

KOMBINATORISCHE ABZÄHLUNG TOPOLOGISCHER INDIZES

J. Will¹⁾, B. Adler*, G. Biess**, S. Dobers*

* Schkopau ** Merseburg

1) Korrespondenzadresse: J. Will, Wiesenweg 1, D-16767 Germendorf, Germany
(received: March 1993)

Zusammenfassung

Angenommen, uns interessieren die Molekülgraphen mit bestimmten Eigenschaften. Durch systematische Konstruktion aller entsprechenden Graphen läßt sich die gewünschte Information erhalten. Oft ist dieses Vorgehen jedoch nicht praktikabel. Die Kombinatorische Abzählungstheorie ermöglicht dann neben der Bestimmung der Gesamtzahl von Molekülgraphen vorgegebener Stoffklasse eine Aufteilung dieser Stoffklasse in Klassen von Graphen mit bestimmten Eigenschaften. Durch diese Klassifikation läßt sich eine genauere Übersicht über die betrachtete Stoffklasse erreichen. In der vorliegenden Arbeit wird gezeigt, wie die Bildung neuer Eigenschaftsparameter bei der Kombination von Eigenschaftsparametern im Abzählungsverfahren berücksichtigt werden kann. Damit lassen sich etwa über das Pólya-Verfahren Topologische Indizes, z.B. Konnektivitäts-Indizes und Charakteristisches Polynom, für Klassen nichtzyklischer Verbindungen ermitteln. Dies stellt eine Erweiterung bisher üblicher Abzählungsverfahren dar. Der Artikel gibt eine kurze Einführung in die Anwendung der gewichteten Form des Satzes von Pólya und beschreibt die rekursive Abzählung von Bäumen und Wurzelbäumen. Am Beispiel der Abzählung von Knotenzahl, Knotengrad und Spektrum aller Weglängen der Alkane wird das Rechenverfahren vorgestellt. Mögliche Anwendungen in der Quantitativen Struktur-Eigenschafts-Simulation (QSPR) werden kurz angedeutet.

Stichworte:

Kombinatorische Abzählungstheorie, Isomerenabzählung, Graphentheorie, Molekülgraph, Satz von Pólya, Topologische Indizes, QSPR

1. Einführung

Die topologische Behandlung von Molekülgraphen spielt in der Chemie eine nicht geringe Rolle [1], [2], obwohl die Topologie des Molekülgerüsts ein bis heute nicht exakt faßbarer Begriff ist [3] und rein topologische Methoden unter Kritik stehen [4] - [8]. Die Situation auf dem Gebiet der Entwicklung Topologischer Indizes für Information und Dokumentation ist in gewisser Weise unbefriedigend. Ständig werden neue Indizes vorgeschlagen, deren Werte nach Möglichkeit die verschiedenen chemischen Strukturformeln eindeutig charakterisieren sollen. Ein Beweis für diese Eigenschaft ist aber derzeit nicht verfügbar. Üblicherweise werden die Molekülgraphen

erzeugt, daraus die Werte des betrachteten Topologischen Indexes berechnet und auf Koinzidenzen hin untersucht. Der Grad der auf diese Weise gefundenen Eindeutigkeit des Indexes ist deshalb von Art und Menge der gewählten Testverbindungen abhängig. Unser Ziel war es, die Topologischen Indizes ohne explizite Erzeugung der Molekülgraphen zu generieren. Die Kombinatorische Abzählungstheorie liefert entsprechende Techniken zur Abzählung von Graphen.

In der Chemie dienen Verfahren der Kombinatorischen Abzählungstheorie der Bestimmung der Anzahl von Skelett-, Stellungen-, Konfigurations- und Stereo-Isomeren [9] - [14], Isomeren mit wechselnden Symmetrien [15], Graphen von Isomerisierungsreaktionen [16], Reaktionsgraphen [17], und finden außerdem in Spektroskopie, Quantenchemie und Statistischer Thermodynamik Verwendung [18]. Dabei läßt sich nicht nur die Gesamtzahl von Objekten ermitteln. Es kann eine abzählende Funktion, die Erzeugende Funktion, gewonnen werden, die die Menge der abzuzählenden Objekte strukturiert, nach vorzugebenden Eigenschaften klassifiziert. Für eine gegebene Stoffklasse haben wir also die Erzeugende Funktion der Werte des gewählten Topologischen Indexes zu ermitteln.

Abzählungsverfahren für Graphen arbeiten meist nach der folgenden Methode [19]. Jeder Graph (Konfiguration) der abzuzählenden Menge wird durch Verteilen von Subgraphen (Figuren) auf die Objekte einer gegebenen Menge mit vorgegebener Symmetrie in dieser Menge erzeugt. Die Charakteristika der Figuren lassen sich gemeinsam mit diesen kombinieren und ergeben die Charakteristika der Konfigurationen. Die Konfigurationen können also bezüglich gewisser Eigenschaften abgezählt werden. Wir müssen uns bei diesem Ansatz allerdings auf Stoffklassen beschränken, deren Molekülgraphen sich über eine Konstruktionsvorschrift aus Subgraphen zusammensetzen lassen. Graphen können nach recht einfachen Parametern, wie z.B. Knoten- und Kantenzahl, abgezählt werden [20]. Diese lassen sich bereits als triviale Topologische Indizes auffassen. Eine Verallgemeinerung der Abzählungsverfahren für diese einfachen Parameter führt auf die Abzählung komplizierterer Charakteristika, der Topologischen Indizes. Im folgenden wird die übliche Darstellung des Molekülgerüsts durch ungerichtete, zusammenhängende Graphen benutzt. Die Alkane werden durch Bäume mit einem Knotengrad kleiner gleich 4, entsprechend der möglichen Bindigkeit des Kohlenstoff-Atoms gegenüber Nicht-Wasserstoff-Atomen, beschrieben. Den Alkylradikale entsprechen dann die Wurzelbäume (knotenmarkierte Bäume) mit einem Wurzelgrad kleiner gleich 3. Aus ihnen werden die Alkane zusammengesetzt. Das Kombinieren von Wurzelbäumen an einem markierten Knoten, der Wurzel, bzw. einer markierten Kante, beide hier kurz Basis genannt, ergibt knoten- bzw. kantenmarkierte Bäume (ABB.1). Durch

Kombinieren aller bereits erzeugten Wurzelbäume unter Berücksichtigung der Symmetrie der hinzugefügten Basis lassen sich alle möglichen Wurzelbäume bzw. kantenmarkierten Bäume erzeugen.

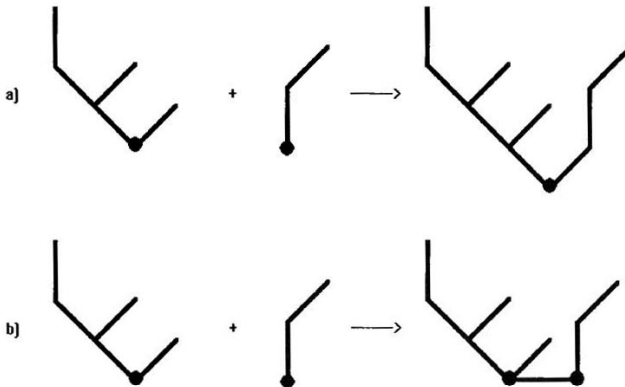


ABB.1 Konstruktion eines Wurzelbaumes mit dem Wurzelgrad $g=2$ (a) und eines kantenmarkierten Baumes (b) aus zwei Wurzelbäumen.

Dieser Schritt kann bequem durch das Pólya-Verfahren [9] beschrieben werden. Die Symmetrie an der Basis wird dabei durch die Permutationsgruppe G der Menge aller möglichen Permutationen an der Basis beschrieben. Da alle Substitutionen an den Positionen der Basis unabhängig voneinander sind, wird die Symmetrische Gruppe S_n , die Gruppe aller Permutationen der n Basis-Positionen, benutzt. Von den einzelnen Permutationen wird die Anzahl der Zyklen entsprechender Länge benötigt. Diese Informationen werden meist in einem Polynom dargestellt, dem Zyklenzeiger $Z(s_1, s_2, \dots, s_n)$ der Permutationsgruppe. Der Index der Variablen s bezeichnet die Zyklenlänge, der Exponent die Zahl des Auftretens eines Zyklus entsprechender Länge in einer Permutation. Der Faktor eines Polynomgliedes gibt die Anzahl der Permutationen entsprechender Zyklen einteilung in der Gruppe an. Für die Menge der Figuren (Wurzelbäume) wird wiederum ein Polynom gebildet, die Erzeugende Funktion der Figuren oder Gewichtsfunktion $c(x_1, x_2, \dots, x_m)$. Jedes Polynomglied darin charakterisiert eine Figur. Der Faktor des Gliedes x_i entspricht der Anzahl der Figuren entsprechender Charakteristik. Die unterschiedlichen Variablen bezeichnen die ein-

zeln interessierenden Eigenschaften, ihre Exponenten die Eigenschafts-Parameter (Gewichte). Der Satz von Pólya besagt nun, daß sich die Erzeugende Funktion $Z(G;c(x_1,x_2,\dots,x_m))$ der Konfigurationen ergibt, wenn die Gewichtsfunktion in den Zyklenzeiger der Gruppe eingesetzt wird, so daß jede Variable s_i des Zyklenzeigers durch $c(x_1^i,x_2^i,\dots,x_m^i)$ ersetzt wird. Die Erzeugende Funktion der Konfigurationen enthält gerade alle benötigten Kombinationen von Gewichten.

ABB.1 zeigt die Verknüpfung zweier Wurzelbäume der Höhe 3 bzw. 2 an einer Wurzel zu einem Wurzelbaum der Höhe $3+1=4$ (a) sowie die Konstruktion eines kantenmarkierten Baumes der Höhe 3 (b). Um die im rechten Teil der Abbildung dargestellten Graphen zu erhalten, ist die folgende Rechnung durchzuführen. Für die Knotenzahl der Teilwurzelbäume ergibt sich das Figurenpolynom $c(x)=x^6+x^3$. Mit $Z(S_2)=1/2(s_1^2+s_2)$ erhalten wir $Z(S_2;x^6+x^3)=1/2((x^6+x^3)^2+(x^6\cdot 2+x^3\cdot 2\cdot 1))=x^{12}+x^9+x^6$. Um das Hinzufügen des Basis-Knotens in ABB.1a zu berücksichtigen, ist noch mit x zu multiplizieren. Die beliebige Kombination zweier Wurzelbäume mit einer Knotenzahl von 6 bzw. 3 ergibt also jeweils einen Wurzelbaum mit den Knotenzahlen 13, 10 und 7 bzw. jeweils einen kantenmarkierten Baum mit den Knotenzahlen 12, 9 und 6, wobei in ABB.1 die unsymmetrischen Kombinationen, die dem mittelsten Term entsprechen, dargestellt sind. Für die rekursive Abzählung der Wurzelbäume mit einem Wurzelgrad kleiner gleich g und einem Knotengrad von höchstens $g+1$ ergibt sich

$$w_{n+1}(x)=\sum_{i=0}^g f_1[Z(S_i;w_n(x))] \quad (1)$$

$w(x)$: Erzeugende Funktion der Wurzelbäume

g : maximaler Wurzelgrad

f_1 : Transformation, die die Verknüpfung über den hinzugefügten Wurzelknoten beschreibt

n : Rekursionsstufe, größte Höhe eines Wurzelbaumes.

Mit $g=3$ erhalten wir die Erzeugende Funktion der Alkylradikale, der einwertigen Alkohole, der Monochloralkane o.ä.. Für die Betrachtung der Knotenzahl, die der Zahl der Kohlenstoff-Atome entspricht, ist $f_1 = \cdot x$ (Multiplikation mit x) und $w_0 = x$ zu setzen. Entsprechende Anzahlen sind z.B. in [21] zu finden. Die durch zweimaliges Ausführen von Gleichung (1) erhaltenen Polynome für die Abzählung von Knotenzahlen sind $w_1 = 1+x^2+x^3+x^4$ und $w_2 = 1+x+x^2+2x^3+3x^4+4x^5+4x^6+5x^7+4x^8+4x^9+3x^{10}+2x^{11}+x^{12}+x^{13}$.

Aus den Wurzelbäumen der n -ten Rekursionsstufe lassen sich die kantenmarkierten Bäume konstruieren (ABB.1b):

$$k_n(x)=f_2[Z(S_2;w_n(x))] \quad (2)$$

$k(x)$: Erzeugende Funktion der kantenmarkierten Bäume

f_2 : Transformation, die die Verknüpfung über die hinzugefügte Basis-Kante beschreibt
 n : größte Höhe eines Wurzelbaumes bzw. kantenmarkierten Baumes.

Für die Abzählung nach Knotenanzahl ist dabei $f_2 = -1$ zu wählen.

Bäume können aus Wurzelbäumen zusammengesetzt werden, was der Konstruktion der Alkane aus den Alkylradikalen entspricht. Während Cayley [22] die Generierung von Bäumen aus Wurzelbäumen durch schrittweises systematisches Konstruieren [23] löst, soll hier ein Konstruktionsverfahren für Bäume benutzt werden, das auf einen Satz von Otter [24] zurückgeht. Danach ist in einem Baum die Zahl voneinander verschiedener Knoten (p) minus der Zahl voneinander verschiedener Kanten (q) plus der Zahl der symmetrischen Kanten (s) gleich eins. Unter einer symmetrischen Kante soll eine Kante verstanden werden, die zwei gleiche Teilbäume miteinander verknüpft (ABB.2). Über alle Bäume aufsummiert ergibt sich $\Sigma 1 = \Sigma p - \Sigma q + \Sigma s$, mit $\Sigma 1$:

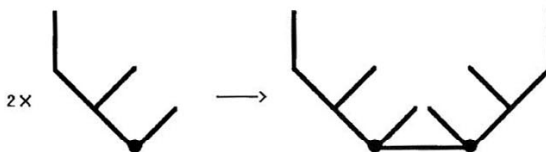


ABB.2 Bildung eines symmetrischen kantenmarkierten Baumes aus zwei identischen Wurzelbäumen.

gesuchte Anzahl der Bäume (bzw. Alkane), Σp : Zahl der knotenmarkierten Bäume (Wurzelbäume), Σq : Zahl der kantenmarkierten Bäume und Σs : Zahl der symmetrischen kantenmarkierten Bäume [11]. Für eine Abzählung der Alkane ist also das folgende Polynom anzusetzen [11]:

$$a(x) = \sum_{i=0}^4 f_3 [\sum_{i=0}^4 f_1 [Z(S_i; w(x))]] - f_2 [Z(S_2; w(x))] + f_2 [w(x)^2] \quad (3)$$

$a(x)$: Erzeugende Funktion der Alkane

$w(x)$: Erzeugende Funktion der Wurzelbäume

f_1 : Transformation, die die Verknüpfung über den hinzugefügten Basis-Wurzelknoten beschreibt

f_2 : Transformation, die die Verknüpfung über die hinzugefügte Basis-Kante beschreibt

f_3 : Transformation, die einen Wurzelbaum bzw. kantenmarkierten Baum in einen freien Baum überführt.

Die drei Terme in Gleichung (3) beschreiben jeweils die Abzählung der knotenmarkierten Bäume, der kantenmarkierten Bäume und der symmetrischen kantenmarkier-

ten Bäume. Die von uns eingeführte gesonderte Formulierung der Transformationen f_1 erleichtert die Ableitung spezieller Abzählformeln. Für die Abzählung der Alkane bezüglich der Zahl der Kohlenstoff-Atome (Knotenzahl) sind die Transformationen $f_1: x$, $f_2: \cdot 1$ und $f_3: \cdot 1$ zu benutzen. Mit den Alkylradikal-Wurzelbäumen $w(x) = x + x^2 + 2x^3 + 4x^4 + 8x^5 + 17x^6 + \dots$ ergibt sich durch Einsetzen in Gleichung (3) für die Alkan-Bäume $a(x) = x + x^2 + x^3 + 2x^4 + 3x^5 + \dots$.

Es wurde ein Computerprogramm geschrieben, das Pólya-Einsetzung, Polynomtransformation und Rekursion für mehrere Variablen in Zyklenzeiger und Gewichtsfunktion gestattet. Darüber hinaus sind Operationen von Gewichtsfunktionen bzw. Zyklenzeigern möglich [25]***.

2. Knotengrad

Die Abzählung einer Klasse von Graphen bezüglich der Knotenzahl kann zur Abzählung von Knotengraden erweitert werden, wenn für jeden Knoten die Zahl der Nachbarknoten, bzw. die Zahl der von diesem Knoten ausgehenden Kanten, mitgeführt wird. Die Abzählungsvariable erhält dazu einen Index, der den Knotengrad bezeichnet. Das Hinzufügen der Basis verändert den Knotengrad der Wurzeln der Teil-Wurzelbäume. Dies wird wiederum durch eine spätere Transformation der Gewichte dieser Wurzeln berücksichtigt. Die Variablen, die die Wurzeln der Teilbäume beschreiben, erhalten deshalb einen zusätzlichen Index. In Erweiterung bisher üblicher Abzählungsverfahren wird damit zwischen Eigenschaften, die durch die Verknüpfung an der Basis geändert bzw. beibehalten werden, unterschieden. Das Gewichtspolynom besteht somit aus Gliedern von Faktoren x_{ij}^n mit i =Knotengrad und

$$j = \begin{cases} 1, & \text{wenn Knoten Wurzel eines Teilbaumes ist} \\ 0, & \text{sonst,} \end{cases}$$

wobei gilt $f_1: x_{i1} \rightarrow x_{i+1,0}$; xx_{i1} , $f_2: x_{i1} \rightarrow x_{i+1,1}$ und $f_3: x_{i1} \rightarrow x_i$, mit $w_0 = xx_{01}$.

Die Transformationen f_1 und f_2 beschreiben die Vergrößerung des Knotengrades der Wurzel um 1 bzw. das Hinzufügen des neuen Wurzelknotens beim Anfügen der Basis. Für ABB.1 erhalten wir $Z(S_2; x^6 x_{10}^3 x_{20} x_{21} x_{30} + x^3 x_{10} x_{11} x_{20}) = x^6 x_{10}^2 x_{11}^2 x_{20}^2 + x^9 x_{10}^4 x_{11} x_{20}^2 x_{21} x_{30} + x^{12} x_{10}^6 x_{20}^2 x_{21}^2 x_{30}^2$. Der mittlere Term entspricht wieder der in ABB.1 dargestellten unsymmetrischen Kombination, so daß sich mit f_1 für den Wurzelbaum ergibt: $x^{10} x_{10}^4 x_{20}^3 x_{21} x_{30}^2$ und mit f_2 für den kantenmarkierten Baum: $x^9 x_{10}^4 x_{20}^2 x_{21} x_{30} x_{31}$. Überführung in die freien Bäume durch f_3 liefert $x^{10} x_1^4 x_2^4 x_3^2$ und $x^9 x_1^4 x_2^3 x_3^2$. Damit haben wir die Gültigkeit der Gleichungen (1) und (2) und der beiden ersten Terme in Gleichung (3), die die Wurzelbäume bzw.

***Das PC-Programm ist beim erstgenannten Autor erhältlich.

Kantenbäume abzählen, auch für unser abgewandeltes Abzählverfahren gezeigt. Der dritte Term in Gleichung (3) $f_2[w(x^2)]$ berechnet die symmetrischen Kantenbäume. Die Gewichtsfunktion $w(x^2)$ drückt allerdings nur die Kombination beider Teilläste des Kantenbaumes und Addition ihrer Gewichte aus. Tatsächlich ist aber die Multiplikation der Gewichte durch die Transformation f_2 zu ersetzen. Wir schreiben deshalb genauer $f_2[Z(G_2;w(x))]$, wobei G_2 diejenige Gruppe ist, die aus $w(x_1, x_2, \dots, x_n)$ die symmetrischen Zweier-Kombinationen bildet: $w((x_1, x_1), (x_2, x_2), \dots, (x_n, x_n))$. Der Zyklenzeiger $Z(G_2) = s_2$ erfüllt gerade diese Bedingung. Wir erhalten für den symmetrischen Kantenbaum aus Abbildung 2 nach dem Satz von Pólya: $Z(G_2; x^6 x_{10}^3 x_{20} x_{21} x_{30}) = x^6 \cdot 2 x_{10}^3 \cdot 2 x_{20}^1 \cdot 2 x_{21}^1 \cdot 2 x_{30}^1 \cdot 2$. Anwendung der Transformation f_2 ergibt $x_{21} \cdot x_{21} \rightarrow x_{31}^2$, woraus f_3 schließlich $x^{12} x_1^6 x_2^2 x_3^4$ liefert. Wir sehen also, daß die Benutzung von Gleichung (4) statt Gleichung (3) unserem erweiterten Abzählverfahren gerecht wird:

$$a(x) = f_3 \left[\sum_{i=0}^4 f_1[Z(S_i; w(x))] \right] - f_2[Z(S_2; w(x))] + f_2[Z(G_2; w(x))]. \quad (4)$$

Damit lassen sich aus den Wurzelbäumen die Alkan-Bäume berechnen.

Für die Abzählung von Bäumen und Wurzelbäumen bezüglich Knotengrad läßt sich eine Formel aufstellen, die von Stammbäumen ausgeht und mit nur einem Variablenindex auskommt [26]. Quintas u. Mitarb. [27] haben auf diese Weise die Valenzen in den Alkan-Isomeren berechnet. Die von uns entwickelte etwas kompliziertere Form soll die Berechnung komplexerer Eigenschaften demonstrieren.

3. Distanzspektrum

Bei der Suche nach graphentheoretischen Invarianten und auf dem Gebiet der QSPR spielt das Distanzspektrum eine gewisse Rolle [28], [29]. Die Anzahl der Knoten im Baum kann als Zahl der Wege der Länge Null aufgefaßt werden. Eine Abzählung der Wege beliebiger Länge läßt sich erreichen, wenn das Gewichtspolynom aus Gliedern von Faktoren der Form x_{ij} ; $i = \text{Weglänge}$;

$j = \begin{cases} 1, & \text{wenn Wurzel Endknoten ist} \\ 0, & \text{sonst} \end{cases}$

besteht, mit $f_1: x_{i,1}^j \cdot x_{k,1}^l \rightarrow x_{i+k+2,0}^j x_{i,1}^j x_{k,1}^l$; $x_{i,1} \rightarrow x_{i,0} x_{i+1,1} x_{i+2,0}$;

$\cdot x_{00} x_{11} g(g-1)/2 x_{20}$ {g: aktueller Wurzelgrad in den Gleichungen (1) bzw. (2)},

$f_2: x_{i,1}^j \cdot x_{k,1}^l \rightarrow x_{i+k+1,0}^j x_{i,1}^j x_{k,1}^l$; $x_{i,1} \rightarrow x_{i,0} x_{i+1,1}$; $\cdot x_{10}$ und $f_3: x_{i,1} \rightarrow x_i$, sowie $w_0 = x_{00}$. Diese Transformationen f_1 und f_2 beschreiben die Kopplung der Wege, die an der Wurzel enden, über die Basis hinweg, sowie die Verlängerung aller dieser

Wege um 1 beim Hinzufügen der Basis. Zum Schluß wird noch das Gewicht der Basis berücksichtigt.

ABB.1 zeigt für die unsymmetrische Kombination in $Z(S_2; x_{00}^6 x_{10}^3 x_{11}^2 x_{20}^3 x_{21}^2 x_{30}^3 x_{31} x_{40} + x_{00}^3 x_{10} x_{11} x_{21})$ das Resultat $x_{00}^9 x_{10}^4 x_{11}^3 x_{20}^3 x_{21}^3 x_{30}^3 x_{31} x_{40}$. Für den ersten Teil von f_1 ergibt sich mit $x_{11}^2 x_{21}^2 x_{31} x_{11} x_{21} : x_{11}^2 x_{11} \rightarrow x_{40}^2$, $x_{11}^2 x_{21} \rightarrow x_{50}^2$, $x_{21}^2 x_{11} \rightarrow x_{50}^2$, $x_{21}^2 x_{21} \rightarrow x_{60}^2$, $x_{31} x_{11} \rightarrow x_{60}$, $x_{31} x_{21} \rightarrow x_{70}$ und für den ersten Teil von f_2 : $x_{30}^2 x_{40}^4 x_{50}^3 x_{60}$. Gemeinsam mit den anderen Teilen von f_1 bzw. f_2 erhalten wir dann $x_{00}^{10} x_{10}^7 x_{11}^2 x_{20}^7 x_{21}^3 x_{30}^7 x_{31}^3 x_{40}^6 x_{41} x_{50}^5 x_{60}^3 x_{70}$ für den Wurzelbaum bzw. $x_{00}^9 x_{10}^7 x_{11} x_{20}^6 x_{21}^3 x_{30}^6 x_{31}^3 x_{40}^5 x_{41} x_{50}^3 x_{60}$ für den kantenmarkierten Baum. Mit f_3 ergibt sich schließlich als Endergebnis $x_{00}^{10} x_{10}^7 x_{11}^2 x_{20}^7 x_{21}^3 x_{30}^7 x_{31}^3 x_{40}^6 x_{41} x_{50}^5 x_{60}^3 x_{70}$ bzw. $x_{00}^9 x_{10}^8 x_{20}^9 x_{30}^9 x_{40}^6 x_{50}^3 x_{60}$.

Zur Illustration soll hier nur die Erzeugende Funktion vorgestellt werden, die sich ergibt, wenn Gleichung (1) nur einmal ausgeführt wird. Eine genauere Diskussion von Distanzspektren muß einer anderen Arbeit [25] vorbehalten bleiben. Es ergeben sich mit $w_0(x) = x_{00}$ alle Wurzelbäume bis zur Höhe eins: $x_{00} + x_{00}^2 x_{11} + x_{00}^3 x_{11}^2 x_{20} + x_{00}^4 x_{11}^3 x_{20}^3$ (ABB.3a). Über Gleichung (4) erhalten wir daraus die in Tabelle 1 aufgeführten Gewichtsvektoren der Alkan-Bäume.

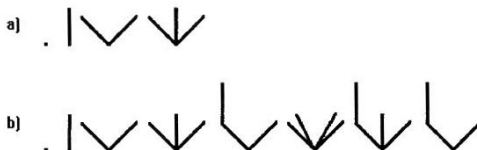


ABB.3 Die Wurzelbäume der ersten Stufe [a] und die daraus zusammengesetzten Alkanbäume mit bis zu 5 Knoten [b].

Tabelle 1 zeigt für die Weglänge 0, also die Knotenzahl der Molekülgraphen bzw. die Zahl der Atome, das Abzählpolynom $x + x^2 + x^3 + 2x^4 + 3x^5 + 4x^6 + 6x^7 + 7x^8 + 8x^9 + 7x^{10} + 7x^{11} + 5x^{12} + 5x^{13} + 3x^{14} + 2x^{15} + x^{16} + x^{17}$. Die volle Erzeugende Funktion der Atomanzahlen in den Alkanen ist dagegen $x + x^2 + x^3 + 2x^4 + 3x^5 + 5x^6 + 9x^7 + 18x^8 + 35x^9 + 75x^{10} + 159x^{11} + 355x^{12} + 802x^{13} + 1858x^{14} + 4347x^{15} + 10359x^{16} + 24894x^{17} + \dots$ [40]. Aus den Wurzelbäumen der ersten Rekursionsstufe der Gleichung (1) lassen sich entsprechend Gleichung (4) alle Wurzelbäume bis zur Höhe 2 bzw. alle kantenmarkierten Bäume bis zur Höhe 1 zusammensetzen, was den Molekülgraphen von Alka-

TAB.1 Distanzspektren der 64 Alkane der ersten Rekursionsstufe, mit Weglänge 0 beginnend.

1	8 7 12 9	11 10 18 18 9	15 14 25 30 36
2 1	8 7 10 9 2	11 10 16 18 11	15 14 24 30 37
	8 7 10 8 3	11 10 16 14 15	
3 2 1	8 7 10 5 6	11 10 15 18 12	16 15 27 33 45
	8 7 9 9 3	11 10 15 18 12	
4 3 3	8 7 9 8 4	11 10 15 14 16	17 16 30 36 54
4 3 2 1	8 7 8 8 5	11 10 14 18 13	
	9 8 13 12 3	12 11 19 21 15	
5 4 6	9 8 13 6 9	12 11 18 21 16	
5 4 4 2	9 8 12 12 4	12 11 18 16 21	
5 4 3 2 1	9 8 12 10 6	12 11 17 21 17	
	9 8 11 12 5	12 11 16 21 18	
6 5 7 3	9 8 11 10 7		
6 5 6 4	9 8 10 12 6	13 12 21 24 21	
6 5 5 4 1	9 8 10 10 8	13 12 20 24 22	
6 5 5 3 2		13 12 19 24 23	
	10 9 15 15 6	13 12 18 24 24	
7 6 9 6	10 9 15 12 9	13 12 21 18 27	
7 6 8 6 1	10 9 14 15 7		
7 6 8 4 3	10 9 13 15 8	14 13 24 27 27	
7 6 7 6 2	10 9 13 12 11	14 13 22 27 29	
7 6 7 4 4	10 9 12 15 9	14 13 21 27 30	
7 6 6 6 3	10 9 12 12 12		

nen bis zur Knotenzahl 5 entspricht (ABB.3b). Nur die ersten 5 Koeffizienten unserer Erzeugenden Funktion sind also richtig. Die Distanzspektren der noch fehlenden Alkane können durch weitere Anwendung von Gleichung (1) berechnet werden.

Wir sehen auch gleich, daß das Distanzspektrum einen Molekülgraphen nicht eindeutig beschreibt. Nach Tabelle 1 ergeben zwei Molekülgraphen, die das Distanzspektrum 11 10 15 18 12 besitzen. Randic [28] gibt dieses Paar isocodaler Graphen ebenfalls an (ABB.4).

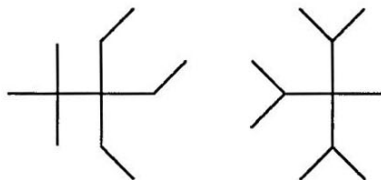


ABB.4 Die Isocodalen Graphen mit dem Distanzspektrum 11 10 15 18 12
[2,2-Dimethyl-3,3-diethylpentan u. 2,3,4-Trimethyl-3-isopropylpentan].

Zur Demonstration haben wir alle Distanzspektren der ersten Stufe berechnet, obwohl hier nur die ersten 5 Glieder der Erzeugenden Funktion interessieren. Während der Rechnung lassen sich aber bei Bedarf Polynomglieder mit Exponenten, die größer als ein vorgegebener Wert sind, verwerfen.

4. Diskussion und Ausblick

Das Ausmultiplizieren des Pólya-Polynoms wird durch eine Transformation ersetzt. Bei der Kombination von Figuren werden die Gewichte nicht einfach zusammengefaßt, sondern entsprechend dieser Transformation miteinander verknüpft. Dadurch wird die Bildung nicht additiver Gewichte berücksichtigt. Dies stellt eine Erweiterung bekannter Abzählungsverfahren dar und ermöglicht die Abzählung komplexerer Eigenschaften. Die Kopplung von Knotengrad und Wegen führt dann z.B. auf Konnektivitätsindizes nach Randic [30], die ja Summen von mit Knotenbewertungen gewichteten Wegen darstellen. Da die Knotengrade am besten über Stammbäume abgezählt werden, ist es zweckmäßig, die Distanzen dabei ebenfalls über Stammbäume zu berechnen. [31] beschreibt die Abzählung der Distanzen in Stammbäumen. Über die Abzählung von Sachs-Graphen [32] können die Koeffizienten des Charakteristischen Polynoms und damit die Eigenwerte der Verknüpfungsmatrix des Molekülgraphen berechnet werden. Die Berechnung aller Kantenfolgen im Graphen ist ebenfalls denkbar [25].

Topologische Indizes scheinen durchaus geeignet, die klassischen QSPR-Methoden zu ergänzen. Im Gegensatz zu vielen physikalisch-chemischen Molekülparametern läßt sich der Wertebereich der mathematisch abgeleiteten Indizes relativ leicht abschätzen. Mit dem von uns entwickelten Verfahren lassen sich Erzeugende Funktionen Topologischer Indizes ohne Generierung der zugehörigen Molekülgraphen gewinnen. Dadurch ist bei bestimmten Fragestellungen eine schnellere Abzählung möglich. Die Verfahren der Kombinatorischen Abzählungstheorie bieten den Vorteil, daß sie mathematisch wohlbegründete Formalismen bereitstellen, andererseits aber sehr flexibel an die jeweiligen Abzählprobleme angepaßt werden können. Abzählpolynome Topologischer Indizes liefern einen sehr ausführlichen Überblick über die Häufigkeit des Vorkommens bestimmter Werte der topologischen Parameter. Mit ihnen läßt sich z.B. die Eindeutigkeit zwischen Index und Strukturformel (Molekülgraph) nachweisen bzw. widerlegen, was im Bereich der Information und Dokumentation von einiger Bedeutung sein dürfte, je nachdem alle Koeffizienten des Polynoms gleich 1 sind oder nicht.

Hat man über QSPR-Untersuchungen eine Moleküleigenschaft in Abhängigkeit von Topologischen Indizes formuliert, so zeigt das Erzeugende Polynom durch ent-

sprechende Koeffizienten ungleich oder gleich Null an, ob es in der betrachteten Stoffklasse Molekülstrukturen mit den für einen gewünschten Eigenschaftswert erforderlichen Ausprägungen Topologischer Indizes gibt.

Bei rekursiver Anwendung, wie bei der Abzählung von Wurzelbäumen demonstriert, läßt sich der Zyklenzeiger als Funktionenreihe auffassen, so daß asymptotische Näherungsformeln für die Anzahlen ermittelt werden können [33]. Die in der vorliegenden Arbeit eingeführten Transformationen sind aber mathematisch recht unhandlich, so daß dieser Vorteil der Kombinatorischen Abzählungstheorie wohl nicht genutzt werden kann.

Es sind Methoden entwickelt worden, die neben den Eigenschaften auch die Molekülgraphen selbst erzeugen [34]. Damit böte sich in der QSPR die Möglichkeit, die Relation $\text{Eigenschaft} = f(\text{Struktur})$ anders herum zu benutzen, was dem Begriff Moleküldesign noch besser gerecht wird. Dadurch ließe sich ein *Konstruktives Moleküldesign* [34b] realisieren.

Gegenwärtig wird die Generierung von Strukturformeln über die Erzeugung lexikographisch geordneter Strukturcodes weiter entwickelt [35]-[37]. Die Generierung Topologischer Indizes auf diesem Wege ist wohl noch nicht versucht worden*. Man hat also die Möglichkeit, Abzählpolynome für topologische Parameter und/oder Molekülgraphen zu entwickeln, oder die Molekülgraphen durch systematische Konstruktion zu erzeugen und die Topologischen Indizes anschließend zu berechnen. Beide Methoden könnten sich sinnvoll ergänzen. Z.B. ist eine Pólya-Abzählung auf einem Strukturcode denkbar.

Schließlich sind Abzählungsverfahren entwickelt worden, die es ermöglichen, einzeln ausgewählte Glieder der Erzeugenden Funktion zu bestimmen [38], so daß sich der Rechenaufwand ganz erheblich verringern ließe.

Ob die angedeuteten Möglichkeiten des Einsatzes der Kombinatorischen Abzählungstheorie für die Chemie nutzbar gemacht werden können hängt wesentlich davon ab, ob effektive Rechner-Algorithmen gefunden werden. Das Gebiet ist offen für Neuerungen, wie die aktuelle Literatur zeigt. Es ist aber noch sehr viel rein mathematische Forschungsarbeit zu leisten.

5. Literatur

- [1] a) I.S.Dmitriev: Moleküle ohne chemische Bindungen?;
Verlag für Grundstoffindustrie, Leipzig, 1982
- b) Molecules without Chemical Bonds: Essays on Chemical Topology;
Central Book Ltd., London, 1982

* [41] beschreibt die Erzeugung aller Molekülgraphen zu einem vorgegeben Wert des Randic-Konnektivitäts-Indizes für Wege der Länge 1.

- [2] P.J.Plath: Diskrete Physik molekularer Umlagerungen; Teubner, Leipzig, 1988
- [3] H.Primas, U.Müller-Herold: Elementare Quantenchemie; Teubner, Stuttgart, 1990, 149ff., 309ff.
Research; Academic Press, New York, 1976
- [4] D.H.Rouvray: The Value of Topological Indices in Chemistry; Match 1 (1975) 125-135
- [5] H.-H.Schmidtke: Some Comments on the Lecture by Dr. D.H.Rouvray; Match 1 (1975) 70
- [6] D.H.Rouvray: A Reply to Some Comments from Prof. Dr. H.-H. Schmidtke; Match 2 (1976) 143-146
- [7] S.B.Elk: The Limitations of Describing Chemical Compounds by Virtue of their Graph Theoretical Representation by a Computer; Match 23 (1988) 9-10
- [8] D.H.Rouvray: The Limitations of Applicability of Topological Indices; J.Mol.Struct. (1989) 187-201
- [9] a) G.Pólya: Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen; Acta Math. 68(1937) 145-254
b) G.Pólya, R.C.Read: Combinatorial Enumeration of Groups, Graphs and Chemical Compounds; Springer, 1987
- [10] F.Harary, E.M., R.W.Robinson, R.C.Read: Pólya's Contributions to Chemical Theory; in A.T.Balaban (Ed.): Chemical Applications of Graph Theory; Academic Press, 1976
- [11] a) R.C.Read: Some Recent Results in Chemical Enumeration; in Lect.Notes Math. 303, Springer, 1972
b) R.C.Read: The Enumeration of Acyclic Chemical Compounds; in A.T.Balaban (Ed.): Chemical Applications of Graph Theory; Academic Press, 1976
- [12] A.T.Balaban: Enumeration of Cyclic Graphs; in A.T.Balaban (Ed.): Chemical Applications of Graph Theory; Academic Press, 1976
- [13] W.Hässelbarth: Substitution Symmetry; Theoret.Chim.Acta 67 (1985) 339-367
- [14] a) A.Kerber, R.Laue, D.Moser: Ein Strukturgenerator für molekulare Graphen; Anal.Chim.Acta 235 (1990) 221-228
b) R.Grund, R.Hager, F.Herrmann, A.Kerber, R.Laue, W.Weber: Groups, Graphs and Isomers; in J.Gmehling (Ed.): Software Development in Chemistry 5; Springer, 1991
c) R.Grund, A.Kerber, R.Laue: MOLGEN, ein Computeralgebra-System für die Konstruktion molekularer Graphen; Match 27 (1992) 87-131
- [15] a) S.Fujita: Enumeration of Non-Rigid Molecules by Means of United Subduced Cycle Indices; Theor.Chim.Acta 77 (1990) 307-321
b) S.Fujita: Symmetry and Combinatorial Enumeration in Chemistry; Springer, 1991
c) S.Fujita: Promolecules with a Subsymmetry of D_{2h} , Combinatorial Enumeration and Stereochemical Properties; J.Chem.Inf.Comput.Sci. 32 (1992) 354-363
- [16] W.G.Klemperer: Enumeration of Permutational Isomerization Reactions; J.Chem.Phys. 56 (1972) 5478-5469
- [17] a) S.Fujita: Description of Organic Reactions Based on Imaginary Transition Structures; J.Chem.Inf.Comput.Sci. 26 (1986) 212-223, 224-230; 27 (1987) 99-104, 104-110
b) S.Fujita: Formulation of Isomeric Reaction Types and Systematic

- Enumeration of Six-Electron Pericyclic Reactions;
J.Chem.Inf.Comput.Sci. 29 (1989) 22-30
- c) S.Fujita: Subduction of Coset Representations. An Application to
Enumeration of Chemical Structures; Theoret.Chim.Acta 76 (1989) 247-268
- [18] a) K.Balasubramanian: Applications of Combinatorics and Graph Theory to
Spectroscopy and Quantum Chemistry; Chem.Rev. 85 (1985) 599-618
b) K.Balasubramanian: Graph Edge Colorings and their Chemical Applications;
Theoret.Chim.Acta 74 (1988) 111-122
c) K.Balasubramanian: Combinatorics of Cluster Enumeration;
J.Chem.Inf.Comput.Sci. 32 (1992) 47-54
d) K.Balasubramanian: Combinatorics of NMR and ESR Spectral Simulations;
J.Chem.Inf.Comput.Sci. 32 (1992) 296-298
- [19] M.Ch.Klin, R.Pöschel, K.Rosenbaum: Angewandte Algebra. Einführung in
gruppentheoretisch-kombinatorische Methoden;
Deutscher Verlag der Wissenschaften, Berlin, 1988
- [20] F.Harary, E.M.Palmer: Graphical Enumeration;
Academic Press, New York, 1973
- [21] K.Ebert, H.Ederer: Computeranwendungen in der Chemie;
VCH, Weinheim, 1983
- [22] E.Cayley: Über die analytischen Figuren, welche in der Mathematik Bäume
genannt werden und ihre Anwendung auf die Theorie chemischer
Verbindungen; Chem.Ber. 8 (1875) 1056-1059
- [23] A.T.Balaban, J.W.Kennedy, L.V.Quintas: The Number of Alkanes Having n
Carbons and a Longest Chain of Length d; J.Chem.Educ. 65 (1988) 4, 304-313
- [24] R.Otter: The Number of Trees; Ann.Math. 49 (1948) 583-599
- [25] J.Will: in Vorbereitung
- [26] F.Harary, G.Prins: The Number of Homeomorphically Irreducible Trees, and
Other Species; Acta Math. 101 (1959) 141-162
- [27] L.V.Quintas, J.Yarmish: Valence Isomers for Chemical Trees;
Match 12 (1981) 65-73
- [28] M.Randic: Characterization of Atoms, Molecules and Classes of Molecules
Based on Paths Enumeration; Match 7 (1979) 5-64
- [29] W.-D. Ihlenfeldt: Computer gestützte Syntheseplanung durch Erkennung
synthetisch nutzbarer Ähnlichkeit von Molekülen;
Dissertation, Technische Universität München, 1991, 45ff.
- [30] M.Randic: On Characterization of Molecular Branching;
J.Am.Chem.Soc. 97 (1975) 6609
- [31] J.Will, B.Adler, G.Biess, S.Dobers: Kombinatorische Abzählung Topologischer
Indizes; Poster, 7. Workshop Chemie - Information - Computer der Gesellschaft
Deutscher Chemiker, 18.-21.10. 1992, Gosen/Berlin
- [32] D.M.Cvetkovic, M.Doob, H.Sachs: Spectra of Graphs - Theory and Application;
Deutscher Verlag der Wissenschaften, Berlin, 1982
- [33] J.Cioslowski: Series Analysis Methods in Enumeration of Chemical Isomers;
Theoret.Chim.Acta 76 (1989) 47-51
- [34] H.Brown, L.Hjelmeland, L.Masinter: Constructive Graph Labeling Using Double
Cosets; Discr.Math. 7 (1974) 1-30
- [35] a) J.V.Knop, W.R.Müller, Z.Jericevic, N.Trinajstić: Computer Enumeration and
Generation of Trees and Rooted Trees;
J.Chem.Inf.Comput.Sci. 21 (1981) 91-99

- b) N.Trinajstić, S.Nikolić, J.V.Knop, W.R.Müller, K.Szymanski: Computational Chemical Graph Theory. Characterization, Enumeration and Generation of Chemical Structures by Computer; Ellis Horwood, 1991
- [36] a) V.Kvasnicka, J.Pospichal: Constructive Enumeration of Acyclic Molecules; Coll.Czech.Chem.Com. 56 (1991) 1777-1802
- b) V.Kvasnicka, J.Pospichal: Constructive Enumeration of Molecular Graphs with Prescribed Valence States; Chemom.Intell.Labor.Syst. 11 (1991) 137-147
- [37] M.L.Contreras, R.Valdivia, R.Rozas: Exhaustive Generation of Organic Isomers. 1. Acyclic Structures; J.Chem.Inf.Comput.Sci. 32 (1992) 323-330
- [38] K.-J.Thürlings: Direkte Berechnung spezieller Anzahlen in Pólya's Abzähltheorie; Match 4 (1978) 161-171
- [39] a) A.Kerber, H.-J. Thürlings: Symmetrieklassen von Funktionen und ihre Abzählungstheorie. I. Die Grundprobleme; Bayreuther Mathemat. Schr. 12 (1983)
- b) A.Kerber: Algebraic Combinatorics via Finite Group Actions; BI Wissenschaftsverlag, Mannheim/Wien/Zürich, 1991
- [40] Jianji Wang, Qiang Wang: The Numbers of Constitutions, Configurations, Chiral and Achiral Configurations of Polyhaloalkanes $C_iH_{2i+2-j}X_j$; Tetrahedron 47 (1991) 2969-2978
- [41] E.V.Gordeeva, M.S.Molchanova, N.S.Zefirov: General Methodology and Computer Program for the Exhaustive Restoring of Chemical Structures by Molecular Connectivity Indexes. Solution of the Inverse Problem in QSAR/QSPR; Tetrahedr.Comput.Method. 3 (1990) 389-415
- [42] K.Jacobs: Einführung in die Kombinatorik; de Gruyter, 1983