = \{\underline{M},\underline{B},\underline{S}\}$ whose elements are also enumerated above. Note d(a) = M.

In addition, an arrow with two shafts will be used between drawings which si marriage structure on a generation which previously had no indicated marriages, while a single shafted arrow will be used to show application of descent maps fro an existing generation to a "new" generation. (See theorems 4.4 and 4.5 for examples.)

Note that because of definitions 4.1 and 4.2, every generation in a sequenc of generations has at least the following minimal statistics: $|M^{(n)}| \geq 1$, $|\underline{B}^{(n)}| \geq 1$ $|\underline{G}^{(n)}| \geq 2$, $|m^{(n)}| \geq 1$, $|f^{(n)}| \geq 1$, for all n. If a particular sequence had only these statistics, it could be illustrated as follows:



DEFINITION 4.3: If $\underline{G}$ and $\underline{G}'$ are a pair of generations with configurations $\underline{C}$ and $\underline{C}'$ respectively, and $d:B' \rightarrow M$ is a descent map from $\underline{B}'$ to $\underline{M}$, and $d(a) = M$, and $p \epsilon a$ and $q \epsilon m$, then q is a parent of p, and p is an offspring of q.

Let $\underline{G}^{(n)}$ be a generation in a (finite or infinite) sequence of generations. Let $p \epsilon \underline{G}^{(n)}$ and suppose $\underline{G}^{(m)}$ is another generation in this sequence and $m<n$, and $q \epsilon \underline{G}^{(m)}$. A descent chain from q to p exists if there exists a sequence $q^{(m)}, q^{(m+1)}, \ldots, q^{(n)}$ such that $q^{(n)} \epsilon \underline{G}^{(n)}$ is an offspring of $q^{(i-1)} \epsilon \underline{G}^{(i-1)}$, $q^{(i-2)}$ is a parent of $q^{(i-1)}, \ldots, q^{(m)}$ is a parent of $q^{(m+1)}$, and $p = q^{(n)}$ and $q = q^{(m)}$.

If there exists a descent chain from q to p, then q is an ancestor of p, and p is a descendant of q.

Note that if q is a parent of p, then q is also an ancestor of p; if p is an offspring of q, then p is also a descendant of q. Note also that for a pair of generations $(\underline{G}, \underline{G}')$ with configurations $(\underline{C}, \underline{C}')$ and a descent map $d:\underline{B}' \rightarrow \underline{M}$, if $d(a) = M$, then there exist, for each $p \epsilon a$, two distinct elements of q M, each of which is an ancestor of p.

DEFINITION 4.4: Suppose $\underline{G}^{(n)}$ is a generation in a sequence of generations, and suppose $a \epsilon \underline{B}^{(n)}$ and $b \epsilon \underline{B}^{(n)}$ and $p \epsilon a$ and $q \epsilon b$. Let $\underline{G}^{(m)}$ be some other generation in the sequence such that $m < n$, and let $j = n-m$. Then p and q are j-removed if there exists $p' \epsilon \underline{G}^{(m)}$ such that both p and q are each descended from p'. (See section 2.1.)

If p and q are j-removed, write $R^j(p,q)$.

Note: $p \equiv q \bmod \underline{B}$ if $R^1(p,q)$.

Definition 4.4 allows discussion of "descent relations" that extend "backward" over several generations, while definition 4.5 below allows discussion of relations which extend "across" a single generation.

DEFINITION 4.5: Suppose $\underline{G}^{(n)}$ is a generation in a sequence of generations, and suppose $p \epsilon \underline{G}^{(n)}$ and $q \epsilon \underline{G}^{(n)}$ and $p \epsilon a \epsilon \underline{B}^{(n)}$ and $q \epsilon b \epsilon \underline{B}^{(n)}$ and $d(a) = M_a \epsilon \underline{M}^{(n-1)}$ and $d(b) = M_b \epsilon \underline{M}^{(n-1)}$.

If $M_a = M_b$ then $a = b$ and $p \equiv q \bmod \underline{B}^{(n)}$. Call this condition "p and q are 0-linked", written $L^0(p,q)$.

If $M_a \neq M_b$ then $a \neq b$. Suppose $p' \epsilon M_a$ and $q' \epsilon M_b$ are ancestors of p and q respectively, and $q' \equiv p' \bmod \underline{B}^{(n)}$. Then p and q are 1-linked, or first linked, written $L^1(p,q)$. If $p' \not\equiv q' \bmod \underline{B}^{(n)}$ but there exist $r' \equiv p' \bmod \underline{B}^{(n)}$ and $t' \equiv q' \bmod \underline{B}^{(n)}$ and $r' \equiv t' \bmod \underline{M}^{(n)}$, then p and q are 2-linked.

Similarly, if p and q are not i-linked for all $i < j$, but a series of $j + 1$ distinct B-equivalence classes linked by j distinct M-equivalence classes, then p and q are j-linked, written $L^j(p,q)$.

The particular sequence of B-sets linked through M-sets of $\underline{G}^{(n-1)}$ from p' to q', inclusive, forms a chain of $j + 1$ B-sets, and a chain of j M-sets.

Remark: The following is easily proved:

Theorem: (Simply extended removal)

Suppose $\underline{G}^{(n)}$ is a generation in a sequence of generations, and assume p and $p \epsilon \underline{G}^{(n)}$ and $R^j(p,q)$. Then $R^{j+1}(p,q)$ for any positive integer n.

Motivation for the next definition should be apparent, except to note that the distinction of several types of isomorphism is necessitated by the structure imposed by definition 4.1.

DEFINITION 4.6: Suppose $\underline{G}$ and $\underline{G}'$ are two generations with configurations $\underline{C}$ and $\underline{C}'$ respectively, and suppose there exists a 1-1 onto map $f:\underline{G} \rightarrow \underline{G}'$.

Then $\underline{C}$ and $\underline{C}'$ are B-isomorphic iff for $p \epsilon \underline{G}$ and $q \epsilon \underline{G}$, if $p \equiv q \bmod \underline{B}$ then $f(p) \equiv f(q) \bmod \underline{B}'$.

$\underline{C}$ and $\underline{C}'$ are M-isomorphic iff for $p \epsilon \underline{G}$ and $q \epsilon \underline{G}$ are $p \equiv q \bmod \underline{M}$ then $f(p) \equiv f(q) \bmod \underline{M}'$.

$\underline{C}$ and $\underline{C}'$ are S-isomorphic iff if $p \epsilon \underline{G}$ and $q \epsilon \underline{G}$ and $p \equiv q \bmod \underline{S}$ then $f(p) \equiv f(q) \bmod \underline{S}'$.

If $\underline{C}$ and $\underline{C}'$ are B-, M- and S- isomorphic, then they are completely isomorphic, or simply isomorphic.

The sign "$=$" will be used between isomorphic configurations (completely isomorphic configurations), the signs "$\underline{B}$", "$\underline{M}$", and "$\underline{S}$" will be used between B-, M- and S- isomorphic configurations, respectively. If $\underline{G} = \underline{G}'$, then $\underline{G}' = \underline{G}$. If $\underline{G} = \underline{G}' = \underline{G}''$ then $\underline{G}'' = \underline{G}$. If $\underline{G}\ \underline{B}\ G'$ then $\underline{G}'\ \underline{B}\ G$, and if $\underline{G} \overset{B}{=} \underline{G}' \overset{B}{=} \underline{G}''$ then $\underline{G} \overset{B}{=} \underline{G}''$, and similarly for S- and M- ismorphisms.

DEFINITION 4.7: Suppose $\underline{G}^{(n)}$ is a generation in a sequence of generations, and $p \epsilon \underline{G}^{(n)}$ and $q \epsilon \underline{G}^{(n)}$ and $p' \equiv q' \bmod \underline{B}^{(n)}$. (Let p' and q' be ancestors in $\underline{G}^{(n-1)}$ of p and q respectively.)

Then if $p' \equiv q' \bmod \underline{S}^{(n-1)}$ then p and q are parallel linked. If $p' \not\equiv q' \bmod \underline{S}^{(n-1)}$ then p and q are cross linked. Notate this "cross $L^1(p,q)$" and "parallel $L^1(p,q)$" respectively.

DEFINITION 4.8: Suppose $\underline{G}^{(n)}$ is a generation in a sequence of generations. A marriage rule A#, where "#" is a number, is satisfied in the sequence $\{\underline{G}^{(n)}\}$ (or, the sequence $\{\underline{G}^{(n)}\}$ satisfies A) if whenever p, $q \epsilon \underline{G}^{(n)}$ for all $\underline{G}^{(n)} \epsilon \{\underline{G}^{(n)}\}$ and p = q mod $\underline{M}^{(n)}$, then A# is true, provided A# is a conjunction of statements drawn from this list:

(1) for some positive integer i, $R^i(p,q)$
(2) for some positive integer i, not $R^i(p,q)$
(3) for some positive integer i, $L^i(p,q)$
(4) for some positive integer i, not $L^i(p,q)$
(5) there exist ancestors p' and q' of p and q respectively in $\underline{G}^{(n-i)}$, and for some positive integer i, cross $L^1(p',q')$
(6) there exist ancestors p' and q' of p and q respectively in $\underline{G}^{(n-i)}$, and for some positive integer i, not -ross $L^1(p',q')$
(7) there exist ancestors p' and q' of p and q respectively in $\underline{G}^{(n-i)}$, and for some positive integer i, parallel $L^1(p',q')$
(8) there exist ancestors p' and q' of p and q respectively in $\underline{G}^{(n-i)}$, and for some positive integer i, not parallel $L^1(p',q')$

A variety of marriage rules with possible English language interpretations are presented below. Note that while the rules are unique, the interpretations may not be unique. Further examples are given in chapter 7.

Rule Name	Rule	Interpretation
A1	None	"Random", sexual mating
A2	(2)1	"Sibling exclusion"
A3	(2)1, (6)1	"Cross cousin exclusion"
A4	(2)1, (8)1	"Parallel cousin exclusion"
A5	(2)1, (4)1	"First cousin exclusion"
A6	(2)1, (1)2	"Double first cousins"

"Contradictory" marriage rules which require both $R^i(p,q)$ and not $R^i(p,q)$, or which require both $L^i(p,q)$ and not $L^i(p,q)$ for the same value or values of i are clearly not satisfied by any sequence of generations.

DEFINITION 4.9: A stable configuration of period k, or simply a k-stable configuration, is a configuration $\underline{C}^{(n)}$ such that there exists a finite sequence $\underline{C}^{(n)}, \ldots, \underline{C}^{(n+k)}$ and $\underline{C}^{(n)} = \underline{C}^{(n+k)}$, and such that if $\underline{C}'^{(n)}, \ldots, \underline{C}'^{(n+i)}$ is another sequence with $\underline{C}'^{(n)} = \underline{C}'^{(n+i)}$, then k i.

Theorem: (cyclic sequence construction)

Suppose a sequence of generations is k-stable, such that $\underline{C}^{(n)} = \underline{C}^{(n+k)}$. T there exists an infinite sequence ..., $\underline{C}_v'$, ..., $\underline{C}_v^{(k)}$, $\underline{C}_v^{(k+1)}$, ... such that $\underline{C}_v^{(i)} = \underline{C}_v^{(i+k)}$ for all i.

4.2 Developments

THEOREM 4.1: Let ..., $\underline{G}^{(n)}$, ... be a sequence which satisfies A2. Then for all $|\underline{G}^{(n-1)}| \geq 4$, $|\underline{B}^{(n-1)}| \geq 2$, $|\underline{M}^{(n-1)}| \geq 2$, $|m^{(n-1)}| \geq 2$, $|f^{(n-1)}| \geq 2$.

Proof: Let p and q be elements of $\underline{G}^{(n)}$ such and p = q mod $\underline{M}^{(n)}$. Let $a \epsilon \underline{B}^{(n)}$ anc $b \epsilon \underline{B}^{(n)}$ pεa and qεb. By hypothesis (A2) d(a) ≠ d(b), hence a ≠ b.

Therefore $|\underline{M}^{(n-1)}| \geq 2$, $|\underline{G}^{(n-1)}| \geq 4$, $|\underline{B}^{(n-1)}| \geq 2$ since each individual in ε M-set must be from different B-sets, and $|m^{(n-1)}| \geq 2$ $|f^{(n-1)}| \geq 2$ from definitior

Q.E.D.

Theorems 4.2 and 4.3 follow immediately from theiorem A.

THEOREM 4.2: Let ..., $\underline{G}^{(n)}$, ... be a sequence of generaticns which satisfies A(Then: $|\underline{G}^{(n)}| \geq 4$, $|\underline{B}^{(n)}| \geq 2$, $|\underline{M}^{(n)}| \geq 2$ for all n.

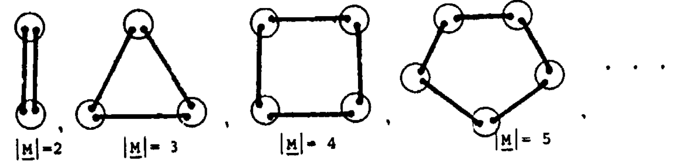
THEOREM 4.3: Let ..., $\underline{G}^{(n)}$, ... be a generation in a sequence of generations which satisfies A2, then for each n there exists at least one p' = $\underline{G}^{(n-1)}$ and o: $q' \epsilon \underline{G}^{(n-1)}$, p' ≠ q', such that $L^1(p',q')$.

Proof: In the proof of theorem 4.1, let p'εM' be an ancestor of p and q'εM'' b: ancestor of q. Note that p' = q' is then prohibited by theorem A' and by the p: bition of $R^1(p,q)$ by A2. Q.E.D.

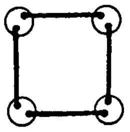
Note that while theorems 4.1, 4.2, and 4.3 specify conditions which must at lea: true for viable sequences under marriage rules, they do not guarantee tha any s' ture exists which satisfies these conditions, nor do they specify particular st: tures.

DEFINITION 4.10: All those configurations which are isomorphic to one another an isomorphism class.

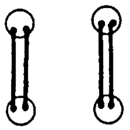
A configurational element is one of the following isomorphism classes:



Notice that an isomorphism class can be unambiguously represented up to an isomorphism of the S-labels, by display of a member of the class. Note also that in the configurations below, both have $|\underline{M}| = 4$, but one consists of a single configurational element, while the other consists of the union of two configurational elements each of which has $|\underline{M}| = 2$. If an isomorphism class is named "M#", where # is the number of M-sets in the class, then figure I above consists of the single configurational element M4, while figure II is the union of two M2 configurational elements. This may unambiguously be notated "2(M2)", describing a single generation $\underline{G}$ whose configuration $\underline{C} = (\underline{M},\underline{B},\underline{S})$ is the union of two separate configurations $\underline{C}_1 = (\underline{M}_1,\underline{B}_1,\underline{S}_1)$, $\underline{C}_2 = (\underline{M}_2,\underline{B}_2,\underline{S}_2)$, i.e., $\underline{M} = \underline{M}_1 \cup \underline{M}_2$, $\underline{B} = \underline{B}_1 \cup \underline{B}_2$, $f = f_1 \cup f_2$, $m = m_1 \cup m_2$.



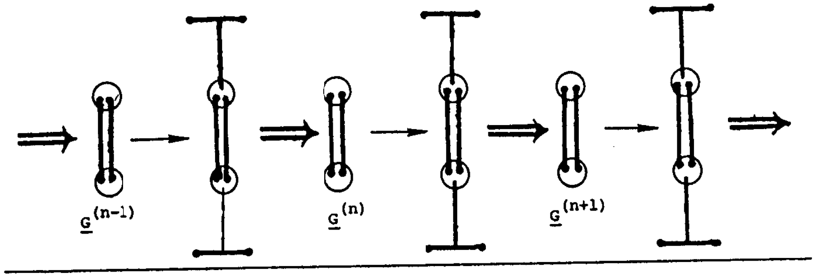
M4
FIGURE I



2(M2)
FIGURE II

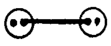
Sequences may be uniquely named if all configurations are isomorphic to a given configuration. For example, a sequence which has $\underline{C}^{(n)} = M2$ for all n may be called $(n)\{\underline{C}^{(n)} = M2\}$. This is suggestive of similar notations for cyclic repetitions of configurations, but these will not be exploited here.

THEOREM 4.4: The sequence $(n)\{\underline{C}^{(n)} = M2\}$ which has $\underline{C}^{(n)} = M2$ for all n satisfies theorem C with equalities in place of inequalities, and any other sequence which thus satisfies theorem C is isomorphic to $(n)\{\underline{C}^{(n)} = M2\}$.



Proof: By inspection, $(n)\{\underline{C}^{(n)} = M2\}$ has $|\underline{M}^{(n)}| = 4$, $|\underline{B}^{(n)}| = 2$, $\underline{G}^{(n)}| = 4$ for all n.

To show all other sequences with this property must be isomorphic to $(n)\{\underline{C}^{(n)} = M2\}$, consider the two examples below:

CLASS I 

CLASS II 

Since each M-class by definition 1 must have $|M| = 2$, these are the only two configurations other than M2 which have $|\underline{M}| = 2$ and also have $|\underline{G}| = 4$. Class I has $|\underline{B}| = 4$, while class II has $|\underline{B}| = 3$. Since M2 has $|\underline{B}| = 2$, the theorem is shown.

Q.E.D.

Note that because of definition 4.5, theorem 4.4 implies theorem 4.4'.

THEOREM 4.4': The sequence $(n)\{\underline{C}^{(n)} = M2\}$ satisfies A2, and satisfies theorem 4.2 with equalities in place of inequalities, and any other sequence satisfying theorem 4.2 with equalities in place of inequalities is isomorphic to $(n)\{\underline{C}^{(n)} = M2\}$.

THEOREM 4.5: Suppose $\underline{G}^{(n)}$ is a generation in a sequence of generations which satisfies A5. Then $|\underline{B}^{(n-1)}| \geq 4$ for all n.

Proof: Let $p,q \epsilon \underline{G}^{(n)}$. Then not $R^1(p,q)$ implies distinct descent chains, from ancestors in $\underline{G}^{(n-1)}$, hence $|\underline{M}^{(n-1)}| > 2$. Let p' and q' in $\underline{G}^{(n-1)}$ be ancestors of p or q respectively. Then not $L^1(p,q)$ implies $p' \neq q'$ mod $\underline{B}^{(n-1)}$. Since each M-set in $M^{(n-1)}$ has 2 individuals, and none are equivalent mod $\underline{B}^{(n-1)}$, $|\underline{B}^{(n-1)}| \geq 4$.

THEOREM 4.5': Suppose $\underline{G}^{(n)}$ is a generation in a sequence of generations which satisfies A5. Then $|\underline{G}^{(n)}| \geq 8$, $|\underline{B}^{(n)}| \geq 4$, $|\underline{M}^{(n)}| \geq 4$ for all n.

Proof: Apply theorem 4.5 to $\underline{G}^{(n)}$ and to $\underline{G}^{n+1)}$, finding respectively that $|\underline{B}^{(n-1)}| \geq 4$ and $|\underline{B}^{(n)}| \geq 4$. Since $|\underline{B}^{(n-1)}| \geq 4$ and descent maps are 1-1, $|\underline{M}^{(n)}| \geq 4$ (since $|\underline{M}^{(n)}| \geq |\underline{B}^{(n-1)}|$). And since $|\underline{G}^{(n)}| \geq 2|\underline{M}^{(n)}|$, $|\underline{G}^{(n)}| \geq 8$.

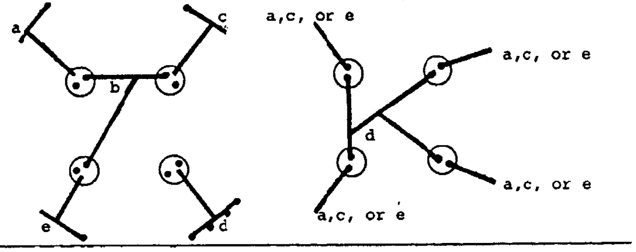
Q.E.D.

THEOREM 4.6: There exists a sequence of generations which satisfies A5, satisfies theorem 4.5' with equalities in place of inequalities, and is k-stable for k = 1.

Proof: The sequence is $(n)\{\underline{C}^{(n)} = 2(M2)\}$.

THEOREM 4.6': (equipartition from A5) Any configuration $\underline{C}^{(n)}$ in a sequence which satisfies A5 and theorem 4.5' with equalities in place of inequalities must have $|b_i| = 2$ for all $b_i \epsilon \underline{B}^{(n)}$, and for all n.

Proof: To show equipartition, note that if there is one B-set b with $|b| = 3$ in a generation. Then three of the M-sets will have 1-linked elements in the next generation. Because of $L^1(p,q)$ prohibition, these three can only marry from the fourth, but in the succeeding generation, all will be 1-linked (illustrated below).



The same illustration shows the consequences of $|b| = 4$ are that L^1-prohibition prevents all marriages. Note that $|b| = 5$ is prohibited, since $8 - 5 = 3$, hence only 3 M-sets at most may be completed, in violation of $|\underline{M}^{(n)}| = 4$. Similarly for b-sets of sizes 6, 7, or 8, while if any B-set is size 1, then some other B-set is size ≥ 3; ($|\underline{B}| = 0$ is meaningless).

Q.E.D.

Call the minimal numbers $|\underline{B}^{(i)}| = \beta_i$, $|\underline{M}^{(i)}| = \mu_i$, and $|\underline{G}^{(i)}| = \gamma_i$ thus discovered structural numbers of A.

Note: Inferences from later generations to earlier ones in a sequence can only be made through ancestors, which are members of M-sets, or to individuals linked to ancestors, which are also members of M-sets. Therefore, all individuals in a step of a minimal representation are members of M-sets. The following is obviously true.

THEOREM 4.7: Suppose A is a k-stable marriage rule with $k = 1$ and with minimal representation @. If $\underline{C}^{(n)}$ is a step in @ and $|\underline{B}^{(n)}| = \beta_n$, $|\underline{M}^{(n)}| = \mu_n$ and $|\underline{G}^{(n)}| = \gamma_n$ then there exist numbers β, μ, and γ where γ is even, such that for all n, $\beta = \beta_n$, $\gamma = \gamma_n$, $\mu = \mu_n$ and $|m_n| = |f_n| = \mu = 1/2\gamma$ and $1 \leq \beta \leq \mu$.

CONJECTURE 4.7': For all $k \geq 1$, theorem 4.7 still holds.

CONJECTURE 4.7": $\beta = \mu$ in every minimal representation.

CONJECTURE 4.7''': Every step in a minimal representation is partitioned into $\beta = s$ equal B-sets of size 2.

A likely method to prove these conjectures would be to generalize the technique of the preceding examples.

If the reader is now convinced that these numbers are "too small" to be meaningful, it is suggested that he turn to chapter 7 where the 1970 U.S. census is predicted from the simple fact that the "average" U.S. structural number is 4.

4.3 Summary of the Theory of Structural Numbers*

What we have done here is to construct a theory which has certain properties, known in advance from empirical studies of "reality", but which perhaps also does more than simply predict known "facts". Here is what we know: Many different "types" of marriage rules exist; these rules may be expressed in terms of "kin" types, which may also differ by society; rules may be efficiently drawn on two and three dimensional diagrams, which appear to have group theoretical properties and characteristic numbers. In addition to this, we suspect that the efficient diagra are "smaller" than actual population sizes, but the size of the diagram should hav some relation to the actual size of population, and to evolutionary properties of the population. Later chapters construct models having the desired statistical pr perties.

Basic Definitions

DEFINITION 1: A generation $\underline{G}$ is a set of elements p, q, ... called individuals. configuration C is a triple ($\underline{B}$, $\underline{M}$, $\underline{S}$) such that: (i) $\underline{B}$ is a partition of $\underline{G}$; (ii) $\underline{M}$ is a collection of two element subsets of $\underline{G}$; (iii) There are subsets of $\underline{G}$ labeled and f such that $|m \cap f| = 0$ and $\{m \cup f\} = \{\underline{G}\}$ and $\underline{S}$ is the partition of $\underline{G}$ imposed by these labelings; (iv) $|m \cap M| = 1$ and $|f \cap M| = 1$ for all $M \epsilon \underline{M}$.

Thus, my model has discrete generations, and monogamy in the structure which illustrates a rule. The sets "M" are married pairs, m, f the sets of males and fe males respectively. Relations between generations will be called descent maps. These will go from sibling groups (B-sets, or "distinct families") of one generat to M-sets of parents in a previous generation.

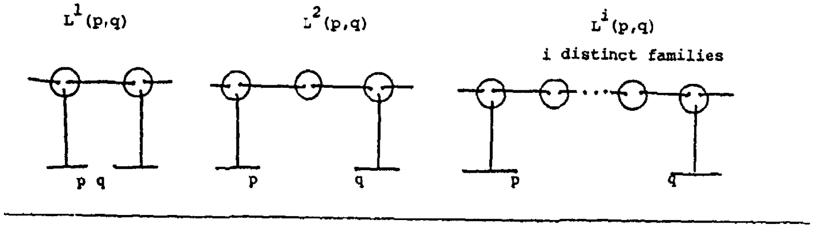
DEFINITION 2: A descent relation d from a generation $\underline{G}$ with a configuration $\underline{C}$ to generation $\underline{G}'$ with configuration $\underline{C}'$ is a map $d:\underline{B}' \rightarrow \underline{M}$, which is 1-1 into $\underline{M}$ and su that the inverse map is 1-1, from a subset of $\underline{M}$ onto $\underline{B}'$. A sequence of generatio $\{\underline{G}^{(n)}\}$ is a sequence of generations ..., $\underline{G}^{(n)}$, ..., with configurations ..., $\underline{C}^{(n)}$, ... and descent maps ..., $d^{(n)}:\underline{B}^{(n+1)} \longrightarrow \underline{M}^{(n)}$, ...

The terms "parent", "offspring", and "ancestor" will have the obvious meani then we can denote "p and q are i-removed" by $R^i(p,q)$. Likewise, if p and q are

---

* Reprinted with permission from Ballonoff (ed.) Genealogical Mathematics, pp 186-193.

"first cousins", then $L^1(p,q)$; if they are linked through "first cousins of first cousins", then $L^2(p,q)$; and likewise, for $L^i(p,q)$. (See illustration.)

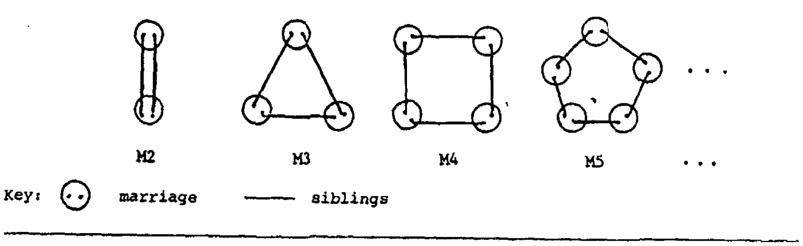


Notice that if $L^1(p,q)$ then $R^2(p,q)$, and if $R^i(p,q)$ then $R^{i+n}(p,q)$ for all integers $n \geq 1$.

It is now possible to define a marriage rule as a conjunction of statements from this list: (1) for some positive integer i, $R^i(p,q)$, and/or; (2) for some positive integer i, not $R^i(p,q)$, and/or; (3) for some positive integer i, $L^i(p,q)$, and/or; (4) for some positive integer i, not $L^i(p,q)$. It is easily possible to extend this list to include cross and parallel cousins of various degrees of removal, but (1) through (4) provide a sufficient notion for summary purposes. Thus, the combination ({(2) i = 1, (4) i = 1} = A) gives a single rule: "brother-sister marriages and first cousin marriages are not possible."

A marriage rule will be called k-stable, if in a sequence $\{\underline{G}^{(n)}\}$, every generation satisfies A, and $C^{(n)} = \underline{C}^{(n+k)}$ (the n^{th} and $(n+k)^{th}$ configurations are isomorphic) and k is the smallest integer value for which this is true. Since there may be several different k-stable sequences (with the same k value) for any particular marriage rule A, we can specify further that a minimal representation of a marriage rule A has smallest k, and also smallest $\underline{B}$, $\underline{M}$, m, f and $\underline{G}$ sizes. These will be called structural numbers, and since it can be shown that in a minimal representation, there exists a unique number s such that $s = |\underline{M}| = |m| = |f| = 1/2|\underline{G}| \geq |\underline{B}|$, call s the (unique) structural number of the marriage rule.

Finally, the state $T_{\underline{C}}$ of a configuration $\underline{C}$ is the set of all configurations isomorphic to $\underline{C}$. The state set T_A of a sequence satisfying A is the set of states of configurations in the sequence. A special class of states are the configurational elements (denoted "M#", according to # = the number of M-sets in the element). Configurational elements are the "cyclic" configurations shown below:

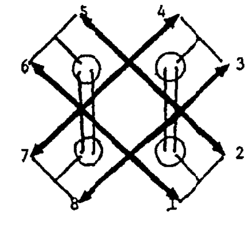


Thus, if one knows s = 4 for a particular minimal configuration $\underline{C}$, both $\underline{C}$ = M4 or $\underline{C}$ = 2(M2) may be possible in the absence of other knowledge.

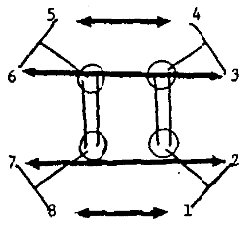
It is possible to make a simple correspondence between a matrix of zeros and ones, and the relations $R^i(p,q)$, $L^i(p,q)$, marriage ($\underline{M}$), sibling ($\underline{B}$), and "same sex" ($\underline{S}$). In particular, if $R^i(p,q)$ is a relation, then R^i is the corresponding matrix, if $L^i(p,q)$ is a relation, then L^i is the matrix, and denote the matrixes of the relations $\underline{M}$, $\underline{B}$, $\underline{S}$, etc. by the letters themselves. Define the composition relation "⊕" as the element by element computation of a logical "or" (i.e., $0 \oplus 0 = 0$, $1 \oplus 0 = 1$, $0 \oplus 1 = 1$, $1 \oplus 1 = 1$). Thus, if

or,

$$X = \begin{bmatrix} 1&1&1&1&1&0&1&0\\ 1&1&1&1&0&1&0&1\\ 1&1&1&1&1&0&1&0\\ 1&1&1&1&0&1&0&1\\ 1&0&1&0&1&1&1&1\\ 0&1&0&1&1&1&1&1\\ 1&0&1&0&1&1&1&1\\ 0&1&0&1&1&1&1&1 \end{bmatrix} \qquad M = \begin{bmatrix} 0&0&0&0&0&1&0&1\\ 0&0&0&0&1&0&1&0\\ 0&0&0&0&0&1&0&1\\ 0&0&0&0&1&0&1&0\\ 0&1&0&1&0&0&0&0\\ 1&0&1&0&0&0&0&0\\ 0&1&0&1&0&0&0&0\\ 1&0&1&0&0&0&0&0 \end{bmatrix}$$



M_1



M_2

The matrix M is the complement of X. M shows the possible marriages under A5, for the configuration illustrated. Where there is a single "1" entry in each row, this is the same "$\underline{N}$" used earlier. As can be seen, there are two apparent choices of partner for each individual in this case. For example, the "1" in positions 6 and 8 of row 1 show that 1 can mate with either 6 or 8. (Notice that all of these matrixes are symmetric.) However, only two sets of choices produce 1-stable sequences. These two are shown below as M_1 and M_2. The reader may wish to verify that any other matrix N is either not premitted in a sequence satisfying A5, or creates M4, which is 2-stable. Both M_1 and M_2 below create 2(M2), which is 1-stable, hence in the sequence.

$$M_1 = \begin{bmatrix} 0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&1&0 \\ 0&1&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \end{bmatrix} \qquad M_2 = \begin{bmatrix} 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{then}$$

$$S \circ B = \begin{bmatrix} b_{11} \oplus s_{11} & b_{12} \oplus s_{12} \\ b_{21} \oplus s_{21} & b_{22} \oplus s_{22} \end{bmatrix} = \begin{bmatrix} 1 \oplus 0 & 0 \oplus 1 \\ 0 \oplus 1 & 1 \oplus 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Using R with no superscript to denote an arbitrary relation, we may define the mating exclusion rule X of any population as

$$X = \sum_{j=1} c_j R_j$$

where j indexes all possible relations R, and if a particular relation R_j is part of the exclusion rule, then $c_j = 1$, otherwise $c_j = 0$, and always $c_s = 1$ for S = the same sex operator. (Definitions thus far follow chapter 3.)

The rule (2) i = 1, (4) i = 1 (to give it a name, call this rule "A5") has the exclusion operator:

$$X_{A5} = S \oplus R^1 \oplus L^1.$$

Since A5 has s = 4, X is a matrix of order $|\underline{G}| = 2s = 8$. Following the "canonical rules" of always placing odd numbers as males, and even as females (which works since $m = f = 1/2|\underline{G}|$ and $|\underline{G}| = 2s$ is even) and always positioning siblings next to each other, the exclusion matrix under A5 showing a minimal configuration is:

$$X = \underset{S}{\begin{bmatrix} 1&0&1&0&1&0&1&0 \\ 0&1&0&1&0&1&0&1 \\ 1&0&1&0&1&0&1&0 \\ 0&1&0&1&0&1&0&1 \\ 1&0&1&0&1&0&1&0 \\ 0&1&0&1&0&1&0&1 \\ 1&0&1&0&1&0&1&0 \\ 0&1&0&1&0&1&0&1 \end{bmatrix}} \oplus \underset{R^1}{\begin{bmatrix} 1&1&0&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 0&0&0&0&1&1&0&0 \\ 0&0&0&0&1&1&0&0 \\ 0&0&0&0&0&0&1&1 \\ 0&0&0&0&0&0&1&1 \end{bmatrix}} \oplus \underset{L^1}{\begin{bmatrix} 0&0&1&1&0&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&1 \\ 0&0&0&0&0&0&1&1 \\ 0&0&0&0&1&1&0&0 \\ 0&0&0&0&1&1&0&0 \end{bmatrix}}$$

Notice that either M_1 or M_2 is a permutation matrix of the first 2s integers {1,2,...,2s}, or in this case the integers 1,2,3,4,5,6,7,8. We can thus recognize the following result:

> If A is marriage rule with 1-stable minimal representation and structural number s, and $\underline{C}^{(n)}$ is the minimal configuration of A, then there is a marriage matrix M_i of A which transforms $\underline{C}^{(n)}$ into an isomorphic image $\underline{C}^{(n+1)}$ of itself, and M_i is a 2s by 2s permutation matrix.

From this, each marriage matrix M_i is therefore uniquely representable as a permutation

$$M_i = \begin{pmatrix} a_1, a_2, \ldots, a_{2s} \\ b_1, b_2, \ldots, b_{2s} \end{pmatrix} \qquad a_i, b_i, i = \{1,2,\ldots,2s\}$$

Notice, however, that not every one of the (2s) ways of permuting 2s integers is a possible permutation (possible marriage matrix). Even ignoring the marriage rules, the restriction $m = f = 1/2|\underline{G}| = s$ means there are only s possible entries for each row, and by symmetry, only s rows to independently permute.

<u>DEFINITION 3</u>: An almost minimal sequence under a k-stable marriage rule A is a sequence $\{\underline{G}^{(n)}\}$ such that for all $\underline{G}^{(n)} \epsilon \{\underline{G}^{(n)}\}$, $|\underline{B}^{(n)}| = |\underline{M}^{(n)}| = |f^{(n)}| = |m^{(n)}| = 1/2|\underline{G}^{(n)}| = s$.

Create the matrix T_{M_A} on $T_A \times T_A$ with all entries zero unless at the rj^{th} entry (the row state T_r, to the column state T_j) a descent map exists in some sequence from a configuration in state T_j to a configuration in state T_r. This follows from the fact that no sequence is possible unless there is some non-zero entry, and each non-zero entry shows an equally possible descent map in that row. In the row of state T_r there are only u_r such entries. Let the value of each entry in a row be = $1/u_r$. Then the fixed points of T_{M_A} show the proportion times spent in each possible state, in all possible almost minimal sequences. If we were to construct sequences where "probability" and "possibility" were not related in this simple way, then the entries in T_{M_A} would not be $1/u_r$, but the fixed points of T_{M_A} would still be the 'average distributions" of states of A.

As an example, we find the fixed points of the state transition matrix of the almost minimal sequences under A5. The only two states which may occur in such sequences are M4 and 2(M2). From M4, only 2(M2) may follow, while 2(M2) may precede either M4 or itself. The state transition matrix is, therefore,

$$T_{M_{A5}} = \begin{array}{cc} \text{M4} & \text{2(M2)} \\ \left[\begin{array}{cc} 0 & 1 \\ 1/2 & 1/2 \end{array}\right] & \begin{array}{l} \text{M4} \\ \text{2(M2)} \end{array} \end{array}$$

Let $a_1 + a_2 = 1$ so that (a_1, a_2) is a probability vector, and we must solve

$$(a_1, a_2)\ T_{M_{A5}} = (a_1, a_2)$$

to find the fixed points of $T_{M_{A5}}$. It is easily seen that $a_1/a_2 = 1/2$, so $a_1 = 1/3$, $a_2 = 2/3$, and, therefore, the fixed point proportions are 1/3 in state M4, and 2/3 in state 2(M2). Note that $T_{M_{A2}}$ is the 1 $\underline{X}$ 1 unit matrix [1].

<u>DEFINITION 3'</u>: Consider the state set T_A of 1-stable rule A. Define the set of a-almost minimal sequences of a marriage rule A with structural number s, as the set $T_{A,a}$ of all sequences satisfying A, and such that for each generation for any configuration in a state of $2s \leq |\underline{G}| \leq 2(s + a)$, where a is a non-negative number. Note that 0-almost minimal is almost minimal.

That is, if a = 0, $T_{A,0}$ is the set of almost minimal sequences of A in the sense of the earlier definition 3, and also note that for all a, and all $\underline{G}\varepsilon\{\underline{G}^{(n)}\} \varepsilon T_{A,a}$, $2(s+a)|\underline{G}|$.

Also note that a may be any non-negative number, not just an integer, which is useful in constructing a more statistically meaningful theory. The restriction to positive numbers is not really needed, since the meaning of "minimal sequence" is that $|T_{A,-a}| = 0$ for $|-a| \leq s$, and if $|\underline{G}| < s$ at any generation in a sequence, then the sequence is terminal, but not necessarily empty. The range $-2 \leq a < 0$ allows study of behavior "near extinction" of sequences, which is not the present purpose.

CHAPTER 5: ELEMENTARY PROBABILISTIC MARRIAGE THEORY

5.0 <u>Introduction</u>

The arguments presented so far have all been "structural", or in the vocabulary of some, "deterministic". In the author's view, "deterministic" is a highly improper description since one need not pretend a system ever followed a rule to study the structure or structures which would allow the rule to function. Indeed, as soon as we discuss anything other than 1-stable minimal systems, the apparent determinism quickly evaporates.

Marriage theory is therefore not a deterministic theory, but neither is it "probabilistic" or "stochastic" as we have so far viewed it. Nonetheless, there ar very important problems of empirical interpretation which appear to demand these approaches. In this and the next chapter, I have arbitrarily isolated two sorts of non-deterministic models: those which resemble sampling with a fixed probability (probabilistic models in the present chapter); and those which may have an underlying probabilistic foundation but appear to be more easily modeled by some other technqiue ("stochastic model", in chapter 6). What we are essentially after in thi chapter is a means to estimate with a probability the "viability" of a population with particular population statistics and particular network structure.

5.1 <u>Family Sizes</u>

The most important and obvious non-deterministic problem in marriage theory is family size. The minimality condition used so far implies that a population ob served over several generations is minimally stable if the size in any later gener tion is equal to or greater than the size of the first generation, but no bigger than it needs to survive. This allows us to write a simple equation which holds i minimal or nearly minimal surviving sequences (or, for that matter, any surviving sequence).

(1) $|\underline{G}^{(t)}| \geq |\underline{G}^{(t-1)}|$

From (1), we can find a new equation. Let $f_i^{(t)}$ be the frequency of family size occurring in the t^{th} generation. Since $\sum_{i=1}^{\infty} f_i^{(t)} = 1$, then

(2a) $\sum_{i=1}^{\infty} if_i^{(t)}$

is the average family size of the t^{th} generation and

(2b) $\sum_{i=1}^{\infty} if_i^{(t-1)}$

is the average family size of the $t-1^{th}$ generation.