

# Ring Cohesion in Marriage and Social Networks

Douglas R. White

July, 2004 (third draft 3.02)

Prepared for a Social Networks special issue edited by Alain Degenne

*Mathématiques, informatique, et science humains*

Journal of the Ecole des Hautes Etudes en Science Sociales, Paris

Structural studies are, in the social sciences, the indirect outcome of modern developments in mathematics which have given increasing importance to the qualitative point of view in contradistinction to the quantitative point of view of traditional mathematics. It has become possible, therefore, in fields such as mathematical logic, set theory, group theory, and topology, to develop a rigorous approach to problems which do not admit of a metrical solution. The outstanding achievements in this connection - which offer themselves as springboards not yet utilized by social scientist - is to be found in J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behaviour*; N. Wiener, *Cybernetics*; and C. Shannon and W. Weaver, *The Mathematical Theory of Communication*. (Lévi-Strauss 1968, Chapter XV, *Social Structure*, section on "Structure and Measure")

## Introduction

What are the basic forms of reciprocity by which strong ties and structural forms of social integration are constructed through marriage? Lévi-Strauss (1969 [1949]) classified forms of elementary marriage cycles created by cousin marriage in terms of their implications for social cohesion, that is, how patterns of marriage integrate social groups. Reciprocity can take the form of a cycle of direct exchange, either delayed, as in  $A \rightarrow B \rightarrow A$ , or immediate, as in  $A \leftrightarrow B$ . Alternately, it can take the form of a cycle of indirect or generalized exchange,  $A \rightarrow B \rightarrow C \rightarrow \dots \rightarrow A$ . This kind of question leads to examining actual patterns by which marital links between families create some of the observable social network patterns of cohesion, especially those involving cycles.

Noting Goldenweiser's (1913) complaint of the impossible complexity of kinship networks, Lévi-Strauss (1969:125) argued that because human beings cannot cognize this "apparent and impossible complexity" in the network patterns of kinship systems, they must formulate their models of social structure, like the analyst of social organization, in terms of rules. Rules, and the strategies for using them, could thus form a game in which social structures are seen to evolve by transformations of rules and how they are employed or applied. "Structure" defines the rules and constraints of the games and strategies are taken accordingly (White 1999). This remains a powerful model of social structure and evolution. This study shows how to use network analysis to account for marriage preferences that are not necessarily based on discrete rules, and how to give a more useful and more powerful statistical account of preferences that incorporates discrete rules as well as probabilistic preferences. It offers a more precise and conceptually intuitive accounting of the complexity that underlies kinship networks and their formation, and of the forms of cohesion that result from kinship dynamics.

Elementary structures were for Lévi-Strauss those formulated by rules that took a closed form: a marriage rule for a class of relations that one should marry. “Semi-complex” structures were those that specified a class of proscriptions for whom once should not marry that were sufficiently broad and organized as to entail as an effect a class of marriageable relations, also of closed form. The “complex systems” were those with merely statistical tendencies, preferences or avoidances as to whom to marry or not.

Studying cycles in a network reopen questions about complexity in a very different way than the presupposition of structural analysis that complexity lies in the nature of the rule, mechanical, statistical, or intermediate. Network analysis applies statistical analysis to the network itself as a complex entity, regardless of how the ‘rules of play’ are apprehended by the anthropologist or articulated by the players. Very different results about types of complexity are the result. This approach – developed by Houseman and White (1998a,b, White and Houseman 2002) in a ‘*theorie de la pratique*’ applied to kinship networks – is also a requirement of any study of statistical effects of different variables on marriage choices.

One way to study the contribution of marriage cycles to social integration and the forms of reciprocity – and to how prior relations affect actual marriages – is by developing a calculus for the occurrence of such cycles in a marriage network. Early attempts to do so, such as the analysis of Purum marriage cycles by Das (1945), failed to develop an adequate calculus, and were critiqued by Schneider (1965).<sup>1</sup> One of the goals of the present paper, embedded in the theory of ring cohesion, is to develop a calculus appropriate to kinship and marriage networks.

## Definitions and Theorems

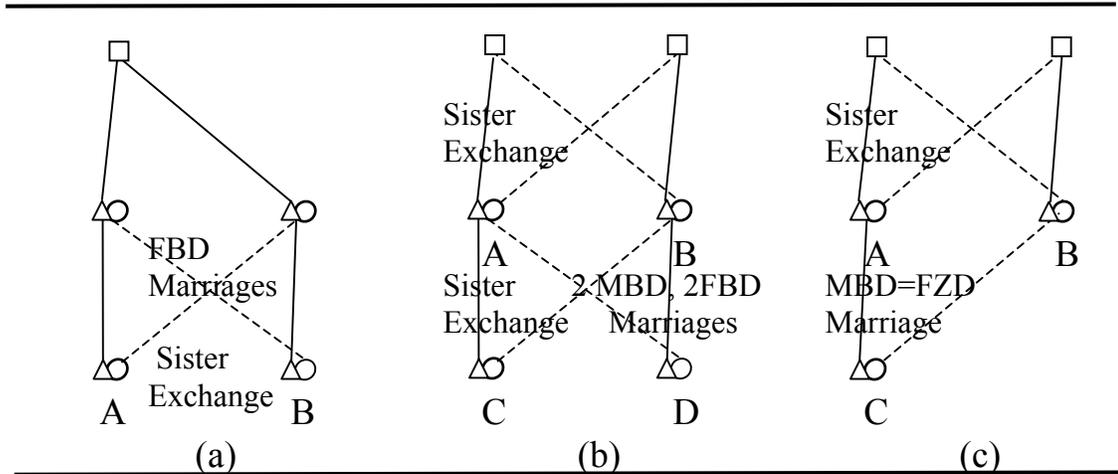
To pose basic research questions, graph theoretic definitions are required.<sup>2</sup> A *multigraph*  $G = \langle V, P, T \rangle$  is a set  $V$  of nodes (vertices) and a set  $P$  of ordered pairs in  $V$  classified by a set  $T$  of types of pairs. The different types  $k=1, t$  of relations in  $P \subset (V \times V)$  are called *edges* if all pairs of this type are undirected ties (e.g., siblings) and as *arcs* if any of its type are directed (e.g., parent/child), with  $e$  the index for types of edges and  $a$  the index for types of arcs. A multigraph  $D$  with at least one arc is called a *digraph*. If  $E$  is the set of all edges of different types  $E_e$  and  $A$  is the set of all arcs of different types  $A_a$  then  $P = (E, A)$ , and  $G = (V, P) = (V, E, A)$ . Let  $R$  consist of all elements in  $P$  with all arcs in  $A$  transformed into edges in  $A^e$ , and let  $G' = \langle V, P^e \rangle = \langle V, E, A^e \rangle$  be the *induced graph* identical to  $G$  except that all elements in  $P^e = \langle E, A^e \rangle$  are edges. A (simple) *graph*  $G = \langle V, P \rangle$  is one in which all relations in  $P \subset (V \times V)$  are edges. A *cycle* in a graph  $G$  is a series of alternating nodes and edges in which each edge connects the two nodes to which it is adjacent in the series, the first and last nodes are identical, and no other nodes are repeated. A *semicycle* of a multi-

graph is identically defined except that arcs are regarded as edges; hence every cycle in the induced graph  $G'$  of a multigraph  $G$  has a corresponding *semicycle* in  $G$ . A *subgraph*  $S = \langle N, P^N \rangle$  of a graph  $G$  is a subset  $N$  of the nodes in  $G$  plus all the edges in  $G$  that are in the induced subset of pairs  $\langle N \times N \rangle$ . The *degree* of a node in a graph  $G$  is the number of edges incident to the node. A *simple cycle* in a graph is a subgraph  $S$  that contains a single cycle, and all nodes in  $S$  have degree two. A (disconnected) *component* of  $G$  is any of its largest connected subgraphs.

Within this graph theoretic, and representing the data of a given kinship and marriage network, a *p-graph* is a multigraph  $G$  in which marriages are taken as nodes and the arcs are those between parents in the ascending generation and their children or children's marriages in the descending generation. Arcs are distinguished by sex of the child.<sup>3</sup> An inventory of types of marriage may be defined on a p-graph, each defined as a *ring*, or simple semicycle. This requires that of each of the types of marriage in the inventory is defined by a distinctive ring capable of identifying an empirical marriage of that type as a subgraph or *fragment* within the empirical network  $G$ . A ring may be composed of several different kinds of elements in  $P$ , arcs or edges, and a corresponding fragment of  $G$  must consist of a subgraph that is isomorphic to the ring. The *null ring* is the marriage type in which there is no prior p-graph relationship for a marriage, i.e., the cyclic closure of the ring is a null link; otherwise an empirical marriage or typological ring that creates a semicycle is called a *relinking* or *relinking marriage*.<sup>4</sup>

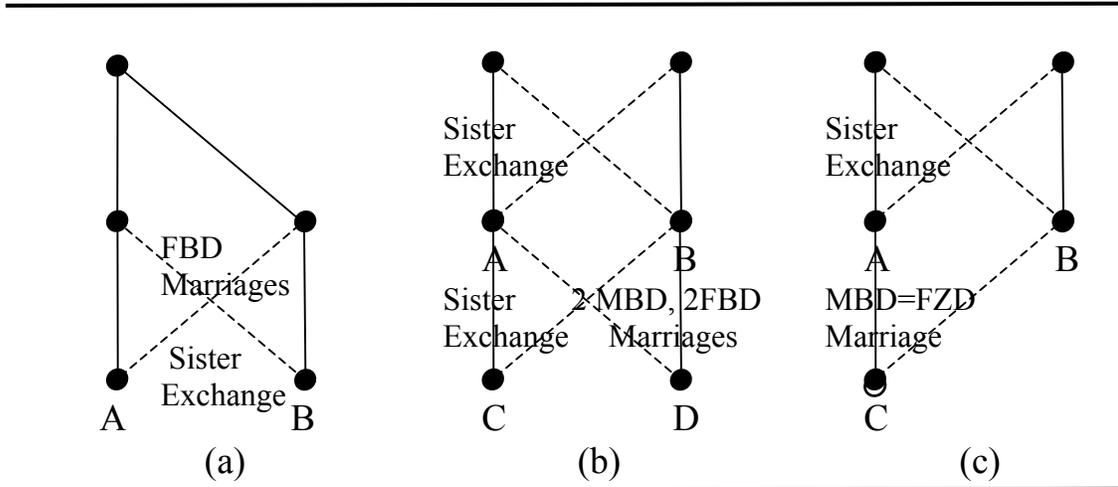
A single empirical marriage may be of one or more ring types and many—but not all—different types of ring can co-occur. An example is shown in Figure 1, a genealogical diagram with triangles for males, circles for females and squares for parents of both or either sex. To convert this to a p-graph, each couple is considered as a node, as in Figure 2. The marriage labeled A in diagram (a) in both figures is of two different types: sister exchange with B and FaBrDa. By implication, marriage B is also of FaBrDa type. This combination of two relinking marriages, one consanguineal and one not, thus entails a second consanguineal marriage for B. Alternately, the two FaBrDa marriages (A, B) entail sister exchange. In p-graph (b), marriage A is a sister exchange for B, and a second marriage C is a repeated sister exchange with D. The consequence is that marriages C and D are also of both types MoBrDa and FaSiDa. In p-graph (c), marriage A is a sister exchange for B, and a second marriage C is of types MoBrDa and FaSiDa, each entailing the other.

**Figure 1. Relinking marriage types with two or three independent cycles**



In Figure 2, all the couples are replaced by single nodes for marriages, and the types of lines—solid for male and dotted for female—differentiate the genealogical relationships. The beauty and simplicity of this graph is that all the marriage types can now be identified strictly in terms of rings or marriage cycles. Sister exchange, for example, is a bow-tie graph for antecedent and precedent generations with links of one sex in the vertical and those of the other sex in the crossover.

**Figure 2. Relinking marriage types with two or three independent cycles**



Different types of graph fragments can now be defined, such as the sister exchange bow-tie pattern of cycle, and these ring fragments can be searched for and enumerated in a database. How this is done efficiently with network software will be discussed below. This produces a frequency distribution for the occurrence of different types of marriage. In the examples of figures (a), (b) and (c) there are three, six and three different types of marriage, respectively.<sup>5</sup>

Graph theory thus simplifies certain problems of analysis prior to posing research questions. In these examples, it is evident that the cycles in a graph are not independent. Figure (a), for example, has three marriage types or cycles: two FaBrDa marriages, and one sister exchange (SiHuSi marriage). When any two of these cycles are present, the third is entailed.

In graph theory, a set of independent cycles is defined as any set of cycles in which each cycle contains at least one edge that is not present in any of the other cycles. I refer to cycles in a p-graph, leaving the time dimension implicit, as semicycles in the underlying directed multigraph. In example (a), any two of the cycles are independent, but all three are not independent because all the edges in the graph are used by any two cycles.

**Independent cycles (ring cohesion) theorem (1).** The *largest possible* number of independent cycles for a graph  $G$  with  $k$  edges,  $n$  nodes, and  $d$  disconnected components is  $k - n + d$ . This applies to any graph  $G$ . Proof: In that  $n-1$  nodes are the minimum required to connect a graph, additional edges will by definition create  $k - (n-1)$  independent cycles, and any arbitrary number of cycles greater than  $k - n + 1$  will be nonindependent. Starting from  $d$  (disconnected) components of a graph,  $n - d$  edges will be required to connect the nodes in each component, so the formula will be derived from  $k - (n - d)$  independent cycles. (see Harary 1969).

In a kinship network represented by a connected p-graph  $G$ , there are a maximum of  $r = k - n + 1$  independent cycles. For a comparable genealogical graph with individuals as nodes and arcs running from parents to children (Ore graph),<sup>6</sup> the formula for the maximum number of independent marriage cycles is  $r = k - n + 1 - \sum_{k=1,s} (n_s - 1)$  where subscript  $s$  is the number of sibling groups with two parents and  $n_s$  is the size of each.

If these formulæ are tested on figure (a),  $r = 2$ , while for (b),  $r = 3$ , and for (c),  $r = 2$ : there are a maximum of two independent marriages in examples (a) and (c) and three in (b). Conversely, while not immediately evident, the minimum number of nonindependent marriages is one in (a) and (c) and three in (b).<sup>7</sup>

Given the way that relinkings are computed, because the fragment frequencies are computed for subgraphs that are simple cycles, nonindependent cycles in a subgraph with  $r = 2$  cannot result from the conjuncture of two consanguineal marriages or from that of two consanguineal marriages but only by the combination of one of each type, as in examples (a) and (c). A given marriage cannot be both a FaBrDa and a MoBrDa, for example, unless (1) one's parents are siblings, or (2) a person has more than one father and/or mother. A man can marry both a FaBrDa and a MoBrDa under these restrictions, for example, only if they are different wives.

Ring cohesion theory uses calculation of independent cycles to solve problems of how to explain a complex marriage structure in terms of preferences whose maximal extension to independent marriage cycles is  $r$ . This is possible because the ex-

cluded marriage cycles are necessarily consequences of the some set of marriage cycles whose maximal extent is  $r$ . Suppose a model in which, for a network with  $r$  independent cycles, precisely  $k = r$  marriages are posited that occur because of a preferential marriage rule in which the preferential marriages create  $k$  cycles. If precisely  $k = r$  such marriages are identified, they are independent in the sense that their presence in the network also accounts for all other cycles.<sup>8</sup>

In graph theory, accounting for other cycles corresponds to taking the union of the subgraphs for two cycles, and then subtracting the edges they have in common. This will generate a new, nonindependent cycle. In figure (a), for example, the graph-addition of the two FaBrDa marriages leaves as the outcome the sister exchange. Similarly, the graph-addition of the one of the FaBrDa marriages with the sister exchange leaves the other FaBrDa marriage, and so forth. Using these procedures allows us to give a reckoning of independent cycles in a graph, provided the model identifies those types of marriages that are either preferential or that are not disallowed by a marriage proscription.

Now suppose that a model in which a certain number  $m$  of marriage types can be ranked in order of preferences. Let  $f_i$  be the empirical frequency of the first and any successive marriage type ( $i=1,m$ ). If we remove each of these marriages from the network, we will typically observe a reduction of the frequency of other types of marriage for cases in which the marriages of this type overlap with marriages of other types. Let  $F = \sum_{i=1,m} f_i$  be the total number of marriages removed by this process. If  $F = r$ , no cycles will remain. This model will have succeeded in accounting for all the cycles in the network: by accounting for the independent cycles, it will account for the nonindependent cycles as well. The removals, that is, reduce the genealogical graph to a approximate a tree with no cycles.<sup>9</sup>

## Research Questions and Further Definitions

### Research Questions

The questions of interest apply not only to individual cases, but to any genealogical network that meets the following conditions: individuals have at most two parents (however defined and thus not necessarily biological), and parental and ancestral relations are temporally ordered (parents preceding children) so that no directed cycles occur where one is one's own ancestor. It is only for terminological convenience that parents of the same children are considered "married"; any other term for this relation (e.g., coparent) will do and will not change the results.

1. What are the sources of marital cohesion in the community?
2. Can ring cohesion calculus be used to help give an account of the sources of cohesion in a kinship network that are due to relinking marriages?

3. Can the frequencies of matrimonial types be accurately enumerated up to the limit of the investigators' knowledge as coded in the database?
4. Is there a valid means for identifying which sets of marriages in any given community are more preferential than the rest? What are the statistical signatures of sets of preferential marriages?
5. What will remain of a given empirical network when those marriages posited as preferential are subtracted and how is this accomplished?

Answers to these questions require further definitions and theorems constructed from the logical basis of ring cohesion theory. Footnotes in the following sections address these questions and refer to operations of the program Pajek (Batagelj and Mrvar 1998, 2002) used for computational purposes in network analysis.

### 1. Overall Cohesion and Sources of Cohesion

To return to the initial theme, that of social cohesion, cycles created by marriage in genealogical networks create the boundaries of *structural endogamy* (White 1997), in which every marriage connects with every other through two or more independent paths. A p-graph representation of marriage networks allows this definition to be operationalized so as to entail social cohesion created by marriage. Two or more paths from one node to another are defined as (*node-*) *independent* if they have no intermediate nodes in common. The *level of cohesion* of a multigraph or p-graph, *measured by an integer k*, is defined as the connectivity number k of nodes that must be removed in order to disconnect it.

**Multiconnectivity (ring cohesion) theorem (2).** If a graph or subgraph has connectivity k, then every pair of its nodes has k node-independent paths between them, and vice versa (for proof see Menger 1927, Harary and White 2001).

A p-graph is *regular* if no individual has more than two parents and no marital node has more than two parental couples, one for a male, one for a female; if married or coupled the male and female are members of the same node in the p-graph. A *bicomponent* of a multigraph or p-graph P is a maximal (largest possible) subgraph in which every pair of nodes is contained in a semi-cycle.<sup>10</sup> No p-graph may have connectivity 3 or higher if it is regular. A bicomponent of a regular p-graph is therefore a maximal unit of structural endogamy and isomorphic to a maximal unit of social cohesion for a genealogical network.

Regular p-graphs thus provide a natural means of recognizing structurally endogamous and genealogically cohesive groups within genealogical networks of kinship and marriage. A sexually reproductive community will typically have a single large bicomponent or structurally endogamous unit. Alternately, a sexually reproductive community may be defined by the local limits of structural endogamy, that is,

discounting the marriages of those who have permanently emigrated from the local area and whose descendants have not returned.

Typically, cohesion analysis is done on a single maximally large bicomponent or structurally endogamous unit of a regular p-graph, referred to as a cohesive ‘community’ subgraph of a genealogical network. It is only within such a community or bicomponent of the kinship graph that rings will be found.

## 2. Ring Cohesion Calculus

Marriages between people who are previously related come in two basic forms, with subvarieties. One is consanguineal marriages. The other is marriage between in-laws, extended generically to all those who are linked, prior to their marriage, by one or more chains that combine blood relations and prior marriages. Both are called relinking marriages. Any type of relinking marriage, when defined by a simple cycle, will involve 1, 2, 3, or a higher number of marriages, including the relinking marriage that is last in the time sequence. The number of marriages involved is identical to the number of families who are relinked, where families are defined as the number of (disconnected) components that remain after the marriages that give rise to cycles are deleted.

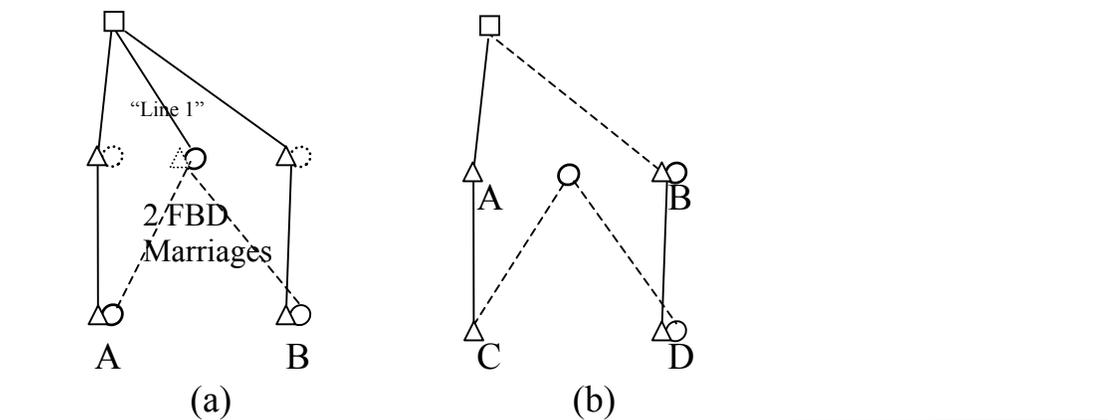
To employ the marriage-removal method of ring cohesion as it applies to genealogical networks, the graph theoretic representation must include marriage as a type of link and, thus, individuals as the nodes of the graph. P-graphs no longer suffice. In the type of graph that is needed, when all marriage links that are embedded in a given type of marriage cycles or ring are removed from the graph, the individuals connected by the marriage must remain.<sup>11</sup> Formats for genealogical networks with this structure also makes possible the following definition, which assumes only that a matrimonial relation (“marriage”) may exist without children, when children of a could exist they are considered for analytic purposes to be “married.”

**Matrimonial rings.**<sup>12</sup> A *matrimonial ring* in a genealogical digraph with edges for marriage links among individuals and arcs from parents to children is a simple (semicycle) such that after removal of the edges all components are rooted trees with a single ancestral node. A tree is *minimal* if it consists of a single node.

The rings in a genealogical digraph together with their trees and branches allow the statement that a matrimonial ring corresponds to a cycle of marriages within or among families defined by members of a rooted tree of consanguineal relatives with a common ancestor or ancestral couple. In figure (a) above there were several rings within a single family: two FaBrDa marriages and one sister exchange. In Figure 3 (a) below there are two rings: two FaBrDa marriages only. But if we removed “line 1” between the man in the center and his parents, would this not be a marriage

between two families? And if we changed the structure only slightly to (b), would this constitute a single ring composed of a marriage between two families?

**Figure 3: Genealogical networks (a) with two rings and (b) no ring**



Using the ring definition to apply to digraph 3(a), however, there are only two separate rings and no additional simple cycle. The semipath on the outer perimeter of figure 3 (a) might describe a FaBrSoWiSi marriage, but as shown, it does not constitute a ring because the subgraph induced by the eight linking individuals is not a simple semicycle. Recall that a subgraph  $S = \langle N, P^N \rangle$  of a digraph  $D$  is a subset  $N$  of nodes in  $D$  plus *all* the edges and arcs in  $D$  that are in the induced subset of pairs  $\langle N \times N \rangle$ . Because *only* those subgraphs of a genealogical network that are isomorphic to a given type of ring are considered matching fragments, and the semidegree (counting ingoing and outgoing arcs and bidirected edges equally) of every node in a simple semicycle (matrimonial digraph fragment) that is isomorphic to ring must be two, subgraph (a) is not a ring since the two nodes connected by "line 1" have semidegree 3.

**Branching Theorem (3).** There are no more than two branches in any of the rooted trees in a matrimonial ring. Proof: Given that a matrimonial ring is a simple semicycle in which a cycle of marriages connects rooted trees, there must be two nodes within each tree connected by marriage. If a marriage occurs within a family, that closes the ring, and the tree must have two branches to reach those nodes. If a marriage occurs between families, each family must have two branches to reach those nodes.

Consider the network in Figure 1(a): here there are two marriages and three branches within a single family. This is not a single ring, however, but multiple rings. The FaBrDa ring has two branches, and the sister exchange ring has a pair of branches in each of two families. Given theorem (3):

*A matrimonial ring corresponds to a cycle of marriages within or among families defined as members of a rooted tree of consanguineal relatives with a common ancestor and at most two branches.*

A minimal matrimonial ring may even might consist of a cycle of marriages with minimal rooted trees consisting of single nodes and no branches. Four individuals, for example, may form a matrimonial ring if a man marries his WiHuWi.

Applying the matrimonial link and ring definition to apply to digraph 3(b), that network violates our criterion for a marriage network because person C has a mother and father, but when the missing “marriage” link between them is supplied, C is not part of the marriage ring. Omitting individual C, however, D is not a true MBD marriage ring but rather a MoBrWiDa marriage ring cycle of marriages between two families, the two rooted trees being a Mo-Da pair and branching tree of MoBr –SiSo.

### 3. Enumeration

Many ethnographers have noted the importance of the full versus half sibling relation in human societies. Given that couples and not solitary individuals typically engender offspring, the ancestral nodes of the rooted trees within matrimonial rings may be either individuals (male, female) or couples (both) without violating the simple cycle criterion of rings. It is the choice of whether to take as an ancestral node an ancestral male, female or couple that determines the potential for half-sibling (paternal, paternal) or full sibling relations. Only if rings of all three types are computed and compared can these types be distinguished. When only one of the potential half-sibling types is found, it must be verified whether or not there the spouse of the ancestor is coded in the database. If the coding rules are always to indicate a parental couple where possible given knowledge of full siblings even when other data on one of the parents (name, ancestry) is missing, then it will be possible to make a complete determination of full and half-sibling relations. This approach is supported by the following theorem:

**Ancestral node theorem (4):** In matrimonial rings, the ancestral node is the only type of node that generates branching. Proof: Because every family in a marriage cycle has an even number of marriage links and thus at most two branches, if the ancestrally rooted tree for a family with a branch that is not generated by two children of the root, then the only way the ancestor can be involved in a matrimonial cycle is by a marriage: if that marriage is within the same family tree, there can only be one branch; if it is with another family then only only a node on a single descent line from the ancestor can be the other marriage involved in the matrimonial ring and hence that tree also has only one branch. In either case a contradiction results, hence the theorem is proven.

By theorems (3) and (4), two branches at most are possible without violating the simple cycle criterion, and these just occur directly under an ancestral node in a marriage ring, hence this is the only juncture at which specification of sibling (full or half) relations are necessary in matrimonial rings.

#### 4. Statistical Signatures of Preferential Marriages

**Short cycles.** Rules of incest prohibition typically apply to sexuality or marriage that would create short cycles: brother-sister incest, mother-son, father-daughter, and their extensions to other relatives. In many communities, there are rules that proscribe marriage with first, second, or third cousins, uncles and aunts, or various types of cousins. Occasionally, we see contact avoidances such as that with WiBrWi, but in the cases observed, this is typically extended to brother-sister behavioral avoidance (White n.d.).

**Preferences.** Among permitted marriages, we would expect to find many communities in which there are preferences for closer relatives, i.e., marriages that involve greater preference for short over long cycle in marriage.

**Raw frequencies.** The frequencies of observed marriages classified by type may provide a first-order indication of marriage preference, but the evidence for preference must be carefully assessed.

**Percentages.** Alternately, we may consider the percentage of given type of relative that are taken in marriage.

**Frequency distributions.** Here, we consider three methods of showing frequency distributions:

4.1. Rank order the marriage types by their *frequencies or percentages*, with frequency on the y-axis of the plot and the **rank order of these quantities** (allowing ties) on the x-axis. Use of ring cohesion calculus may show that only a limited number of high-frequency types (with few ties among their frequencies) are needed to account for the independent cycles in the marriage network.

4.2. Nominate the x-axis as a *frequency descriptor for types*, starting with 1 for those types with lowest frequency or percentage and running up to the largest value. In this case the *number of types that have this nominated value* is put on the y-axis. Hence, if there are 200 types with frequency 1, the number of marriages involved is  $1 \times 200 = 200$ , while if there are 100 types with frequency 2 these also represent  $200 = 2 \times 100$  marriages. At the extreme there will be very few types with high frequency, but these will represent many marriages. When the ring cohesion calculus is used, and only a certain number of high-frequency types are sufficient to account for the independent cycles, this method will result in discontinuities in the graph of x-y values. To look for continuous relationships between the x-y values, we may use a cumulative measure of frequency as follows.

4.3. Nominate the x-axis as the cumulative frequency of all types having *successively less frequent occurrence* and the y axis at the *cumulative number of types*. This plot allows continuous interpolation among the plotted x,y values and consideration of the shapes and slopes of the curve under various transformations of the x-y axes (linear, logged on one axis, and log-log plots).

When the distributions are constructed in one of these three ways, but preferably the third, we may observe various types of empirical relationships among the pairs of values on the x-y axes as follows:

4a) ***Discrete staircase relationship***. Here, in plot method 1, several of the most frequent types will show as roughly tied in rank, with possibly a second or third echelon of tied ranks.

4b) ***Linear relationship***. One or several of the plot methods shows a linear relationship between unlogged x and y values. This indicative of a preference for “more” of the most frequent items, with less frequent items as partial substitutes.

4c) ***Exponential or logarithmic relationship***. Here the x-y relationship is linear when one but not both of the axes are logged. The clearest interpretation will result when method 3 is used. This is not necessarily indicative of a preferential order because this may also occur with randomly generated frequencies for different types. The question here, as with a) and b) is whether the ordering is predicted by a certain preferential logic that can be predicted in advance from some independent characteristics of individuals in the community.

4d) ***Power-law relationship***. Here the x-y relationship is linear when both of the axes are logged. The clearest interpretation will result when method 3 is used. This suggests that there are differentially self-amplifying elements in the ordering. The appropriate model is to look for independent components that go into the ordering, each weighted differentially to produce both the power relationship (the higher the product or sum of weights the more the self-amplification) and the ordering itself.

## 5. Reductive analysis and format

The definition and methods for enumeration of marriage types allow us to identify all marriages in a genealogical network of each given type, and these frequency distributions can be analyzed for evidence of preferential marriages. Once the actual identification of marriages by type is completed, a new question arises from ring cohesion theory: If we remove marriages in order of presumed preferences, regarding them as within the set of independent cycles, what happens to the frequencies of the remaining marriage types? The independent cycles theorem (1) requires that these will include types that are nonindependent of the preferential types, and hence these frequencies will diminish. When few or no cycles remain after a series of

removals, it is valid to say that an account of the total ensemble of cycles in the network has been rendered in terms of the preferential and independent cycles.

In the marriage-removal method operationalized by Hamberger et al. (2004), a minimum of five relations need to be defined, and a sixth is optional.<sup>13</sup> Five are primitive in the sense that they are not the logical result of any concatenation of other primitives. The sixth, siblingship, is a relationship that is not primitive but derived from having common parents. These six relations in the reductive analysis format are:

F-D (arc)	M-D (arc)	Marriage (arc)
F-S (arc)	M-S (arc)	Siblings (edge)

If only the five primitive kinship relations are used to define search fragments for use with a kinship network coded in the same format, the ring fragments may be coded to include individual ancestors or ancestral couples in the ‘family’ components of the ring. Rings of complementary types may be designed to include only female ancestors, only male ancestors, or only ancestral couples. These distinctions will be important when sibling branches occur from the ancestral nodes.

An alternative network format introduces the sibling relation and deletes the ancestral node parental to a sibling branch in a family tree. This introduces cycles in the graph that are not due to marriage, but occur within nuclear families with two or more children. The appropriate formula for the maximum number of independent marriage cycles given above is  $r = k - n + 1 - \sum_{k=1,s} (n_s - 1)$  where subscript  $s$  is the number of sibling groups with two parents and  $n_s$  is the size of each.<sup>14</sup>

## Methods and Measures

**Samples.** Hundreds of genealogical samples are available for analysis (White, Houseman and Schweizer 1993). Project TIME (see Hamberger et al. 2004, Barry 2004) has converted many of these and other genealogical networks into the reductive analysis format.

**Software.** From the first year of its implementation in 1996, the Pajek program for large networks analysis has provided and updated a series of algorithms for kinship networks and genealogies (Batagelj and Mrvar 1998, 2002, Mrvar and Batagelj 2004, White, Batagelj and Mrvar 1999). The index of relinking, a measure of the extent to which a graph with  $k$  independent ancestors has the maximal possible number of relinkings, was implemented in 1998. Pajek is now in its definitive version (1.01, 2004), released with an instructional manual (de Nooy, Mrvar, and Batagelj 2004), and for use with the reductive method for studying ring cohesion. Macro commands for kinship analysis come with the Pajek installation.

**Fragments.** Search and identification of fragments in graphs was implemented in Pajek in 1997. In 2003, I developed a set of kinship fragments for identifying dif-

ferent types of consanguineal marriages in p-graphs. The Pajek count of fragments eliminates isomorphisms (e.g., two brothers marrying two sisters is counted as a single fragment), which allows counting of marriage types by marriages, or as isomorphic sets of subgraphs. Knowing the number of isomorphisms of each graph, e.g., the two isomorphisms for sister exchange, these may be easily converted to frequency counts by individuals. Hamberger et al. (2004) developed an improved set of kinship fragments in reductive analysis format for identifying different types of relinking, including blood marriages. This approach uses software by Jürgen Pfeffer (FAS-Research, Vienna) to convert a kinship file in Excel format (Ego, Sex, Fa, Mo, Spouse) into a Pajek genealogical file in the new format. Sibling links must then be added and resultant loops removed, which are simple operations in Pajek.<sup>15</sup> To find marriage types as fragments in the network, the genealogical file is scanned for successive fragment types, now a standard Pajek option.<sup>16</sup> The fragments found of a given type may be saved without renumbering nodes so that marriages of a given type can be subtracted from a given network.<sup>17</sup> This feature is used for the reduction method of ring cohesion analysis.<sup>18</sup>

### Hypotheses

White and Houseman (2002:78-79) posited a tricotomy of community types with respect to sources of marital cohesion:

H1. Communities with many blood marriages may have power-law preferences on blood marriages, but will be neutral on multifamily relinkings, that is, with differences that are exponentially distributed. The appropriate ring cohesion model to test here is that the independent cycles that generate structural endogamy are the blood marriages.

H2. Communities with few blood marriages may have power-law preferences on multifamily relinkings. These may exhibit one of two subtypes:

- a. Power-law preferences on two-family relinkings.
- b. Power-law preferences on three-family relinkings

**Partition (ring cohesion) principle (3).** Ring cohesion theory helps to explain the basis for White and Houseman's (2002) hypotheses 1 and 2. When uniqueness of parents and restrictions against sibling marriage apply for a given genealogical database, it is convenient to think of either the entire set of consanguineal marriages as constituting the bulk of a set  $S$  of independent cycles, with nonconsanguineal marriages constituting the bulk of the complementary set of nonindependent cycles, or vice versa. This is because among the consanguineal marriage set, for any subgraphs with  $r = 2$ , there will be no nonindependent cycles.<sup>19</sup> Nonindependence for  $r = 3$  subgraphs will involve combinations of consanguineal and of nonconsanguineal relink-

ings: two MoBrDa marriages entail two FaBrDa marriages, for example, only if there is sister exchange in the first generation, as in figure (b). Thus, the major restrictions on the types of relinking that can co-occur are within the distinct classes of relinking: consanguineal or nonconsanguineal (White 1997). Complex cases such as example 1(b) may involve combinations of one nonconsanguineal and two consanguineal types, but may also be seen as a combination of two nonconsanguineal types that can occur only where  $r \geq 2$ .

What the independent cycles theorem (1) and partition principle jointly imply is that if a community has preferences oriented towards types of consanguineal relatives, a preferential ring distribution of consanguineal marriages will imply an exponential (nonpreferential) distribution of nonindependent affinal relinking cycles. The second distribution cannot be preferential because (1) by definition these cycles are not independently formed, and (2) it is impossible or highly unlikely that both distributions could be jointly optimized in choice behavior. The converse is the case if a community has preferences oriented towards types of nonconsanguineal relinking. Thus, the third hypothesis as a summary derived from first principles.

H3. No community will have power-law preferences on both consanguineal and affinal relinkings.

White and Houseman's (2002:78) hypotheses (H1 and H2) were initially a report of an empirical pattern of investigative results concerning kinship networks generally. Ring cohesion theory identifies the derivation of these hypotheses from network and mathematical principles that hold for kinship systems. H3 holds because kinship has a hereditary component (descent) and an elective component (affinity). Biological and other kinds of networks also have internal contrasts of this sort for certain subtypes of network relations, so further generalization of H3 to other sciences would be expected from first principles. More generally:

H4: Whatever the method of detecting marriage preferences, there will be some statistical tendencies for preferences of the consanguineal sort to reduce those of nonconsanguineal relinkings, and vice versa.

This takes into account the possibility that a preference ordering on a small subset of consanguineal marriages would not rule out additional affinal preferences, and vice versa. It is only when a preference ordering is over a larger range of marriage types that H3 is likely to be operative.

### **Illustrative Results for Kinship Network Analysis**

The ring cohesion approach is exemplified from the Turkish Aydınli nomad data studied by White and Johansen (2004, Johansen and White 2002). Results in the following tables and figures are illustrative of tests of hypotheses about the Turkish

nomads, with an Arabic type of kinship system and rights to FaBrDa marriage.

## Preferences

Evidence for consanguineal marriages as preferential and thus as the appropriate set of independent cycles in the network is considered first. If both these conditions were true, it is possible that the affinal relinkings, considered as nonindependent cycles, would require no further explanation in terms of preference. The supporting evidence that consanguineal marriages are preferential (and nonconsanguineal relinkings are not) comes from White and Johansen (2004:275-278), as shown in the following figures, some of which are discussed in White and Houseman (2002).

Figure 4 uses plotting method 2 to break out the frequency gradient for all types of consanguineal marriages, regardless of kinship distance.<sup>20</sup> Here the  $x$  axis is a variable for the number of spouses whose marriage fits one of the 234 types of blood kinship within the range of fifth cousins (7 generations to a common ancestor). The  $y$  axis is the number of those types of marriage with exactly the number  $x$  of related spouses.

From the graph we can read that 156 of the 234 types of marriage have a marriage frequency of one in the dataset. This number drops to 36 of the 234 types for which there are two marriages. If the graph were exponential it would keep dropping by a constant fraction, such as from 156 to 36 to 10 to  $\sim 2$  and quickly to zero. Instead, the graph follows a power law, and drops from 156 to 36 to  $\sim 18$  to  $\sim 10$  to  $\sim 5$ , showing the extended tail of a power law distribution. The long tail is indication of a few types with much higher frequencies than would occur if types of marriage partner were chosen randomly. The extreme outlier in this breakdown of frequencies by type is FaBrDa marriage, for which there are 32 instances, nearly twice that of the next most frequent marriage, of the MoBrDa marriage type. The distribution has a fixed exponent ca. 2, an inverse square power law whose equation is approximately  $y=156/x^2$ , with  $y$  the number of types whose frequency  $x$  is between 1 and 32. The fit of the power-law curve to the data and is  $r^2=.83$  with an estimated slope of  $1.97 \sim 2$ . Because a power-law distribution is not what we expect at random, we call this a preferential distribution, and power laws of this sort are suggestive of networks that operate as self-organizing systems having fractal properties, where a basic pattern of behavior is repeated at different scales, and more diffuse forms of the patterns operate at lower frequencies. When we examine the types of marriage with the highest frequencies (FaBrDa, MoBrDa, FaSiDa) they also seem to follow an order that reflects kinship distance reckoned from a perspective of patriline.

**Figure 4: Power Law Fractality of Marriage Frequencies**

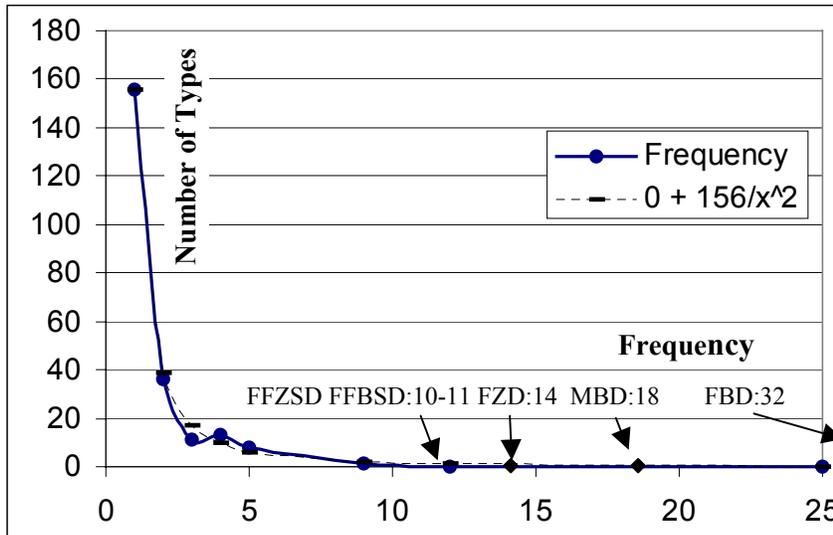
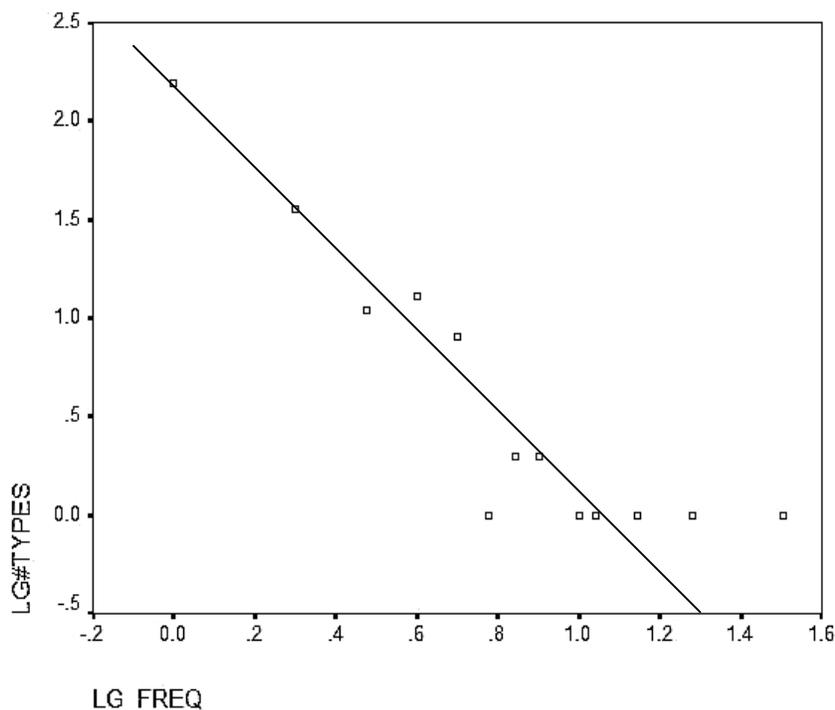


Figure 5 repeats the analysis in Figure 4, this time taking the log of values on the x and y axes, and fitting a straight line to the plot. The fit approximates a power-law distribution. The statistical signature of this relationship is discussed under that heading, 4d) – the hypothesis being that of differentially self-amplifying elements in the ordering. The amplification process resulting in these frequencies is greatest the closer the link-distance when male links are weighted higher than female, so that FaBrDa comes out highest in the ordering.

Fractal marriage patterns function rather like Granovetter’s (1973) strong and weak ties, which have complementary strengths at complementary distances. The stronger and more frequent ties (of many fewer types) work at closer distances, in this case concentrically oriented toward close and patrilineal relatives, while the weaker ties of each type are individually less frequent but work as an ensemble in a distributed manner over longer distances. The fractal distribution of a strong/weak tie pattern of this sort. Unlike the way that marriage preferences are usually formulated as discrete rules, the fractal pattern is continuously scaled rather than a simple dichotomy of types of marriage.

**Figure 5: Power Law for Marriage Frequencies - Log-log plot for Figure 4 with fitted line, slope  $\sim 2$**

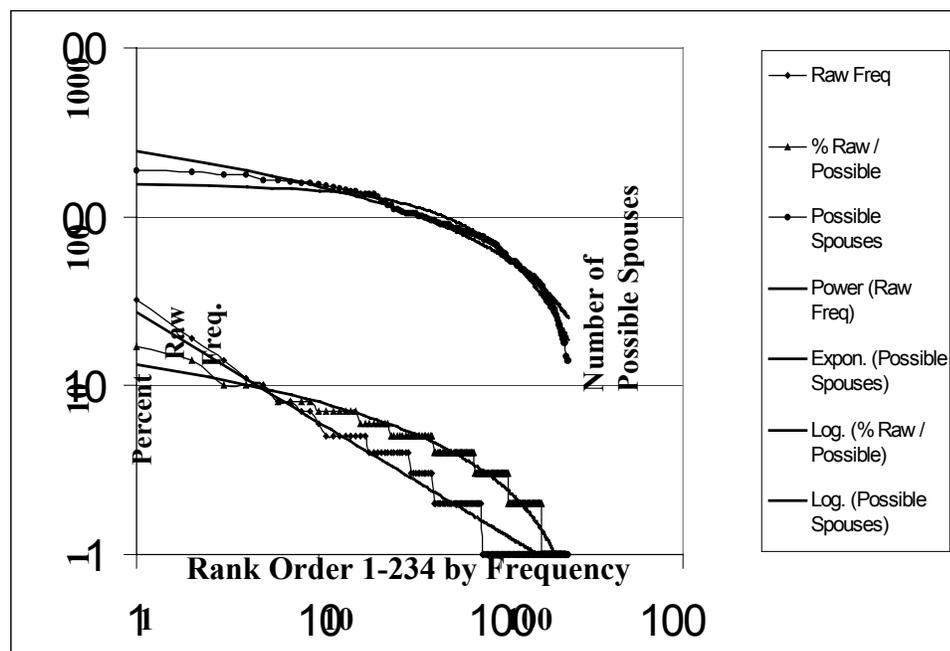


These results can be compared with those of plotting method in 4.1, as shown in Figure 6. Here the raw frequency line shows the outcome of graphing the frequencies of all the 234 types of consanguineal marriages up to seventh cousins ordered by rank with logged frequencies on the y axis and the logged number of observations for this type and frequency on the x axis. The distribution of raw frequencies is linear in the log-log graph and thus approximates a power-law distribution, fitting our overall observation about a fractal marriage pattern. The other aspects of consanguineal marriage distributions graphed by this method are the number of possible spouses of each type in the upper part of Figure 6 and the percentage married of each type in the lower part.

The upper curve for frequencies of types of possible spouses (all those available in a given category) shows an exponential decay or logarithmic distribution. Here FaBrDa is the most frequently available type of relative, and MoBrDa the next. The curve for percentage married of each type of those available is also a logarithmic distribution, unlike Figure 4, again with FaBrDa as the highest percentage and MoBrDa the next. The logarithmic shape is due to the fact that there are many fewer *types* of consanguines at each kinship distance as we move closer to ego (four types of first cousins). Also, in a limited network, as we move to very distant relations that thin out if there are few apical ancestors. This is because many of the vast number of combinatorial possibilities do not occur, and the closer relationships have already

used up many of the relatives in the network. Only the raw frequencies fit the power-law distribution that is characteristic of fractality. This is as it should be in a system that is behaviorally self-organizing. In this society, FaBrDa is not only a preferred marriage but a relative whose proximity is more likely because of the preference for brothers and their children to live together, stay together and work together.

**Figure 6: The Fractality of Consanguineal Marriage Frequencies**



Analyses of relinking marriages that are nonconsanguineal for the case of the Aydınlı nomad clan, comparable to those of consanguineal marriages, do not show power-law tendencies, consistent with H1 and H3.

### Ring Cohesion Reduction Results for the Aydınlı Turkish Nomads

The ring cohesion results on preferential distributions for the nomad clan are that although FaBrDa marriage is the most frequent consanguineal marriage, it is not essential to the cohesion of the clan. It is the class of consanguineal marriages that generate this cohesion; they are hypothesized to be the set of cycles we should consider as independent.

**H5 (Aydınlı):** The elimination of FaBrDa marriage does not reduce structural endogamy.

In Table 1, the structurally endogamous bicomponent was computed prior to adding sibling links as edges. The ring cohesion calculus was used to make Table 2 by find-

ing FaBrDa marriages in the reductive analysis database and deleting only the FaBrDa marriage links while retaining the rest of the graph. The bicomponent is then recomputed after the removal of sibling links. These comparisons make a single important point: no marriage in the bicomponent is lost by removing the FaBrDa marriage links.

Table 1

nodes in bicomponent = 960	Arcs	Edges
Number of lines with value=1	248	2573 (p-tur-bico)
Number of lines with value#1	1749	0

Table 2

FaBrDa: 34 marriages found and deleted  
After removal of FaBrDa and sibling edges

nodes in bicomponent = 960	Arcs	Edges
Number of lines with value=1	214	2573
Number of lines with value#1	1749	0

What if all consanguinal marriages are eliminated?

**H6 (Aydınlı):** The elimination of all consanguineal marriages also eliminates almost all of the two-family and other affinal relinkings.

H6 holds for all but 11 marriages between pairs of siblings: once the consanguineal marriages between first and second cousins and the consanguineally linking relatives are eliminated, the three types of sibling relinkage are reduced from frequencies of 18 – 42 – 75 to frequencies of 3 – 3 – 5.<sup>21</sup> None of the other two-family relinkings have any significant frequency.

Table 3 shows calculations of the independent cycles in the marriage graph for close blood marriages compared to the total genealogical network. There are 116 independent cycles accounted for by these consanguineal marriages out of a total of 266, leaving 150 unaccounted for.

Table 3. Independent Cycles in the First and Second Cousin Marriage Graph compared to the Total graph

**In the cousin fusion graph**

Number of lines                    315 - (no edges)  
Number of nodes                    176  
-----                    139 - 11 - 11 independent cycles = 116

**In the total graph**

Number of lines                    1997 (no edges)  
Number of nodes                    960  
-----                    1037 - 1 - 771 independent cycles = 266

There are, however, as in Figure 4, nearly a hundred and fifty other consanguineal marriages with frequencies of one or two per type, and consanguineal marriages plus the 11 sibling set marriages are sufficient in number to account for all the structural endogamy.

The ring cohesion calculus in the illustrative case of Turkish nomads is consistent with consanguineal marriages as the source of relinking between and almost fully accounts for the structural endogamy in the network without having to take into account multifamily relinking. FaBrDa marriage per se accounts for none of the structural endogamy, which is entirely due to a variety of other consanguineal marriages, most of which are between lineages. In this case, these results speak to the fact that there are two major poles to the marriage structure, one being widely distributed consanguineal marriage choices within the clan, many of them with distant consanguineals. The other is that close marriages that reinforce lineage cohesion, such as FaBrDa, are easily available with the high frequency of such coresident women available, and these also account for the highest percentage of marriages with those in any given marriage type. Given that the size of the consanguineal marriage cycles is sufficient to account for all the independent marriage, the nonconsanguineal relinkings may be considered for the most part as nonindependent cycles and by-products of consanguineal marriages, as in Figure 1.

### Conclusions

For the analysis of networks generally, ring cohesion theory provides a solution to problems of studying cycle formation and the relation between local cohesion through behavior that leads to relational cycles and more global properties of network cohesion. For the analysis of marriage networks in particular, ring cohesion theory provides a solution to problems that have long plagued this field (Schneider 1965): How to count cycles? How do marriage cycles contribute to social cohesion? How to measure cohesion in a genealogical network? How to define the boundaries of a cohesive community? How to determine preferential marriages? How and to what extent do preferential marriages contribute to structural endogamy? What is the relation between social cohesion and structural endogamy?

The Pajek network program provides operational procedures to address these issues, some of which have been exemplified and examined in the Turkish nomad example of an Arabic type of kinship and marriage system. In that case a preferential marriage gradient is evident for consanguineal but not for nonconsanguineal relinking. Removing the consanguineal marriages as a set of marriages large enough to constitute the total independent set of marriage cycles gives a confirmatory result in that nonconsanguineal relinking cycles are removed as well, as the theory stipulates.

The theory is highly useful, then, in bounding the problems of explanation and prediction in developing social theory for marriage networks, and may be easily extended to networks in general for the study of sources of cohesion.

As a more general theory applied to social networks, the problem of ring cohesion is isomorphic to that of behavioral gradients: what are the local contexts in which individuals make choices that affect the larger issues of social cohesion by the formation of cycles in networks? This is a micro-macro problem: choices made locally in a network to form links either do or do not form cycles. Pairs of cycles, when concatenated, form indirectly, out of the view or intent of the local actor, and not as a direct result of local action. What ring cohesion offers is a set of measures and theorems that allow independent sets of cycles to be hypothesized for any given network and matched against actual behavior generating them, while the nonindependent set of cycles (typically longer) that necessarily occur in the larger network are only a byproduct of that behavior. The nonindependent cycles, however, while formed ‘behind the backs’ of local actors, are equally a part of the structural cohesion of the larger structures of embedding in the network. These larger macro structures and their nonindependent cycles do not require separate explanation except as a necessary byproduct of micro-macro linkage. Further, because they are not statistically independent of the independent set of cycles they do not bear statistical analysis alongside of that set. Except for very unusual circumstances, it is toward the independent set of cycles that analysis of cycle formation must be focused. Which set of cycles should be considered independent, however, is a matter of identifying preferential or gradient-driven social action. The guiding hypothesis is that the rings of cohesion that form as a result of local action are local rings, and easily recognizable locally (see Brudner and White 1997). These locally definable contexts are the first places one should look from gradient-driven or preferential behavior. Failing that, or for exceptional individuals who are exploiting structural holes in networks (Burt, 1995) to form long but simple cycles, one might look for intentional behavior in the outer rings of cohesion, which are typically nonindependent of the closer rings. It is the inner/outer ring differentiation that helps to understand how history – in the sense of the effects of cohesive groupings – is made behind one’s back, often as a result of micro-macro linkages.

### **Acknowledgements**

Partial funding for this project was supported by ISCOM, Information Society as a Complex System, an EU grant to David Lane, Sander van der Leeuw and Geoffrey West, with a subcontract to the author, and research assistance from Joseph Wehbe. The first presentation of the ideas of ring cohesion theory in summer 2003 benefited from support of the Santa Fe Institute for the annual ISCOM meeting. Andrej Mrvar graciously responded to my subsequent request to enlarge the fragment-finding ca-

pabilities of Pajek by automating the ring-cohesion procedures, extending the macros, commands and options available, and finally programming a repeat-macro function. The second presentation of ring cohesion theory, specifically in the context of kinship analysis with Pajek, was made possible by an invitation from Sorbonne historians, specifically Cyril Grange and his colleagues, to give a three day Pajek workshop in Paris in June 2004 on kinship network analysis. That workshop was held jointly with members of the TIME (Traitement Informatique des Matériaux Ethnographiques) research group in anthropological kinship analysis headed by Laurent Barry and Michael Houseman. Subsequent meetings of that group in Paris provided the final breakthroughs needed for this method to be applied to kinship, and leading to the paper by Hamberger et al. (2004).

## Bibliography

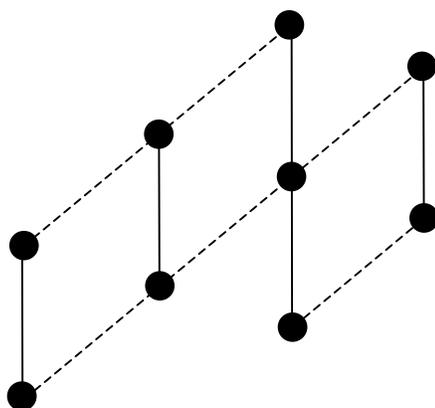
- Batagelj, Vladimir, and Andrej Mrvar. 1998. Pajek—A program for large networks analysis. *Connections* 21 (2): 47-57. University of Ljubljana. <http://vlado.fmf.uni-lj.si/pub/networks/pajek>. Program descriptions are at <http://vlado.fmf.uni-lj.si/pub/networks/pajek/pajek.pdf> <http://vlado.fmf.uni-lj.si/pub/networks/pajek/sunbelt.97/pajek.htm> <http://vlado.fmf.uni-lj.si/pub/networks/doc/sitges.pdf>.
- . 2001. A subquadratic triad census algorithm for large sparse networks with small maximum degree. *Social Networks* 23: 237-243
- . 2002. Pajek—Analysis and visualization of large networks. *LNCS 2265: Graph Drawing, 9th International Symposium*, 477-478. Berlin: Springer.
- . 2003. Pajek analysis and visualization of large networks. Preprint, to be published in *Graph Drawing Software*. Edited by M. Junger and P. Mutzel, Springer series *Mathematics and Visualization*. <http://www.ijp.si/ftp/pub/preprints/ps/2003/pp871.pdf>
- Brudner, Lilyan A., and Douglas R. White. 1997. Class, Property and Structural Endogamy: Visualizing Networked Histories. *Theory and Society* 25:161-208.
- Bertin, Jacques. 1983. *Semiology of Graphics*, Trans. William J. Berg; Madison, WI: University of Wisconsin Press.
- Burt, Ronald S. 1995. *Structural Holes: The Social Structure of Competition*. Harvard University Press, paper back edition.
- Das, Tarakchandra. 1945. *The Purum*.
- de Nooy, Wouter, Andrej Mrvar, and Vladimir Batagelj. 2004. *Exploratory Social Network Analysis with Pajek*. New York: Cambridge University Press.
- Goldenweiser, Alexander A. 1913. The origins of totemism. *American Anthropologist* N.S. 14: 530-562.
- Hamberger, Klaus, Douglas R. White, Michael Houseman, Isabelle Dailent, and Raymond Barry. 2004. Second order networks of marriage patterns. Ms.
- Hèran, François. 1995. *Figures et légendes de la parenté*. Paris: INED.
- Houseman, Michael, and Douglas R. White. 1998a. Network mediation of exchange structures: Ambilateral sidedness and property flows in Pul Eliya. In *Kinship, Networks and Exchange*. Edited by Thomas Schweizer and Douglas R. White, 59-89. Cambridge: Cambridge University Press.
- . 1998b. Taking sides: Marriage networks and Dravidian kinship in lowland South America. In *Transformations of Kinship*. Edited by Maurice Godelier, Thomas Trautmann, and F. Tjon Sie Fat, 214-243. Washington, D.C.: Smithsonian Institution Press.
- Johansen, Ulla C., and Douglas R. White. 2002. Collaborative Long-Term Ethnography Study and Longitudinal Social Analysis of a Nomadic Clan in Southeastern Turkey. In *Chronicling Cultures: Long-Term Field Research in Anthropology*. Edited by Robert Kemper and Anya Peterson Royce, pp. 81-99. Walnut Creek, CA: AltaMira Press.
- Jola, Tina, Yvonne Verdier and Françoise Zonabend. 1970. Parler famille, *L'Homme* 10(3): 5-26.
- Jorion, Paul. 1984. L'inscription dans la structure de parenté: Omicar? *Revue du Champ Freudean* 31:56-97.
- Lévi-Strauss, Claude. 1958. *Structural Anthropology*. Allen Lane, The Penguin Press., 1968. (reprinted 1968, Allen Lane, The Penguin Press).
- . 1963. Linguistics and Anthropology. *Structural Anthropology*, 67-97. New York: Basic Books.
- Moody, James, and Douglas R. White. 2003. Social cohesion and embeddedness: A hierarchical conception of social groups. *American Sociological Review* 68: 1-25. <http://www.santafe.edu/sfi/publications/Working-Papers/00-08-049.pdf>.
- Mrvar, Andrej, and Batagelj, Vlado. 2004. Relinking Marriages in Genealogies. *Metodoloski zvezki - Advances in Methodology and Statistics*, 1, 407-418. Edited by Anuška Ferligoj and Andrej Mrvar. Ljubljana: FDV.
- Ore, Oystein. 1960. Sex in graphs. *Proceedings of the American Mathematical Society* 11:533-539.
- Schneider, David. 1965. Some muddles in the models: Or how the system really works. In *The Relevance of Models for Social Anthropology*. Edited by Michael Banton, 25-85. ASA Monograph No. 1. London: Tavistock.
- Weil, André. 1969[trans. of 1949]. On the Algebraic Study of Certain Types of Marriage Laws, Appendix to Part One, Claude Lévi-Strauss, *Elementary Structures of Kinship*.
- White, Douglas R. 1997. Structural endogamy and the network *graphe de parenté*. *Mathématiques, informatique et sciences humaines* 137: 101-125. <http://eclectic.ss.uci.edu/nsh/str-endo.htm>
- . 1999. Controlled simulation of marriage systems, *Journal of Artificial Societies and Social Simulation* 3(2). <http://jasss.soc.surrey.ac.uk/2/3/5.html>
- . n.d. (accepted for publication and undergoing author revision). The Logic of Avoidance. *Social Networks*.

- White, Douglas R., Vladimir Batagelj and Andrej Mrvar. 1999. Analyzing large kinship and marriage networks with Pgraph and Pajek. *Social Science Computer Review* 17: 245-274.
- White, Douglas R., and Frank Harary. 2001. The cohesiveness of blocks in social networks: connectivity and conditional density. *Sociological Methodology* 2001, Vol. 31: 305-359. Boston: Blackwell Publishers.
- White, Douglas R., and Michael Houseman. 2002. The navigability of strong ties: small worlds, tie strength and network topology. *Complexity* 8(1): 72-81. <http://eclectic.ss.uci.edu/~drwhite/Complexity/K&C-a.pdf>
- White, Douglas R., and Michael Houseman., and Thomas Schweizer. 1993. Kinship, social biography and social change database (serially updated).. <http://eclectic.ss.uci.edu/~drwhite/PDATASET.htm>
- White, Douglas R., and Ulla Johansen. 2004. *Network Analysis and Ethnographic Problems: Process Models of a Turkish Nomad Clan*. Lanham, MD: Lexington Press.
- White, Douglas R., and Paul Jorion. 1992. Representing and analyzing kinship: A network approach. *Current Anthropology* 33: 454-462.
- \_\_\_\_\_. 1996 Kinship networks and discrete structure theory: Applications and implications. *Social Networks* 18: 267-314.

Source: Ftp/Sorbonne/RingCohesionMarriage.doc

all links are live in pdf

Save this image for use with footnote graphic



## Notes

<sup>1</sup> Empirically oriented Anglo-American anthropologists tended to shy away from such studies following Schneider's critique, while French anthropologists continued to study kinship and marriages networks using the vocabulary of a logico-deductive framework of structural thinking and without adopting a network framework for the empirical analysis.

<sup>2</sup> Terms given formal definition are in italics. It is the definitions that lend precision to graph theoretic and network analysis, and that allow theorems to be proven.

<sup>3</sup> The p-graph convention of marriages as nodes was adopted by Weil (1949) for algebraic representation of kinship models, generalized by Bertin (1983) to genealogies, continued by Jorion (1984) for broader classes of kinship models, and regeneralized by White and Jorion (1992, 1996) to genealogies. P-graphs do not distinguish half-siblings because when the same parent is in two different marriages the half-sibling relationship resembles that of cousins. To distinguish half-siblings it is preferable to use the bipartite p-graph format in which individuals are one set of nodes and couples another. The alternative is to use a more conventional graph with individuals as nodes. The Ore graph (Ore 1960)

---

is one such graph, with arcs from parents to children, and no marriages, but individual nodes may be further distinguished by sex.

<sup>4</sup> Jola, Verdier and Zonabend (1970).

<sup>5</sup> Sister exchanges, like many symmetric 2-family relinkings, could also be counted as two marriages of the same type. The counts used here eliminate such symmetries.

<sup>6</sup> In the Ore (1960) graph, individual nodes are distinguished by sex.

<sup>7</sup> The computation of the total number of cycles in a graph is complicated, but can be determined from how pairs of cycles overlap in a maximal set of cycles. If two independent cycles share an edge, they create one dependent cycle. If they share two nodes but no edge they generate four dependent cycles, and so forth. There is no easy formula.

<sup>8</sup> In a bicomponent of a p-graph with  $m$  marriages, every marriage with two parental nodes is a relinking, and  $m \geq r$ , but as (b) in Figure 3 shows,  $m$  does not always equal  $r$ , and the total number of cycles may be much greater than  $m$ .

<sup>9</sup> The proof is obvious: remove  $r = k - n + d$  edges from a graph with  $k$  edges, and  $k - r = n + d$  edges remain; if both the initial and final graphs are connected, then the final graph is a tree if  $d = 1$  and a set of  $d$  trees otherwise.

<sup>10</sup> Pajek computation of bicomponents is done by Net/Component/Bicomponents. The results are posted in the Hierarchy window that must be clicked to see the hierarchy of possibly overlapping bicomponents. Two bicomponents may share at most one node in common and no set of bicomponents may be connected by a cycle of overlapping bicomponents. To see the numbers of nodes nested in different bicomponents, click the root that appears in the Hierarchy window. The Net/Component/Bicomponents command also generates two partitions in the partitions window. The first identifies vertices belonging to exactly one bicomponent by their number, with nodes that are not in any of the bicomponents assigned to partition zero and with nodes of intersection (articulation points) between bicomponents assigned to partition 99998. The second partition identifies articulation points, assigning each the number of bicomponents in which they are members. To select the subgraph consisting of nodes in any of the bicomponents, Operations/Extract from network/Partition, entering a minimum value of 1 and a sufficiently large maximum value will result in the reduced subgraph shown in the network window.

<sup>11</sup> In Pajek, Nets/First network is used to select the main network and Nets/Second network the network with edges or arcs to be subtracted. Nets/difference will subtract the arcs and edges in the second from the first, leaving all the nodes.

<sup>12</sup> This section and the final definition of matrimonial rings is a joint work with Klaus Hamberger and others in the ongoing collaboration of the Parisian research group TIME.

<sup>13</sup> A bipartite p-graph is an alternative network formalization for this purpose since it has individuals as well as marriage nodes, and removal of a marriage consists of replacing the marriage node with descent lines from individual parents to individual children. As this is done, however, cycles are created which do not correspond to marriage cycles, so a variant of the correction formula for Ore (1960) graphs must be applied to compute the number of remaining marriage cycles.

<sup>14</sup> An approximation of this number is computed in an Ore graph by deleting sibling edges and spousal links and then computing the indegree of each node, which is saved as a partition, and then computing sibling links using Net/Transform/Add/Sibling Edges/Input. Then delete all arcs using Net/Transform/Remove/all Arcs. Now use the previously saved (and active) indegree partition to select Operation/Extract from Network/Partition/values from 1 up, and use Net/Partition/Degree/Input degree to compute the numbers of sibling links on the edges and again Operation/Extract from Network/Partition/values from 1 up and again. These sibling sets must be counted for each size group, labeling them  $k=1$  to  $(s-1)$ , where  $s$  is the largest number of siblings, and adding the products of each frequency by its index  $k$ . The reason this is approximate is that there are some half-siblings.

<sup>15</sup> The procedure is to make an \*.xls file with Ego#, Name, Sex (H=homme, F=femme), Fa#, Mo#, Spouse#, sort by sex, delete all the spouse numbers for males (to make the spousal relation directed), then run the Gen2Pajek program to create a Pajek \*.net file. This file requires the addition of sibling links, which is a Net/Transform/Add operation (Input option for equivalence of siblings) in Pajek.

<sup>16</sup> The procedure is to read the fragments as a \*.paj file (File/Pajek Project file/Read), then load the genealogy file made with Gen2Pajek, and use option Nets/Second network to load it for comparison. The network point is then reset to the first fragment, option Nets/First network enables it to be found in the first by subsequent option Nets/Find (1 in 2). Options must be set in Nets/Find to [x] check values of lines [x] Extract subnetwork and [x] Retain all vertices after extraction – this will allow the found marriages to be subtracted from the genealogy. If rings are wanted then the option must be checked for [x] induced subgraph. Otherwise, if [ ] induced subgraph is left unchecked then the cycle in the fragment will be found whether or not the subgraph identified as containing this cycle has additional links that are not in the fragment.

<sup>17</sup> Because the marriages have line value 1, it is only necessary in the extracted subnetwork to use option Net/Transform/Remove/lines with value/more than [1] and Net/Transform/Remove/all edges in order to create a net-

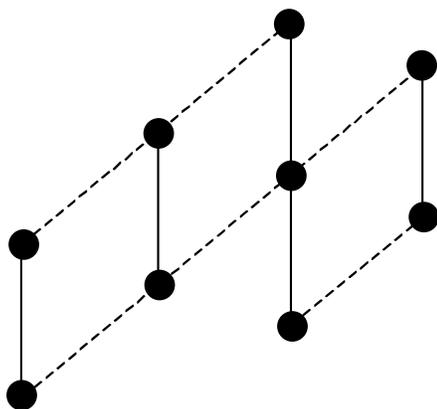
---

work file with the marriage links only. This is then subtracted by loading the genealogy file with option Nets/First network and load the marriage file to be subtracted with Nets/Second network, so that Nets/difference subtracts the marriages.

<sup>18</sup> Hamberger developed an additional series of Pajek macros to calculate from this file into a 3-mode dataset with individuals, marriages, and marriage types. This will be treated in a separate article (Hamberger et al. 2004).

<sup>19</sup> While this is a technical point that bears on kinship structures, an explanation is needed both for the technical side of kinship analysis and for the larger framework of ring cohesion analysis in social networks. The figure below exemplifies the point: here there are three MoBrDa marriages, two that share an edge and two that share a node. There are three independent cycles, and while graph addition for the two that share an edge would form a MoMoBrDaDa marriage ring, this ring would not be found as a ring-fragment of the graph because of the requirement that the search is for a simple cycle, not a pair of cycles.

---



---

<sup>20</sup> Kinship distance can be reckoned in various ways, one of which is the shortest path from individual A to B in the reductive analysis kinship format. Other measures are more culturally specific.

<sup>21</sup> These types involve two brothers with two sisters, and sister exchange.