Acta Botanica Hungarica 30 (3—4), pp. 363—402 (1984)

SPATIAL DEPENDENCE OF PLANT POPULATIONS

PART 2. A FAMILY OF NEW MODELS*

P. JUHÁSZ-NAGY

DEPARTMENT OF PLANT TAXONOMY AND ECOLOGY, EÖTVÖS L. UNIVERSITY, BUDAPEST, HUNGARY

(Received: 20 March 1984)

This paper attempts to show that equivalence analysis belongs to a family of models, where certain basic concepts and corresponding statistical estimates play a key role in translating propositions of one construction to another. This kind of semantic translation is important, because it guarantees the more or less unified treatment of such different basic phenomena as diversity, resemblance, preference, etc., important in vegetation research.

VIII. Introductory notes

VIII.1 In the first part of this series of papers, an operational extension of *the attribute duality principle* and a related methodology, called *equivalence analysis* have been outlined (JUHÁSZ-NAGY 1976). The main purpose of such analyses is to study "*orderliness*" of different kinds.

VIII.2 As it has already been emphasized, however, orderliness of any kind depends on a number of conditions (e.g. due to sampling, evaluation, etc.) to a considerable extent. Clearly, several further aspects of study are to be clarified, with special reference to *spatial processes* (see VII.). The main aim of this paper is to present a collection of constructions (models), rich enough to cope with at least a part of our difficulties due to "context-dependence", viz. model-dependence (see 0.).

VIII.2.1 The primary purpose of the present paper is to show that even if *the simplest possible conditions* are applied (e.g. *binary* variables are used, *classical* probability fields are assumed, etc.), several types of arguments are possible to model different aspects of study (like *diversity*, *resemblance*, etc.). Note further that, due to strange historical reasons, *the mutual interpretation* of the basic phaenomena in most cases is still very difficult. (For instance, during the development of *numerical syntaxonomy*, some type of *resemblance* has always been "overemphasized", while *diversity* has been practically neglected for a long period.) Therefore, the secondary purpose of the present paper is to show how these simple constructions are related inherently to each other, forming a family (or, sometimes, even a system) of models, where the meaning of one model can safely be translated to the meaning of the other model.

VIII.2.2 This kind of mutual interpretability is important, because at the present state of vegetation research *some "semantic gaps"* seem to be even more alarming than beforehand. There exists, for instance, a huge "*terra incognita*" between syntaxonomy and synmorphology (e.g. pattern analysis). If we can say that the most difficult and pertinent problems of vegetation research (such as a deeper understanding of succession and degradation) lie *in the interface of syntaxonomy, synmorphology, and syndynamics*, represented by the inter-

* See Part 1. in Acta Bot. Acad. Sci. Hung. **22:** 61–78 (1976).

section of Fig. 1, then all efforts should be made to clarify that interface as intensely as we can. In order to sharpen this statement even further, suppose that we want to understand in the future *dendrogenesis* [in terms of ERDŐS and RÉNYI (1960): a type of "*graph evolution*"], where *a dendrite* (say, a cluster dendrite) changes in space, or in time, or in a proper spatio-temporal referential system (see Fig. 2). Because any dendrite (a topological tree) is a certain

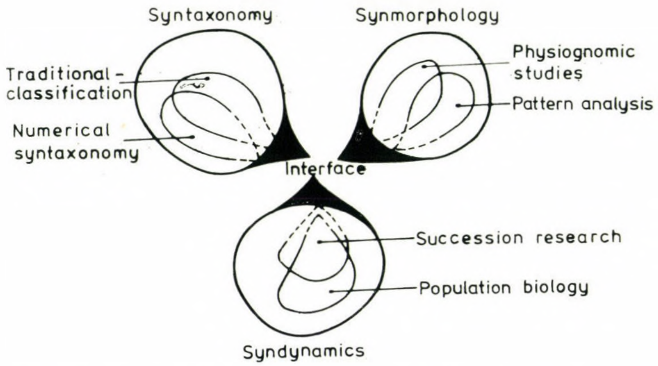

Fig. 1. The importance of three subdisciplines and their interface (see VIII.2.2)

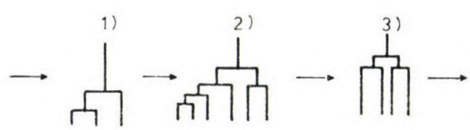

Fig. 2. A possible dendrogenesis (see VIII.2.2)

type of structure, and because a cluster dendrite results in some taxonomic algorithms, the study of a cluster dendrite belongs simultaneously to the realms of both synmorphology and syntaxonomy. Clearly, any prudent interpretation of some dendrogenesis can be made possible, if these realms will be properly interconnected with the third realm of syndynamics (i.e. *the "processing aspect"* of the problem). At the present stage of ignorance, it is very difficult to answer the simple question: at what size of plots a certain (say, maximum) degree of interconnectedness might be expected?

VIII.3 This paper deals with problems simpler than dendrogenesis, and has some carefully chosen other limitations as well (see part 1; VIII.2.1; IX.; etc.). Only models interpretable in terms of information theory are used, and by the concept "information" is always meant *contingency information* (*sensu* KULLBACK 1959). The terminology tries to follow some of the best foundations in this field (e.g. KHINCHIN 1959, RÉNYI 1962, ACZÉL and DARÓCZY 1975). The attention of the reader is called to some remarkable works (like PIELOU 1974, 1975, 1977; ORLÓCI 1978a), and to a book under preparation (FEOLI, LAGONEGRO and ORLÓCI 1984) that are concerned with several important problems of such modelling in ecology, or in vegetation research.

This paper is dedicated to the memory of prof. A. RÉNYI, who had first encouraged the author to play with information theory functions.

[A] INITIAL FRAMEWORK OF CONCEPTS

IX. Classes of concepts: basic concepts

IX.1 Let Q be a set of plant populations (labelled usually, but not necessarily as "species"), $Q = \{q_1, q_2, \ldots, q_i, \ldots, q_s\}$, living in a τ topographicum. Let T_j be a set of planar sampling units (plots) with j-specification (say, for shape and size), $T_j = \{t_{j1}, t_{j2}, \ldots, t_{jg}, \ldots, t_{jm}\}$, to be laid-out in the area τ', $\tau' \subset \tau$, $t_{jg} \subset \tau'$. By projecting the bodily points of q_i into τ', and gaining the projection set q_i', $q_i' \subset \tau'$, the binary function (an *elementary floristic function*) can be defined as

$$n_{ijg} = \begin{Bmatrix} 1, \text{ if } q_i' \cap t_{jg} \neq \emptyset, \\ 0, \text{ if } q_i' \cap t_{jg} = \emptyset, \end{Bmatrix} \qquad \text{(IX; 1)}$$

where $n_{ijg} = 1$ is called *coincidence*, and $n_{ijg} = 0$ is called *incoincidence*. [Clearly, (IX; 1) can be regarded as *an event*, like that of a BERNOULLI trial, with two elementary outputs. The usual notion of "presence — absence" is avoided, because these concepts should be reserved for special cases, like for "presence in a stand".]

IX.1.1 From a theoretical point of view, Q and τ' are regarded as *universal sets* (sets of *all important points*), while sets $q_1', q_2', \ldots$, and sets $t_{j1}, t_{j2}, \ldots$ are *sets of sampling points* by which some statistical relations can be obtained (say, with respect to the DESCARTES space, $Q \times Q$). In this way of reasoning, $q_i \in Q$ is a *universal relation* (presence), while $q_i \in \Phi_{ij}$ (coincidence) is a *sampling relation*, where φ_{ij}, $\varphi_{ij} \subset Q$, is *the flora* (or, florula, if you like it better) of plot t_{jg}.

IX.1.2 Since elements of T_j correspond to the elements of the class (set of subsets) Φ_j, $\Phi_j = \{\varphi_{j1}, \varphi_{j2}, \ldots, \varphi_{jm}\}$, the results of survey sampling can be arranged in *a binary contingency table*, to be called interchangeably either *an elementary table*, [ET], or a K_j (as in part 1., or in X. and the sequel, if a series of such tables is needed). (IX; 2) gives a very simple illustration of an [ET], where $s = 4$, $m = 8$, $Q = \{u, v, w, z\}$,

u	0	0	1	1	1	0	0	1	4
v	1	0	1	1	1	1	0	1	6
w	0	1	0	0	0	0	1	1	3
z	0	0	0	0	0	1	1	0	2
	1	1	2	2	2	2	2	3	15

(IX; 2)

IX.2 It is fairly important to regard even such a primitive thing as (IX; 2) from the point of view of *classes of concepts*.

IX.2.1 For the sake of simplicity, suppose that we distinguish now only three classes of concepts. The class of *fundamental concepts* (like set, relation, vector, function, graph, etc.) consists of such mathematical terms that are all-important, everywhere in science. The class of *methodological concepts* (like sampling unit, survey sampling, a statistical test, curve fitting, etc.) consists of such terms that make workable (operational) fundamental concepts in a given context. The class of *basic concepts* (like flora, vegetation, floral diversity, interlocally random composition, etc.) consists of terms of a discipline of its own (in our case: vegetation research). We regard a basic concept *operational*, if it is resulted in a proper interconnection of some fundamental *and* methodological concepts. (Of course, it is perfectly true that such "descriptive terms" as formation or habitat, might be as powerful in their own methodological place and role as operational concepts.)

IX.2.2 A basic concept of the operational type has a definite value, if it is somehow *a member of an operational framework of concepts*. In order to illustrate such a framework, let us set up the following series or *ordered concepts:*

$\langle 1\rangle$ *local vectors* (row vectors) of an [ET],

$\xrightarrow{1}$

$\langle 2\rangle$ *interlocal relations* between row vectors of an [ET],

$\xrightarrow{2}$

$\langle 3\rangle$ *interlocal dependence* between two populations of set Q,

$\xrightarrow{3}$

$\langle 4\rangle$ *interlocal association* between two populations of Q (or, pairwise comparison of all elements of Q),

$\xrightarrow{4}$

$\langle 5\rangle$ *interlocally significant association* based on conditions of $\langle 4\rangle$,

$\xrightarrow{5}$

$\langle 6\rangle$ *plexus graph*, a random graph whose points (elements of Q) are interconnected only, if conditions for significant association are satisfied at given probability level(-s),

$\xrightarrow{6}$

$\langle 7\rangle$ *plexus group*, a subset of Q, generated by *a clicque*, a stochastically separate subgraph of the plexus graph.

IX.2.2.1 It can be seen at once that these concepts are really ordered, because any concept is meaningful only with respect to the former ones. It is true even for $\langle 1\rangle$, for which the concepts of IX.1 give meaning [i.e. *arrangement* of values of (IX; 1) in a proper order]. Note, however, that this chain of mutual interpretability is realized because of 1, . . ., 6, a number of operations, or, algorithms. 1 means only to set up *some* relations on $Q\times Q$, or, even, on $\underset{i=1}{\overset{n}{\times}} Q_i$. 2 implies the essential novelty of *independence hypothesis*, by which interlocal dependence is defined by such relations as $<$, $>$, $\leq$, etc. (in general, by "$\neq$"). 3 gives *a function* (e.g. chi^2, contingency information, etc.) whose estimated values are supposed to characterize the *deviations* from hypothetically independent situations. (4) evaluates these deviational estimates in given conditions (e.g. at certain *probability levels* of confidence). (5) sets up a number of rules (in particular, if the "plexus" is *a coloured graph*) in order to make the construction possible. Finally, (6) gives some "*cutting algorithm*" for a stochastic partitioning of the graph into more or less separable subgraphs.

IX.2.2.2 Order and correspondence in IX.2 shows clearly that such a series of concepts and operations is almost always needed. It is obvious again that the hard-boiled *dominance of traditional names*, such as the devaluated misuse of "*interspecific correlation*" for any kind of interlocal relations (even if, *not r*, correlation coefficient, was estimated) might be extremely misleading. Note, further, that the adjective "interspecific", in a sense, would have been attached to the other type of relations (i.e. *interfloral* relations); the present connotation is due to a rather unfortunate convention. It is slightly ridiculous that *attribute duality* of [ET] is usually stated by means of such dubious jargon-terms (like "number of species", "RAUNKIAER frequecy", Q-technique, R-technique etc.) that *have no true suggestion whatever* — in particular, for outsiders, or beginners — *on the true nature of dual relations.* (When the author has recently asked a dozen of his friends, working as applied mathematicians in the field of multivariate analysis, nobody was aware of the mysterious meaning of "Q", or "R"; one mathematician called them "*obscured free-masonic symbols* originated from the proto-chaos of science".)

IX.3 This is the main reason why the author has decided to substitute some traditional terms by such basic concepts which are, at least potentially, able to cope with the difficulties of interpreting dual relations.

IX.3.1 Without repeating the details of notation in II., (IX; 3) gives the highlights, the simplest list of terms for an [ET], or for a K_j of $s \times m$ size:

		plots		Σ 1	Σ 0
populations		local vectors n_{ijg}	floral vectors	(1) local valences	(2) local invalences
Σ	1	(3) floral valences		N_j	n_j
Σ	0	(4) floral invalences		n_j	sm

(IX; 3)

IX.3.2 First, note the difference among the three adjectives: (a) "*local*", (b) "*floral*", (c) "*floristic*". (a) and (b) is related always to some *row* and *column properties* of a K_j, respectively, while (c) involves *a joint relational meaning* between some row *and* column attributes, simultaneously. This is why K_j itself (as a whole binary data structure) can be regarded as a binary representation of some *floristic composition*, and why frequency distributions (1)–(4) in (IX; 3) *altogether* can safely be called *floristic marginals*, or, in brief, *f-marginals*. (Note, further, that, *mutatis mutandis*, a similarly simple semantics may be applied to such adjectives as "local", "*faunal*", "*faunistic*" as well.)

IX.3.3 Using the numbers in (IX; 3), *f-marginals* are symbolized as (1) $V_j^{(q)}$, (2) $v_j^{(q)}$, (3) $V_j^{(t)}$, (4) $v_j^{(t)}$. The symbols (5) N_j and (6) n_j refers to *total valence* and *total invalence* of a K_j, respectively. The entropy *estimators* for these quantities are, in order: (1) A_j, (2) a_j, (3) B_j, (4) b_j, (5) C_j, (6) c_j, where, for instance, $c_j = n_j \log n_j$. (Without further comment, "log" always means "$\log_2$", and numerical values will be given in Bits, i.e. properly *weighted bits*.) In the sequel, we need the following simple generalizations:

$$\left.\begin{aligned} A_j + a_j &= \alpha_j, \\ B_j + b_j &= \beta_j, \\ C_j + c_j &= \gamma_j. \end{aligned}\right\} \qquad \text{(IX; 4)}$$

IX.4 As beforehand, relations between *local* and *floral* vectors are called *interlocal* and *interfloral* relations, respectively.

IX.4.1 In order to illustrate some main points, let us consider two exceedingly simple examples, where $s = 3$, $m = 2^s = 8$, $Q = \{u, v, w\}$,

$$\begin{array}{c|cccccccc|c} u & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 4 \\ v & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 4 \\ w & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 4 \\ \hline & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 12 \end{array} \qquad \text{(IX; 5)}$$

$$\begin{array}{c|cccccccc|c} u & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 4 \\ v & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 4 \\ w & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 4 \\ \hline & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 12 \end{array} \qquad \text{(IX; 6)}$$

IX.4.1.1 It is easy to see that all the three pairs of populations in (IX; 5) are *interlocally independent*, therefore, (IX; 5), or, any similar construction, like (V; 8), might be called *interlocally random composition*, [IRC]. On the other hand, in (IX; 6) no pair of populations is independent, in an interlocal sense, and this kind of construction might be called an *interlocally dependent composition*, [IDC].

IX.4.1.2 It goes without saying that, while an [IRC] is *unique* (for a given s), [IDC] may be of very different kinds. We guess at first sight that (IX; 6) has, in many respects, a greater degree of orderliness than (IX; 1).

IX.4.2 This sort of comparison is possible at a first instant by estimating *floral diversity* (JUHÁSZ-NAGY 1976), or, as JUHÁSZ-NAGY and PODANI (1983) call it, by *florula diversity*, i.e. by comparing *possible* and *realized floral vectors*.

IX.4.2.1 Let π_q be the *power set* of floral universe Q, i.e. the set of all subsets of Q, $\Pi_q = \{\pi_{j1}, \pi_{j2}, \ldots, \pi_{jk}, \ldots, \pi_{j\omega}\}$, where π_{jk} is a *potential flora*, and $\omega = 2^s$. Let F_j be the frequency distribution for Π_q generated by T_j, $F_j = \{f_{j1}, \ldots, f_{jk}, \ldots, f_{j\omega}\}$. The elements of Φ_j (see IX.1.2) are those element Π_q that are represented in the survey-sampling by non-zero frequency values (*manifested floras*). By supposing that the sampling conditions are satisfactory ($m > s$, $m \to \infty$), and the probability estimate $\hat{p}_{jk}$, $\hat{p}_{jk} = f_{jk}/m$, is more or less unbiased, we can set the up *an empirical finite scheme*,

$$\Pi_{qj} = \begin{bmatrix} \pi_{j1}, \pi_{j2}, \ldots, \pi_{jk}, \ldots, \pi_{j\omega} \\ \hat{p}_{j1}, \hat{p}_{j2}, \ldots, \hat{p}_{jk}, \ldots, \hat{p}_{j\omega} \end{bmatrix} \tag{IX; 8}$$

with estimated probability distribution $\hat{p}_j$, related to Π_q, *a complete system of events*. If by letting U_j to be a suitable *estimator* such as $U_j = \sum_k f_{jk} \log f_{jk}$, $\sum_k f_{jk} = m$, then *floral diversity*, $m\widehat{H}_j^{(\varphi)}$,

$$m\widehat{H}_j^{(\varphi)} = m \log m - U_j, \tag{IX; 9}$$

a weighted entropy-estimate of the SHANNON-type can be defined.

IX.4.2.2 It is to be noted that $m\widehat{H}_j^{(\varphi)}$ is not only one estimate of the the many similar estimates; it has a number of fairly important properties. The most remarkable feature of $mH_j^{(\varphi)}$ is that floral entropy, as *an "overall uncertainty bound"*, or, more precisely, as *a joint entropy* for *all* elements of Q,

$$m\widehat{H}_j^{(\varphi)} = ([q_1, q_2, \ldots, q_i, \ldots, q_s]), \tag{IX; 10}$$

contains *all* possible "*regions*", where any kind of "interaction" among populations may take place (see Fig. 3). In other words, $m\widehat{H}_j^{(\varphi)}$ is served as an *upper bound* for association estimates, or even, for associatum (see IX.5).

IX.4.2.3 Disregarding at the moment suitable conditions for a satisfactory survey sampling, consider the estimated values of $m\widehat{H}_j^{(\varphi)}$ for (IX.1), (IX; 5), and (IX; 6), which are: 19.25, 24, and 8 Bits. It is easy to relate these values to the ratio of "*posse — esse*"; e.g., for any kind of [IRC], like (IX; 5), $m\widehat{H}_j^{(\varphi)} = \max mH_j^{(\varphi)} = m \log m$, because all kinds of potential floras are realized. (If, in an [IRC], where $f_{jk} > 1$, then $\max \widehat{H}_j^{(\varphi)} = sm$ Bits.)

IX.4.3 If we do not distinguish *all* possible floral vectors, only those vectors whose floral marginal values are the same, we can speak of *floral multiplets*, characterized by a multiplet varable, R, $R = \{0, 1, \ldots, r, \ldots, s\}$ (see II.3). The frequency distribution of R values of a K_j is $F'_j = \{g_{j0}, g_{j1}, \ldots, g_{jr}, \ldots, g_{js}\}$, $\sum_r g_{jr} = m$, and the frequency distribution of the weighted R values

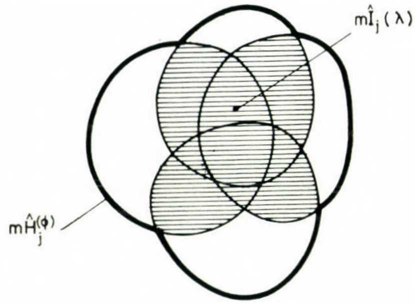

Fig. 3. Associatum $(m\hat{I}_j(\lambda))$ is a measurable subset of the intersection in a VENN-diagram, whose bound is $m\hat{H}_j(\Phi)$, floral diversity

is F''_j, $F''_j = \{g'_{j1}, g'_{j2}, \ldots, g'_{jr}, \ldots, g'_{js}\}$, $g'_{jr} = rg_{jr} = n._{jr}$, $\sum rg_{jr} = N_j$. By introducing suitable estimators, G_j for F'_j, and E_j for F''_j, $G_j = \sum_r g_{jr} \log g_{jr}$, $E_j = \sum_r n._{jr} \log n._{jr}$, we can define

$$m\hat{H}_j(R) = m \log m - G_j, \tag{IX; 11}$$

$$N_j\hat{H}_j(V_R) = C_j - E_j. \tag{IX; 12}$$

Both estimates, in particular, $N_j\hat{H}_j(V_R)$, *entropy of floral multiplets for valence values*, will be used in the sequel (see e.g. XI.1).

IX.5 Finally, the reader is reminded of *associatum* (cf. JUHÁSZ-NAGY et al., 1973, JUHÁSZ-NAGY 1976, JUHÁSZ-NAGY and PODANI 1983),

$$m\hat{I}_j(\lambda) = sm \log m - \alpha_j - m\hat{H}_j^{(\varphi)}, \tag{IX; 13}$$

a multiple contingency information, i.e. contingency information estimate of a $2 \times 2 \times \ldots^{s} \times 2$ table. Since, with respect to (IX; 14),

$$m\hat{I}_j(\lambda) = ([q_1, q_2, \ldots, q_i, \ldots, q_s]), \tag{IX; 14}$$

the relation between $m\hat{H}_j^{(\varphi)}$ and $m\hat{I}_j(\lambda)$ can easily be shown by a VENN-diagram (see fig. 3).

X. Some basic relations: spatial processes

X.1 Suppose that elements of T_j are circular plots, to be randomized over τ'. Suppose, further, that we have a class of plots, Θ_h, $\Theta_h = \{T_0, T_1, \ldots, T_j, \ldots, T_u\}$, where elements of T_0 are approximately "points", and the size of plots increases from T_0 to T_u, in a strictly monotonous way. More precisely, if the size of plots are labelled as $t_0, t_1, \ldots, t_j, \ldots$, then the increasing series $t_0 < t_1 < \ldots < t_j < \ldots$ is the first feature of the survey sampling below. (It should be noted that, while the symbol t_{jg} indicates an element, $t_{jg} \in T_j$, as *an entity*, the symbol t_j refers to the geometric size of this element.)

X.1.1 Of course, this increasing series may be of very different kinds. If we define the difference between t_j and t_{j-1} as an interval, $| t_j, t_{j-1} |$, or, give rule to generate such differences, we can speak of several types of *articulation* of sampling. If the articulation is "rough", then the strictly monotonous series, $N_0 < N_1 < \ldots < N_j < \ldots$, and, $n_0 > n_1 > \ldots > n_j > \ldots$ are likely to be expected. If the articulation is "fine" enough, then the relations $N_j \leq N_{j+1}$, $n_j \geq n_{j+1}$, might occur, i.e. the series are *not* strictly monotonous.

X.1.2 The choice of articulation should depend on the particular object (vegetation-type), and on the purpose of the investigator. (Unfortunately enough, *a priori*, before making a survey sampling, we do not know in most cases the "resolutional properties" of our object, which shows clearly the need of *an iteration*, viz. *successive approximation*.) Experience suggests that a "good articulation" must not be "monotonously regular", because of the existence of some "critical regions".

X.2 In order to illustrate some primary points of this paper, it is worthwhile to consider a concrete example, and derive certain primary conclusions from it. Similar examples are given by JUHÁSZ-NAGY and PODANI (1983).

X.2.1 The particular type of vegetation is *a meadow*, in NE-Hungary (near the village Beregdaróc) which has been described by the author, on the top of his SIGMA-tist period, as *Anthoxantho-Festucetum pseudovinae, festucetosum sulcatae* (JUHÁSZ-NAGY, 1958), where $s = 83$, $m_1 = 256$, $m_2 = 512$, due to two survey samplings. The long and tedious sampling procedures have been established twice (May, 1965 and 1966); the second survey has been aimed at correcting some results of the first one. (This is why $m_2 > m_1$.) Some details about survey sampling are considered later on.

X.2.2 Let $R_j^{(1)}(\lambda)$ be *a redundancy estimate*, called *simple relative associatum*,

$$R_j^{(1)}(\lambda) = 1 - m\hat{I}_j(\lambda)/m\widehat{H}_j^{(\varphi)}, \qquad (X; 1)$$

which has a fairly good intuitive meaning with respect to fig. 3. Roughly speaking, (X; 1) relates associatum to an "overall uncertainty bound". (The upper index (1) in (X; 1) distincts *this* redundancy estimate from better ones, that can be considered later, on when the boundary relations of $m\hat{I}_j(\lambda)$ will be clarified.)

X.2.3 The curve of Fig. 4 shows *change of the redundancy estimate in space* with respect to our meadow of X.2.1. We can see at once two peaks of the curve, indicating two local maximum points of $R_j^{(1)}(\lambda)$. The rough-and-ready interpretation of such curves is possible, if we suppose that while the first peak is due to the "dominance of *negative* associations", the position of the second peak may due to the "dominance of *positive* associations". Indeed, the position of the first peak corresponds surprisingly well to the first maximum of $m\hat{I}_j(\lambda)$, and to the local maximum of pooled negative interlocal associations. The second peak of fig. 4, however, has no such well-defined correspondences, except that it is located fairly closely to A_{comp}, *conpensatory area*, where

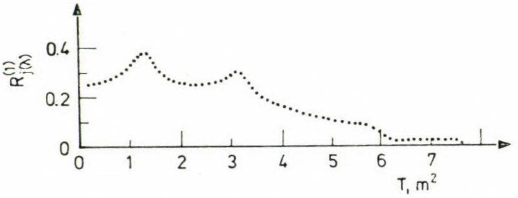

Fig. 4. An $R_j^{(1)}(\lambda)$ *versus* T curve (see X.2.3)

$N_j \approx n_j$. Besides some arguments to be explained later on, it is not too difficult to recognize that such a tranformation in most cases is necessary, because $m\hat{I}_j(\lambda)$ converges more quickly to zero than $m\widehat{H}_j^{(\varphi)}$, and, therefore, $R_j^{(1)}(\lambda)$ reaches zero before *a minimum area*, $A_{\min}$, where $N_j = ms$, $n_j = 0$. It is to be observed that *maximum area*, $A_{\max}$, local maximum of $m\widehat{H}_j^{(\varphi)}$, is located between the two peaks of Fig. 4.

X.3 In the case of Fig. 4, or in a number of similar cases, we deal with *a spatial process* of the simplest type, i.e. with a discrete approximation of an almost continuous graphicon. (The second survey was made to control wheather the curve is "smooth" enough, there is no third peak, etc.) In order to discuss such spatial processes of the floristic type, the following *terms* are proposed for further use.

X.3.1 Let us call such functions as $m\widehat{H}_j^{(\varphi)}$, $m\hat{I}_j(\lambda)$, $R_1^{(1)}(\lambda)$ etc. *characteristic functions.* (Note that here the adjective "characteristic" is used, of course, in a special *methodological sense* of the word, i.e. the present meaning should be kept clearly distinct from the meaning of "characteristic functions" of the mathematical analysis, *alias*, "calculus".) The *extremal values* of characteristic functions (maxima, minima) indicate certain *characteristic points* of T, corresponding in our case to a number of *characteristic areas* (e.g. $A_{\max}$, A_{comp}, $A_{\min}$, etc.). A proper inequality, such as

$$\ldots A_{\mathrm{ass}}^{(1)} < A_{\max} < A_{\mathrm{comp}} < \ldots \qquad (X; 2)$$

shows some *characteristic ordering* of characteristic points or areas. A characteristic ordering can be made more efficient by defining a number of *characteristic intervals*, such as $|A_{\max}, A_{\mathrm{comp}}|$, etc.

X.3.2 It is quite obvious that the characteristic interval $| t_0, A_{\min} | = T$, called the *floristic scale* (in brief, *the scale*) has the important property that T contains (by relation $\in$, or by $\subset$) *all* relevant characteristic points or intervals. $A_{\min}$, in many cases exists, but it may be *trivial*, if $A_{\min} \approx \mu(\tau')$, where the symbol $\mu(\tau')$ indicates the size of the study area. A non-trivial $A_{\min}$ exists, if $A_{\min} \ll \mu(\tau')$. (In our case, for instance, where the study area, viz. *stand* area of the meadow is about 60 000 m^2 in size, and since this area is surrounded by forests and swamps, we can safely say that a non-trivial $A_{\min}$ exists.) There are cases, where there is no well-defined $A_{\min}$, but a scaling point t_u can be found from which all information estimates (defined floristically) converge stochastically to zero.

X.4 During a survey sampling, that is operating with such *scaling points* as t_j, $t_j \in T$, our intention is to make *a characteristic scaling*, or, at a more advanced level, *a characteristic screening*. Both *searching procedures* attempt to locate certain characteristic values (points, or intervals) in the scale, but while in a scaling procedure we want to study *some* synmorphological features of a community, during a screening procedure we have some specific *desideratum* in mind, e.g. we would like to locate *the optimal interval* of the scale for a certain classificatory algorithm. (The personal guess of the author is that some interval around the point given by the first local maximum of $m\hat{I}_j(\lambda)$, $A^{(1)}_{\mathrm{ass}}$, might be optimal for a number of classificatory procedures.)

X.4.1 The primary concern of any characteristic scaling is to consider some bounds of the estimators, e.g.

$$\left.\begin{aligned} N_j \log \overline{m}_j &\leq A_j \leq N_j \log m \\ N_j \log \overline{s}_j &\leq B_j \leq N_j \log s \\ n_j \log (m - \overline{m}_j) &\leq a_j \leq n_j \log m \\ n_j \log (s - \overline{s}_j) &\leq b_j \leq n_j \log s \end{aligned}\right\} \qquad (\mathrm{X}; 3)$$

where $\overline{m}_j = N_j/s$, $\overline{s}_j = N_j/m$, or,

$$\left.\begin{aligned} (\gamma_j - ms \log s) &\leq \alpha_j \leq sm \log m \\ (\gamma_j - sm \log m) &\leq \beta_j \leq ms \log s \end{aligned}\right\} \qquad (\mathrm{X}; 4)$$

These simple bounds will be frequently used in the sequel (see e.g. XII.2).

X.4.2 As for the primary meaning of such bounds, we can reconsider some statements made in IV. If a K_j is in an l-state, then A_j has its value at the lower bound; if a K_j is in an L-state, then A_j has its maximum value. (*Mutatis mutandis*, similar relations are true for all the estimators.) From the point of view of scaling, some states, relatively close to an l-state, can be expected somewhere in certain "mid-regions" (critical intervals) of the scale (except, of course, the irrelevant $A_{\min}$-state). Some object (our meadow, for instance) ~~has~~ the property, however, of approaching some l-state more than once. The presence or absence of this property ~~depends~~ on ~~the~~ "structural complexity" of the community, or, speaking in simpler terms, on the *scoring* (floristic) *diversity states* of the object that can be traced by means of some functions defined on a series of $K_0, K_1, \ldots, K_j, \ldots$ (see XII.3.2; XIII., etc.).

X.4.3 These relations indicate again the need for a better understanding of such concepts as *orderliness, equivalence information*, etc. This need is stressed further by the fact that, contrary to the naive belief of many investigators, some *maximum* degree of entropy *and* information estimates might be very closely located in the scale. Without repeating notation and arguments used in part 1, the reader is reminded of $\langle 1 \rangle$-table and $\langle 0 \rangle$-table, for wich $N_j\hat{I}_j(E)$ and $n_j\hat{I}_j(e)$ can be related.

XI. Equivalence information – revisited

XI.1 In order to study again $N_j\hat{I}_j(E)$ more closely, let

$$\left.\begin{array}{ll} (1) & N_j\hat{H}_j(V_q) = C_j - A_j, \\ (2) & N\,\hat{H}_j(V_R) = C - E_j, \\ (3) & N_j\hat{H}_j([V_q, V_R]) = C_j - D_j, \end{array}\right\} \qquad \text{(XI; 1)}$$

where (1) is *entropy of local valences* (cf. XII.1), (2) is *entropy of floral multiplets for valence values* (cf. IX.4.3), and (3) is *equivalence entropy* (a special kind of *joint entropy*), by which

$$\begin{aligned} N_j\hat{I}_j(E) &= N_j\hat{H}_j(V_q) + N_j\hat{H}_j(V_R) - N_j\hat{H}_j([V_q, V_R]) = \\ &= C_j - A_j - E_j + D_j \end{aligned} \qquad \text{(XI; 2)}$$

as in (II; 5). Using this new notation in hand, it is straightforward to introduce the corresponding pair of *conditional entropy functions*:

$$\left.\begin{array}{l} N_j\hat{H}_j(V_q \mid V_R) = E_j - D_j, \\ N_j\hat{H}_j(V_R \mid V_q) = A_j - D_j. \end{array}\right\} \qquad \text{(XI; 3)}$$

This pair of function almost always play a fairly important role in understanding orderliness of different types, because

$$N_j\hat{H}_j(V_q) - N_j\hat{H}_j(V_q \mid V_R) = N_j\hat{H}_j(V_R) - N_j\hat{H}_j(V_R \mid V_q) = N_j\hat{I}_j(E). \quad \text{(XI; 4)}$$

XI.2 In the case of an [IRC] (cf. IX.4.1.1), where, according to the proof given in V.2.4, the equality $C_j - A_j = E_j - D_j$ is always true, and, therefore, the equations of (XI; 5)

$$\begin{aligned} N_j\hat{H}_j(V_q) &= N_j\hat{H}_j(V_q \mid V_R) \\ N_j\hat{H}_j(V_R) &= N_j\hat{H}_j(V_R \mid V_q) \\ N_j\hat{H}_j(V_q) + N_j\hat{H}_j(V_R) &= N_j\hat{H}_j([V_q, V_R]) \end{aligned} \tag{XI; 5}$$

make clear that $N_j\hat{I}_j(E) = 0$ for any [IRC], i.e. the stochastic independence of events $[q]$ and $[R]$. Note that conditions of (XI; 5) correspond naturally to the maximum value of $m\hat{H}_j^{(\varphi)}$ and to the minimum (zero) value of $m\hat{I}_j(\lambda)$.

XI.2.1 More precisely, by applying the proper Pascal-triangle rule (cf. V.4), and by letting $s_r = \binom{s-1}{r}$, $\Delta_s = \sum_r \sigma_r \log \sigma_r$, $s^* = \log s$, a fairly concise description of estimators is possible,

$$\begin{aligned} C_j &= s(s + s^* - 1)\, 2^{s-1}, \\ A_j &= s(s - 1)\, 2^{s-1}, \\ D_j &= s\Delta_s, \\ E_j &= s(s^* 2^{s-1} + \Delta_s), \end{aligned} \tag{XI; 6}$$

and, using (XI; 6), we get (XI; 7),

$$\left.\begin{aligned} N_j\hat{H}_j(V_q) = N_j\hat{H}_j(V_q \mid V_R) &= ss^*\ 2^{s-1}, \\ N_j\hat{H}_j(V_R) = N_j\hat{H}_j(V_R \mid V_q) &= s[(s-1)2^{s-1} - \Delta_s], \\ N_j\hat{H}_j(V_q \mid V_R) + N_j\hat{H}_j(V_R \mid V_q) = N_j\hat{H}_j([V_q, V_R]) &= \\ = s[(s + s^* - 1)2^{s-1} - \Delta_s], \end{aligned}\right\} \tag{XI; 7}$$

i.e. the detailed verification of (XI; 5). Since $N_j\hat{I}_j(E) = 0$, thus $\hat{Y}_j = N_j\hat{I}_j(K)$; $N_j\hat{I}_j(K) = ss^*2^{s-1} - B_j$.

XII.2.2 For an [IRC], where $s = 3$, as in (IX; 5), $C_j = 43.0188$, $A_j = 24$, $E_j = 25.118$, $D_j = 6$, and so $N_j\hat{H}_j(V_q) = N_j\hat{H}_j(V_q \mid V_R) = 19.0188$, $N_j\hat{H}_j(V_R) = N_j\hat{H}_j\ (V_r \mid V_q) = 18$, $N_j\hat{H}_j\ (V_q \mid V_r) + N_j\hat{H}_j(V_r \mid V_q) = N_jH_j\ ([V_q, V_R]) = 37.0188$, $\hat{Y}_j = N_j\hat{I}_j(K) = 8.2641$. Note that as s increases, the difference of $N_j\hat{H}_j(V_q)$ and $N_j\hat{H}_j(V_R)$ increases as well. Note, further, that by dropping the notation of (XI; 6), we can express $\hat{Y}_j = N_j\hat{I}_j(K)$ as $N_j \log s - B_j$ (see XIV.3.2).

XI.3 On the other hand, if a K_j represents an [IDC] of some sort, where $m\hat{I}_j(\lambda) > 0$, the particular K_j might be of very different kinds, according to the conditions for orderliness involved. For further use, let us distinguish here only two kinds of compositions.

XI.3.1 In an at least locally *well-ordered case*, where $A_j = D_j = N_j \log s$, $E_j = B_j + D_j$, that is, $N_j\widehat{H}_j(V_q \mid V_R) = C_j - D_j = N_j \log m_j$, $N_j\widehat{H}_j(V_R \mid V_q) = 0$, $N_j\widehat{H}_j(V_q \mid V_R) = N_j\widehat{H}_j([V_q, V_R])$, and, therefore, $\widehat{Y}_j = 0$, $N_j\hat{I}_j(E) = N_j\hat{I}_j(K) = C_j - E_j$. In words, this means that the measurable subset of joint entropy that represents intersection will be identical with the whole set of $N_j\widehat{H}_j(V_R)$; this is why $N_j\widehat{H}_j(V_R \mid V_q) = 0$.

XI.3.2 In an at least locally *ill-ordered case*, where $C_j = E_j$, $A_j = D_j = N_j \log s$, $B_j = N_j \log m$, that is, $N_j\widehat{H}_j(V_q \mid V_R) = E_j - D_j$, $N_j\widehat{H}_j(V_R) = N_j\widehat{H}_j(V_R \mid V_q) = 0$, and, therefore, $\widehat{Y}_j = N_j\hat{I}_j(K) = N_j \log \overline{N}_j$ (see e. g. XIV.2.2), consequently, $N_j\hat{I}_j(E) = 0$. In words, this means that because one of the entropy estimates in (XI; 2) completely disappears, the intersection should be "empty"; this is why the value of equivalence information is zero.

XI.4 Although our understanding of orderliness became perhaps a bit more advanced, a branch of new problems arise, if, *ad analogiam* of (XI; 2), we introduce such new estimates as

$$\left.\begin{aligned} n_j\widehat{H}_j(v_q \mid v_R) &= e_j - d_j, \\ n_j\widehat{H}_j(v_R \mid v_q) &= a_j - d_j, \\ n_j\widehat{H}_j([v_q, v_R]) &= c_j - d_j, \end{aligned}\right| \qquad \text{(XI; 8)}$$

and start asking questions about such additive relations as

$$\left.\begin{aligned} N_j\hat{I}_j(E) + n_j\hat{I}_j(e) &= ms\hat{I}_j(\varepsilon) = \gamma_j - \alpha_j - \varepsilon_j + \delta_j, \\ N_j\hat{I}_j(K) + n_j\hat{I}_j(k) &= ms\hat{I}_j(X), \\ \widehat{Y}_j + \widehat{y}_j &= \widehat{\eta}_j = \varepsilon_j - \delta_j - \beta_j, \end{aligned}\right\} \qquad \text{(XI; 9)}$$

where $\varepsilon_j = E_j + e_j$, $\delta_j = D_j + d_j$. We might guess that a proper interpretation of (XI; 8) is possible within the framework of *syncretic models*, where coincidences *and* incoincidences of *all* elements of a K_j are considered altogether,

[B] SOME SYNCRETIC MODELS

XII. Marginal diversity functions

XII.1 Let us define the following *weighted entropy functions* on the f-marginals of a K_j, referring to marginal uncertainty estimates at t_j,

$$\left.\begin{aligned} &(1) \quad N_j\widehat{H}_j(V_q) = C_j - A_j, \\ &(2) \quad N_j\widehat{H}_j(V_t) = C_j - B_j, \\ &(3) \quad n_j\widehat{H}_j(v_q) \;= c_j - a_j, \\ &(4) \quad n_j\widehat{H}_j(v_t) \;\;= c_j - b_j. \end{aligned}\right\} \qquad \text{(XII; 1)}$$

XII.1.1 If, for the sake of brevity, we label here entropy function by [EF], then (1) is [EF] of *local valences* (cf. XI.1), (2) is [EF] of *floral valences*, (3) is [EF] of *local invalences*, and, (4) is [EF] of *floral invalences*. Functions (1) and (2) are monotonously *increasing*, functions (3) and (4) are monotonously *decreasing* in the spatial process described in X.

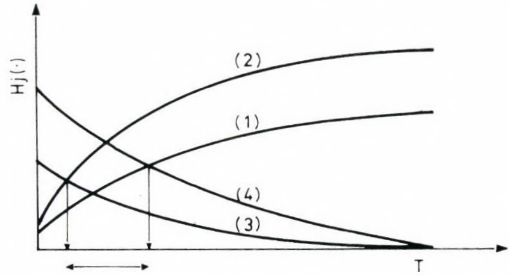

Fig. 5. A quartet of "species area curves", according to (XII; 1) (see XII.1.2)

XII.1.2 More precisely, if a non-trivial $A_{\min}$ exists, then (1) and (2) reaches (or, at least, stochastically converges to) its maximum of $ms \log s$ and $sm \log m$, respectively, while (3) and (4) tends to zero *within* the size of a τ'. Because of this property (monotony), functions of (XII; 1) cannot be used as "diversity indeces" (see GRASSLE et al. 1979). These functions, however, might be the optimal substituents of the classical "*species-area curves*" used so far in many respects. It is remarkable that in the majority of cases the interval of scale, whose extremal points are located, where the approximate equalities of $N_jH_j(V_t) \approx n_j\widehat{H}_j(v_q)$ and $n_j\widehat{H}_j(v_t) \approx N_j\widehat{H}_j(v_q)$ are satisfied, includes practically *all the important characteristic points* (except, of course, $A_{\min}$) that may be interesting in vegetation research (see Fig. 5). The search for such *a maximum interval* (as *a characteristic interval*) is perhaps the simplest way of indicating the *size-classes* within which a *reasonable survey sampling* can be taken. Besides, the curves of Fig. 5 can frequently be linearized as special "*synallometric*" (log-linear) *curves*. (This topic should be treated elsewhere.)

XII.2 The best property of functions in (XII; 1) is that they are all bounded by the the following simple inequalities [cf. (X; 3) etc.],

$$
\begin{array}{lllll}
(1) & N_j \log \bar{s}_j & \leq N_j\widehat{H}_j(V_q) & \leq N_j \log s & \\
(2) & N_j \log \bar{m}_j & \leq N_j\widehat{H}_j(V_t) & \leq N_j \log m & \qquad (\text{XII; } 2)\\
(3) & n_j \log (s - \bar{s}_j) & \leq n_j\widehat{H}_j(v_q) & \leq n_j \log s & \\
(4) & n_j \log (m - \bar{m}_j) & \leq n_j\widehat{H}_j(v_t) & \leq n_j \log m &
\end{array}
$$

where, again, $\bar{s}_j = N_j/m$ (i.e. local mean density), $\bar{m}_j = N_j/s$ (floral mean density) as in IV.

XII.3 According to inequalities in (XII; 2), both *upper deviates*,

$$\Delta N_j\widehat{H}_j(V_q) = \max N_jH_j(V_q) - N_j\widehat{H}_j(V_q) = N_j \log s + A_j - C_j,$$
$$\Delta N_j\widehat{H}_j(V_t) = N_j \log m + B_j - C_j,$$
$$\Delta n_j\widehat{H}_j(v_q) = n_j \log s + a_j - c_j, \qquad \text{(XII; 3)}$$
$$\Delta n_j\widehat{H}_j(v_t) = n_j \log m + b_j - c_j,$$

and *lower deviates*,

$$\nabla N_j\widehat{H}_j(V_q) = N_j\widehat{H}_j(V_q) - \min N_jH_j(V_q) = N_j \log m - A_j,$$
$$\nabla N_j\widehat{H}_j(V_t) = N_j \log s - B_j,$$
$$\nabla n_j\widehat{H}_j(v_q) = n_j \log m - a_j, \qquad \text{(XII; 4)}$$
$$\nabla n_j\widehat{H}_j(v_t) = n_j \log s - b_j,$$

as empirical *gain of information* estimates at t_j (see KHINCHIN 1959) indicate how diverse *f*-marginals are. Contrary to (XII; 1), functions in (XII; 3) and (XII; 4) represent *useful diversity functions* (called *marginal diversity functions*) based on the simple reasoning as follows.

XII.3.1 If K_j is in *a monovalence state* of either kind, or both kinds, then proper *upper deviate* (or deviates) has (or have) zero value. On the other hand, in the same case, the proper *lower deviates* are represented by their *maximum* values. If K_j is in *an oligovalence state*, then, *vice versa*, *upper* and *lower* deviates take their *minimum* and *maximum* values, respectively.

XII.3.2 This shows clearly that *all* the eight deviates (in short, *the octet*) should be used *altogether*, in order to characterize *the marginal diversity state* of a K_j. In addition, *f*-deviates, i.e. members of octet play an important role in composing information estimates, because each contingency information can be gained as the sum of the proper gain of information estimates (see XIII—XIV.).

XIII. Three partners modelling

XIII.1 It is sometimes very useful to think of a K_j as an output of a situation resulted in *a three-ways selection*. Because any selection is *a repetitive choice* of some kind, our elementary concept should be (IX; 1), *a random choice*, which henceforth will always be called *an intersection*. (Note that the very general meaning of set operation $\cap$ is substituted here by a particular *methodological meaning* of the word.) Let us define three "*abstract partners*", *an average population* (q), *an average plot* of j-size (t), and *an average intersection* (i), related to the ordered pair $\langle q, t \rangle$. All partners are endowed with selective property, namely,

- q selects *plots* from the elements of T_j,
- t selects *populations* from set Q,
- i selects *ordered pairs* from ms possibilities.

XIII.1.1 In this *selective situation*, a fairly good intuitive framework might be introduced. For instance, besides *choice* and *selection* (represented by the "inner" scalars and vectors of a K_j), we can rightly think of "*selectional powers*" of partners (shown by the f-marginals of a K_j), or of *diversity of selectional powers* (characterized by an *octet*). Since selection of some sort is always related to a certain degree of *preference*, we might think of our selective situation as of *a preferential situation* as well. For instance, in an [IRC], where no kind of local, or interlocal preference is present, all properties of an [IRC] (l-state, max $mH_j^{(\varphi)}$, zero-valued $m\hat{I}_j(\lambda)$, etc.) are due to this lack of preference. Note, further that a huge number of old notions (e.g. ubiquity, commonness, rarity, constancy, fidelity etc.) is closely related to such an intuitive framework of concepts. These concepts, however, have a strong "historical load", and this is responsible for a number of misunderstanding (JUHÁSZ-NAGY 1964).

XIII.1.2 For the reasons why is it plausible to allow for either t or i to have *selectional powers*, we may turn to several generalizations of our situation, e.g. where Q can be a set of *animals*, *enzymes*, etc., T_j can be a set of some *foods*, *substrates*, etc., at a given j-condition, say, at a given degree of temperature. Although elements of Q, without doubt, are "more active" than elements of T_j, *plots*, *foods*, *substrates* (simply, by virtue of "quality" or j-state they represent) also selects *plants*, *animals*, *enzymes*; this is why such concepts as *enzyme specificity*, *substrate specificity*, and *reaction specificity*, notwithstanding *coplimentary*, but *are by no means related trivially to each other* (see QUASTLER 1953).

XIII.2 According to KOLMOGOROV, assume *a classical probability field* for our selective situation, i.e. $p(q) = 1/s$, $p(t) = 1/m$, $p(i) = 1/ms$, and thus, in terms of information theory, we have

$$\left.\begin{aligned} H(q) &= \log s, \\ H(t) &= \log m, \\ \hat{H}_j(i) &= \log ms - \gamma_j/ms, \end{aligned}\right\} \qquad \text{(XIII; 1)}$$

i.e. quantities of the HARTLEY type (see RÉNYI, 1962; REZA, 1961; LUCE, 1960). Note that even in this utterly simple (say, "hyperprimitive") formalism, the entropy $\hat{H}_j(i)$ can be given as an estimate, only. The primary explanation of this is due to *a particular transition*; namely, the relation $t_{jg} \in T_j$ is transferred here to the selectional relation of partner i. The same is true for for the corresponding joint entropy functions,

$$\left.\begin{aligned}\hat{H}_j([q, i]) &= \log ms - \alpha_j/ms,\\ \hat{H}_j([t, i]) &= \log ms - \beta_j/ms,\end{aligned}\right\} \qquad \text{(XIII; 2)}$$

$$H([q, t]) = \log ms, \qquad \text{(XIII; 3)}$$

where (XIII.3), *not* containing i, is of constant value, but functions of (XIII; 2) are estimates, related to the actual (j-dependent) values of the estimators involved.

XIII.3 It is fairly easy to see that, having (XIII; 1)—(XIII; 3), a whole system of equations can be obtained, and *some* of these equations have a definite bearing on certain aspects of our selective or preferential situation.

XIII.3.1 According to the simplest kind of a Boolean reasoning, the particular event space is partitioned by (XIII; 1) into $2^3 = 8$ molecular events, and we might consider these events as singletons, doubletons etc., so we have altogether $2^8 = 256$ kinds of functions. (Fortunately enough, the vast majority of these functions is quite irrelevant and immaterial for our present purpose.) This family of functions includes e.g. *conditional entropy functions* (like $H(q \mid t) = H([q, t]) - H(t) = \log s$), *information functions* (e.g. $\hat{I}_j(q, \mathrm{i})$, $\hat{I}_j(t, i)$, etc.), *multiple informations functions* (e.g. $\hat{I}_j(i, [q, t])$, etc.), *partial information functions* (like $\hat{I}_j([q, t] \mid i)$, etc.), and so on.

XIII.3.2 The simplest possible information is

$$I(q, t) = H(q) + H(t) - H([q, t]) = \log s + \log m - \log ms = 0 \quad \text{(XIII; 4)}$$

which has the definite meaning that partners q and t, *without partner* i, have *no information* whatsoever with respect to each other. This indicates clearly that we should concentrate on such functions which include partner i as well. For the sake of convenience, it is much better to use weighted forms (i.e. each function is multiplied by ms).

XIII.4 It is worthwhile investigating *a triplet of functions* (in itself a simple *linear model*) in a little bit more detailed way, and turning afterwards to other relevant functions of the three partner modelling.

XIII.4.1 According to XIII.2,

$$ms\hat{I}_j(q, i) = msH(q) + ms\hat{H}_j(i) - ms\hat{H}_j([q, i]) = ms \log s + ms \log ms - \gamma_j - \\ - (ms \log ms - \alpha_j) = ms \log s + \alpha_j - \gamma_j. \qquad \text{(XIII; 5)}$$

It is to be noticed that (XIII; 5) is identical either with contingency information of the little (*local*) contingency table" on the right side of an [ET], or with the sum of the proper *upper* deviates in (XII; 3):

$$ms\hat{I}_j(q, i) = \Delta N_j\widehat{H}_j(V_q) + \Delta n_j\widehat{H}_j(v_q) \qquad \text{(XII; 6)}$$

It is to be stressed again that *all contingency information estimates* are composed linearly from certain *gain of information estimates.* (This important point is practically missed in the ecological literature, because of the lack of *systematic* treatment, and strange "potpourism" with "indeces" and names, where simple entropy estimates are frequently called "total information content" and similar rubish nicknames.) We might guess that the sum of the corresponding *lower* deviates play a similar role in composing another relevant information estimate. Indeed,

$$\left.\begin{aligned} \nabla N_j\widehat{H}_j(V_q) + n_j\widehat{H}_j(v_q) &= ms\widehat{H}_j([t, i] \mid q) - \\ - ms\widehat{H}_j(t \mid [q, i]) - ms\widehat{H}_j(i \mid [q, t]) &= \\ = sm \log m - \alpha_j &= ms\hat{I}_j([t, i] \mid q). \end{aligned}\right\} \qquad \text{(XIII; 7)}$$

(XIII; 7) is *a partial information function* referring to information of partners t and i, if q is neglected (or, say, in an experimental situation, if q is held *constant*).

XIII.4.2 We can see at once that there exists a simple additive relation between (XIII; 5) and (XIII; 7),

$$\left.\begin{aligned} ms\hat{I}_j(q, i) + ms\hat{I}_j([t, i] \mid q) &= \\ = ms \log ms - \gamma_j &= ms\hat{I}_j(i, [t, q]), \end{aligned}\right\} \qquad \text{(XIII; 8)}$$

where $ms\hat{I}_j(i, [t, q])$ is *a multiple information estimate* of partner i, with respect to the pair of other partners. It is obvious that

$$ms\hat{I}_j(i, [t, q]) = \Delta N_j\widehat{H}_j(V_q) + \ldots + \nabla n_j\widehat{H}_j(v_q), \qquad \text{(XIII; 9)}$$

i.e. (XIII; 8) is the sum of all the four *local* deviates. The main question is how can one interpret (XIII; 8) from the point of view of vegetation science.

XIII.4.3 A primary explanation should start with the statement that both (XIII.5) and (XIII; 7) are bounded in such a way that

$$\left.\begin{aligned} 0 \leq ms\hat{I}_j(q, i) &\leq msI_j(i, [q, t]), \\ 0 \leq msI_j([t, i] \mid q) &\leq msI_j(i, [q, t]). \end{aligned}\right\} \qquad \text{(XIII; 10)}$$

Secondly, because of the composition of both information estimates of proper *local* deviates, we might guess tha $ms\hat{I}_j(q, i)$ approaches its upper bound if

K_j approaches an L-state, and $ms\hat{I}_j([t, i] \mid q)$ has an opposite "trend". For the sake of illustration, let us consider the local valence and invalence values of two simple cases, where $s = 4$, $m = 8$, and where $\langle A \rangle$ and $\langle B \rangle$ represents an l-state and an L-state, respectively.

$$\begin{array}{cc} \langle A \rangle & \langle B \rangle \\ \begin{array}{c|c} 4 & 4 \\ 4 & 4 \\ 4 & 4 \\ 4 & 4 \\ \hline 16 & 16 \end{array} & \begin{array}{c|c} 8 & 0 \\ 8 & 0 \\ 0 & 8 \\ 0 & 8 \\ \hline 16 & 16 \end{array} \end{array} \qquad \text{(XIII; 11)}$$

Very simple calculation shows that

		$\langle A \rangle$	$\langle B \rangle$
(1)	$ms\hat{I}_j(q, i)$	0	32
(2)	$+ms\hat{I}_j([t, i] \mid q)$	32	0
(3)	$msI_j(i, [q, t])$	32	32

(XIII; 12)

i.e. (1) and (2) has its maximum value in an L-state and an l-state, respectively, within the interval permitted by (3). This kind of complementary relation is shown by data of (IX; 2) as well, where (1) 3.295 + (2) 28.617 = (3) 31.912.

XIII.4.4 If elements of a Q are labelled alphabetically such as $Q = \{a, b, c, \ldots\}$, and *the particular* (elementary) *finite schemes* are labelled by capital letters, $A = \{a, \bar{a}\}$, $B = \{b, \bar{b}\}$, ..., then we can denote $m\hat{H}_j(A)$, $m\hat{H}_j(B)$, ..., *particular local entropy functions* (in brief, *local entropy*), related to each element of Q. It is noteworthy that

$$m\hat{H}_j(A) + \ldots + m\hat{H}_j(S) = m\hat{H}_j([L]) = ms\hat{I}_i([t, i] \mid q), \quad \text{(XIII; 13)}$$

where $mH_j([L])$, *pooled local entropy*, equivalent to partial information estimate $ms\hat{I}_j([t, i] \mid q)$, is called by Juhász-Nagy and Podani (1983) *local distinctiveness*. [For simple data of (IX; 2), $mH_j([L]) = 8 + 6.4906 + 7.6358 + 6.4906 = 28.617$.] According to (IX; 13), *associatum* is given by a relation which in itself is a good illustration of a better understanding of *attribute duality*; in words, associatum is equal to *local* distinctiveness minus *floral* entropy. [For data of (IX; 2), $m\hat{I}_j(\lambda) = 28.617 - 19.245 = 9.362$.] The meaning of local distinctiveness, however, is linearly complementary to that of $ms\hat{I}_j(q, i)$ which might be called *local evenness*. Both functions are meaningful only with respect to $ms\hat{I}_j(i, [q, t])$, *floristic insaturation*. Because γ_j is a biased U-shaped estimator (along the scale), floristic insaturation has a reversed U-shaped curve, with a local maximum at A_{comp}, compensatory area (where

$N_j \approx n_j$). and with a global minimum at $A_{\min}$ (if it exists). Note that approaching $A_{\min}$, all these functions tend to zero.

XIII.5 The arguments used in XIII.4 can safely be transferred into *floral relations*, gaining

$$\begin{array}{ll} (1) & ms\hat{I}_j(t, i) = sm \log m \ + \beta_j - \gamma_j \\ (2) & +ms\hat{I}_j([q, i] \mid t) = ms \log s \ - \beta_j \\ \hline (3) & ms\hat{I}_j(i, [q, t]) = ms \log ms - \gamma_j \end{array} \qquad \text{(XIII; 14)}$$

where (1) is *floral evenness*, (2) is *floral distinctiveness*, and (3) is again *floristic insaturation*. (1) is composed linearly as $\Delta N_j\widehat{H}_j(V_t) + \Delta n_j\widehat{H}_j(v_t)$, (2) can be composed either as $\nabla N_j\widehat{H}_j(V_t) + \nabla n_j\widehat{H}_j(v_t)$, or as *a pooled floral entropy estimate*, $s\widehat{H}_j([F])$, analogous to (XIII; 13).

XIII.6 It is relevant to introduce at least two further information estimates, namely,

$$ms\hat{I}_j([q, t] \mid i) = ms\widehat{H}_j([q, t] \mid i) - ms\widehat{H}_j(q \mid [t, i]) - ms\widehat{H}_j(t \mid [q, i]) =$$
$$= \gamma_j - \alpha_j - \beta_j, \qquad \text{(XIII; 15)}$$

and

$$ms\hat{I}_j([q, t, i]) = ms\widehat{H}_j([q, t, i]) - ms\widehat{H}_j(q \mid [t, i]) -$$
$$- ms\widehat{H}_j(t \mid [q, i]) - ms\widehat{H}_j(i \mid [q, t]) =$$
$$= ms \log ms - \alpha_j - \beta_j. \qquad \text{(XIII; 16)}$$

(XIII; 15) is *a partial information estimate* of partners q and t, with respect to partner i. In this case, the meaning of (XIII; 15) has the good interpretation of a contingency information estimate for q and t, if condition labelled by i is excluded, or, if i is held constant. Note that although $msI(q, t) = 0$, $ms\hat{I}_j([q, t] \mid i) \neq 0$, except special conditions (e.g. at $A_{\min}$). (XIII; 16) is *the joint information* for our trio of partners. This quantity should be kept clearly distinct from $ms\hat{I}_j(\langle q; t; i\rangle) = \alpha_j + \beta_j - \gamma_j$ which estimate (as $[msH(q) \cap \cap\, ms\widehat{H}(t) \cap msH_j(i)]$) is distinguished in the terminology of Gardner (1962) as *strictly defined interaction*. We might call (XIII.15) and XIII; 16) *floristic efficiency* and *floristic interaction*, respectively.

XIII.6.1 By using (XIII; 15) and (XIII; 16), some new additive relations can be gained, e.g.

$$\left.\begin{array}{r|l} ms\hat{I}_j(q, i) & \text{local evenness} \\ +msI_j([q, t] \mid i) & \text{floristic efficiency} \\ \hline ms\hat{I}_j([q, i] \mid t) & \text{floral distinctiveness} \end{array}\right\} \qquad \text{(XIII; 17)}$$

and

$$\left.\begin{array}{r|l} ms\hat{I}_j(t, i) & \text{floral evenness} \\ +ms\hat{I}_j([q, t] \mid i) & \text{floristic efficiency} \\ \hline ms\hat{I}_j([t, i] \mid q) & \text{local distinctiveness} \end{array}\right\} \qquad \text{(XIII; 18)}$$

XIII.6.2 (XIII; 17) and (XIII; 18) illustrate again the main point of IX.3.2, showing how can one gain some *floral* inference from some *local* quantities (and *vice versa*) *via* some proper *floristic* estimate. The same is true for (XIII; 19),

$$\left.\begin{array}{r|l} ms\hat{I}_j([t, i] \mid q) & \text{local distinctiveness} \\ +ms\hat{I}_j([q, i] \mid t) & \text{floral distinctiveness} \\ \hline ms\hat{I}_j([q, t, i]) & \text{floristic interaction} \end{array}\right\} \qquad \text{(XIII; 19)}$$

where the equality holds, even if (XIII; 19) implies only the sum of two *partial* estimates, because in this primitive system of models (as a consequence of our probability field is being a *classical* one) several different quantities may have the same value, e.g. $ms\widehat{H}_j(i \mid q) = ms\hat{I}_j(t, [q, i]) = ms\hat{I}_j([t, i] \mid q) = sm \log m - \alpha_j$. (Similar correspondences have been already taken into account.) Floristic estimates, of course, have their own reference with respect to each other, e.g.

$$\left.\begin{array}{r|l} ms\hat{I}_j([q, t] \mid i) & \text{floristic efficiency} \\ +ms\hat{I}_j(i, [q, t]) & \text{floristic insaturation} \\ \hline ms\hat{I}_j([q, t, i]) & \text{floristic interaction} \end{array}\right\} \qquad \text{(XIII; 20)}$$

XIII.7 Without going any further into some important details of the three partners modelling, let us consider some *close relations* between *associatum* and certain *estimates* treated above. Suppose we are interested here in the boundary relations of $m\hat{I}_j(\lambda)$.

XIII.7.1 Starting with *f*-marginals (viz. marginal idealizations), it is not too difficult to see that

(1) if a K_j is in an *F-l*-state, then $m\hat{I}_j(\lambda) = 0$, because $sm \log m = \alpha_j$, and $m\widehat{H}_j^{(\varphi)} = 0$,

(2) if a K_j is in an *L*-state, then $m\hat{I}_j(\lambda) = ms\hat{I}_j(i, [q, t]) - m\widehat{H}_j^{(\varphi)}$,

(3) if, in addition, an *L*-state is an *f-L* state, then, since $m\widehat{H}_j^{(\varphi)} = m\hat{I}_j(i, [q, t])$, $m\hat{I}_j(\lambda) = (s-1)m\hat{I}_j(i, [t, q]) = (s-1)mH_j^{(\varphi)}$,

(4) and, if our *f-L*-state occurs at A_{comp}, then $m\widehat{H}_j^{(\varphi)} = m$, and $m\hat{I}_j(\lambda) = (s-1)m$, which is the *global maximum* of associatum.

XIII.7.2 In a more realistic sense (i.e. in *a topologically "local" sense* with respect to the scale),

$$\max \hat{I}_j(\lambda) = m\bar{I}_j(\lambda) = ms\hat{I}_j(i, [q, t]) - m\widehat{H}_j^{(\varphi)}, \qquad \text{(XIII; 21)}$$

and, therefore,

$$m\bar{I}_j(\lambda) - m\hat{I}_j(\lambda) = ms\hat{I}_j(q, i), \qquad \text{(XIII; 22)}$$

that is, the difference between the *maximal* and *actual* values of associatum is equal to *local evenness.* (In other words, local evenness indicates the "*niveau-interval*" within which associatum can move.)

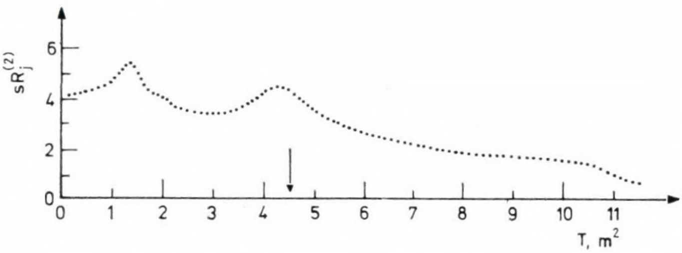

Fig. 6. An $sR_j^{(2)}(\lambda)$ *versus* T curve (see XIII.7)

XIII.7.3 According to (XIII; 22), it is plausible to introduce (XIII; 23), *a redundancy estimate,*

$$s\widehat{R}_j^{(2)}(\lambda) = 1 - m\hat{I}_j(\lambda)/ms\hat{I}_j(q, i), \qquad \text{(XIII; 23)}$$

in order to characterize a better *relative associatum* than (X; 1). Fig. 6 shows the curve of (XIII; 23) for the same object as fig. 4 does. The present curve is in many respects similar to that of fig. 4, but the second peak has a transformation toward $A_{\min}$ which property might make (XIII; 23) more "realistic" than (X; 1). For instance, the position of this second peak is quite close to the point, indicated in fig. 6, where the pooled positive association has its maximum.

XIII.8 A short comment is needed (at least, in the form of a few allusions to be developed later on) to call attention to some important relations between *equivalence analysis* and *three partners modelling.*

XIII.8.1 One of the key estimates is $ms\hat{I}_j([q, t] \mid i)$, *floristic efficiency.* It is easy to recognize that if a K_j is well-ordered, then

$$\left.\begin{aligned} &N_j\hat{I}_j(E) + n_j\hat{I}_j(e) = ms\hat{I}_j(\varepsilon) = \\ &= \gamma_j - \alpha_j - (\varepsilon_j - \delta_j) = msI_j([q, t] \mid i), \end{aligned}\right\} \qquad \text{(XIII; 24)}$$

because $(\varepsilon_j - \delta_j) = N_j\widehat{H}_j(V_q \mid V_R) + n_j\widehat{H}_j(v_q \mid v_r) = \beta_j$. Since the equation

$$\left.\begin{aligned} &N_j\hat{I}_j(K) + n_j\hat{I}_j(k) = ms\hat{I}_j(X) = \\ &= \gamma_j - \alpha_j - \beta_j = ms\hat{I}_j([q, t] \mid i) \end{aligned}\right\} \qquad \text{(XIII; 25)}$$

is satisfied for any K_j, if a K_j is not ordered, then the joint estimate

$$\widehat{Y}_j + \hat{y}_j = \hat{\eta}_j = \varepsilon_j - \delta_j - \beta_j \qquad \text{(XIII; 26)}$$

is to be taken into account.

XIII.8.2 This has the consequence that a number of equations of this chapter can be rewritten in a suitable form, e.g.

$$ms\hat{I}_j([q, i] \mid t) = ms \log s - (\varepsilon_j - \delta_j) + \hat{\eta}_j, \qquad \text{(XIII; 27)}$$

which relates floral distinctiveness to some degree of orderliness. If K_j approaches orderliness, then $(\varepsilon_j - \delta_j) \to \beta_j$, and, $\hat{\eta}_j \to 0$.

XIV. Two partners modelling

XIV.1 Suppose that, contrary to the reasoning of XIII.1, partner i is neglected on the whole, and only partner q and t are regarded as "*true partners*", in exactly the same way as before (e.g. assuming classical probability field). These two partners are characterized at the first instant by f-marginals. (This kind of argument is perhaps "fitting" better to the traditional thinking of the "trade" than XIII.) It should be recognized, however, that we have already used (say, in an implicit form) *two kinds of two partners modelling*, associated with $N_j\hat{I}_j(E)$ and $N_j\hat{I}_j(K)$. Therefore, the present task is to make relevant relations more explicit, and try to show the proper interconnections between these constructions.

XIV.2 First of all, it is necessary to reconsider *coincidence information*, $N_j\hat{I}_j(K)$, and *incoincidence information*, $n_j\hat{I}_j(k)$, more closely, with particular reference to *extremal values* and *boundary relations*.

XIV.2.1 If joint entropy, $N_j\widehat{H}_j([V_q\ V_t]) = C_j - 0 = C_j$, is introduced, and $N_j\widehat{H}_j(V_q)$, $N_j\widehat{H}_j(V_t)$ are regarded as "familiar quantities" (see XII.1), then $N_j\hat{I}_j(K)$ can be redefined such as

$$\left.\begin{aligned} N_jH_j(V_q) + N_jH_j(V_t) - N_jH_j([V_q, V_t]) = \\ = C_j - A_j - B_j, \end{aligned}\right\} \qquad \text{(XIV; 1)}$$

with $N_j\widehat{H}_j(V_q \mid V_t) = B_j$, and $N_j\widehat{H}_j(V_t \mid V_q) = A_j$. *Mutatis mutandis*, similar relations are true for $n_j\hat{I}_j(k)$ as well.

XIV.2.2 If K_j is in an L-f-state, then $N_j\widehat{H}_j(V_q \mid V_t) = N_j \log \bar{s}_j$, $N_j\widetilde{H}_j(V_t \mid V_q) = N_j \log m$, and, consequently, $N_j\hat{I}_j(K) = 0$. If K_j is either in a single f-state, or in a single l-state, without being in an oligovalence state on the part of the other type of marginals (like in the case of an [IRC]), then both $N_j\hat{I}_j(K)$ and $n_j\hat{I}_j(k)$ are greater than zero. If K_j is in an f-l-state, then

$N_j\widehat{H}_j(V_q \mid V_t) = N_j \log \bar{s}_j$, $N_j\widehat{H}_j(V_t \mid V_q) = N_j \log \overline{m}_j$, and so coincidence information has its *maximum*, $N_j\bar{I}_j(K)$,

$$N_j\bar{I}_j(K) = N_j \log \overline{N}_j \tag{XIV; 2}$$

where $\overline{N}_j = ms/N_j$ (mean *total densiy* for coincidences), and, consequently, the same is true for incoincidence information,

$$n_j\bar{I}_j(k) = n_j \log \bar{n}_j, \tag{XIV; 3}$$

where $\bar{n}_j = ms/n_j$ (mean *total density* for incoincidences).

XIV.2 For interpreting the pairwise bound,

$$\left.\begin{array}{l} 0 \leq N_j\hat{I}_j(K) \leq N_j\bar{I}_j(K) \\ 0 \leq n_j\hat{I}_j(k) \quad \leq n_j\bar{I}_j(k) \end{array}\right\} \tag{XIV; 4}$$

both the equations of (XIV; 5) and (XIV; 6),

$$\left.\begin{array}{l} \Delta N_j\widehat{H}_j\,(V_q) + \nabla N_j\widehat{H}_j\,(V_q) = \\ = \Delta N_j\widehat{H}_j\,(V_t) + \nabla N_h\widehat{H}_j\,(V_t) = N_j\bar{I}_j(K), \end{array}\right\} \tag{XIV; 5}$$

$$\left.\begin{array}{l} \Delta n_j\widehat{H}_j\,(v_q) + \nabla n_j\widehat{H}_j\,(v_q) = \\ = \Delta n_j\widehat{H}_j\,(v_t) + \nabla n_j\widehat{H}_j\,(v_t) = n_j\bar{I}_j(k), \end{array}\right\} \tag{XIV; 6}$$

and the additive relation of (XIV; 7),

$$\left.\begin{array}{l} N_j\hat{I}_j(K) + n_j\bar{I}_j(k) = ms\bar{I}_j(X) = \\ \quad = ms\hat{I}_j(i, [q, t]), \end{array}\right\} \tag{XIV; 7}$$

can be taken into account. (XIII; 25) and (XIV; 7) show clearly that in the present case there exist similar subdivisions, composed of the same estimates (namely, *f-deviates* of XII.), than in (XIII; 8), or in (XIII; 14). Further, by virtue of (XIII; 20), (XIII; 25), and (XIV; 7), we can see the intimate relations between three and two partners modellings, for instance, in the form of (XIV; 8),

$$\left.\begin{array}{l} N_j\hat{I}_j(K) + n_j\hat{I}_j(k) + ms\hat{I}_j(i, [q, t]) = \\ \quad = ms\hat{I}_j([q, t, i]), \end{array}\right\} \tag{XIV; 8}$$

where, in essence, we have the same subdivision than in (XIII; 20). If we define two special gain of information estimates of the second order,

$$\left.\begin{array}{l} \Delta N_j\hat{I}_j(K) = N_j\bar{I}_j(K) - N_j\hat{I}_j(K) \\ \Delta n_j\hat{I}_j(k) \quad = n_j\bar{I}_j(K) \quad - n_j\hat{I}_j(k) \end{array}\right\} \tag{XIV; 9}$$

then it is to be recognized that

$$\Delta N_j\hat{I}_j(K) + \Delta n_j\hat{I}_j(k) - ms\hat{I}_j(i, [q, t]) = ms\hat{I}_j(\langle q; t; i\rangle), \quad \text{(XIV; 10)}$$

where $ms\hat{I}_j(\langle q; t; i\rangle) = \alpha_j + \beta_j - \gamma_j$, strictly defined *interaction*, mentioned in XIII.6. (XIV; 10) means that the broadly and strictly defined interactions play central role in interpreting *f-marginal states* by means of two partners modelling.

XIV.3 In order to understand $N_j\hat{I}_j(E)$ even better than in part 1, or in XI., we can use the results of III. and XIV.2 for *another revision of equivalence information.*

XIV.3.1 Starting with the well-known equations of (XIV; 11),

$$\left.\begin{aligned} N_i\hat{I}_j(E) &= N_j\hat{I}_j(K) - \hat{Y}_j, \\ \hat{Y}_j &= E_j - D_j - B_j = N_j\hat{H}_j(V_q \mid V_R) - B_j, \end{aligned}\right\} \quad \text{(XIV; 11)}$$

it is to be recognized that *subcoincidence information*, $\hat{Y}_j$, can be regarded *by definition* as the sum in (XIV; 12),

$$\hat{Y}_j = \hat{y}_{j0} + \hat{y}_{j1} + \ldots + \hat{y}_{jr} + \ldots + \hat{y}_{js}, \quad \text{(XIV; 12)}$$

where the values of the multiplet variable R, $R = 0, 1, \ldots, s$, are implied in members of (XIV; 12) as indexing numbers at t_j (see III.). Note that $\hat{y}_{j0} = \hat{y}_{js} = 0$, for any [ET], because $\binom{s}{0} = \binom{s}{s} = 1$. If we have, for instance, an [IRC], where, as in (V; 8), $s = 4$, $m = s^s = 16$, then we can set up easily the three non-zero valued *subcontingency tables*, like (III; 3), such as

$$\begin{array}{cccc|c} \multicolumn{5}{c}{(R=1)} \\ 1&0&0&0&1\\ 0&1&0&0&1\\ 0&0&1&0&1\\ 0&0&0&1&1\\ \hline 1&1&1&1&4 \end{array} \qquad \begin{array}{cccccc|c} \multicolumn{7}{c}{(R=2)} \\ 1&1&1&0&0&0&3\\ 1&0&0&1&1&0&3\\ 0&1&0&1&0&1&3\\ 0&0&1&0&1&1&3\\ \hline 2&2&2&2&2&2&12 \end{array} \qquad \begin{array}{cccc|c} \multicolumn{5}{c}{(R=3)} \\ 1&1&1&0&3\\ 1&1&0&1&3\\ 1&0&1&1&3\\ 0&1&1&1&3\\ \hline 3&3&3&3&12 \end{array} \quad \text{(XIV; 13)}$$

for which $Y_j = 8 + 12 + 4.9812 = 24.9812$. Let us call one table in (XIV; 3), or any such table in a similar series of tables, *an elementary contingency subtable*, [EST], and try to understand relations of (IX; 11) on the basis of (XIV; 12) by means of (XIV; 13).

XIV.3.2 It is fairly easy to conceive that for (XIV; 13), or, even for any [IRC], the pair of estimates, $\hat{Y}_j = \overline{Y}_j$, $\hat{y}_j = \overline{y}_j$, are of *maximum* values, because *any [EST] is in a double monovalence state*, and this *interlocal property* of an [ET] is responsible for such relations as $N_j\hat{I}_j(K) - \overline{Y}_j = N_j\hat{I}_j(E) = 0$, $n_j\hat{I}_j(k) - \overline{y}_j = n_j\hat{I}_j(e) = 0$. Generally speaking, and using the notation of

(III; 4), for any [EST] of an [IRC], where $0 < r < s$, the conditions of (XIV; 14)

$$
\begin{array}{c|l}
 & \left.\begin{array}{l} r\, g_{jr}/s \\ \cdot \\ \cdot \\ \cdot \\ r\, g_{jr}/s \end{array}\right\} s \\
\hline
\underbrace{r \,.\,.\,.\,.\,.\, r}_{g_{jr} = \binom{s}{r}} & r\, g_{jr}
\end{array}
\qquad \text{(XIV; 14)}
$$

are always satisfied; thus, $\overline{y}_{jr}$, contingency information estimate for [EST] in (XIV; 14) can be expressed as

$$\overline{y}_{jr} = g_{jr} r \log (s/r), \qquad \text{(XIV; 15)}$$

and, therefore,

$$
\left.\begin{array}{l}
\overline{Y}_j = \sum\limits_{r=1}^{r-1} g_{jr} r \log (s/r) = \\
= N_j \log s - B_j = \nabla N_j \widehat{H}_j(V_t) \,.
\end{array}\right\} \qquad \text{(XIV; 16)}
$$

For illustrating these extremely simple relations, we can write down the values of (XIV; 15) for (XIV; 13) in order, gaining $\hat{y}_{j1} = 4 \cdot 1 \log (4/1) = 8$, $\hat{y}_{j2} = 6 \cdot 2 \log (4 \cdot 2) = 12$, $\hat{y}_{j3} = 4 \cdot 3 \log (4/3) = 4.9812$, and so $\widehat{Y}_j = 24.9812$, as before.

XIV.3.3 Contrary to the conditions of XIV.3.2, if a K_j is interlocally well-ordered, i.e. if it represents an [IOC], then each [EST] is in *an oligovalence state*, like (XIV; 17),

$$
\begin{array}{c|l}
 & \left.\begin{array}{l} g_{jr} \\ \cdot \\ \cdot \\ \cdot \\ g_{jr} \end{array}\right\} r \\
 & \left.\begin{array}{l} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{array}\right\} (s - r) \\
\hline
\underbrace{r \,.\,.\,.\, r}_{g_{jr}} & r\, g_{jr}
\end{array}
\qquad \text{(XIV; 17)}
$$

and, therefore, $\hat{y}_{jr} = 0$, $\hat{Y}_j = 0$. There are, of course, certain cases, like in (XIV; 18),

$$\begin{array}{c|cccccccc|c} u & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ v & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 5 \\ w & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 5 \\ z & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 5 \\ \hline & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 16 \end{array} \qquad \text{(XIV; 18)}$$

where we have a mixture of (XIV; 14) and (XIV; 18), and, therefore, $\hat{Y}_j = 8$ for (XIV; 18). We can rightly suppose that a still better understanding of orderliness is possible with respect to these extreme cases.

XIV.4 Using reasoning of XIV.2, (XIV; 16) implies *the double bound*

$$\left.\begin{aligned} Y_j &\leq \nabla N_j\hat{H}_j(V_t), \\ \hat{\eta}_j &\leq msI_j([q, i] \mid t), \end{aligned}\right\} \qquad \text{(XIV; 19)}$$

which shows how equivalence analysis is related to both three partners and two partners modellings. If a K_j is in an [IRC], then the upper bound of (XIV; 19) is satisfied which means that floral distinctiveness of (XIII; 27) is uniquely defined by $\hat{\eta}_j$, that is, $\Delta N_j\hat{H}_j(V_t) + \Delta n_j\hat{H}_j(v_t) = \hat{\eta}_j = ms\hat{I}_j(X) = msI_j([q, t] \mid i)$. If a K_j in an [IOC], then $\hat{Y}_j$ and $\hat{\eta}_j$ disappear, and, therefore, $ms\hat{I}_j(\varepsilon)$ is uniquely defined by $ms\hat{I}_j(X)$, and, by relations of (XIV; 10).

XIV.5 Of course, we might be interested in the problem of how orderliness of some kind is composed of some *elementary relations* (e.g. $q_i \in Q$), or, of some *coalitional relations* (e.g. $Q_k \subset Q$). This problem leads us to the family of diacretic models whose certain implications and importance will be discussed later on (see XVII.).

[C] SOME DIACRETIC MODELS

XV. Decomposition of associatum

XV.1 If we would like to study the populations one by one without "loosing the common touch" (and, roughly speaking, this is implied in the term "*dia*cretic modelling"), then a plausible way of doing this is shown by the following simple formalism.

XV.2 Starting with XIII.4.4, in particular, with (XIII.13), let us consider population "a". By letting

$$m\hat{H}_j(A \mid [B, C, \ldots, S]) = m\hat{H}_j(\{A\}), \qquad \text{(XV; 1)}$$

where

$$m\hat{H}_j(\{A\}) = m\hat{H}_j(A) - m\hat{I}_j(\langle A\rangle), \tag{XV; 2}$$

where

$$m\hat{I}_j(\langle A\rangle) = m\hat{I}_j(A, [B, C, \ldots, S]) \tag{XV; 3}$$

$$= m\hat{I}_j(\lambda) - m\hat{I}_j([\bar{A}]), \tag{XV; 4}$$

and where

$$m\hat{I}_j([\bar{A}]) = m\hat{I}_j([B, C, \ldots, S]), \tag{XV; 5}$$

it is possible to introduce the following decomposition (to be called more precisely, *an elementary decomposition*) of associatum, with respect to one single population, namely, "*a*":

$$\left.\begin{array}{lll}
(1) & m\hat{H}_j(A) & \text{local entropy} \\
(2) & -m\hat{H}_j(\{A\}) & \text{complete dissociation} \\
\hline
(3) & m\hat{I}_j(\langle A\rangle) & \text{complete association} \\
(4) & +m\hat{I}_j([\bar{A}]) & \text{subassociatum} \\
\hline
(5) & m\hat{I}_j(\lambda) & \text{associatum}
\end{array}\right\} \tag{XV; 6}$$

A similar decomposition is straightforward for all elements of Q.

XV.2.1 First, an explanation of terms is in order (see Fig. 6). (2) in (XV; 6) is *a multiple conditional entropy* estimate for "*a*" with respect to all other elements of Q. In other words, $m\hat{H}_j(\{A\})$ is a measurable subset of $m\hat{H}_j(A)$, local entropy, that is not "shared" by any other population. (Hence the name "*complete dissociation*" refers to the "unique interlocal behaviour" of the population concerned.) On the other hand, (3) in (XV; 6) is *complete association* of "*a*", because it is the measurable subset of $m\hat{H}_j(A)$ which is in *some* "interaction" with others, in the interlocal sense of the word (i.e. it is "shared" by at least one other population). (3), in the form of (XV; 3), is *a multiple information* estimate; that is, the difference of two closely related quantities, *associatum and subassociatum.*

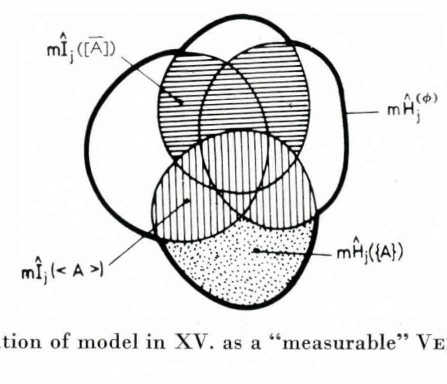

Fig. 7. Graphical explanation of model in XV. as a "measurable" VENN-diagram (see XV.2.1)

XV.2.2 Secondly, note that in the first part of (XV; 6), namely in (1)—(3), the main purpose of the construction is *to distinct* "*a*"; in the second part /(4)—(5)/ of it, a special kind of *elimination of* "*a*" is aimed at. The key concept of this elimination is *subassociatum*, related to $[\bar{A}]$, $[\bar{A}] = [B, C, \ldots, S]$, a joint scheme *not* containing "*a*". Therefore,

$$m\hat{I}_j([\bar{A}]) = m\widehat{H}_j([L \mid A]) - m\widehat{H}_j^{(\varphi|A)}, \qquad \text{(XV; 7)}$$

where $mH_j([L \mid A])$ is the pooled local entropy with the omission of "*a*"-data (local distinctiveness without "*a*"), and

$$m\widehat{H}_j^{(\varphi|A)} = m\widehat{H}_j([B, C, \ldots, S]), \qquad \text{(XV; 8)}$$

where $m\widehat{H}_j^{(\varphi|A)}$ is *floral subdiversity* of an [ET], if local vector belonging to "*a*" is eliminated. In other words, subassociatum is joint (total) information of a $2 \times 2 \times \ldots^{s-1} \times 2$ table, instead of being a joint information estimate of a $2 \times 2 \times \ldots^{s} \times 2$ table.

XV.3 In a more detailed way,

$$\begin{aligned} m\hat{I}_j(\langle A \rangle) &= m\hat{I}_j(\lambda) - m\hat{I}_j([\bar{A}]) = \\ &= m\widehat{H}_j([L]) - m\widehat{H}_j^{(\varphi)} - m\widehat{H}_j([L \mid A]) + m\widehat{H}_j([\bar{A}]) = \\ &= \{m\widehat{H}_j[L] - m\widehat{H}_j([L \mid A])\} - \{m\widehat{H}_j^{(\varphi)} - m\widehat{H}_j([\bar{A}])\}, \end{aligned} \qquad \text{(XV; 9)}$$

where (XV; 9) is *a difference* of *two kinds* of *gain of information* estimates. The first member of (XV; 9) refers to *local quantities*, only; the second member of (XV; 9) reflects to *interlocal quantities*, which implies that

$$m\widehat{H}_j(\{A\}) = m\widehat{H}_j^{(\varphi)} - m\widehat{H}_j([B, C, \ldots, S]). \qquad \text{(XV; 10)}$$

(XV; 10) is practically useful, because, in most cases, total dissociation can optimally estimated by means of subdiversity values. "*Subdiversity*", as a concept, has a special interest of its own; the present meaning is simple, because of the combinatorial interpretation of $m\widehat{H}_j^{(\varphi)}$.

XV.4 Due to the meaning of terms in (XV; 6), it is possible to introduce the concepts of *associativity* and *dissociativity* of population in the following sense of the words.

XV.4.1 It is plausible to state that (*a*) *associativity* of population "*a*" is of *maximum* value if and only if $m\widehat{H}_j(\{A\}) = 0$, $m\hat{I}_j(\langle A \rangle) = m\widehat{H}_j(A)$, and (*b*) associativity of "*a*" is of *minimum* value, if and only if $m\hat{I}_j(\langle A \rangle) = 0$, $m\widehat{H}_j(\{A\}) = m\widehat{H}_j(A)$.

XV.4.2 For reasons to be considered later on (see XVI.), it is better to follow another way of thinking of defining dissociativity. First, it is to be noted that complete dissociation estimates for all elements of Q are "independent" from each other, and therefore, *additive*,

$$m\widehat{H}_j(\{A\}) + m\widehat{H}_j(\{B\}) + \ldots + m\widehat{H}_j(\{S\}) = m\widehat{H}_j\{\delta_\lambda^{(s)}\}, \quad \text{(XV; 11)}$$

where $m\widehat{H}_j\{\delta_\lambda^{(s)}\}$ is called *total dissociatum* of Q. (Total dissociatum is a measurable subset of $m\widehat{H}_j^{(\varphi)}$ which is excluded from all kinds of "interactions", i.e. manifested interlocal relations.) Secondly, by supposing that $m\widehat{H}_j\{\delta_\lambda^{(s)}\} > 0$, we can define for population "a"

$$\widehat{R}_j(\{A\}) = m\widehat{H}_j(\{A\})/m\widehat{H}_j\{\delta_\lambda^{(s)}\} \quad \text{(XV; 12)}$$

which redundancy estimate makes possible *to rank* populations according to their degrees of dissociativity ($0 \leq \widehat{R}_j(\{A\}) \leq 1$).

XV.5 For the sake of a primary illustration, let us consider an [ET] in (XV; 13), where $s = 3$, $m = 8$, $m\widehat{H}_j^{(\varphi)} = 16$, $m\hat{I}_j(\lambda) = 8$, $m\widehat{H}_j\{\delta_\lambda^{(s)}\} = 8$,

a	0	0	0	0	1	1	1	1	4
b	0	0	1	1	0	0	1	1	4
c	0	0	0	0	1	1	1	1	4
	0	0	1	1	2	2	3	3	12

(XV; 13)

The values of estimates (XV; 6) are shown by (XV; 14).

	a and *c*	*b*
(1)	8	8
(2)	−0	−8
(3)	8	0
(4)	+0	+8
(5)	8	8

(XIV; 14)

We can see at once that "*a*" and "*c*" have a maximum degree of associativity, and a minimum degree of dissociativity, while "*b*" has an opposite qualification (with maximum degree of dissociativity, and minimum degree of associativity). This observation is consistent with proper interlocal relations as well. Since the pairs (a, b) and (b, c) are stochastically independent, but the pair (a, c) is *not* independent (i.e. interlocally dependent), the special interlocal role of "*b*" is explained by the relation that "*b*" *is* element of both independent pairs, but it *is not* an element of the single dependent pair. This kind of simple reasoning can be applied to an [IRC], as (IX; 5), where, since all the possible pairs are independent, and therefore all elements have a maximum degree of dissociativity, and a minimum degree of associativity.

XV.5.1 Turning to a less trivial case, consider (IX; 2), where $s = 4$, $m = 8$, and where $m\widehat{H}_j^{(\varphi)} = 19.2453$, $m\hat{I}_j^{(\lambda)} = 9.3717$, $m\widehat{H}_j\{^{(4)}_{\lambda}\delta\} = 10.3263$. The estimates of (XV; 6) for (IX; 2) is shown by (XV; 15):

	u	v	w	z
(1)	8.0000	6.4906	7.6358	6.4906
(2)	-3.2353	-0.0000	-3.2453	-4.0000
(3)	4.7547	6.4906	4.3905	2.4906
(4)	$+4.6170$	$+2.8811$	$+4.9812$	$+6.8811$
(5)	9.3717	9.3717	9.3717	9.3717

(XV; 15)

According to (XV; 15), (*a*) *the associativity ranking* of elements in (IX; 2) is: $v > u > w > z$; while (*b*) *the dissociativity ranking* of these elements is: $z > | u \sim w | > v$. Note that (*a*) is not related trivially to (*b*), *et vice versa*; for instance, we cannot explain, without further considerations, why the tie "$u \sim w$" exists in (*b*), or, why v has in (XV; 15) its zero-valued complete dissociation. Note, further, that the values of total dissociation can easily be estimated by (XV.10), as it is shown by (XV; 16),

	u	v	w	z
$\langle 1\rangle$	19.2453	19.2453	19.2453	19.2453
$-\langle 2\rangle$	16.0000	19.2453	16.0000	15.2453
$\langle 3\rangle$	3.2453	0.0000	3.2453	4.0000

(XV; 16)

where $\langle 1\rangle$ is floral diversity, $\langle 2\rangle$ is floral subdiversity, and $\langle 3\rangle$ is total dissociation.

XV.5.2 Ranking procedures are fairly important in understanding of how elements are related to each other (see Orlóci 1978b). There are, however, different types ranking, and simple "cumulative ranks" (i.e. ranks by means of pooled information estimates) might be rather misleading in several cases.

XVI. Notes on further modelling

XVI.1 Clearly, (XV; 6) represents an *elementary* decomposition of associatum, because this simple additive model refers to one single element of Q, $q_i \in Q$, only. Suppose that we would like to generalize the reasoning of XV., in order to cover *coalitional relations* (or, speaking in more traditional terms, "group relations") of a Q as well, related to some Q_k coalition of Q, $Q_k \subset Q$. This intention is justitiable for a number of good motives. For instance, this kind of generalization opens up some new possibilities to a better interconnection of synmorphology and syntaxonomy (say, clustering methods) in the future. Similarly, if we were able to attach analogous estimates to coalitions (and, say, to rank these coalitions according to their complete association or dissociation of some kinds), then our knowledge would be more advanced as to the "collective structural relations" of the communities.

XVI.2 The first obstacle of making such generalizations is that $m\widehat{H}_j^{(\varphi)} \neq \neq m\hat{I}_j(\lambda) + m\widehat{H}_j\{\delta_\lambda^{(s)}\}$, because of the existence of "multiple overlaps". Taking again the simple example of (IX; 2), where $m\widehat{H}_j^{(\varphi)} = 19.2453$, $m\hat{I}_j(\lambda) = = 9.3717$, $m\widehat{H}\{\delta_\lambda^{(4)}\} = 10.3263$, the inequality indicated before is due to a quantity, 0.453 in value, which is to be called "*interassociatum*" in this chapter. Interassociatum may be, in a number of cases, of small value that can be occasionally neglected. This negligence, however, should be an *a posteriori* operation; *a priori*, it is at least to be clarified what is to be meant by interassociatum. Our present aim cannot go any further than to study briefly some relevant estimates and basic relations whose better understanding and usage will be treated in some next parts of this series of papers (or, elsewhere).

XVI.3 Suppose that, in analyzing an [ET], we distinguish some multiplets, say, *v-test of elements*. Suppose, further, that we confine here our attention to *local v-tets*, $2 \leq v \leq (s-1)$. A local multiplet of v order can be studied by means of *an interlocal scheme* of the same order; the best known example is a set of 2×2 table, if $v = 2$. If we want to generalize the traditional "pairwise reasoning" of association analysis to any v, then we have to introduce a general notation.

XVI.3.1 Let $m\widehat{H}_j\{\Sigma_\lambda^{(v)}\}$ be *the pooled interlocal entropy estimate* of the v-order, i.e. for the whole set of the joint schemes of the v-order. Let $m\hat{I}_j\{\Sigma_\lambda^{(v)}\}$ be corresponding *pooled information estimate* of the same order such as

$$
\begin{aligned}
m\hat{I}_j\{\Sigma_\lambda^{(v)}\} &= \binom{s-1}{v-1}[sm \log m - \alpha_j] - m\widehat{H}_j\{\Sigma_\lambda^{(v)}\} = \\
&= \binom{s-1}{v-1} ms\hat{I}_j([t, i] \mid q) - m\widehat{H}_j\{\Sigma_\lambda^{(v)}\},
\end{aligned}
\qquad \text{(XVI; 1)}
$$

where $ms\hat{I}_j([t, i] \mid q)$ is again *local distinctiveness* (see XIII.4).

XVI.3.2 In order to illustrate (XVI; 1), let us turn again to (IX; 2), where $ms\hat{I}_j([t, i] \mid q) = 28.617$, $m\widehat{H}_j\{\Sigma_\lambda^{(2)}\} = 76.0263$, $m\widehat{H}_j\{\Sigma_\lambda^{(3)}\} = 66.4906$, and thus

$$
\left.
\begin{aligned}
m\hat{I}\{\Sigma_\lambda^{(2)}\} &= \binom{3}{1} 28.617 - 76.0263 = 9.8247, \\
m\hat{I}_j\{\Sigma_\lambda^{(3)}\} &= \binom{3}{2} 28.617 - 66.4906 = 19.3604.
\end{aligned}
\right\}
\qquad \text{(XVI; 2)}
$$

The sums gained in (XVI; 2) have to correspond to the sums of proper interlocal association estimates, e.g. $m\hat{I}_j\{\Sigma_\lambda^{(2)}\} = m\hat{I}_j([U, V]) + \ldots + m\hat{I}_j([W, Z]$ where $m\hat{I}_j([U, V]) = 2.4906$, $m\hat{I}_j([W, Z]) = 0.1264$. By listing information estimates for triplets, $m\hat{I}_j([U, V, W]) = 6.8811$, $m\hat{I}_j([U, V, Z]) = 4.9812$, $m\hat{I}_j([U, W, Z]) = 2.8811$, $m\hat{I}_j([V, W, Z]) = 4.617$, and by adding these values, we gain again the sum 19.3604, as in (XVI; 2).

XVI.3.3 Having these primary pooled quantities, we can estimate *interassociatum of the* (v + 1) *order* such as

$$\begin{aligned} mI_j([\lambda_{(v+1)}]) = \left[\binom{s-1}{v-1} - (v-1)\right] ms\hat{I}_j([t, i] \mid q) - m\widehat{H}_j\{\Sigma_\lambda^{(v)}\} + \\ + (v-1)\, m\widehat{H}_j^{(\varphi)} = m\hat{I}_j\{\Sigma_\lambda^{(v)}\} - (v-1)\, m\hat{I}_j(\lambda) \end{aligned} \qquad \text{(XVI; 3)}$$

For instance, (XVI; 3) has the pair of values for our data of XVI.3.2,

$$\left.\begin{aligned} m\hat{I}í([\lambda_{(4)}]) &= 19.3604 - 18.7434 = 0.617, \\ m\hat{I}í([\lambda_{(3)}]) &= 9.8247 - 9.3717 = 0.453. \end{aligned}\right\} \qquad \text{(XVI; 4)}$$

XVI.3.4 The meaning of interassociatum of some order corresponds to *multiple interaction*, exactly in the sense of KULLBACK (1959). Speaking perhaps in more picturesque terms, interassociatum measures *the multiple overlaps* of local entropy estimates of an [ET] at a given order. As KULLBACK explains it properly, such interaction terms may be *of negative values*. The author has the experience that, during a secondary succession, the progress of degradation is frequently resulted in the increase of negative interassociatum estimates, sometimes for the majority of orders.

XVI.4 The concept of interassociatum is closely and complementarily related to that of *dissociatum*. The primary point is that if we do not consider *total dissociatum* (as in XVI.4.2), then, as in XVI.3, we should tackle dissociatum estimates of different orders.

XVI.4.1 Using the same reasoning and symbolism as before, *dissociatum of the* $(v+1)$ *order* can be estimated as

$$\left.\begin{aligned} m\widehat{H}_j\{\delta_\lambda^{(v+1)}\} = \left[\binom{s-1}{v-1} - v\right] ms\hat{I}_j([t, i] \mid q) - m\widehat{H}_j\{\Sigma_\lambda^{(v)}\} + \\ + (v+1) m\widehat{H}_j^{(\varphi)} = ms\hat{I}_j([t, i] \mid q) - [(v+1) mI_j(\lambda) - m\hat{I}_j\{\Sigma_\lambda^{(v)}\}]. \end{aligned}\right\} \qquad \text{(XVI; 5)}$$

If $v = (s-1)$, then (XVI; 5) is happily reduced to (XVI; 6),

$$m\widehat{H}_j\{\delta_\lambda^{(s)}\} = sm\widehat{H}_j^{(\varphi)} - m\widehat{H}_j\{\Sigma_\lambda^{(s-1)}\}, \qquad \text{(XVI; 6)}$$

where total dissociatum can be regarded as a special kind of gain of information estimate. For data of (XI; 2), the value of (XVI; 6) can be given as $m\widehat{H}_j\{\delta_\lambda^{(4)}\} = 4 \cdot 19.2453 - 66.4906 = 10.4906$, as before. Note that $sm\widehat{H}_j^{(\varphi)}$ in (XVI; 6) is specifically weighted floral diversity.

XVI.4.3 The main importance of dissociatum is shown by (XVI; 7),

$$m\widehat{H}_j^{(\varphi)} = mH_j\{\delta_\lambda^{(v+1)}\} + m\hat{I}_j(\lambda) - mI_j([\lambda_{(v+1)}]), \qquad \text{(XVI; 7)}$$

where we have a well-defined additive subdivision of floral diversity as an "overall uncertainty bound" for interlocal relations. The particular subdivisions for data of (IX; 2) are given by (XVI; 8),

$$\left.\begin{array}{ll} (v = 3) & 19.2453 = 10.4906 + 9.3717 - 0.617 \\ (v = 2) & 19.2453 = 10.3266 + 9.3717 - 0.453 \end{array}\right\} \qquad \text{(XVI; 8)}$$

As it has been stressed before, interaction terms (i.e. *interassociata*) might be neglected, if the particular values are small, but it is always relevant to know how such an additive composition is built up as (XVI; 7).

XVI.5 (XVI; 7), however, is by no means the only additive subdivision which can be constructed from our present, our closely related estimates. Some examples are given for further use.

XVI.5.1 By simple rearrangement of proper terms, it is obvious that

$$m\hat{I}_j([\lambda_{(v)}]) = m\widehat{H}_j^{(\varphi)} - m\widehat{H}_j\{\delta_\lambda^{(v+1)}\} \qquad \text{(XVI; 9)}$$

We may define a new estimate, called *manifest associatum*,

$$m\hat{I}_j\{\lambda_{(v)}\} = vm\hat{I}_j(\lambda) - m\hat{I}\{\Sigma_\lambda^{(v)}\} \qquad \text{(XVI; 10)}$$

which play a fairly important role in understanding of how associatum and interassociatum are related to each other. Using (XVI; 9) and (XVI; 10), we can obtain an interesting subdivision of $ms\hat{I}_j(i, [q, t])$, *floristic insaturation* (see XIII.4),

$$\begin{array}{r|l} ms\hat{I}_j(q, i & \left.\begin{array}{r} 3.2952 \\ 9.3717 \end{array}\right\} m\hat{I}_j(\langle\lambda\rangle) = 12.6669 \\ +m\hat{I}_j(\lambda) & \\ +m\hat{I}_j\{\lambda_{(3)}\} & \left.\begin{array}{r} 8.7547 \\ 10.4906 \end{array}\right\} mH_j^{(\varphi)} = 19.2453 \\ +mH_j\{\delta_\lambda^{(4)}\} & \\ \hline ms\hat{I}_j(i, [q, t]) & 31.9122 \end{array} \qquad \text{(XVI; 11)}$$

where $m\hat{I}_j(\langle\lambda\rangle)$ is *hyperassociatum*, possible local maximum for given data.

XVI.5.2 By defining two new estimates,

$$ms\hat{I}_j([t, i] \mid q^{(v-1)}) = (v + 1)\, m\hat{I}_j(\lambda) - m\hat{I}_j\{\Sigma_\lambda^{(v)}\}, \qquad \text{(XVI; 12)}$$

called *manifest local distinctiveness* at a given order, and,

$$m\hat{I}_j\{\Pi_\lambda^{(v)}\} = m\hat{I}_j(\lambda) - (s-1)\, m\hat{I}_j([\lambda_{(v+1)}]), \qquad \text{(XVI; 13)}$$

called *local particulatum* at a given order, it is possible to introduce several other subdividions. A particularly important one is shown by (XVI; 14),

$$m\hat{I}_j\{\lambda_{(v)}\} = m\hat{I}_j\{\pi_\lambda^{(v)}\} + m\hat{I}_j([\lambda_{(v+1)}]) \qquad \text{(XVI; 14)}$$

which is a subdivision of manifest associatum into two component estimates, particulatum and interassociatum. The relevance of the concept "particulatum" is shown by the triple comparison:

- *dissociatum* indicates "*no interaction*"
- *particulatum* indicates "*partial interaction*"
- *interassociatum* indicates "*full interaction*"

at a given order and in a given context. In other words, particulatum is the measurable subset of $m\widehat{H}_j^{(q)}$ which is *not shared* by *all elements* of Q at a given order. Using these new estimates, we can produce another subdivision of floristic insaturation, similar to (XVI; 11),

$$\begin{array}{l|r}
ms\hat{I}_j(q, i) & 3.2952 \\
+m\hat{I}_j(\lambda) & 9.3717 \\
+mI_j\{\Pi_\lambda^{(3)}\} & 7.5207 \\
+mI_j\{\lambda_{(4)}\} & 1.2340 \\
+mH_j\{\delta_\lambda^{(4)}\} & 10.4906 \\
\hline
ms\hat{I}_j(i, [q, t]) & 31.9122
\end{array}
\quad \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} ms\hat{I}_j([t, i] \mid q^{(3)} = 18.1264 \qquad \text{(XVI; 15)}$$

XVI.6 Without listing further a number of interesting additive relations, let us turn back to our primary concern (XVI.1), i.e. possible coalitional subdivisions of relevant estimates.

XVI.6.1 First, it is to be noted that, having a number of new estimates, the pooled data of (XV; 15) make sense

$$\left.\begin{array}{llr}
(1) & ms\hat{I}_j([t, i] \mid q) & 28.6170 \\
(2) & -m\widehat{H}\,\{\delta_\lambda^{(s)}\} & -10.4906 \\
\hline
(3) & ms\hat{I}_j([t, i] \mid q^{(s-1)}) & 18.1264 \\
(4) & +m\hat{I}_j\{\Sigma_\lambda^{(s-1)}\} & +19.3604 \\
\hline
(5) & sm\hat{I}_j(\lambda) & 37.4868
\end{array}\right\} \qquad \text{(XVI; 16)}$$

where, for instance, (3) is the proper manifest local distinctiveness, or, (4), sum of subassociatum estimates, appears in (XVI.16) as the pooled information estimates for triplets, as in (XVI; 2).

XVI.6.2 Suppose we are interested here in coalition Q_k, $Q_k \cap \bar{Q}_k = Q$. In a more complicated case, we have a partition of Q, $Q_1, Q_2, \ldots, Q_k, \ldots$, i.e. $\bigcup_k Q_k = Q$, $Q_j \cap Q_k = \emptyset$, $j \neq k$.) Then, *ad analogiam* of (XV; 6), we can set up (XVI; 17) for Q_k

(1)	$m\hat{H}_j([Q_k])$	coalitional entropy	(XVI; 17)
(2)	$+m\hat{I}_j([Q_k])$	coalitional association	
(3)	$m\hat{I}_j\{Q_k\}$	special distinctiveness	
(4)	$-m\hat{H}_j(\Phi : Q_k)$	elimination entropy	
(5)	$m\hat{I}_j(\langle Q_k \rangle)$	complete association	
(6)	$+m\hat{I}_k([Q_k])$	subassociatum	
(7)	$m\hat{I}_j(\lambda)$	associatum	

Special distinctiveness can be estimated as

$$m\hat{I}_j\{Q_k\} = km \log m - \alpha_j^{(k)}, \tag{XVI; 18}$$

where $\alpha_j^{(k)}$, $\alpha_j^{(k)} < \alpha_j$, is the sum of local estimates of those elements that are elements of Q_k. *Elimination entropy*,

$$m\hat{H}_j(\Phi : Q_k) = m\hat{H}_j^{(\varphi)} - m\hat{H}_j([\bar{R}_k]), \tag{XVI; 19}$$

can be regarded as a special kind of gain of information estimate, with respect to Q_k. (XVI; 20) gives a primary illustration of (XVI; 17), only for three pairs of elements, and using only the numbers of (XVI; 17):

(XVI; 20)

	(u, v)	(u, w)	(v, w)
(1)	12.00	15.24	10.39
+(2)	2.49	0.39	3.73
(3)	14.49	15.63	14.12
−(4)	5.24	6.85	7.24
(5)	9.25	8.78	6.88
+(6)	0.12	0.59	2.49
(7)	9.37	9.37	9.37

XVI.6.3 Fortunately enough, (XVI; 17) and (XVI; 20) are "redundant", because

$$\left.\begin{aligned} m\hat{I}_j(\lambda) &= m\hat{H}_j([Q_k]) + m\hat{H}_j([\bar{Q}_k]) - m\hat{H}_j^{(\varphi)} - m\hat{I}_j([Q_k]) - m\hat{I}_j([\bar{Q}_k]) = \\ &= m\hat{I}_j([Q_k], [Q_k]) + m\hat{I}_j([Q_k]) + m\hat{I}_j([\bar{Q}_k]), \end{aligned}\right\} \tag{XVI; 21}$$

and, by (XVI; 21), the relations of (XVI; 17) are reduced to the form of (XVI; 22)

	(u, v)	(u, w)	(u, z)
$m\hat{I}_j([Q_k])$	2.4906	0.3905	2.4906
$+m\hat{I}_j([Q_k], [\bar{Q}_k])$	6.7547	8.3905	3.1452
$m\hat{I}_j(\langle Q_k \rangle)$	9.2453	8.7810	5.6358
$+m\hat{I}_j([\bar{Q}_k])$	0.1264	0.5907	3.7379
$m\hat{I}_j(\lambda)$	9.3717	9.3717	9.3717

(XVI; 22)

Note that because $m\hat{I}_j([Q_k], [\bar{Q}_k])$, *collective association*, is a symmetric quantity (since $Q_k \cup \bar{Q}_k$ is a symmetric relation in itself),

$$\left.\begin{aligned} & m\hat{I}_j([Q_k], [\bar{Q}_k]) = m\hat{I}_j([\bar{Q}_k], [Q_k]) = \\ & = m\widehat{H}_j([Q_k]) + m\widehat{H}_j([\bar{Q}_k]) - m\widehat{H}_j^{(\varphi)}, \end{aligned}\right\} \qquad \text{(XVI; 23)}$$

(XVI; 22) is partly informative for the other three pairs as well. [This is why the pair (u, z) is chosen here, instead of (v, w), as in (XVI; 20).]

XVI.6.4 Note, further, that the sum of the collective association estimates is equal to the value of manifest local distinctiveness, i.e.

$$\left.\begin{array}{l} sm\hat{I}_j([t, i] \mid q^{(2)}) + m\hat{I}_j\{\Sigma_\lambda^{(2)}\} = 3 \cdot m\hat{I}_j(\lambda) \\ \quad 18.2904 \quad + 9.8247 \quad = 28.1151 \end{array}\right\} \qquad \text{(XVI; 24)}$$

which has the implication that

$$\left.\begin{array}{l} 2 \cdot m\hat{I}_j(\lambda) - sm\hat{I}_j([t, i] \mid q^{(2)}) = m\hat{I}_j([\lambda_{(2)}]) \\ \quad 18.7434 - \quad 18.2904 \quad = 0.453 \end{array}\right\} \qquad \text{(XVI; 25)}$$

as before.

XVII. Some interim remarks

XVII.1 These models can be useful in several ways. For instance, we might be interested in ranking coalitions, according to their *complete association.* (During a process of some kind, we should face, of course, the difficult problems of "trans-ranking".) This argument may lead us to redefine the old but still unsolved problems of "*dominance structures*" *of communities* (see MAY, 1981). It is perhaps not too far-fetched to believe that the most dominant, the second most dominant etc. elements or coalitions are not defined *only* by their abundance relations but, *prima facie*, by their "*associational powers*" by which they truly "dominate" over other elements or coalitions.

XVII.2 Let us make a distinction between (a) *autophenetic* and (b) *synphenetic patterns*. Concept (a) is related to one single element of a set, like the Bangorian "species pattern" representations. Concept (b) is related to all (or, at least, conventionally all) elements of a set, like a number of diversity pattern representations (see e.g. GREIG-SMITH 1964; PATIL and ROSENZWEIG 1979).

XVII.2.1 The efficient translation of (a) to (b), and *vice versa*, is still very limited. The aim of the present paper was to describe some tools to make this kind of communication, or, translation more efficient. The *syncretic* models of [B] and the *diacretic* models of [C] refer, of course, to *some synphenetic* and *autophenetic patterns*, respectively.

XVII.2.2 We should be aware of the fact that some seemingly simple problems are "classical" but somehow neglected. KYLIN, for instance, had made clear and operational distinction between the "minimiareals" of a species and a community (see GOODALL 1952). Instead of forgetting important results, it is much better to try to follow the tracks of our ancestors.

XVII.3 If we generalize the reasoning of KYLIN and others, then all concepts of X. (e.g. characteristic points, intervals, scaling, screening etc.) have to be interpreted for elements and coalitions as well. We might guess that there exist some empirically "good" (although probably not simple) relations between certain autophenetic and synphenetic parameters. For instance, in the the majority of cases there is an accumulation of particular maxima of complete dissociation for most populations around A_{diss}, the characteristic point, where total dissociatum reaches its maximal value. The detection of such "multi-pattern" relations can be made only by having a reasonably rich repertoir of models. The next parts of this series will keep this aim in view.

REFERENCES

Aczél, J., Daróczy, Z. (1975): *The Measures of Information Characterizations*. Academic Press, London.
Erdős, P., Rényi, A. (1960): On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5:** 17–61.
Feoli, E., Lagonegro, M., Orlóci, L. (1984): *Information Analysis in Vegetation Research.* [Preprint of a book under preparation; Trieste–London (Ont.).]
Gardner, W. R. (1962): *Uncertainty and Structure as Psychological Concepts.* Wiley, New York.
Goodall, D. W. (1952): Quantitative aspects of plant distribution. *Biol. Rev.* **27:** 194–245.
Grassle, J. F., Patil, G. P., Smith, W. K., Taillie, C. (eds) (1979): *Ecological Diversity in Theory and Practice.* Internat. Co-Operative Publ. House, Fairland, Maryland.
Greig-Smith, P. (1964): *Quantitative Plant Ecology.* 2nd ed. Butterworth, London.
Juhász-Nagy, P. (1958): A Beregi-Sík rét-legelőtársulásai (Hung., with French resumé). *Acta Univ. Debrecina* **4:** 195–228.
Juhász-Nagy, P. (1964): Some theoretical problems of cenological fidelity. *Acta Biol. Debrecina* **3:** 33–43.
Juhász-Nagy, P., Dévai, I., Horváth, K. (1973): Some problems of model-building in synbiology. Part II. Associatum processes in a simple situation. *Annal. Univ. Sci. Budapest, Sect. Biol.* **15:** 39–51.
Juhász-Nagy, P. (1976): Spatial dependence of plant populations. Part 1. Equivalence analysis (An outline of a new model). *Acta Bot. Acad. Sci. Hung.* **22:** 61–78.

Juhász-Nagy, P., Podani, J. (1983): Information theory methods for the study of spatial processes and succession. *Vegetatio* **51:** 129–140.

Khinchin, A. I. (1957): *Mathematical Foundations of Information Theory*. Dover, New York.

Kullback, S. (1959): *Information Theory and Statistics*. Wiley, New York.

Luce, D. (1960): The theory of selective information and some of its behavioural applications. In: Luce, D. (ed.): *Developments in Mathematical Psychology*. The Free Press, Glencoe, 5–119.

May, R. M. (ed.) (1981): *Theoretical Ecology*. (Principles and Applications.) Sinauer Ass., Sunderland, Mass.

Orlóci, L. (1978a): *Multivariate Analysis in Vegetation Research*. 2nd ed. Junk, The Hague.

Orlóci, L. (1978b): Ranking species based on the components of equivocation information. *Vegetatio* **37:** 123–125.

Patil, G. P., Rosenzweig, M. (eds) (1979): *Contemporary Quantitative Ecology and Related Ecometrics*. Internat. Co-Operative Publ. House, Fairland, Maryland.

Pielou, E. C. (1974): *Population and Community Ecology: Principles and Methods*. Gordon and Breach, New York.

Pielou, E. C. (1975): *Ecological Diversity*. Wiley, New York.

Pielou, E. C. (1977): *Mathematical Ecology*. Wiley, New York.

Quastler, H. (1953): The measure of specificity. In: Quastler, H. (ed.): *Essays on the Use of Information Theory in Biology*. Univ. of Ill. Press, Urbana, 41–71.

Rényi, A. (1962): *Wahrscheinlichkeitsrechnung, mit einem Anhang über Informationstheorie*. VEB Deutscher Verlag der Wissenschaften, Berlin.

Reza, F. M. (1961): *An Introduction to Information Theory*. McGraw-Hill, New York.