

## Die Konstruktion von elektronischen Superiorstrukturen\*

**Hans-Georg Bartel**

Sektion Chemie, Humboldt-Universität zu Berlin, DDR-1080 Berlin,  
Deutsche Demokratische Republik

### Construction of Electronic Superior Structures

**Summary.** Using the matrix  $C$  of a quantumchemical problem  $\Phi = C\sigma$ , hierarchic classification methods and the connex-tableau-algorithm [20, 21] based on graph theory, superior structure of an electronic system were constructed. These structures are fragments with maximal electronic similarity in relation to the information given by  $C$ . It is possible to study the relative stability and the inner structure of the superior fragments. The method is demonstrated for three hydrocarbons with symmetry  $C_{1h}$  using results of HMO calculations [24].

**Keywords.** Substructures of an electronic system; Classification; Graph theory.

### Einleitung

Es ist das Anliegen dieses Aufsatzes, Professor Dr. Oskar E. Polansky, dessen Ableben am Anfang dieses Jahres (1989) zu beklagen war, ehrend zu gedenken. Will ein solcher Beitrag wenigstens in einem bescheidenen Maß versuchen, die Leistung dieses Förderers der theoretischen Chemie in ihrem gesamten Umfang erahnen zu lassen, so müssen als Minimum drei Aspekte berücksichtigt werden: die halbempirische Quantenchemie, die Schaffung und Verwendung von Ähnlichkeitsmaßen und die Graphentheorie. Die ersten beiden sind z. B. mit den Begriffen Parsorbital-Methode und Charakterordnung (s. [1–3] u. v. a.) verbunden, der letztere mit der Begründung der Zeitschrift „match“ und zahlreichen Arbeiten, die die Verbindung von Chemie und Graphentheorie zum Inhalt haben und von denen stellvertretend die beiden Beiträge [4, 5] im ersten Heft dieses Organs der mathematischen Chemie genannt seien.

Dem genannten Ziel entsprechend wird im vorliegenden Beitrag das Auffinden von elektronischen Superiorstrukturen beschrieben. Von der quantenchemischen Seite schließt er damit an Methoden an, die darauf abzielen, ein größeres elektronisches System in Bruchstücke zu zerlegen, um damit eine Reduzierung des Problems zu erreichen und um zu Aussagen über innere strukturelle Beziehungen zu gelangen. Neben der Parsorbital(PO)-Methode [1] und deren Ausbau als LCPO-MO-Verfahren [6] sind es die auf FSGO-Basis geschaffenen Ansätze von Christoffersen [7–10], die MIM-(„molecules in molecule“)- [11–15], die auf Ausnutzung von

\* In memoriam Professor Dr. Oskar E. Polansky

Symmetrien beruhende [16, 17] und die von Dewar [18, 19] entwickelte LCMO-Methode.

Das Auffinden der Bruchstücke erfolgt unter Benutzung mathematischer Methoden der Klassifikation, deren Grundidee es ist, daß solche Objekte zu einer Klasse vereinigt werden, die untereinander möglichst große Ähnlichkeit bzw. geringe Unähnlichkeit (Distanz, Abstand) aufweisen. Auf diese Weise werden auch hier Ähnlichkeits- bzw. Distanzmaße mit dem quantenchemischen Problem verknüpft. Derartige Klassifikationsverfahren enthalten stets eine gewisse Mehrdeutigkeit, die durch Anwendung des im Zusammenhang mit der Archäometrie und Chemometrie entwickelten Konnex-Tafel-Algorithmus [20, 21] eingeschränkt wird. Der letztere nutzt graphentheoretische Vorstellungen.

## Ergebnisse und Diskussion

### *Grundgedanke und Definition*

Es sei ein elektronisches System  $\Sigma = (\Phi, \Gamma)$  gegeben, wobei es für die Objektmenge  $\Phi$  eine Verknüpfung mit einer zweiten Menge  $\sigma$  gibt, die ebenfalls bekannt ist und sich in vektorieller bzw. Matrixschreibweise als

$$\Phi = \mathbf{C} \sigma \quad (1)$$

darstellen läßt. Die quantenchemische Beziehung (1) kann als LCAO-MO-, CI- oder ähnliche Entwicklung verstanden werden oder das jeweils Inverse davon zum Ausdruck bringen. Die Koeffizientenmatrix  $\mathbf{C}$  enthält in jedem Fall Information über die elektronischen Verhältnisse in  $\Sigma$ . Ebenso soll die hier nicht weiter zu spezifizierende Menge von Operationen  $\Gamma$  sich auf die elektronischen Beziehungen im System beziehen.

Es ist nun vorstellbar, daß das System  $\Sigma$  durch eine a priori vollzogene Klassenzerlegung aus einer umfassenderen Menge, etwa der Umgebung, abgesondert wurde. Als Beispiele ließen sich die Aussonderung eines Moleküls aus einem Gas, einer Elementarzelle aus einem Kristall u. v. a. denken. Hier kann die Frage angeschlossen werden, ob sich dieser Prozeß nicht in der Hinsicht fortsetzen läßt, daß  $\Sigma$  unter Ausnutzung der „elektronischen“ Information, die mit Gleichung (1) und insbesondere mit  $\mathbf{C}$  verbunden ist, weiterhin sinnvoll in Klassen zerlegt werden kann. Die Elemente dieser Zerlegungsmenge  $\theta = \{\Phi_1, \dots, \Phi_n\}$  stellen zusammen mit entsprechenden Operationen  $\Gamma_1, \dots, \Gamma_n$  Unterstrukturen von  $\Sigma$  dar. Da die Operationen im weiteren nicht interessieren, werden diese Unterstrukturen durch die Elemente von  $\theta$  gekennzeichnet.

Die durch Klassenzerlegung unter Zuhilfenahme von (1) gefundenen Unterstrukturen  $\Phi_i \in \theta$  nennen wir elektronische Superiorstrukturen des Systems  $\Sigma$ . Der Name erklärt sich aus der Bedeutung des lateinischen Komperativs „superior“: „obere, höhere, überlegen, ...“. Hierbei liegt folgende Blickrichtung zugrunde: Eine triviale Zerlegung von  $\Phi$  ist die in die einzelnen Elemente dieser Menge, eine „höhere, überlegene“ wird dagegen diejenige in die Teilmengen  $\Phi_i \in \theta$  von  $\Phi$  sein.  $\theta$  selbst und eine Menge von Operationen, die sich auf dessen Elemente beziehen, bilden dann eine gewonnene Superiorstruktur.

Auch mit einem Anliegen der Informatik ist diese Benennung zu vereinbaren: Aus einer auf irgendeine Art erhaltenen Grundinformation ist mit den Methoden

der Informatik eine „höhere“, in gewissem Sinn „veredelte“ Information zu gewinnen. Die ursprüngliche Information ist jetzt mit der aus quantenchemischen Berechnungen stammenden Matrix  $\mathbf{C}$  verbunden. Die Anwendung von Klassifikations- bzw. Cluster-Analyse-Verfahren, welche zu den Methoden der explorativen Datenanalyse zählen, liefert im Sinne einer Mustererkennung die Superiorstrukturen, aus deren Kenntnis sich eine weiterführende Information über das untersuchte System  $\Sigma$  ableiten läßt. Unter Umständen kann auf diese Weise auch eine Datenreduktion, die für große quantenchemische Systeme von praktischem Wert ist, durchgeführt werden. Das letztere soll aber in dieser Arbeit nicht behandelt werden.

### *Die mathematischen Methoden*

Die mathematischen Klassifikationsmethoden lassen sich in hierarchische und partitionierende unterteilen [22]. In dieser Arbeit beziehen wir uns nur auf die ersteren, wobei sieben agglomerative Verfahren nach [22] zur Anwendung kommen und als Distanzmaß in allen Fällen der Euklidische Abstand benutzt wird. Die oben erwähnte Mehrdeutigkeit besteht darin, daß die verschiedenen Algorithmen im allgemeinen unterschiedliche Ergebnisse liefern, „ohne daß jedoch für eine spezielle Klassenanzahl eine Partition besser als eine andere bezeichnet werden kann, da nur die Hierarchiebildung optimiert wird“ [23] (S. 15). Um die Mehrdeutigkeit teilweise aufzuheben, wurde schon zuvor [20, 21] – allerdings auch im Zusammenhang mit partitionierenden Verfahren – die Konnex-Tafel-Methode entwickelt, die gestattet, aus einer gegebenen Anzahl von Klassenerlegungen (Partitionen) eine verallgemeinerte Partition zu bilden. Für eine vorgegebene, feste Klassenanzahl, in welche die ursprüngliche Zerlegung erfolgte, liefert das Verfahren eine eindeutige verallgemeinerte Partition.

Der Algorithmus der Konnex-Tafel-Methode wurde in [20] bereits beschrieben. Er wird hier in strengerer mathematischer Form wiederholt, wobei insbesondere eine graphentheoretische Darstellung Anwendung finden soll. Die Anzahl der zu betrachtenden verschiedenen Partitionen sei  $p$ , in jeder sind  $m$  Objekte in  $c$  Klassen zerlegt worden. Die Objekte selbst bezeichnen wir durch die ihnen zugeordnete Nummer  $i$ . Sie lassen sich somit als Elemente einer Menge  $M = \{1, 2, \dots, m\}$  auffassen, die eine endliche Teilmenge der natürlichen Zahlen ist. Im graphentheoretischen Sinne ist  $M$  die Knotenmenge. Eine (ungerichtete) Kante zwischen den Knoten  $i$  und  $j$  wird durch  $[i, j]$  bzw.  $[j, i]$ , eine Schlinge am Knoten  $i$  entsprechend durch  $[i, i]$  symbolisiert.

Um die Existenz einer Kante in dem Graphen  $G$ , mit dessen Hilfe wir die verallgemeinerte Partition GP konstruieren wollen, festzulegen, wird die Konnex-Tafel  $\mathbf{T} = (t_{ij})$  benötigt. Der Wert eines Elementes  $t_{ij}$  bringt zum Ausdruck, in wie vielen Fällen die Objekte  $i$  und  $j$  in den  $p$  zu betrachtenden Partitionen zu einer Klasse gehören. Analog zählt ein Hauptdiagonalwert  $t_{ii}$ , wie oft das Objekt  $i$  als einziges (isoliertes) Element einer Klasse auftritt. Im weiteren wollen wir Schlingen  $[i, i]$  und die Werte  $t_{ii}$  als Spezialfälle von  $[i, j]$  bzw.  $t_{ij}$  mit  $i = j$  ansehen und diese Besonderheit nicht hervorheben. Für die Elemente von  $\mathbf{T}$  gilt somit

$$0 \leq t_{ij} = t_{ji} \leq p. \quad (2)$$

Bezeichnen wir mit  $E$  die Menge der Kanten im Graphen  $G = \{M, E\}$ , so wird festgelegt, daß

$$t_{ij} \neq 0 \Rightarrow [i, j] = [j, i] \in E \quad (3)$$

ist. Andererseits sind die von Null verschiedenen Werte  $t_{ij}$  Bewertungen der Kanten  $[i, j]$ .

Für die Bestimmung der GP werden noch die Zahl  $q = \max(t_{ij}) \leq p$  und die Teilmenge von  $M$

$$H(s) = \{i \in H(s) \mid t_{ij} \geq s\} \text{ mit } i \in M. \quad (4)$$

benötigt. Schließlich sei  $r$  der maximale von denjenigen  $s$ -Werten, für welche  $H(s) = M$  gilt.

Im folgenden wird nun eine Konstruktion beschrieben, nach welcher  $G$  in  $n$  vollständige Teilgraphen  $G_1, \dots, G_n$  zerlegt wird. Die Knoten eines dieser Graphen  $G_i$  sind die Objekte der  $i$ -ten von den  $n$  Klassen der gesuchten GP. Die Bestimmung von GP läßt sich an folgendem Algorithmus verdeutlichen:

(a) Aus  $G$  wird der Graph  $F_q = \{M, \emptyset\}$  durch Löschen aller Kanten gebildet. Ein Laufindex  $l$  wird  $l = q + 1$  gesetzt.

(b)  $l = l - 1$

(c) Der Graph  $F_{l-1}$  wird gebildet, indem zu  $F_l$  alle Kanten  $[i, j]$  mit  $t_{ij} = l$  hinzugefügt werden.

(d) Alle vollständigen Graphen, die sich in  $F_l$  finden lassen, werden markiert.

(e) Für alle gefundenen vollständigen Graphen werden folgende Untersuchungen durchgeführt:

(e1) Gehört ein Knoten mehreren vollständigen Graphen mit gleichgroßer, maximaler Kantenzahl an und ist auch die Summe der Kantenbewertung dieser Graphen gleich, so wird die Markierung vorerst wieder gelöscht.

(e2) Gehört ein Knoten mehreren vollständigen Graphen mit gleichgroßer, maximaler Kantenzahl an, ist aber die Summe der Kantenbewertungen genau eines dieser Graphen größer als die der anderen, so werden die Markierung dieser letzteren wieder gelöscht.

(e3) Inzidiert ein Knoten eines vollständigen Graphen mit einer Kante, die eine größere Bewertung besitzt als diejenigen, die zu dem vollständigen Graphen gehören, so wird dessen Markierung gelöscht.

(f) Ist  $l > r$ , so wird bei (b) fortgesetzt, anderenfalls ist die Festlegung der vollständigen Graphen beendet. Die Anzahl  $n$  der Klassen in der GP entspricht der Anzahl der nichtzusammenhängenden vollständigen Graphen, die eine Markierung aufweisen. Die Knoten eines solchen Graphen sind die Objekte der entsprechenden Klasse der GP.

Zugleich mit dieser Konstruktion der GP läßt sich ein Zusammenhangsdiagramm (ZD) aufstellen, aus welchem sich Aussagen über die innere Struktur und die Stabilität der gefundenen Klassen ablesen lassen. Auf der Senkrechten eines ZD werden die  $t_{ij}$ -Werte aus der Konnex-Tafel aufgetragen, bei denen sich Knoten zu markierten vollständigen Graphen vereinigt haben. Die Knoten werden in der Waagerechten dann aufgetragen, wenn sie im Verlaufe der Konstruktion das erste Mal auftreten. Handelt es sich nicht um einen isolierten Knoten, so werden sie durch einen Strich mit dem vollständigen Graphen verbunden, dessen Knotenmenge sie nun zu einem erweiterten vollständigen Graphen ergänzen. Die Vereinigung von

vollständigen Graphen wird durch Zusammenführen von Linien dargestellt, wobei der Schnittpunkt bei dem  $t_{ij}$ -Wert eingetragen wird, bei dem der erweiterte markierte vollständige Graph gebildet wurde. Jeder Klasse entspricht dann ein Baum im ZD. Ihre Stabilität ist umso höher, je größer die Summe der Vereinigungs- $t_{ij}$ -Werte bezogen auf die Objekt- bzw. Knotenanzahl ist. Das Analoge gilt bezogen auf vollständige Teilgraphen und auf einzelne Knoten im Zusammenhang mit ihrer Zugehörigkeit zur Klasse. Die Homogenität und innere Struktur einer Klasse ist aus der Anzahl der unterschiedlichen Vereinigungs- $t_{ij}$ -Werte ablesbar. Auf die Definition bzw. Berechnung entsprechender Stabilitäts- und Homogenitätsindizes wollen wir hier verzichten.

Es sei betont, daß die Aufstellung der Konnex-Tafel und die angegebene Konstruktion der GP und des ZD nicht auf die hierarchischen Klassifikationsverfahren beschränkt ist, wie mit dem Hinweis auf [20, 21] schon angedeutet wurde.

Dieses für die allgemeine explorative Datenanalyse geschaffene und verwendbare Verfahren kann nun für die Ermittlung der oben definierten elektronischen Superiorstrukturen verwendet werden. Die Objekte sind die Elemente des Vektors  $\Phi$  in Gleichung (1). Die Anzahl der Eigenschaften entspricht der Dimension von  $\sigma$ , die Matrix  $C$  der Datenmatrix.  $c_{ij} \in C$  ist somit der Wert der  $j$ -ten (elektronischen) Eigenschaft des  $i$ -ten Objektes  $\Phi_i \in \Phi$ . Da die Elemente von  $C$  reelle Zahlen vergleichbarer Größenordnung sind, läßt sich zur Berechnung des Abstandes  $d_{ij}$  zwischen den Objekten  $\Phi_i$  und  $\Phi_j$  sinnvoll das Euklidische Maß verwenden:

$$d_{ij} = \left( \sum_h (c_{ih} - c_{jh})^2 \right)^{1/2} \quad (5)$$

Mit Hilfe der Distanzmatrix  $D = (d_{ij})$  lassen sich die Dendrogramme der sieben agglomerativen hierarchischen Cluster-Analyse-Verfahren mit dem Programm HIERCL nach [22] ermitteln und aus diesen bei Vorgabe der gewünschten Klassenanzahl  $c$  die zur Aufstellung der Konnex-Tafel notwendigen Partitionen. Die mit dem beschriebenen Algorithmus erhaltenen Klassen der GP sind die gesuchten Superiorstrukturen, über deren Stabilität, innere Struktur usw. die Bäume des zugehörigen ZD Auskunft geben.

Eine Datenreduktion ist z. B. dadurch möglich, daß die Objekte einer Klasse der GP, d. h. einer Superiorstruktur auf ein typisches Objekt oder auf den Klassenschwerpunkt eingeschränkt wird. Der letztere kann eventuell unter Verwendung von Gewichten berechnet werden, die aus Angaben im ZD ermittelt werden. Als

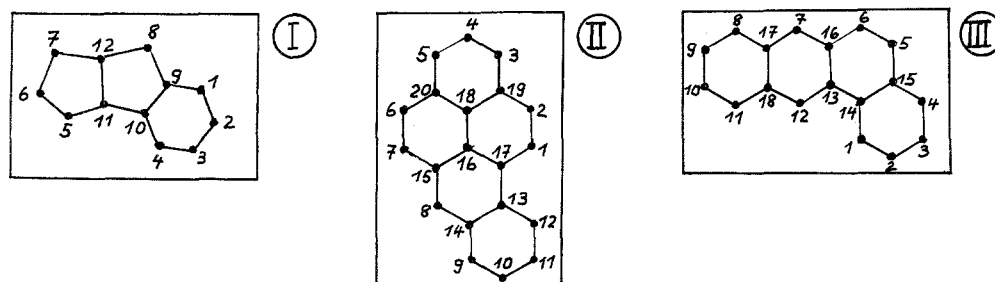


Abb. 1. Die Moleküle Benzpentalen (I), 1,2-Benzpyren (II) und 1,2-Benzanthracen (III) mit Nummerierung der C-Atome bzw. der  $\pi$ -Atomorbitale

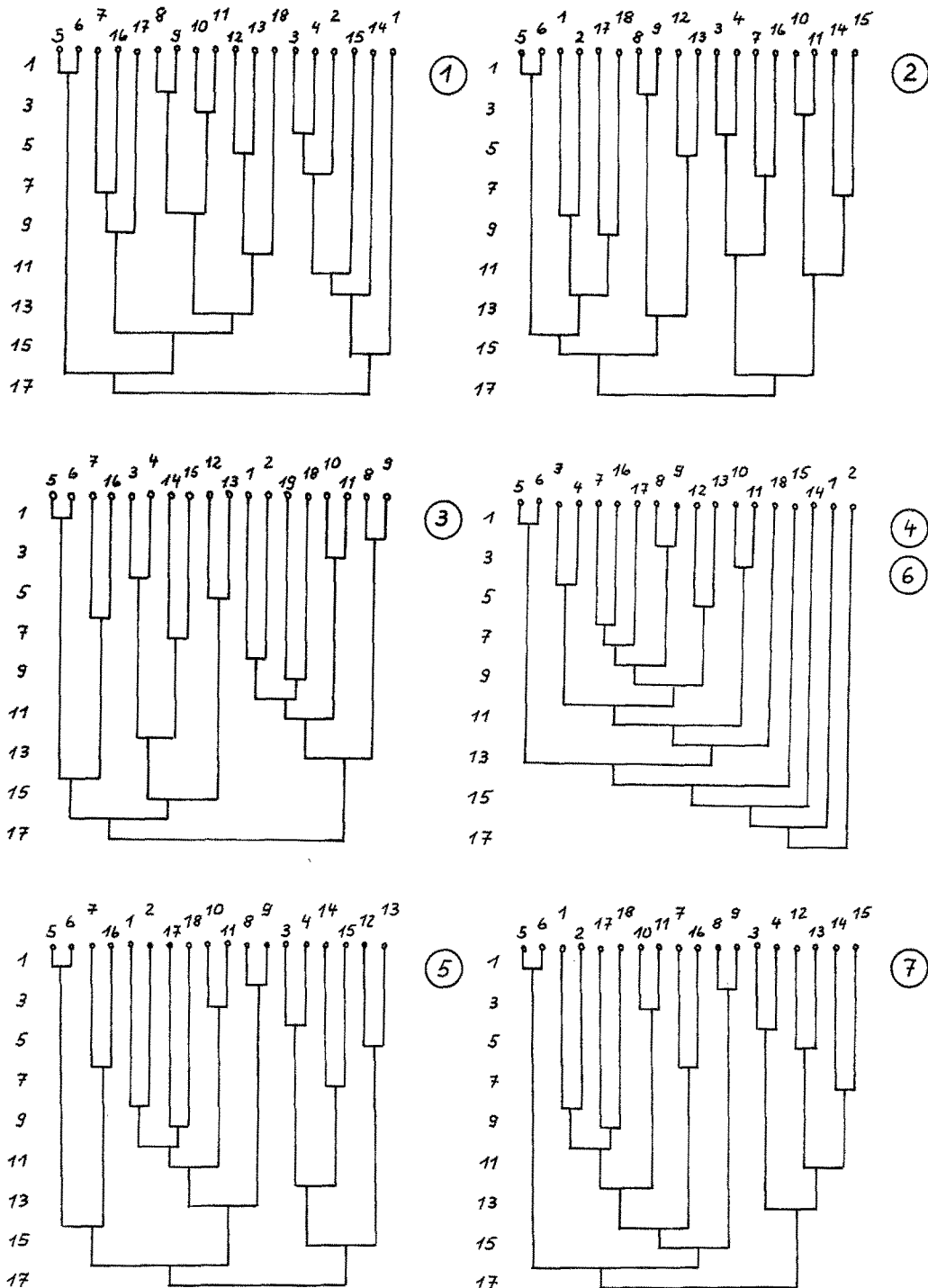


Abb. 2. Die Dendrogramme für das 1,2-Benzanthracen

typisches Objekt kann z. B. dasjenige angesehen werden, das den geringsten Abstand zum Klassenschwerpunkt besitzt. Auch auf die Probleme der Datenreduktion bzw. auf ihre Ausnutzung für die quantenchemische Untersuchung größerer Systeme wird in dieser Arbeit nicht weiter eingegangen.

*Ein Beispiel*

Als Beispiel wollen wir die drei molekularen Systeme Benzpentalen (I), 1,2-Benzpyren (II) und 1,2-Benzanthracen (III) (s. Abb. 1) betrachten. Die Gleichung (1) wird als inverse LCAO-MO-Entwicklung aufgefaßt, d. h.,  $\Phi$  stellt den AO- und  $\sigma$  den MO-Vektor dar. In den letzteren sollen nur besetzte Zustände eingehen. Für die Demonstration genügt das HMO-Niveau. Es hat außerdem den Vorteil der Anschaulichkeit, da jedem Kohlenstoffatom genau ein  $\pi$ -AO entspricht, was neben der Möglichkeit einer übersichtlichen graphischen Darstellung einen (für einen späteren Zeitpunkt vorgesehenen) Anschluß an eine graphentheoretische Behandlung erlaubt. Die drei genannten Kohlenwasserstoffe wurden ausgewählt, da sie  $C_{1h}$ -Symmetrie aufweisen, so daß es keine Klassifikation vermittelt der Symmetrie gibt. Die C-Matrizen wurden aus [24] entnommen. Die Numerierung der AOs bzw. der C-Atome kann der Abb. 1 ersehen werden. Benzpentalen stellt also ein System mit  $m = 12$ , 1,2-Benzpyren mit  $m = 20$  und 1,2-Benzanthracen mit  $m = 18$  Objekten dar. Die Anzahl der Eigenschaften beträgt jeweils  $m/2$ .

Von den ermittelten Dendrogrammen sind als Beispiel die sechs verschiedenen,

**Tabelle 1.** Zugehörigkeit der Orbitale des Benzpentalens zu den Klassen \*, + und ■ (Orbitalnr. s. Abb. 1)

| Verf.-Nr. | Orbital/Objektnr. |   |   |   |   |   |   |   |   |    |    |    |
|-----------|-------------------|---|---|---|---|---|---|---|---|----|----|----|
|           | 1                 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1         | *                 | * | + | * | * | * | ■ | * | * | *  | *  | *  |
| 2         | *                 | + | * | ■ | * | * | * | + | ■ | *  | *  | +  |
| 3=5       | +                 | * | + | + | * | ■ | * | * | * | *  | ■  | *  |
| 4=6       | *                 | * | * | * | * | * | + | * | * | *  | ■  | *  |
| 7         | *                 | + | * | * | * | * | + | ■ | * | *  | +  | ■  |

**Tabelle 2.** Zugehörigkeit der Orbitale des 1,2-Benzpyrens zu den Klassen \*, + und ■ (Orbitalnr. s. Abb. 1)

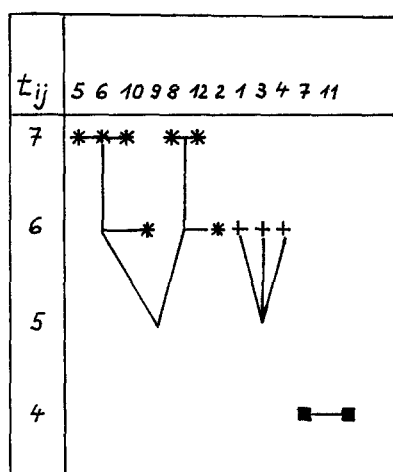
| Verf.-Nr. | Orbital/Objektnr. |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |
|-----------|-------------------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|
|           | 1                 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 1         | +                 | + | + | + | + | * | * | ■ | ■ | ■  | ■  | ■  | ■  | ■  | ■  | *  | *  | *  | *  | *  |
| 2         | ■                 | ■ | + | + | ■ | + | + | ■ | + | ■  | ■  | *  | *  | ■  | +  | +  | *  | *  | ■  |    |
| 3=5=7     | ■                 | ■ | * | * | * | * | * | + | + | +  | +  | +  | +  | +  | +  | ■  | ■  | ■  | ■  | *  |
| 4         | +                 | + | + | + | + | + | + | + | + | +  | +  | +  | *  | ■  | +  | +  | +  | +  | +  | +  |
| 6         | +                 | + | + | + | + | + | + | + | + | +  | +  | +  | +  | +  | +  | *  | +  | ■  | +  | +  |

**Tabelle 3.** Zugehörigkeit der Orbitale des 1,2-Benzanthracens zu den Klassen \*, + und ■ (Orbitalnr. s. Abb. 1)

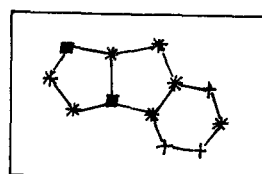
| Verf.-Nr. | Orbital/Objektnr. |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |
|-----------|-------------------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|
|           | 1                 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 1         | ■                 | ■ | ■ | ■ | + | + | * | * | * | *  | *  | *  | ■  | ■  | *  | *  | *  | *  |
| 2         | +                 | + | * | * | + | + | * | + | + | ■  | ■  | +  | +  | ■  | ■  | *  | +  | +  |
| 3=5       | ■                 | ■ | * | * | + | + | ■ | ■ | ■ | ■  | *  | *  | *  | *  | +  | ■  | ■  | ■  |
| 4=6       | *                 | ■ | + | + | + | + | + | + | + | +  | +  | +  | +  | +  | +  | +  | +  | +  |
| 7         | *                 | * | ■ | ■ | + | + | * | * | * | *  | *  | ■  | ■  | ■  | ■  | *  | *  | *  |

**Tabelle 4.** Konnex-Tafel für das Benzpentalen (die durch Fettdruck hervorgehobenen  $t_{ij}$ -Werte kennzeichnen die Kanten  $[i, j]$ , die für die Konstruktion der GP und des ZD benutzt werden)

| $ij$ | 1 | 2 | 3        | 4        | 5        | 6        | 7 | 8        | 9        | 10       | 11       | 12       |
|------|---|---|----------|----------|----------|----------|---|----------|----------|----------|----------|----------|
| 1    | 0 | 3 | <b>6</b> | <b>6</b> | <b>5</b> | <b>5</b> | 1 | 3        | <b>4</b> | <b>5</b> | 2        | 3        |
| 2    |   | 0 | 2        | 3        | <b>5</b> | <b>5</b> | 1 | <b>6</b> | <b>5</b> | <b>5</b> | 2        | <b>6</b> |
| 3    |   |   | 1        | <b>5</b> | <b>4</b> | <b>4</b> | 1 | 2        | 3        | <b>4</b> | 1        | 2        |
| 4    |   |   |          | 0        | <b>4</b> | <b>5</b> | 0 | 3        | <b>5</b> | <b>4</b> | 1        | 3        |
| 5    |   |   |          |          | 0        | <b>7</b> | 1 | <b>5</b> | <b>6</b> | <b>7</b> | 2        | <b>5</b> |
| 6    |   |   |          |          |          | 0        | 1 | <b>5</b> | <b>6</b> | <b>7</b> | 2        | <b>5</b> |
| 7    |   |   |          |          |          |          | 3 | 0        | 0        | 1        | <b>4</b> | 0        |
| 8    |   |   |          |          |          |          |   | 0        | <b>5</b> | <b>5</b> | 1        | <b>7</b> |
| 9    |   |   |          |          |          |          |   |          | 0        | <b>6</b> | 1        | <b>5</b> |
| 10   |   |   |          |          |          |          |   |          |          | 0        | 2        | <b>5</b> |
| 11   |   |   |          |          |          |          |   |          |          |          | 2        | 1        |
| 12   |   |   |          |          |          |          |   |          |          |          |          | 0        |



Ⓘ



**Abb. 3.** Zusammenhangsdiagramm und verallgemeinerte Partition für das Benzpentalen

die zum 1,2-Benzanthracen gehören, in Abb. 2 wiedergegeben. Die jeweilige Partition bei einer Zerlegung in  $c = 3$  Klassen ist aus einer der Tabellen 1–3 zu entnehmen. Die Klassen werden hier durch die Symbole \*, + und ■ unterschieden.

Tabelle 4 zeigt die Konnex-Tafel des Benzpentalens (hier ist  $q = 7$  und  $r = 4$ ). Die aus ihr ermittelte GP bzw. Superiorstruktur mit dem zugehörigen ZD ist aus Abb. 3 zu entnehmen. Sie enthält wiederum drei Klassen (\*, +, ■). Im C-Atomgerüst des Systems hängen nicht alle Objekte einer Klasse topologisch zusammen, so daß keiner der drei Ringe vollständig zu einer Superiorstruktur gehört. Bis auf die beiden ohnehin schwach zusammenhängenden Objekte der ■-Struktur besetzen diejenigen der beiden anderen Klassen in den Ringen im wesentlichen immer jeweils die Begrenzung einer Kante und die diagonal gegenüberliegende Position. Die +-Struktur des Sechsrings ist homogen und einigermaßen stabil.



II

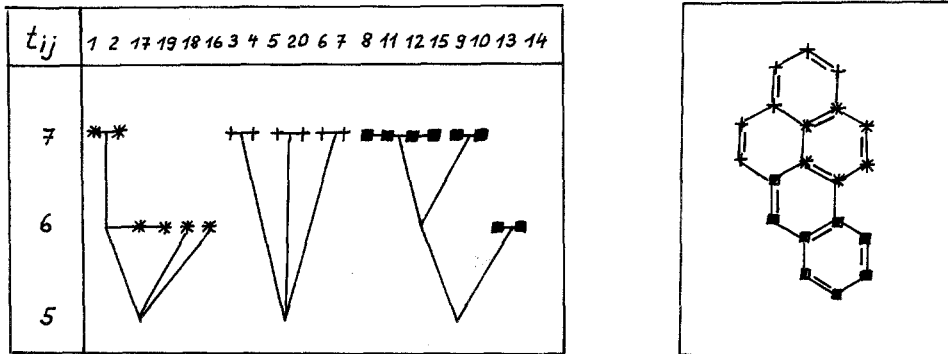


Abb. 4. Zusammenhangsdiagramm und verallgemeinerte Partition für das 1,2-Benzopyren

III

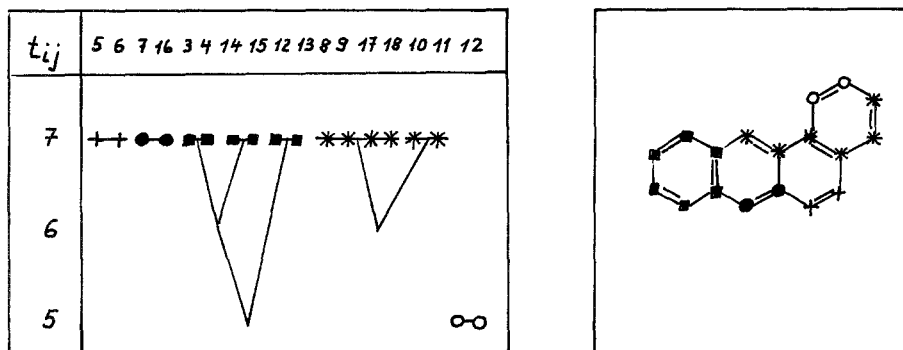


Abb. 5. Zusammenhangsdiagramm und verallgemeinerte Partition für das 1,2-Benzanthracen

Die entsprechenden Ergebnisse für das 1,2-Benzopyren sind in Abb. 4 dargestellt. Hier bilden alle drei Klassen (\*, +, ■) zusammenhängende Superiorstrukturen mit konjugierten  $\pi$ -Bindungen: die +-Kette [Hexatri(1,3,5)en], den \*-Benzenring und das ■-Styren. Interessant ist die Inhomogenität des \*-Ringes im Gegensatz zur insgesamt schwachen, in den drei Ethen-Unterstrukturen aber sehr stabilen +-Kette.

Auch die fünf Klassen (\*, +, ■, ○, ●) im Falle des 1,2-Benzanthracen (s. Abb. 5) bilden zusammenhängende Strukturen mit  $\pi$ -Konjugation: drei isolierte Ethenbruchstücke (+, ○, ●), die ■-Hexatri(1,3,5)en-Kette und den \*-Benzenring. Der letztere ist wiederum nicht homogen und zerfällt in einen Ethen- und einen 1,3-Butadien-Anteil, die jeweils sehr stabil sind. Die drei sehr stabilen Ethen-Unterstrukturen der ■-Kette sind im Gegensatz zum Analogon im 1,2-Benzopyren nicht so homogen vereinigt. Die drei isolierten Ethen-Superiorstrukturen schwanken zwischen sehr stabil (+, ●) und sehr schwach (○).

Die Ergebnisse der drei Beispiele zeigen, daß Untersuchungen quantenchemischer Berechnungen mit Hilfe von Methoden der explorativen Datenanalyse, ins-

besondere der Konstruktion von Superiorstrukturen unter Verwendung der Konnex-Tafel zu interessanten, chemisch verwert- und korrelierbaren Aussagen führen. Damit erscheint auch eine Zusammenführung der Quantenchemie und der Chemischen Informatik sinnvoll.

### *Ausblick*

Natürlich ist das hier Dargestellte nur als ein Anfang zu werten.

So wurde begonnen, wie in [20, 21] ebenfalls partitionierende Verfahren der Cluster-Analyse zur Anwendung bringen. Es wurde oben schon darauf hingewiesen, daß bei Benutzung des HMO-Ansatzes leicht ein Anschluß der hier erzielten Ergebnisse an solche, die mit rein graphentheoretischen Überlegungen gewonnen wurden, erreicht werden kann. Auch hierzu sind im Zusammenhang mit dem von John und Sachs [25] entwickelten Algorithmus erste Untersuchungen begonnen worden.

Zu den weiterhin absehbaren Problemen, deren Bearbeitung sinnvoll sein könnte, gehören die Anwendung von faktoriellen Methoden (z. B. Hauptkomponentenanalyse). Sie werden sich außer für strukturelle und klassifikatorische Untersuchungen besonders für die Datenreduktion eignen. Überhaupt sollte die letztere, oben auch schon angesprochene Aufgabe eine Bearbeitung erfahren. Dasselbe gilt für die Erarbeitung von Korrelationen der z. B. aus den ZD gewonnenen Angaben mit der chemischen Reaktivität und mit anderen Struktur-Eigenschaft-Zusammenhängen. Es werden sich Fragen ergeben, wenn ein anderes als das HMO-Modell als quantenchemische Ausgangsbasis benutzt wird.

Diese Aufzählung deutet an, daß die Zahl der Möglichkeiten zur Bearbeitung einer eventuell lohnenden Aufgabe relativ groß ist. Es wäre daher wünschenswert, wenn sich der eine oder andere interessierte Forscher entschließen würde, zu ihrer Lösung einen Beitrag zu leisten. So kann denn mit G. W. Rabener (1714 – 1771) gesagt werden: „Für mich allein ist dieses Werk viel zu groß und wichtig. Vielleicht bin ich zu offenherzig, daß ich dieses Bekenntnis von mir selbst tue.“ (zit. nach [26])

### **Literatur**

- [1] Polansky O. E., Derflinger G. (1967) Internat. J. Quantum Chem. **1**: 379
- [2] Sofer H., Polansky O. E., Derflinger G. (1968) Monatsh. Chem. **99**: 1879
- [3] Sofer H., Derflinger G., Polansky O. E. (1968) Monatsh. Chem. **99**: 1895
- [4] Polansky O. E. (1975) match **1**: 11
- [5] Polansky O. E. (1975) match **1**: 183
- [6] Mientus W., Haberditzl W., Bartel H.-G. (1979) Monatsh. Chem. **110**: 351
- [7] Christoffersen R. E., Maggiora G. M. (1969) Chem. Phys. Letters **3**: 419
- [8] Christoffersen R. E., Genson D. W., Maggiora G. M. (1971) J. Chem. Phys. **54**: 239
- [9] Christoffersen R. E. (1971) J. Amer. Chem. Soc. **93**: 4104
- [10] Christoffersen R. E., Spangler D., Hall G. G., Maggiora G. M. (1973) J. Amer. Chem. Soc. **95**: 8526
- [11] Longuet-Higgins H. C., Murrell J. N. (1955) Proc. Phys. Soc. **A 68**: 961
- [12] Murrell J. N. (1955) Proc. Phys. Soc. **A 68**: 967
- [13] Suzuki H. (1967) Electronic Absorption Spectra and Geometry of Organic Molecules. Academic Press, New York London

- [14] Giovanelli K. H., Hohlneicher G., Straub P. A. (1971) Ber. Bunsenges. Phys. Chem. **75**: 857
- [15] Giovanelli K. H., Dehler J., Hohlneicher G. (1971) Ber. Bunsenges. Phys. Chem. **75**: 864
- [16] Schuster P. (1965) Theoret. Chim. Acta **3**: 278
- [17] Derflinger G., Sofer H. (1968) Monatsh. Chem. **99**: 1866
- [18] Dewar M. J. S. (1969) The Molecular Orbital Theory of Organic Chemistry. McGraw-Hill, New York
- [19] Dewar M. J. S., Proc. Cambridge Philos. Soc. **45**: 638 (1949).
- [20] Bartel H.-G. (1990) Meroitica **12**: 33
- [21] Bartel H.-G., Nofz M., Stößer R., Kupfer M. (1990) Z. Phys. Chem. (Leipzig) **271** (im Druck)
- [22] Späth H. (1975) Cluster-Analyse-Algorithmen zur Objektklassifizierung und Datenreduktion. Oldenbourg, München Wien
- [23] Späth H. (1983) Cluster-Transformation- und -Analyse, Theorie – FORTRAN-Programme. Oldenbourg, München Wien
- [24] Coulson C. A., Streitwieser jr. A. (1965) Dictionary of  $\pi$ -Electron Calculations. Pergamon Press, Oxford
- [25] John P., Sachs H. (1985) Calculating the Number of Perfect Matchings and Pauling's Bond Orders in Hexagonal Systems Whose Inner Dual Is a Tree. In: Sachs H. (ed.), Graphs, Hypergraphs and Applications (Proceed. Conf. Graph Theory, Eyba 1984, Teubner-Texte zur Mathematik, Bd. 73). Teubner, Leipzig, pp. 80–91
- [26] Rabener G. W. (1984) Versuch eines deutschen Wörterbuchs. In: Hofmannsthal H. v. (Hrsg.), Deutsches Lesebuch. Reclam, Leipzig, S. 19

*Eingegangen 10. September 1989. Angenommen 29. September 1989*