match no. 31 pp. 153-203 1994

Erzeugung, Abzählung und Konstruktion von Stereoisomeren*

Thomas Wieland,
Lehrstuhl II für Mathematik
Universität Bayreuth, 95440 Bayreuth, Deutschland

(received: August 1994)

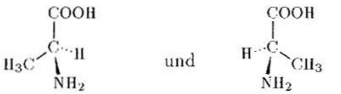
Einleitung

In den letzten zwei Jahrzehnten erregten Probleme der Strukturaufklärung insbesondere organischer Substanzen immer mehr Interesse. Dabei lag das Augenmerk vordringlich auf der Konstruktion von Bindungsisomeren. Eine Reihe von Computerprogrammen ([5], [10], [18]) wurde entwickelt, die in der Lage sind, zu einer gegebenen Summenformel unter Berücksichtigung verschiedener Nebenbedingungen wie Teilstrukturen oder Ringgrößen alle Strukturisomere vollständig und redundanzfrei zu berechnen. Diese leisten heute dem Chemiker im Labor wertvolle Hilfe.
Ein Nachteil dieser Programme besteht allerdings darin, daß sie nur ein graphentheoretisches Modell der Moleküle verwenden, räumliche Eigenschaften aber ganz außer acht lassen. Auch die räumliche Lage der einzelnen Atome eines Moleküls bestimmt einige chemische und physikalische Eigenschaften dieser Substanz. Wichtig ist dafür aber die Wahl eines geeigneten Abstraktionsmodells, um nicht die große Datenmenge aller reellen Koordinaten verarbeiten zu müssen.

Bereits im 19. Jahrhundert erkannten van't Hoff u.a., daß polarisiertes Licht von zwei Substanzen, deren Bindungsgraphen übereinstimmen, in verschiedene Richtungen, aber um denselben Betrag gedreht werden kann. Diese Erscheinung, die z.B. auch Zucker zeigt, nennt man *optische Aktivität*. Sie gehört zu den bekanntesten physikalischen Eigenschaften, die auf der räumlichen Gestalt des Moleküls beruhen. In organischen Substanzen tritt sie dann auf, wenn das Molekül ein Kohlenstoffatom mit vier verschiedenen Liganden enthält. Vernachlässigt man Drehungen, so gibt es nur zwei verschiedene Möglichkeiten, wie sich die Liganden eines solchen sog. asymmetrischen Atoms räumlich anordnen können. Treten mehrere asymmetrische Kohlenstoffe in einem Molekül auf, so muß darüber hinaus noch die Symmetrie des Moleküls berücksichtigt werden.
In der Natur tritt optische Aktivität recht häufig auf. Bei vielen körpereigenen Stoffen des Menschen entscheidet die räumliche Struktur über die Reaktivität. Ein

*Gefördert durch das BMFT unter 211-4003-03-KE7BAY

Beispiel ist die Aminosäure *Alanin*, α-Amino-propionsäure mit der Formel CH_2-$CH(NH_2)$-COOH ([3]). Diese tritt in zwei verschiedenen Formen auf:



L(+)-Alanin D(-)-Alanin

Alanin fungiert als Baustein von Proteinen und ist in modifizierter Form Bestandteil der Pantothensäure, einem Wachstumsfaktor, der zu den hydrophilen Vitaminen des B-Komplexes zählt. Von diesem ist nur die D(+)-Form biologisch aktiv.

Im folgenden wird der Algorithmus von Nourse u.a. ([21], [22]) vorgestellt, der in der Lage ist, alle Isomere, die sich aufgrund asymmetrischer Atome und anderer lokaler Effekte ergeben, vollständig und redundanzfrei zu berechnen. Die mathematische Umsetzung dieses Problems gelingt mit Hilfe von Gruppenoperationen. Berechnet werden Symmetrieklassen von Abbildungen, die jedem für Stereoeffekte verantwortlichen Atom eine der beiden Ausprägungen zuordnen. Im Rahmen dieser Theorie wird es auch möglich, Ergebnisse chemischer Überlegungen formal zu beweisen.
Nach der Vorstellung des Algorithmus werden die Möglichkeiten der Darstellung der Resultate im Rahmen der Nomenklatur und anhand von dreidimensionalen Modellen erörtert. Daneben wird auch die Erweiterung des Verfahrens auf anorganische Elemente mit einer Valenz von fünf oder größer diskutiert.
Doch zunächst soll die Problemstellung in der für die Mathematik unverzichtbaren formalen Weise definiert werden, da der Sprachgebrauch seitens der Chemie für eine algorithmische Behandlung oft zu vage und mehrdeutig ist.

1 Definitionen und Problemstellung

1.1 Modelle chemischer Moleküle

Bevor die räumlichen Eigenschaften chemischer Strukturen erläutert werden, muß geklärt werden, von welchem Abstraktionsmodell die Rede ist. Für die nötigen Begriffe *Konstitution*, *Konfiguration* und *Konformation* sei auf die Literatur (etwa [4] und [17]) verwiesen.
Damit sind wir in der Lage, die wesentlichen Abstraktionsmodelle chemischer Moleküle zu untersuchen (nach [28]):

1.1 Modellebenen chemischer Moleküle

1. *Stöchiometrische Ebene:* Auf dieser Ebene wird nur die *Summenformel* einer Substanz betrachtet. Substanzen mit derselben Summenformel werden als gleich betrachtet. Zur Darstellung genügen die Elementsymbole.

II. *Topologische Ebene:* Relevant ist hier neben der Summenformel die Art der Bindungen und die gegenseitige Verknüpfung der Atome. Unterschieden werden *Konstitutionsisomere*. Die Darstellung gelingt am besten über einen beschrifteten Multigraphen durch Nachbarschaftsliste oder Adjazenzmatrix.

III. *Geometrische Ebene:* Hier wird die *Konfiguration* betrachtet. Daraus können auch gewisse sterische Eigenschaften abgeleitet werden. Lassen sich zwei Moleküle durch Drehungen eines ganzen Moleküls oder um Einfachbindungen zur Deckung bringen, gelten sie als gleich. Zur Darstellung werden neben Nachbarschaftslisten auch Konfigurationslisten benötigt; graphisch können Moleküle auf dieser Ebene durch dreidimensionale bzw. perspektivische Zeichnungen oder durch Projektionen unter Verwendungen von Symbolen zur Andeutung der Räumlichkeit wiedergegeben werden.

IV. *Elektro-mechanische Ebene:* Auf dieser Ebene werden auch intramolekulare Bewegungen wie Translationen oder Rotationen in Betracht gezogen. Unterschieden werden *Konformere*. Auch eine weitere Untergliederung dieser Ebene in starre, sozusagen „eingefrorene" und in freier Bewegung befindliche Moleküle ist möglich. Elektronische Eigenschaften wie Ladungsverteilung oder Polarisierbarkeit können untersucht werden. Die Darstellung braucht dreidimensionale reelle Koordinaten aller Atome oder alle Bindungslängen, -winkel und Torsionswinkel. Graphisch kann dies dann mit Hilfe von Konformationsenergiediagrammen oder Computergraphiken veranschaulicht werden.

V. *Quantenmechanische Ebene:* Die Elektronen werden als Orbitale aufgefaßt; diese sind Linearkombinationen der Wellenfunktionen der Elektronen der Einzelatome. Auf dieser Ebene lassen sich die physikalischen Grundlagen der chemischen Vorgänge darstellen. Rechnerisch wird das Molekül durch seine Hamiltonfunktion und deren (nur näherungsweise angebbaren) Lösungen repräsentiert, graphisch durch Elektronenwolken.

Strikt trennen lassen sich die einzelnen Ebenen natürlich nicht immer; intramolekulare Rotationen können u.U. Einfluß auf die Konfiguration nehmen usw. Doch sollte bei der Bewertung von Resultaten der mathematischen Chemie stets das Modell im Auge behalten werden, das den mathematischen Methoden zugrunde liegt. Denn nur im Rahmen dieses Modells kann die Mathematik Aussagen machen.

1.2 Stereoisomerie und Stereozentren

Allgemein sagt man: Zwei Moleküle heißen *stereoisomer*, falls sie sich hinsichtlich der räumlichen Anordnung ihrer Atome unterscheiden, sonst aber gleiche Konstitution haben.
Diese sehr vage Definition, die in der Literatur auch wieder häufig kontextabhängig verwendet wird, ist für die mathematische Analyse unbrauchbar. Wir wollen daher für unsere Zwecke festhalten:

1.2 Konvention In dieser Arbeit ist mit Stereoisomerie stets *Konfigurationsisomerie* gemeint.

Für einen Überblick über die Stereoisomerie sind noch einige weitere Begriffe nötig (nach [17]): Verhalten sich zwei Stereoisomere wie ein Gegenstand und sein Spiegelbild, so nennt man sie *Enantiomere* oder zueinander *chirale Moleküle*. Dagegen werden Moleküle, zu denen kein Enantiomer existiert, als *achiral* bezeichnet. Stereoisomere, zwischen denen eine solche Beziehung nicht besteht, heißen *Diastereomere*. Das bedeutet, zwei Moleküle können nicht gleichzeitig enantiomer und diastereomer zueinander sein. Zu einem Molekül existieren nur zwei Enantiomere, aber möglicherweise mehrere Diastereomere.

Im Mittelpunkt unserer Betrachtungen wird die von einem Zentrum ausgehende Stereoisomerie stehen:

1.3 Definition *Ein Atom heißt (wirkt als)* Stereozentrum, *wenn es möglich ist, an ihm seine Liganden räumlich auf zwei oder mehr verschiedene Weisen anzuordnen, die nicht durch Drehungen ineinander übergeführt werden können.* •

Diese Definition stellt den Ausgangspunkt unserer Untersuchungen dar und wird später äquivalent umformuliert werden.
Zunächst sollen aber die verschiedenen Arten von Stereozentren betrachtet werden. Man unterscheidet nach der *Koordinationszahl*, der größtmöglichen Anzahl von Nachbarn.

1.4 Arten der Stereozentren ([28])

i) *Koordinationszahl 3:* Die Liganden eines 3-fach koordinierten Stereozentrums ordnen sich im Raum pyramidal an. Das Atom wird zum Stereozentrum, falls alle drei Liganden verschieden sind.

   Bei dreifach koordinierten Atomen spielt die Stabilität der räumlichen Anordnung eine wichtige Rolle. Stabil sind Zentralatome der zweiten Reihe des Periodensystems (z.B. Sulfoxide und Phosphine). Bei Elementen der ersten Reihe wie C, O oder N geht obige Umwandlung so rasch vor sich, daß nicht mehr von einer Konfiguration gesprochen werden kann[1]. Insbesondere beim Stickstoff hängt die Fähigkeit zur Inversion stark von den Liganden ab. Daher wurde im vorliegenden Programm dem Benutzer die Entscheidung offengelassen, ihn als Stereozentrum zu akzeptieren oder nicht (gem. [23]).

ii) *Koordinationszahl 4:* Diese Art von Stereozentrum ist die häufigste, vor allem deshalb, weil der die organische Chemie bestimmende Kohlenstoff sie zeigt. Die Liganden sind an den Ecken eines Tetraeders angeordnet, in dessen Mittelpunkt sich das Zentralatom befindet:

   Wie sich später noch ergeben wird, ist das Atom genau dann Stereozentrum, wenn alle vier Liganden verschieden sind. Da die Umwandlungsenergie relativ hoch ist, kann diese räumliche Anordnung als stabil angesehen werden.

[1] Kohlenstoff ist hier in sp^3-Hybridisierung als Ion gemeint. Die Doppelbindung in sp^2-Hybridisierung ist planar.

iii) *Koordinationszahl 5:* Bei den fünffach koordinierten Atomen herrscht als räumlich Struktur die trigonale Bipyramide vor. Hier sind die stereochemischen Verhältnisse merklich komplizierter. Stellt man sich beispielsweise die Tetraeder-Struktur mit vier verschiedenen Liganden vor, so kann ein fünfter Ligand an 10 verschiedenen Stellen angreifen. Insgesamt ergeben sich bei fünf unterschiedlichen Liganden zwanzig Stereoisomere.

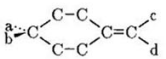
Das häufigste Beispiel eines 5-fach koordinierten Zentrums ist Phosphor.

iv) *Koordinationszahl 6 oder höher:* Mit wachsender Koordinationszahl gibt es immer mehr Möglichkeiten der räumlichen Anordnung. Bei 6-fach koordinierten Zentren sind z.B. Sechseck, Dreieckprisma und Oktaeder möglich. Eine Behandlung ihrer Stereochemie gelingt mit [20] und wird in [31] und [32] diskutiert.

1.5 Chirale Strukturen ohne Stereozentrum ([17])

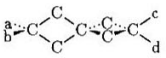
i) *Gerade Anzahl konjugierter Doppelbindungen:* Folgt eine gerade Anzahl konjugierter Doppelbindungen mit Liganden $a = c \neq b = d$ hintereinander, so ist das Molekül chiral, obwohl es kein ausgewiesenes Stereozentrum besitzt.

ii) *Moleküle mit Doppelbindung an alkanischem Ring:* An die Stelle einer Doppelbindung aus i) kann auch ein Ring aus Einfachbindungen treten, z.B.:



Hier kann auch $a \neq c$ oder $b \neq d$ sein.

iii) *Spiroverbindungen:* Eine Verbindung mit zwei Ringen aus Einfachbindungen, die ein Atom gemeinsam haben, nennt man *Spiran.* Auch hier kann ein ähnlicher Effekt wie in i) auftreten, falls wieder $a = c \neq b = d$ ist, z.B.:



1.6 Weitere Formen von Diastereoisomerie ([28])

i) *π-Diastereoisomerie:* Ein Molekül mit einer Doppelbindung und jeweils unterschiedlichen Liganden ist zwar achiral, tritt aber in zwei diastereomeren Formen auf.

ii) *Allenisomerie:* Systeme mit ungerader Anzahl konjugierter Doppelbindungen sind zwar nicht mehr chiral (vgl. 1.5.i), zeigen aber gleichwohl Konfigurationsisomerie.

Alle aus 1.4-1.6 hervorgehenden Stereoisomere sind auf der geometrischen Modellebene beschreibbar und können durch das vorliegende Programm erzeugt werden. Daneben gibt es noch einige weitere Formen der Stereoisomerie; bei diesen spielen elektro-mechanische Kräfte eine deutliche Rolle. Sie können nicht durch diesen Isomergenerator erzeugt werden, seien aber der Vollständigkeit halber hier vorgestellt:

1.7 Weitere Typen der Stereoisomerie

i) *Atropisomerie* ([17]): Normalerweise ist eine freie Drehung um Einfachbindungen möglich. Es kann aber vorkommen, daß voluminöse Substituenten oder Brückenbindungen die Drehung einschränken bzw. unmöglich machen. Bei dieser speziellen Form der Konformationsisomerie entstehen Enantiomere. Ein Beispiel sind Biphenyl-Derivate wie 6,6'-Dinitro-diphensäure.

ii) *Helixstrukturen:* Helicale aromatische Systeme oder Makroproteine haben als Enantiomere die beiden Drehrichtungen der Schraube. Streng gesehen handelt es sich hier wieder um Konformationsisomerie.

iii) *Topologische Isomerie* ([19]): Große zyklische Moleküle können Verschlingungen mit sich selbst zeigen (Kleeblatt, Knoten etc.). Unterschiedliche Konfigurationen sind dabei aber nicht feststellbar.

1.3 Gruppenoperationen und Symmetrieklassen

Nun wollen wir uns den mathematischen Hilfsmitteln zuwenden (nach [15]): Es sei G eine endliche, multiplikativ geschriebene Gruppe und X eine endliche Menge. Eine *Operation von G auf X* (von links) ist eine Abbildung:

$$G \times X \to X,\ (g,x) \mapsto gx \text{ mit } g(g'x) = (gg')x \text{ und } 1x = x.$$

Man schreibt dafür kurz ${}_GX$. Ferner heiße $G(x) := \{gx \mid g \in G\}$ die *Bahn* des Punktes $x \in X$, $G\backslash\backslash X$ die *Menge der Bahnen*, $T(G\backslash\backslash X)$ eine *Transversale der Bahnen* mit $X = \dot{\bigcup}_{t\in T} G(t)$, die sich aus der Äquivalenzklasseneigenschaft der Bahnen ergibt, $G_x := \{g \in G \mid gx = x\}$ der Stabilisator von x, und $X_g := \{x \in X \mid gx = x\}$ die *Menge der Fixpunkte von g*. Das fundamentale Werkzeug für Abzählungen ist das

1.8 Lemma von Cauchy-Frobenius $|G\backslash\backslash X| = \frac{1}{|G|} \sum_{g\in G} |X_g|$

1.9 Definition *Die Operation $G \times G \to G$, $(g,x) \mapsto gxg^{-1}$ heißt* Konjugation. *Die Bahnen $C^G(x) := G(x) = \{gxg^{-1} \mid g \in G\}$ nennt man* Konjugiertenklassen, *den Stabilisator $C_G(x) := G_x = \{g \mid x = gxg^{-1}\}$* Zentralisator. •

Ist nun Y eine weitere endliche Menge, so kann man setzen: $Y^X := \{f \mid f : X \to Y\}$. Für eine Operation wollen wir einführen:

1.10 Definition *Sei H eine weitere endliche Gruppe, so heißt*

$$H\wr_X G := H^X \times G = \{(\psi, g) |\ \psi : X \to H,\ g \in G\}$$

das Kranzprodukt *von H mit G bezüglich X. Die Verknüpfung lautet $(\psi, g)(\psi', g') = (\psi\psi'_g, gg')$ mit $(\psi\psi'_g)(x) := \psi(x)\psi'(g^{-1}x)$* •

1.11 Bemerkung Operiert H auf Y, so operieren G und H auf Y^X auf natürliche Weise: $G \times Y^X \to Y^X$, $(g, f) \mapsto f \circ g^{-1}$ und $H \times Y^X \to Y^X$, $(h, f) \mapsto h \circ f$. Damit operiert das Kranzprodukt folgendermaßen:

$$H\wr_X G \times Y^X \to Y^X,\ ((\psi, g), f) \mapsto \tilde{f},\ \text{mit } \tilde{f}(x) := \psi(x) f(g^{-1}x)$$

Die Bahnen der Operationen von G, H, $H\wr_X G$ und auch von $H \times G$ heißen *Symmetrieklassen von Abbildungen.*

Wir wollen nun $\underline{n} := \{1, 2, \ldots, n\}$ setzen. Die Abbildungen $Y^{\underline{n}}$ lassen sich dann in Form einer *Liste* schreiben: $f = (f(1), f(2), \ldots, f(n))$. Damit können wir festlegen:

1.12 Definition *Sei $Y \subset \mathbb{N}$. Es heiße $f \leq g$ genau dann, wenn der Bildvektor von f lexikographisch kleiner oder gleich dem Bildvektor von g ist. Für $f \in M \subseteq Y^{\underline{n}}$ heiße es $f \leq M$ oder $f = \min(M)$, falls für alle $g \in M$ gilt: $f \leq g$.* •

$S_{\underline{n}} := \{\pi \in \underline{n}^{\underline{n}} |\ \pi \text{ bijektiv}\}$ heißt *symmetrische Gruppe.* Die Elemente von $S_{\underline{n}}$ heißen *Permutationen.* Die Funktion

$$\operatorname{sgn}(\pi) := \prod_{1 \leq i < j \leq n} \frac{\pi j - \pi i}{j - i}$$

wird als *Signum-Funktion* bezeichnet. Die Gruppe $A_{\underline{n}} := \{\pi \in S_{\underline{n}} |\ \operatorname{sgn}(\pi) = 1\}$ heißt *alternierende Gruppe.* Elemente $\pi \in S_{\underline{n}} \cap A_{\underline{n}}$ nennt man *gerade Permutationen*, andere $\pi \in S_{\underline{n}} \backslash A_{\underline{n}}$ *ungerade.*
Wir wollen nun einige Eigenschaften der symmetrischen Gruppe zusammenstellen. Für die Details sei wieder auf [15] verwiesen.

1.13 Bemerkungen

- $|S_{\underline{n}}| = n!$.
- Für die Elemente $\pi \in S_{\underline{n}}$ gibt es zwei Schreibweisen: Zum einen die Listenschreibweise
$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix},$$
wobei die erste Zeile auch weggelassen werden kann, wenn Unklarheiten ausgeschlossen sind.

Ist π andererseits von der Form

$$\begin{pmatrix} i_1 & i_2 & \dots & i_{r-1} & i_r & i_{r+1} & \dots & i_n \\ i_2 & i_3 & \dots & i_r & i_1 & i_{r+1} & \dots & i_n \end{pmatrix},$$

so heißt π *r-Zyklus*, und man schreibt: $\pi = (i_1, i_2, \dots, i_r)$. Jede Permutation kann als Produkt paarweise disjunkter Zyklen geschrieben werden. Dabei bezeichne $c(\pi)$ die Anzahl der Zyklen, und l_i die Länge des i-ten Zykels. Zyklen der Länge 2 heißen *Transpositionen*.

- Die geordneten Längen $\alpha_i(\pi)$, $\pi \in S_{\underline{n}}$, der zyklischen Faktoren von π bilden eine eindeutig bestimmte (eigentliche) Partition von n, d.h. $\forall i: \alpha_i \geq \alpha_{i+1}$ und $\sum_i \alpha_i = n$, in Zeichen $\alpha(\pi) := (\alpha_1(\pi), \dots, \alpha_{c(\pi)}(\pi)) \vdash n$, die *Zykelpartition* von π.

- Definiere für $i \in \underline{n}$, $\pi \in S_{\underline{n}}$:

$$a_i(\pi) := |\{j \mid \alpha_j(\pi) = i\}|$$

Dann nennt man das Tupel $a(\pi) := (a_1(\pi), \dots, a_n(\pi))$ den *Zykeltyp von* π. Dabei gilt: $\sum i \cdot a_i = n$, und man schreibt: $a \vdash n$.

- Seien $\pi, \sigma \in S_{\underline{n}}$. Dann gilt: $C^S(\pi) = C^S(\sigma) \Leftrightarrow \alpha(\pi) = \alpha(\sigma) \Leftrightarrow a(\pi) = a(\sigma)$. Somit lassen sich die Konjugiertenklassen der $S_{\underline{n}}$ mit Hilfe der Partitionen von n parametrisieren.

- Ferner ist $|C^S(\pi)| = n! / \prod_i i^{a_i(\pi)} a_i(\pi)!$.

1.4 Das Konstruktionsproblem

Der Algorithmus von Nourse wurde entwickelt, um die Stereoisomere von Molekülen der Typen 1.4.i) und ii), 1.5 und 1.6 zu konstruieren. Im wesentlichen geht das Verfahren aber von Zentren der Koordinationszahl 4 aus (1.4.ii); alle anderen Effekte werden auf diesen Fall zurückgespielt. Daher erscheint es vernünftig, das Konstruktionsproblem zunächst nur hinsichtlich tetravalenter Stereozentren zu untersuchen. Die erste Frage, die sich dabei stellt, ist: Wie müssen die Liganden[2] beschaffen sein, um verschiedene Konfigurationen hervorzurufen? Dazu vorab noch eine

1.14 Definition

- *Ein Paar* (G, P) *aus einer Menge* G *von Punkten und einer Teilmenge* $P \subseteq G$ *heißt* Gerüst. *Dabei nennt man die Elemente von* P Gerüstplätze *und* $G \backslash P$ Rumpf.

- *Zu* (G, P) *und einer weiteren Menge* L, *der* Ligandenmenge *heißt* $m \in L^P$ *eine* Isomerisation.

Es seien alle Mengen endlich. •

[2] Mit *Ligand* ist stets nicht nur das Atom, das mit einem Zentrum o.ä. verbunden ist, gemeint, sondern auch die weiteren Atome, die außerdem damit verbunden sind (der „Schwanz"). Will man nur ein Atom allein betrachten, spricht man von *Nachbarn*.

Eine Isomerisation ist also die Vorstufe eines Isomers. Denn das Gesamtmolekül wird nur hinsichtlich *eines* Rumpfes und dessen Liganden betrachtet. Ein Atom, das einmal den Rumpf bildete, kann aus anderer Sicht Teil eines Liganden sein und umgekehrt. Somit kann ein Molekül durch Untersuchung verschiedener Rümpfe zu mehreren Isomerisationen führen.
Man kann quantenmechanisch ableiten (s. [7]), daß sich die Valenzen eines vierwertigen Atoms so ausrichten, daß sie maximalen Abstand haben. Stellt man sich die Nachbarn am Ende dieser Valenzen vor, so sitzen sie an den Ecken eines regelmäßigen Tetraeders, in dessen Mittelpunkt sich das Zentralatom befindet.

1.15 Bemerkung Zu einem Gerüst (G, P) heißt

$$S_G := \{f \in O(\mathbb{R}^3) \mid f(G) = G, f(P) = P\}$$

die Symmetriegruppe; alle $f \in S_G$ mit $\det(f) = 1$ bilden die Gruppe D_G. Falls $S_G \neq D_G$, so ist nach dem Homomorphiesatz für Gruppen D_G ein Normalteiler vom Index 2.
Ist $|P| = k$, so existiert ein Homomorphismus $\varphi : S_G \to S_{\underline{k}}$, $f \mapsto \overline{f}$. Dabei heißt $\overline{S}_G := \varphi(S_G)$ die induzierte Permutationsgruppe. Auf diese Weise kann man Symmetricabbildungen durch Permutationen darstellen. Im Fall des Tetraeder ist $\overline{S}_G \cong S_{\underline{4}}$. Dies ist leider aber eher die Ausnahme. Im allgemeinen ist φ weder injektiv noch surjektiv; $\overline{S}_G$ ist dann eine echte Untergruppe der $S_{\underline{k}}$. Will man Symmetriebetrachtungen anstellen, muß man sich daher für jede Figur die induzierte Drehgruppe herleiten. Diese kann im Spezialfall, wie etwa dem ebenen Dreieck im Raum, auch die ganze $S_{\underline{k}}$ sein, also auch ungerade Permutationen beinhalten.

Wir wollen nun zeigen, daß nur eine bestimmte Ligandenabbildung auch Konfigurationsisomere hervorrufen kann:

1.16 Satz *Eine Isomerisation m mit einem vierwertigen Atom als Rumpf hat genau dann mehr als eine Konfiguration, wenn m injektiv ist.*

Beweis: Der Beweis soll mit Hilfe der Theorie der gewichteten Abzählung geführt werden (nach [14] und [24]).
Sei $|P| = 4$ und $L = \{a, b, c, d\}$. Die induzierte Drehgruppe des Tetraeders ist $A_{\underline{4}}$. Da Moleküle, welche durch Drehungen in einander übergeführt werden können, gleiche Konfiguration haben, entsprechen den verschiedenen Konfigurationen die verschiedenen Bahnen der Operation von $A_{\underline{4}}$ auf L^4,
Diese lassen sich mit Hilfe des Satzes von Pólya ([24]) gewichtet abzählen. Dabei ergibt sich für den Zykelindex:

$$C(A_{\underline{4}}, \underline{4} \mid a+b+c+d) = \frac{(a+b+c+d)^4}{12} + \frac{(a^2+b^2+c^2+d^2)^2}{4} + \frac{(2a+2b+2c+2d)(a^3+b^3+c^3+d^3)}{3}$$

$$
\begin{aligned}
= \; & bc^2d + bcd^2 + ac^2d + acd^2 + ab^2d + a^2cd + a^2bd + \\
& a^4 + b^4 + c^4 + b^3c + b^2c^2 + bc^3 + a^3b + a^2bc + \\
& a^3c + b^2cd + a^2b^2 + a^3d + a^2c^2 + ab^3 + ac^3 + \\
& a^2d^2 + ad^3 + b^2d^2 + d^4 + bd^3 + b^3d + abc^2 + abd^2 + \\
& cd^3 + c^3d + c^2d^2 + ab^2c + 2\,abcd
\end{aligned}
$$

Daraus ergibt sich, daß alle Ligandenzuordnungen nur zu jeweils einer Konfiguration führen – bis auf die Zuordnung mit vier verschiedenen Liganden, welche zwei Konfigurationen hat. Diese Isomerisation wird aber genau dann erreicht, wenn jedem Gerüstplatz ein anderer Ligand zugeordnet wird, die Abbildung also injektiv ist.

□

Mit diesem Satz haben wir formal die dem Chemiker wohlbekannte Tatsache bewiesen, daß ein vierwertiges Atom nur dann asymmetrisch sein und zu einem chiralen Molekül führen kann, wenn alle vier Liganden sich unterscheiden (vgl. 1.4.ii).

Zur Erzeugung aller Stereoisomere, die sich aus vierfach koordinierten Atomen ergeben, muß man also „lediglich" jedem Stereozentrum eine dieser beiden Konfigurationen, die wir mit 0 und 1 bezeichnen wollen, zuordnen – unter Berücksichtigung von intramolekularen Symmetrien. Damit können wir als Ausgangsproblem für Nourses Algorithmus formulieren:

1.17 Stereochemisches Konstruktionsproblem

i) Finde alle vierfach koordinierten Atome, deren Liganden alle verschieden sind.

ii) Bilde alle Kombinationen aus der Zuordnung von 0 *oder* 1 *zu jedem dieser Atome unter Berücksichtigung der Symmetrien des Moleküls.*

Die Lösung dieses Problems wird im Mittelpunkt des folgenden Kapitels stehen. Während ii) i.w. auf die Berechnung von Bahnenrepräsentanten führt, bedarf i) einer genaueren Untersuchung; insbesondere muß geklärt werden, was mit „verschiedene Liganden" gemeint ist.

1.5 Konfigurationen und Doppelnebenklassen

Bevor wir die Details des Algorithmus betrachten, wollen wir noch einen Blick auf die mathematischen Hintergründe werfen. Dabei werden sich Doppelnebenklassen als das geeignete mathematische Konzept zur Beschreibung von Isomerisationen herausstellen.

Sei G eine Gruppe und A und B Untergruppen von G. Dann operiert $A \times B$ auf G: $(A \times B) \times G \to G,\ ((a,b),g) \mapsto agb^{-1}$. Die Bahnen dieser Operation heißen *Doppelnebenklassen* von A und B in G. Man schreibt $A\backslash G/B := \{AgB \mid g \in G\} = A \times B\backslash\backslash G$. Eine Bahn AgB heißt *Doppelnebenklasse* von A und B bezüglich g.

1.18 Lemma *Für die Anzahl der Doppelnebenklassen gilt:*

$$|A\backslash G/B| = \frac{|G|}{|A||B|} \sum_{g \in \mathcal{C}} \frac{|C^G(g) \cap A||C^G(g) \cap B|}{|C^G(g)|},$$

wobei $\mathcal{C}$ eine Transversale der Konjugiertenklassen sei.

Für den Beweis s. [15].

1.19 Definition

i) Das m-Tupel natürlicher Zahlen $\lambda_i \in \mathbb{N}$ $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m)$ mit $\sum_{i=1}^m \lambda_i = n$ heißt uneigentliche Partition *von n, kurz $\lambda \models n$.*

ii) Sei $\underline{n}^\lambda$ eine Mengenzerlegung von $\underline{n}$ mit $\underline{n}_i^\lambda \subseteq \underline{n}$, wobei $|\underline{n}_i^\lambda| = \lambda_i \, \forall i \in \underline{m}$ und $\underline{n} = \dot{\bigcup}_{i \in \underline{m}} \underline{n}_i^\lambda$. Der Stabilisator von $\underline{n}^\lambda$

$$S(\underline{n}^\lambda) := S_{\underline{n}_1^\lambda} \otimes S_{\underline{n}_2^\lambda} \otimes \ldots \otimes S_{\underline{n}_m^\lambda} = \{\pi \in S_{\underline{n}} \mid \pi \underline{n}_i^\lambda = \underline{n}_i^\lambda \; \forall i \in \underline{m}\}$$

heißt Young-Untergruppe.

•

Die Younggruppe ist eine Untergruppe der symmetrischen Gruppe: $S(\underline{n}^\lambda) \leq S_{\underline{n}}$. Zudem gilt: $S(\underline{n}^\lambda) \cong S_{\lambda_1} \times \ldots \times S_{\lambda_m}$.

Wir wollen nun Isomerisationen betrachten, die einen Rumpf der Koordinationszahl k haben, d.h. $|P| = k$. Dieser kann – chemisch gesehen – aus einem oder mehreren Atomen bestehen.

1.20 Isomerisationen k-fach koordinierter Zentren

i) Numeriert man die Liganden mit $l_1, \ldots, l_k$ unter Vernachlässigung möglicher chemischer Identitäten, so zeigt sich, daß die Menge aller Isomerisationen dieses Zentrums bijektiv zur $S_{\underline{k}}$ ist:

$$\begin{pmatrix} p_1 & p_2 & \ldots & p_k \\ l_{i_1} & l_{i_2} & \ldots & l_{i_k} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & \ldots & k \\ i_1 & i_2 & \ldots & i_k \end{pmatrix}$$

ii) Die Gerüstplätze zeigen stets eine gewisse Drehsymmetrie, die durch eine Rotationsgruppe R ausgedrückt werden kann. Da wir Isomerisationen unabhängig von Drehungen der Gerüstplätze betrachten wollen, müssen wir m und ρm mit $\rho \in R$ als gleich ansehen, uns also den Bahnen der Operation

$$R \times L^P \to L^P, \; (\rho, m) \mapsto m \circ \rho^{-1}$$

zuwenden (vgl. 1.11).

iii) Sind Liganden chemisch identisch, so ändert es nichts an der Isomerisation, wenn sie beliebig an den Gerüstplätzen vertauscht werden. Es gebe $l \leq k$ verschiedene Sorten von Liganden und von der Sorte $i \in \underline{l}$ jeweils λ_i. Damit ist $\lambda \models k$ und S_λ o.E. die Gruppe der Vertauschungen. Auch diese operiert auf L^P:

$$S_\lambda \times L^P \to L^P,\ (\sigma, m) \mapsto \sigma \circ m.$$

Nimmt man die chemische Ununterscheidbarkeit als Kriterium, so muß man auch hier nur die Bahnen betrachten.

Damit haben wir erhalten (vgl. [12], [25], [26]):

1.21 Satz *Die wesentlich verschiedenen Isomerisationen eines k-fach koordinierten Zentrums entsprechen einer Transversalen der Doppelnebenklassen $R\backslash S_{\underline{k}}/S_\lambda$, wobei R die Drehgruppe der Gerüstplätze und S_λ die Younggruppe der Ligandenvertauschungen mit $\lambda \models k$ sind.*

Mit Hilfe dieses Ergebnisses können wir nun auch Definition 1.3 neu fassen:

1.22 Definition *Ein Atom der Koordinationszahl k heißt* Stereozentrum, *wenn es eine Ligandenpartition $\lambda \models k$ gibt, so daß $|R\backslash S_{\underline{k}}/S_\lambda| > 1$ ist.*

•

Damit stehen uns für Konstruktionsprobleme bei Zentren mit einer Koordinationszahl größer als 4 effektive Algorithmen zur Verfügung, die Doppelnebenklassentransversalen berechnen, etwa [8]. Auch zur Abzählung ist dieses Resultat nützlich; wir können nämlich nun 1.18 anwenden:

1.23 Satz *Sei R die Drehgruppe eines k-fach koordinierten Zentrums, $\lambda \models k$ eine Zerlegung der Ligandenmenge in chemisch ununterscheidbare und o.E. S_λ die Younggruppe zu λ. Dann existieren an diesem Zentrum*

$$|R\backslash S_{\underline{k}}/S_\lambda| = \frac{1}{|R| \cdot \lambda_1! \cdot \ldots \cdot \lambda_l!} \sum_{a \vdash k} \left(\prod_i i^{a_i} a_i! \right) \cdot |\{\rho \in R \mid a(\rho) = a\}| |\{\sigma \in S_\lambda \mid a(\sigma) = a\}|$$

wesentlich verschiedene Isomerisationen.

Beweis: Die Formel ergibt sich mit 1.13 durch Einsetzen in 1.18. □

Eine ähnliche Formel, ohne Ausnutzung von 1.13, wird in [20] angegeben. Dort sind aber auch dynamische Rotationen der mit Liganden behafteten Gerüstplätze zugelassen, was auf eine etwas kompliziertere Version führt.
Wir wollen die Leistungsfähigkeit dieser mathematischen Beschreibung nun veranschaulichen:

1.24 Beispiel Wie mehrfach erwähnt, ist die (induzierte) Drehgruppe des Tetraeders die $A_{\underline{4}}$, also

$$R = \{1, (12)(34), (13)(24), (14)(23), (123), (124), (134), (234), (132), (142), (143), (243)\}$$

Damit ergibt sich beispielsweise für die Ligandenpartition $(2,1,1)$, daß nur über die Zykeltypen $(4,0,0,0)$ und $(2,1,0,0)$ summiert werden muß:

$$\frac{1}{12 \cdot 2 \cdot 1}(4! \cdot 1 \cdot 1 + (2! \cdot 2) \cdot 0 \cdot 1) = 1$$

Insgesamt erhalten wir:

λ	$(1,1,1,1)$	$(2,1,1)$	$(3,1)$	$(2,2)$	(4)
$\lvert R\backslash S_{\underline{4}}/S_\lambda \rvert$	2	1	1	1	1

Weitere Beispiele finden sich in [32].

Wir haben also mit Hilfe von Doppelnebenklassen die Chiralität des Tetraeders mit verschiedenen Liganden ebenso erhalten wie mit der gewichteten Abzählung in 1.16. Wir wollen nun zeigen (nach [11] und [25]), daß dies kein Zufall ist, sondern daß ein enger Zusammenhang besteht. Dazu zunächst ein

1.25 Lemma *Sei G eine endliche Gruppe und $U \leq G$ eine Untergruppe. Es operiere ${}_GX$, wobei X endlich, und es sei $x \in X$. Ist dann T eine Transversale der Doppelnebenklassen $U\backslash G/G_x$, so ist $\{tx \mid t \in T\}$ eine Transversale von $U\backslash\backslash X$.*

Beweis: Wegen ${}_GX$ operiert auch U auf X. Wir können ferner o.E. annehmen, daß $|G\backslash\backslash X| = 1$. (Die Operation heißt dann auch *transitiv.*)
Sei nun $x \in X$ mit Stabilisator G_x. Dann gibt es für alle $x' \in X$ ein $g \in G$ mit $x' = gx$. Zwei Elemente g_1x und g_2x liegen in derselben U-Bahn, wenn ein $u \in U$ existiert mit $g_1x = ug_2x$. Zudem gilt $g_2x = g_3x$, falls es ein $h \in G_x$ gibt mit $g_2 = g_3h$. Damit folgt:

$$g_2x \in U(g_1x) \Longleftrightarrow \exists u \in U, h \in G_x : g_2 = ug_1h \Longleftrightarrow g_2 \in Ug_1G_x$$

□

Damit sind wir in der Lage, den Zusammenhang zwischen Bahnen gegebenen Gewichts und Doppelnebenklassen zu zeigen. Dabei wollen wir sagen, ein $m \in L^P$ habe den *Inhalt* λ, falls für alle $i \in \{1, \ldots, |L|\}$ gilt $|\{m^{-1}(l_i)\}| = \lambda_i$, bei entsprechender Numerierung der Elemente von L. Der Einfachheit halber sei $P = \underline{k}$ für $k = |P|$.

1.26 Satz *Es operiere $R \leq S_{\underline{k}}$ auf L^P. Ist $m \in L^P$ vom Inhalt λ, so existiert eine Bijektion zwischen den Doppelnebenklassen $R\backslash S_{\underline{k}}/S_\lambda$ und den Bahnen $R\backslash\backslash L^P$ des Inhaltes λ:*

$$R\backslash S_{\underline{k}}/S_\lambda \to R\backslash\backslash L^P_\lambda,\ R\pi S_\lambda \mapsto R(m \circ \pi^{-1})$$

Beweis: Aus 1.11 ist bekannt, wie R auf L^P operiert, nämlich $R \times L^P \to L^P$, $(\rho, m) \mapsto m \circ \rho^{-1}$. Betrachten wir statt R die ganze $S_{\underline{k}}$, so haben alle Elemente einer $S_{\underline{k}}$-Bahn jeweils gleichen Inhalt, da die Operation nur die Reihenfolge, nicht aber die Werte der Bildvektorelemente ändert.
Nun setzen wir in 1.25 $G := S_{\underline{k}}$ und $U := R$. Somit ist eine Bijektion zwischen $R\backslash\backslash L^P_\lambda$ und $R\backslash S_{\underline{k}}/(S_{\underline{k}})_m$ hergestellt. m wiederum ist unter *den* Elementen von $S_{\underline{k}}$ stabil, die nur gleiche Bilder vertauschen. Diese sind genau die Elemente von S_λ. Die Gestalt der Abbildung ist ebenfalls mit 1.25 klar.

□

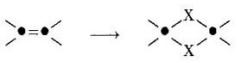
Es gibt also in der Abzählung nach Gewicht mit Hilfe des Satzes von Pólya und in der Bestimmung der Anzahl von Doppelnebenklassen zwei äquivalente Methoden, um die Zahl der wesentlich verschiedenen Isomerisationen an einem k-fach koordinierten Rumpf zu bestimmen. Wollen wir aber Repräsentanten angeben, so erweist sich die gewichtete Abzählung – wie schon der Name sagt – als ungeeignet. Hierfür können Algorithmen zur Konstruktion von Doppelnebenklassentransversalen hingegen erfolgreich eingesetzt werden. Diese Anwendung wird in [31] und [32] diskutiert.

2 Der Algorithmus von Nourse

Nun wollen wir uns den Details des Algorithmus zur Generierung von Stereoisomeren zuwenden, der in den späten siebziger Jahren von J.G. Nourse, D.H. Smith, R.E. Carhart und C. Djerassi entwickelt wurde und seitdem nicht durch einen umfassenderen abgelöst werden konnte. Die Darstellung folgt den Veröffentlichungen [21] und [22].
Die grundsätzlichen Gedanken wurden weitgehend übernommen, insbesondere in den Algorithmen 2.2, 2.4 und 2.32. Bei den Verfahren 2.14 und 2.21 wurde die bekannten Methoden präzsiert und verfeinert, während die weiteren Algorithmen – soweit es sich nicht um klassische handelt – sowie alle Sätze selbst entwickelt wurden.

2.1 Suche nach aromatischen Systemen

Der eigentliche Konstruktionsalgorithmus nutzt nur die Chiralität des Gerüstes (s. 1.4) aus. Doppelbindungen zwischen Atomen, die in der Lage sind, Stereoisomerie der Typen 1.5 und 1.6 hervorzurufen, werden auf diese Situation zurückgespielt, indem an die Stelle der Doppelbindung zwei fiktive Knoten X eingesetzt werden:



Um den Konstruktionsprozeß effektiv zu gestalten, ist es nötig, im voraus die Zahl potentieller Stereozentren möglichst stark zu reduzieren. Im Hinblick auf die Doppelbindungen sind es vor allem aromatische Systeme, die eine größere Menge nichtstereorelevanter Zweifachbindungen beitragen. Diese gilt es, in einem ersten Schritt zu entdecken.

2.1 Definition *[17] Unter einem* aromatischen System *wollen wir Strukturen verstehen, die aus mehreren Ringen des Benzol zusammengesetzt sind. Dabei unterscheidet man zwischen* linear anellierten *Systemen und* angular anellierten *Systemen.* •

Durch die völlige Delokalisation der π-Elektronen ist eine freie Beweglichkeit um die in der Zeichnung einfach dargestellten Bindungen nicht mehr möglich. Derartige Systeme zeigen daher keine Konfigurationsisomerie. Darüber hinaus gibt es auch sog. nicht-benzoide Aromaten wie Annulene oder Pyridin. Aufgrund des nicht ganz eindeutigen Begriffs und der vereinzelten Möglichkeit zur Bildung von Diastereomeren berücksichtigt das vorliegende Programm derartige Strukturen nicht. Um die Atome, die an einem aromatischen System beteiligt sind, zu identifizieren, wird ein einfacher Algorithmus zur Substruktursuche verwendet, auf den hier nicht näher eingegangen werden muß (s. etwa [8]).

2.2 Behandlung von Doppelbindungen

Nachdem die aromatischen Systeme identifiziert sind, können wir nun alle die Atome suchen, die als Stereozentren in Frage kommen.

2.2 Algorithmus (Suche nach potentiellen Stereozentren)

i) Durchlaufe alle Atome und markiere diejenigen, die Valenz 3 (falls erlaubt) oder 4 haben, höchstens einen Wasserstoff als Nachbarn haben und die nicht in einem aromatischen System vorkommen.

ii) Untersuche nun die Nachbarschaft jedes markierten Atoms.

   (a) Ist es an einer Stelle drei- oder vierfach[3] gebunden, so lösche die Markierung.

   (b) Ist es mit zwei Partnern doppelt gebunden, dann lösche die Markierung und beginne den Durchlauf erneut, falls einer der beiden Nachbarn kein potentielles Stereozentrum ist.

   (c) Ist es nur mit einem Partner doppelt gebunden, lösche die Markierung, falls der doppelt gebundene Nachbar kein potentielles Stereozentrum ist.

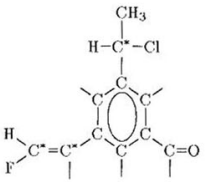
◇

[3] Die vierfache Bindung wird nur für den Spezialfall C_2 berücksichtigt.

Durch die Unterscheidung der Nachbarschaft wird folgendes erreicht:

- Kohlenstoffe (o.ä.), die mit Sauerstoff (o.ä.) doppelt gebunden sind, werden korrekt als Stereozentren verworfen.
- Alle Partner konjugierter Doppelbindungen werden als potentielle Stereozentren erkannt, sofern die Atome an den Enden der Kette verschiedene weitere Liganden tragen (vgl. 1.5).

2.3 Beispiel Im folgenden Beispiel sind die potentiellen Stereozentren mit einem * gekennzeichnet.



Weitere Beispiele werden bei der Diskussion um die Eleminierung potentieller Zentren vorgestellt werden.

Sind alle Atome entsprechend markiert, können die Doppelbindungen weiterverarbeitet werden:

2.4 Algorithmus (Entfernen von Doppelbindungen) Ermittle die Anzahl der einzufügenden Atome. Für jede Doppelbindung zwischen zwei potentiellen Stereozentren sind zwei Atome hinzuzufügen. Durchlaufe sodann alle Atome und füge neue Atome anstelle von Doppelbindungen ein. ◇

Damit hat jedes für die Stereochemie relevante Atom nur einfach gebundene Nachbarn. Ingesamt gesehen werden mit dem Erweitern des Moleküls zwei Ziele verfolgt:

i) Die algorithmische Behandlung wird vereinfacht. Jedes mehrfach gebundene Stereozentrum kann nun auf gleiche Weise wie ein einfach gebundes behandelt werden. Dies erweist sich besonders dann als Vorteil, wenn auch die Liganden Stereozentren enthalten.

ii) Die Enantiomere von Strukturen mit mehrfach aneinander liegenden Doppelbindungen können leicht errechnet werden, indem man nämlich die Konfigurationen aller beteiligten Stereozentren umkehrt.

2.3 Die Konfigurationssymmetriegruppe

Wir wollen uns nun der Symmetriegruppe zuwenden, die in den folgenden Abschnitten benötigt wird. Dazu betrachten wir ein Konstitutionsisomer als Multigraphen (s. dazu auch [10]).

2.5 Definition *Ein* numerierter m-Graph *mit p Punkten ist eine Abbildung*

$$\gamma : \binom{\underline{p}}{2} \to \{0, 1, \ldots, m\},$$

wobei $\binom{\underline{p}}{2}$ alle zweielementigen Teilmengen von $\underline{p}$ seien. Wird einem Punktepaar $\{i, j\}$ der Wert r zugeordnet, so bedeutet dies, daß die Punkte durch eine r-fache Kante miteinander verbunden sind. (Ist $r = 0$, sind die Punkte unverbunden.) •

Insgesamt gibt es $\binom{p}{2}$ Punktepaare. Die $S_{\underline{p}}$ operiert auf den Punkten und damit auch auf den Graphen:

$$S_{\underline{p}} \times \{0, \ldots, m\}^{\binom{\underline{p}}{2}} \to \{0, \ldots, m\}^{\binom{\underline{p}}{2}},\ (\pi, \gamma) \mapsto \gamma \circ \pi^{-1} \tag{2.6}$$

Dabei ist das Bild die Abbildung $\gamma' = \pi\gamma = \gamma \circ \pi^{-1}$ mit $\gamma'(\{i, j\}) = \gamma(\pi^{-1}\{i, j\}) = \gamma(\{\pi^{-1}i, \pi^{-1}j\})$. Von besonderer Bedeutung sind die Permutationen, welche die Bindungsverhältnisse des Graphen unverändert lassen:

2.7 Definition *Der Stabilisator eines numerierten m-Graphen $\gamma \in \{0, \ldots, m\}^{\binom{\underline{p}}{2}}$ heißt* Automorphismengruppe *von γ: $A(\gamma) := \{\pi \in S_{\underline{p}} \mid \pi\gamma = \gamma\}$* •

Entscheidend sind neben den Bindungsverhältnissen eines Moleküls selbstverständlich aber auch noch die Typen der einzelnen Atome, d.h. das Element[4] (des Periodensystems), aus dem das jeweilige Atom besteht. Mathematisch bedeutet dies, daß jedem Knoten ein zusätzlicher Wert zugeordnet wird, er „beschriftet" wird, wie man sagt.

2.8 Definition *Besteht ein Molekül aus n verschiedenen Elementen E_i, so heißen die Abbildungen $\beta \in \{E_1, \ldots, E_n\}^{\underline{p}}$ oder einfacher $\beta \in \underline{n}^{\underline{p}}$* Beschriftungen *des numerierten Graphen γ, für die für jedes $i \in \underline{p}$ gilt, daß die Gesamtzahl der von i ausgehenden Kanten gleich der Valenz des Atomtyps $\beta(i)$ ist.* •

Ein Molekül gegebener Konstitution, also ein molekularer Graph, kann daher durch ein Paar $(\gamma, \beta) \in \{0, \ldots, m\}^{\binom{\underline{p}}{2}} \times \underline{n}^{\underline{p}}$ von Abbildungen beschrieben werden.
Auch auf den Beschriftungen operiert $S_{\underline{p}}$ auf natürliche Weise

$$S_{\underline{p}} \times \underline{n}^{\underline{p}} \to \underline{n}^{\underline{p}},\ (\pi, \beta) \mapsto \beta \circ \pi^{-1} \tag{2.9}$$

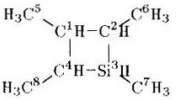
Insbesondere operiert damit auch $A(\gamma)$ auf den Beschriftungen. Fassen wir die Stabilisatoren der Operationen 2.6 und 2.9 zusammen, so erhalten wir die

2.10 Definition *Der Stabilisator eines molekularen Graphen (γ, β) heißt* Automorphismengruppe *von (γ, β): $A(\gamma, \beta) := \{\pi \in A(\gamma) \mid \pi\beta = \beta\}$* •

[4] Aufgrund der Verzahnung von Chemie und Mathematik läßt es sich leider nicht vermeiden, daß in diesem Abschnitt der Begriff „Element" mit unterschiedlicher Bedeutung verwendet wird. Es sollte sich jedoch dem aufmerksamen Leser aus dem Kontext erschließen, ob die mathematische Bedeutung – einzelnes Teil einer Menge – oder die chemische – Atomsorte im Periodensystem – gemeint ist.

Im folgenden Beispiel werden wir nur die Permutationen von Nichtwasserstoffen betrachten. Auch in der programminternen Repräsentation werden Wasserstoffe nicht als Knoten des Graphen sondern als zusätzliche Beschriftung angegeben, um Speicher zu sparen und um die Komplexität nicht unnötig zu erhöhen.

2.11 Beispiel Wir betrachten (als Graph und als Molekül) Tetramethylsilacyclobutan



Die verbundenen Punktepaare sind

$$\{1,2\},\{1,4\},\{1,5\},\{2,3\},\{2,6\},\{3,4\},\{3,7\},\{4,8\}.$$

Hat C den Typ 1 und Si den Typ 2, so kann die Beschriftung durch den Vektor $(1,1,2,1,1,1,1,1)$ dargestellt werden. Die Automorphismengruppe des Graphen hat die Ordnung 8 und lautet (läßt man Fixpunkte weg):

1	(13)(57)
(24)(68)	(13)(24)(57)(68)
(12)(34)(56)(78)	(14)(23)(58)(67)
(1234)(5678)	(1432)(5876)

Wären alle Beschriftungen gleich, wie etwa bei Tetramethylcyclobutan, so wäre dies auch die Automorphismengruppe des Moleküls. Wegen des Siliziums aber fallen eine Reihe von Abbildungen weg. So sind etwa die Spiegelungen an den Seitenhalbierenden gegenüberliegender Seiten keine selbsterhaltenden Abbildungen mehr. Als Automorphismengruppe bleibt lediglich $\{1,(24)(68)\}$.

Eine größere Automorphismengruppe bedeutet in der Sprache des Chemikers eine größere Symmetrie des Moleküls. Diese wirkt sich natürlich auch auf die Zahl der Stereozentren aus. Als Faustregel kann man sagen: Je größer die Automorphismengruppe, desto weniger Stereoisomere gibt es. So existieren zum Beispiel zu 1,2,3,4-Tetramethylcyclobutan vier Stereoisomere, während das gleiche Molekül mit einem Silizium bereits acht besitzt.
Für die Berechnung der Automorphismengruppe wird im vorliegenden Programm eine Routine aus dem Programmsystem **MOLGEN**$^+$ eingesetzt (s. [16]). Zunächst werden die Knoten des Graphen in Automorphismenpartitionen eingeteilt, d.h. es kann danach keine Automorphismen geben, die Knoten verschiedener Partitionen aufeinander abbilden.
Die Automorphismengruppe selbst wird nun mit Hilfe einer sog. *Sims-Kette* berechnet und gespeichert. Dies ist eine Kette von Untergruppen

$$G \geq C(\underline{1}) \geq \ldots \geq C(\underline{l}) = \mathrm{id}, \tag{2.12}$$

wobei $C(\underline{r}) := \{\pi \in G \mid \pi i = i \; \forall i \in \underline{r}\}$ der punktweise Stabilisator von $\underline{r}$ ist (vgl. [27]). Damit können Permutationsgruppen sehr effizient verwaltet werden. Es genügt nämlich, Transversalen der Linksnebenklassen $C(\underline{r-1})/C(\underline{r})$ für $r \in \underline{l}$ zu speichern; alle Elemente von G sind damit eindeutig darstellbar. Auf diese Weise wird vermieden, jedes einzelne Element von G zu speichern, was eine unnötige Einschränkung der Reichweite der Algorithmen zur Folge hätte.

Gibt es in einem Molekül s Stereozentren, so kann man diese gemäß der Ausgangsnumerierung des Moleküls numerieren. Die entsprechende Abbildung heiße ν. Dabei sei $\nu \in \underline{s}^{\underline{p}}$ und nur auf den Stereozentren definiert. Ist der Punkt i das j-te Stereozentrum, so ist $\nu(i) = j$.
Unter den Elementen der Automorphismengruppe bleiben Stereozentren entweder fest oder werden in Stereozentren überführt. Damit induziert ν eine Abbildung

$$\overline{\nu} : A(\gamma, \beta) \to S_{\underline{s}} \tag{2.13}$$

Das Bild von $\overline{\nu}$ sind die molekülerhaltenden Permutationen der Stereozentren. Zu bemerken ist noch, daß $\overline{\nu}$ i.a. nicht injektiv ist.

Wir setzen nun $S_2 := \{(0)(1), (01)\}$. Die Konfigurationssymmetriegruppe wird eine Untergruppe von $S_2 \wr_{\underline{s}} \overline{\nu}(A(\gamma, \beta))$ sein. Jedem $\pi \in A(\gamma, \beta)$ wird eine Abbildung $\epsilon \in (S_2)^{\underline{s}}$ zugeordnet. Diese gilt es zunächst zu ermitteln. Wir wollen sie *Inversion* nennen.

2.14 Algorithmus (Bestimmung der Inversionen) Durchlaufe alle Automorphismen und für jede Permutation $\pi \in A(\gamma, \beta)$ alle Stereozentren. Unterscheide für das Zentrum $i \in \underline{s}$ folgende Fälle:

i) $\overline{\nu}(\pi)(i) = i$, d.h. das Stereozentrum bleibt fest unter π. Die Liganden des Zentrums $(l_1, \ldots, l_k)$ (gemäß der Ausgangsnumerierung des Moleküls) werden durch π auf $(l_{i_1}, \ldots, l_{i_k})$ abgebildet. Bestimme die Permutation

$$\begin{pmatrix} 1 & \ldots & k \\ i_1 & \ldots & i_k \end{pmatrix}.$$

Ist sie gerade, so setze $\epsilon_\pi(i) = (0)(1)$, sonst setze $\epsilon_\pi(i) = (01)$.

ii) $\overline{\nu}(\pi)(i) = j \neq i$. Zusammen mit dem Stereozentrum werden auch die Liganden $(l_1, \ldots, l_k)$ abgebildet, etwa auf $(\tilde{l}_{j_1}, \ldots, \tilde{l}_{j_k})$. Bestimme die Permutation

$$\begin{pmatrix} 1 & \ldots & k \\ j_1 & \ldots & j_k \end{pmatrix}.$$

Ist sie gerade, so setze $\epsilon_\pi(j) = (0)(1)$, sonst setze $\epsilon_\pi(j) = (01)$.

$\diamond$

Damit können wir festlegen:

2.15 Satz und Definition *Die Menge*

$$C(\gamma,\beta) := \{(\epsilon_\pi, \overline{\nu}(\pi)) \,|\, \pi \in A(\gamma,\beta)\}$$

ist eine Untergruppe der $S_2 \wr_{\underline{s}} \overline{\nu}(A(\gamma,\beta))$ *und heißt* Konfigurationssymmetriegruppe.

Beweis:

i) Nach Konstruktion 2.14 ist $C(\gamma,\beta) \subseteq (S_2)^{\underline{s}} \times \overline{\nu}(A(\gamma,\beta))$ klar.

Zudem ist auch wegen $((\mathrm{id}_2,\ldots,\mathrm{id}_2),\mathrm{id}_{\underline{s}}) \in C(\gamma,\beta)$ die Menge $C(\gamma,\beta) \neq \emptyset$.

ii) Seien $\pi,\sigma \in A(\gamma,\beta)$ und $\overline{\pi} := \overline{\nu}(\pi)$ und $\overline{\sigma} := \overline{\nu}(\sigma)$. Nach dem Untergruppenkriterium für endliche Gruppen ist zu zeigen, daß dann auch $(\epsilon_\pi,\overline{\pi})\cdot(\epsilon_\sigma,\overline{\sigma}) \in C(\gamma,\beta)$. Gemäß der Verknüpfung für Kranzprodukte gilt: $(\epsilon_\pi,\overline{\pi})\cdot(\epsilon_\sigma,\overline{\sigma}) = (\epsilon_\pi(\epsilon_\sigma)_{\overline{\pi}},\overline{\pi}\,\overline{\sigma})$ mit $(\epsilon_\pi(\epsilon_\sigma)_{\overline{\pi}})(i) = \epsilon_\pi(i)\,\epsilon_\sigma(\overline{\pi}^{-1}i)$. Da die Unnumerierung ν ein Homomorphismus ist, gilt: $\overline{\pi}\,\overline{\sigma} = \overline{\pi\sigma}$.

Sei nun $i \in \underline{s}$. Unter π werden die Liganden $(l_1,\ldots,l_k)$ des Zentrums i auf $(\pi(l_1),\ldots,\pi(l_k))$ abgebildet. Das Signum der Indexvertauschungen bestimmt $\epsilon_\pi(\overline{\pi}i)$. Bei σ erhalten wir $(\sigma(l_1),\ldots,\sigma(l_k))$ und wieder aus dem Signum $\epsilon_\sigma(\overline{\sigma}i) = (\epsilon_\sigma)_{\overline{\pi}}(\overline{\pi}i)$. Aus Symmetriegründen gilt zudem:

$$\epsilon_\sigma(\overline{\sigma}i) = \epsilon_{\sigma^{-1}}(\overline{\sigma}^{-1}i) \tag{2.16}$$

Zur Bestimmung von $\epsilon_{\pi\sigma}(\overline{\pi}i)$ müssen wir die Liganden des Zentrums $\overline{\sigma}^{-1}i$, also etwa $(\tilde{l}_1,\ldots,\tilde{l}_k)$ betrachten. Diese werden durch $\pi\circ\sigma$ auf $(\pi\sigma(\tilde{l}_1),\ldots,\pi\sigma(\tilde{l}_k))$ abgebildet. Diese Abbildung von $\overline{\sigma}^{-1}i$ auf $\overline{\pi}i$ läßt sich über i faktorisieren:

$$\overline{\sigma}^{-1}i \xrightarrow{\overline{\sigma}} i \xrightarrow{\overline{\pi}} \overline{\pi}i$$

Da die Signum-Funktion ein Homomorphismus bezüglich der Hintereinanderausführung von Abbildungen ist, folgt zusammen mit 2.16: $\epsilon_\pi(\overline{\pi}i)\,(\epsilon_\sigma)_{\overline{\pi}}(\overline{\pi}i) = \epsilon_{\pi\sigma}(\overline{\pi}i)$. Da mit i auch $\overline{\pi}i$ alle Elemente von $\underline{s}$ durchläuft, ergibt sich daraus die Behauptung.

□

Es sollte noch bemerkt werden, daß es durch eine geeignete Wahl der Ausgangsnumerierung des Moleküls immer möglich ist, ν gleich der Identität zu setzen. Die Gruppe $\overline{\nu}(A(\gamma,\beta))$ ist dann nichts anderes als die Einschränkung von $A(\gamma,\beta)$ auf die ersten s Elemente.

Der recht formal beschriebene Prozeß der Ermittlung der Inversionen wird nun durch Beispiele deutlicher werden.

2.17 Beispiele

i) Aus 2.11 ist bereits 1,2,3,4-Tetramethylcyclobutan mit seiner Automorphismengruppe bekannt. Die Stereozentren sind die Atome 1, 2, 3 und 4; daher ist hier ν die Identität. Wir wollen nun die Inversionen ϵ_π bestimmen.

(a) Unter der Identität bleiben alle Liganden fest:

$$\epsilon_{\pi_1} = ((0)(1), (0)(1), (0)(1), (0)(1)).$$

(b) Unter $\pi_2 := (24)(68)$ werden die Liganden von 2 folgendermaßen abgebildet: $(1,3,6) \mapsto (1,3,8)$. Da die Reihenfolge fest bleibt, ist $\epsilon_{\pi_2}(4) = (0)(1)$. Für das Zentrum 4 gilt dies analog. Bei den Liganden von 1 gilt: $(2,4,5) \mapsto (4,2,5)$ Hier wird eine Transposition benötigt, um die aufsteigende Reihenfolge wiederherzustellen, also $\epsilon_{\pi_2}(1) = (01)$. Analog bei 3. Insgesamt bedeutet dies:

$$\epsilon_{\pi_2} = ((01), (0)(1), (01), (0)(1))$$

(c) Bei $\pi_3 := (12)(34)(56)(78)$ gilt an 1: $(2,4,5) \mapsto (1,3,6)$, d.h. die Reihenfolge bleibt erhalten, $\epsilon_{\pi_3}(1) = (0)(1)$. Für 2, 3 und 4 gilt dies analog, so daß

$$\epsilon_{\pi_3} = ((0)(1), (0)(1), (0)(1), (0)(1)).$$

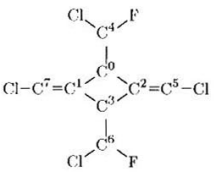
(d) Am Zentrum 1 haben wir bei $\pi_4 = (1234)(5678)$ die Abbildung: $(2,4,5) \mapsto (3,1,6)$. Die Reihenfolge ist umgedreht, daher $\epsilon_{\pi_4}(2) = (01)$. Bei 2 wiederum gilt: $(1,3,6) \mapsto (2,4,7)$, keine Umkehrung der Reihenfolge, $\epsilon_{\pi_4}(3) = (0)(1)$. Für 3 und 4 bestimmen sich die Inversionen analog, womit wir erhalten:

$$\epsilon_{\pi_4} = ((0)(1), (01), (0)(1), (01))$$

Auf die gleiche Weise läßt sich dies fortführen. Die vollständige Konfigurationssymmetriegruppe lautet (Zur besseren Übersichtlichkeit steht $e := (0)(1)$ und $i := (01)$.):

(e,e,e,e),	$(1)(2)(3)(4)$
(i,e,i,e),	(24)
(e,e,e,e),	$(12)(34)$
(e,i,e,i),	(1234)
(e,i,e,i),	(13)
(i,i,i,i),	$(13)(24)$
(i,i,i,i),	$(14)(23)$
(i,e,i,e),	(1432)

ii) Ein weiteres Beispiel soll ein halogeniertes Cyclobutan-Derivat sein. (Nur die Stereozentren sind numeriert. Freie Valenzen seien durch Wasserstoffe besetzt.)



Wie auch im Programm üblich, beginnt hier die Numerierung bei 0.

Die Konfigurationssymmetriegruppe lautet:

(e,e,e,e,e,e,e,e),	1	(e,i,i,e,e,e,e,e),	(03)(46)
(e,e,i,e,e,i,e,e),	1	(e,i,e,e,e,i,e,e),	(03)(46)
(e,i,e,e,e,e,e,i),	1	(e,e,i,e,e,e,e,i),	(03)(46)
(e,i,i,e,e,i,e,i),	1	(e,e,e,e,e,i,e,i),	(03)(46)
(i,e,e,i,e,e,e,e),	(12)(57)	(i,i,i,i,e,e,e,e),	(03)(12)(46)(57)
(i,i,e,i,e,e,e,i),	(12)(57)	(i,e,i,i,e,e,e,i),	(03)(12)(46)(57)
(i,e,i,i,e,i,e,e),	(12)(57)	(i,i,e,i,e,i,e,e),	(03)(12)(46)(57)
(i,i,i,i,e,i,e,i),	(12)(57)	(i,e,e,i,e,i,e,i),	(03)(12)(46)(57)

iii) Die Konfigurationssymmetriegruppe muß aber nicht immer so klein sein, wie in den letzten Beispielen. Bei dem Molekül in Abb. 1 hat sie die Ordnung 384.

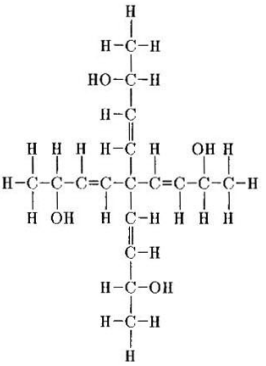


Abbildung 1: 2,8-Hydroxy-5,5-Di(3-Hydroxy-1-Propenyl)-3,6-Nonadien

Es gibt hier insgesamt 13 Stereozentren und 36 Stereoisomere.

Wozu müssen die Inversionen aber nun eigentlich bestimmt werden? Eine Antwort liefert

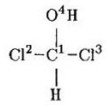
2.18 Bedeutung der Konfigurationssymmetriegruppe

- Die Automorphismengruppe des Moleküls beschreibt nur die Bindungsverhältnisse und die Atomnamen. Nicht in Betracht gezogen wird, welche räumlichen Ausprägungen dieses Molekül haben kann und wie diese ineinander umgewandelt werden können. Die Konfigurationssymmetriegruppe berücksichtigt den Wechsel zwischen verschiedenen Konfigurationen, der in der räumlichen Darstellung einem „Umklappen" ähnelt.

- Um die Liganden als (geordnete) Tupel schreiben zu können, stellt man sich die Umgebung der Stereozentren lokal räumlich plaziert vor. Mit dieser Methode, die sich algorithmisch durch die Numerierung der Knoten bei der Eingabe völlig natürlich ergibt, kann dann die Änderung der Reihenfolge als Konfigurationsänderung (je nach Vorzeichen) interpretiert werden.

- Durch die Inversionen wird festgestellt, daß es Ligandenpermutationen gibt, die imstande sind, die Konfiguration eines Stereozentrums zu ändern, d.h. – in der Sprache der Doppelnebenklassen (s. 1.21) – daß durch diese Abbildung die räumliche Anordnung von einer Doppelnebenklasse in die andere überführt wird.

  Gibt es nur ein potentielles Stereozentrum, so wird dies besonders augenfällig. Die Automorphismengruppe z.B. des Dichlormethanols



  ist $\{1,(23)\}$, die zugehörige Konfigurationssymmetriegruppe $\{(e,(1)),(i,(1))\}$. Da eine Inversion existiert, können nicht zwei Konfigurationen unterschieden werden; es gibt also nur ein Stereoisomer.

  Bei mehreren Stereozentren, wie in den letzten Beispielen, wird durch die Inversionen die Auswirkung der Permutation auf die Konfigurationen aller Stereozentren berücksichtigt. Liegt ein Stereozentrum beispielsweise in einem Liganden eines anderen Zentrums, so ändert sich seine Konfiguration, wenn der Ligand permutiert wird.

- Die Konfigurationssymmetriegruppe erlaubt, eine Reihe von Untersuchungen des stereochemischen Verhaltens eines Moleküls anzustellen, ohne eine räumliche Plazierung zu benötigen. In den folgenden Abschnitten werden wir sehen, wie sich die Bahnen der $C(\gamma,\beta)$ auf $\{0,1\}^{\underline{s}}$ als die verschiedenen Stereoisomere

erweisen. Dabei kann man dann entweder die Transversale für die Generierung berechnen oder die Bahnenzugehörigkeit als Kriterium zur Bestimmung der Konfiguration nutzen. Darüber hinaus stehen auch die Sätze der kombinatorischen Abzähltheorie damit offen, was die Ermittlung von Anzahlen gestattet.

2.19 Bemerkung Bereits G. Pólya ([24]) erkannte die Wichtigkeit, stereochemische Probleme mit gruppentheoretischen Mitteln zu behandeln. Unsere Konfigurationssymmetriegruppe heißt bei ihm *Stereoformelgruppe* – allerdings mit dem Unterschied, daß die Inversionen nicht in der hier vorgestellten Weise berücksichtigt werden. Pólya betrachtete vielmehr das Gerüst als ganzes und die Abbildungen, die das Gerüst erhalten. Auf diesem Weg gelangte er etwa zur Aussage, daß es bei n Stereozentren maximal 2^n Stereoisomere geben kann. Seine *erweiterte Stereoformelgruppe* läßt auch Spiegelungen des Gerüstes zu, wodurch Enantiomere in dieselbe Äquivalenzklasse fallen.
Der hier diskutierte Ansatz hat sicherlich Pólya als einen seiner geistigen Väter; das Konzept der Stereoformelgruppe ist indessen nicht für weiterreichende Untersuchungen geeignet, da stets die genaue Kenntnis des räumliche Gerüstes des ganzen Moleküls vorausgesetzt wird.

2.4 Verifizierung der Stereozentren

Nachdem die potentiellen Stereozentren entdeckt sind und die Automorphismengruppe berechnet ist, muß noch untersucht werden, ob alle Atome, die als potentielles Stereozentrum erkannt wurden, auch wirklich welche sind. Dies stellt sich zwar auch beim Konstruktionsalgorithmus heraus; der Aufwand, diese Erkenntnis auf diesem Weg zu gewinnen, ist aber ungleich höher. Um die Generierung so effektiv wie möglich zu gestalten, sollte daher die Zahl der Stereozentren vorher auf das tatsächlich Nötige reduziert werden.
Dabei hilft folgende Überlegung:

2.20 Lemma *Gibt es eine Permutation in der Automorphismengruppe, die eingeschränkt auf die Liganden eines potentiellen Stereozentrums ungerade ist, so ist dieses Atom nur dann wirklich ein Stereozentrum, wenn die Liganden, die nicht fest bleiben, auch mindestens ein Stereozentrum enthalten.*

Beweis: Enthalten die Liganden selbst Stereozentren, so ändert ihre Vertauschung zwar nicht die Konstitution, wie die Automorphismengruppe zeigt, wohl aber die Konfiguration. Daher sie müssen als stereochemisch verschieden angesehen werden. Sind aber alle Liganden verschieden, so ist das Atom nach 1.16 (oder analog) ein Stereozentrum. □

Ein Beispiel für eine solche Struktur ist 2.17.iii.
Nach 1.16 würde man vermuten, daß dies nicht nur für gerade Permutationen, sondern für alle außer der Identität gilt. Tatsächlich lassen sich aber Moleküle finden, die ein anderes Verhalten zeigen (s. [22]).
Zur Verfizierung der Stereozentren muß man also in zwei Schritten vorgehen: Zuerst müssen die Umgebungen der potentiellen Zentren daraufhin untersucht werden, ob es Automorphismen gibt, die auf den Liganden ungerade permutieren. Dann müssen die Liganden der so identifizierten Atome auf Stereozentren durchsucht werden – und zwar möglichst von „außen" nach „innen", um kein Zentrum zu übergehen.
Ein erster Algorithmus markiert die permutierten Nachbarn der Stereozentren unter entsprechenden Automorphismen. Nun gilt es, die so markierten Nachbarn zu durchsuchen. Wurde auf diese Weise ein potentielles Stereozentrum als nicht stereorelevant erkannt, muß die Suche von vorn beginnen, da das Streichen eines Zentrums Auswirkungen auf die anderen haben kann. Um diesen Prozeß nicht zu oft wiederholen zu müssen, sollte mit den Atomen begonnen werden, die – anschaulich gesprochen – im Außenbereich des Moleküls liegen, und später dann die weiter inneren untersucht werden. Diese bildliche Vorstellung setzt ein Algorithmus um, der die Längen der längsten Wege im Graphen ausgehend von den Stereozentren ermittelt. Nun können die Stereozentren gemäß 2.20 überprüft werden.

2.21 Algorithmus (Verifizierung der potentiellen Stereozentren)

i) Durchlaufe alle Weglängen absteigend, beginne mit der maximalen Länge aller Wege, die von relevanten Zentren ausgehen.

ii) Durchlaufe alle Stereozentren mit der aktuellen als längster Weglänge. Ermittle die Zahl der Stereozentren in allen markierten Liganden mit folgender rekursiver Unterfunktion. Gehe aus von dem Stereozentrum als Zentralatom, das als Vorgänger angesehen werden soll, und seinem Nachbarn als aktuelles Atom.

   (a) Ist das aktuelle Atom gleich dem Zentralatom, so gib 0 zurück.

   (b) Durchlaufe alle Nachbarn des aktuellen Atoms außer dem Vorgänger. War einer bereits in einer höheren Ebene berücksichtigt, so gib 1 zurück. Anderenfalls summiere über die Rückgabewerte der Funktion auf den darunterliegenden Ebenen mit dem aktuellen Atom als Vorgänger und den Nachbarn als aktuell.

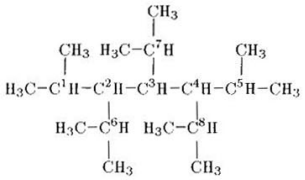
   (c) Ist die Summe größer als 0 oder ist das aktuelle Atom selbst Stereozentrum, so gib 1 zurück; ansonsten 0.

   Falls die Funktion bei allen permutierenden Liganden 0 liefert, ist das gegenwärtige Atom kein Stereozentrum; lösche also die Markierung als Stereozentrum, setzte die äußere Schleife wieder auf den Anfang und gehe zu i).

iii) Gibt es nach Durchlauf dieser Schleifen keine Stereozentren mehr, so verlasse das Programm.

◇

2.22 Beispiel Ein etwas größeres Beispiel ist



Nur die potentiellen Stereozentren sind numeriert. Als maximale Weglänge ermittelt sich für 1,5,6,8 jeweils 5, für 2,4 und 7 jeweils 4 und für 3 den Wert 3. Beim Durchlauf in 2.21 wird zunächst für 1 erkannt, daß die beiden Methylgruppen keine Stereozentren enthalten, dann das gleiche bei 5, bei 6 und 8. Die beiden permutierenden Liganden der 2 enthalten zwar die Atome 1 und 6, diese wurden aber bereits als Stereozentren verworfen. Daher wird auch das Atom 2 als nicht stereo-relevant erkannt, ebenso 4 und 7. Der Kohlenstoff 3 schließlich ist auch kein Stereozentrum, da die ehemaligen Zentren 1,2,6 sowie 4,5,8 bereits weggefallen sind. Damit enthält das Molekül überhaupt keine potentiellen Stereozentren mehr, ist also achiral.

2.5 Generierung der Stereoisomere

Nachdem nunmehr alle Vorarbeit erledigt ist – die Stereozentren verifiziert, die Konfigurationssymmetriegruppe berechnet, können jetzt alle wesentlich verschiedenen Stereoisomere bestimmt werden. Nach der Aufspaltung der Doppelbindungen durch 2.4 kann jedes der s Stereozentren nach 1.16 (oder analog) in zwei Konfigurationen vorkommen, die 0 und 1 genannt werden können. Die Menge aller möglichen Stereoisomere ist somit $\{0,1\}^{\underline{s}}$. Von diesen können aber einige „gleich" sein. Als „gleich" sehen wir Stereoisomere an, für die es einen Isomorphismus gibt, der sie aufeinander abbildet und dabei die Änderungen der Konfigurationen unter dieser Abbildung berücksichtigt. Die Menge aller dieser Isomorphismen ist aber genau $C(\gamma,\beta)$, so daß die wesentlich verschiedenen Stereoisomere genau den Bahnen von $C(\gamma,\beta)$ auf $\{0,1\}^{\underline{s}}$ entsprechen. Damit haben wir erhalten:

2.23 Satz *Zu einem Molekül mit Konfigurationssymmetriegruppe $C(\gamma,\beta)$ und s Stereozentren ist eine Transversale von $C(\gamma,\beta)\backslash\backslash\{0,1\}^{\underline{s}}$ die Menge aller wesentlich verschiedenen Stereoisomere der Typen 1.4.i) und ii), 1.5 und 1.6.*

Wir stehen nun vor der Aufgabe, die Bahnen $C(\gamma,\beta)\backslash\backslash\{0,1\}^{\underline{s}}$ zu konstruieren. Nourse schlägt in [21] dazu vor, immer einen Vektor aus $\{0,1\}^{\underline{s}}$ zu nehmen, der nicht Mitglied einer bereits berechneten Bahn ist. Diese Methode erfordert es aber, jeden neuen Repräsentanten mit allen bereits bekannten zu vergleichen. Die Zahl der dafür notwendigen Tests kann sehr groß werden, so daß dieses Vorgehen algorithmisch nicht empfehlenswert ist. Vorteilhafter ist folgendes Verfahren (nach [9], vgl. dazu 1.12):

2.24 Kanonischer Bahnenalgorithmus

i) Starte mit der leeren Menge $T := \emptyset$.

ii) Berechne in lexikographisch aufsteigender Reihenfolge die Vektoren $\{0,1\}^{\underline{s}}$ und untersuche dabei jedes $f \in \{0,1\}^{\underline{s}}$ folgendermaßen:

- Existiert ein $(\epsilon, \overline{\pi}) \in C(\gamma, \beta)$ mit $\tilde{f} < f$, wobei $\tilde{f}(i) := \epsilon(i) f(\overline{\pi}^{-1} i)$, so ist f kein kanonischer Bahnrepräsentant (d.h. nicht der kleinste seiner Bahn).
- Andernfalls speichere $T := T \cup \{f\}$.

$\diamond$

Auf diese Weise muß nicht für jedes f die ganze $C(\gamma, \beta)$ durchlaufen werden; denn hat man ein $\tilde{f} < f$, so kann mit dem nächsten Repräsentanten der Test fortgesetzt werden. Der entscheidende Vorteil ist aber, daß der Test jedes Repräsentanten nur von $C(\gamma, \beta)$ und nicht von den bereits berechneten Repräsentanten abhängt.

Im einzelnen gliedert sich dieser Algorithmus in den Test auf Minimalität und die Transversalenkonstruktion.

2.25 Algorithmus (Minimaltest)

i) Durchlaufe alle $(\epsilon_\pi, \overline{\pi}) \in C(\gamma, \beta)$ mit $(\epsilon_\pi, \overline{\pi}) \neq (\mathrm{id}_2, \mathrm{id}_{\underline{s}})$ und berechne $\tilde{f}(i) = \epsilon_\pi(i) \cdot f(\overline{\pi}^{-1}(i))\ \forall i \in \underline{s}$.

ii) Suche das kleinste $j \in \underline{s}$ mit $\tilde{f}(j) \neq f(j)$.

iii) Gilt $\tilde{f}(j) < f(j)$, so brich den Durchlauf ab und gib 0 zurück; sonst gehe zu i).

iv) Gib 1 zurück.

$\diamond$

Damit können wir die kanonischen Bahnrepräsentanten erzeugen:

2.26 Algorithmus (Konstruktion der Repräsentanten)

i) Beginne mit $f = (0, \ldots, 0)$.

ii) Gibt der Minimaltest 2.25 den Wert 1 zurück, so ist der aktuelle Repräsentant minimal. Speichere ihn in einer Liste.

iii) Erhöhe f nach folgender Regel:

(a) Setze $\kappa := s$.

(b) Ist $f(\kappa) = 0$, so setze $f(\kappa) := 1$ und springe zurück.

(c) Setze $f(\kappa) := 0$ und vermindere κ um 1. Falls $\kappa > 0$, gehe zu (b).

iv) Gehe zu ii), falls $\kappa > 0$; sonst beende die Funktion.

$\diamond$

2.27 Beispiel Wir wollen das Tetramethylcyclobutan aus 2.11 betrachten, dessen Konfigurationssymmetriegruppe wir in 2.17 berechnet haben.
Wir starten mit $f = (0,0,0,0)$. Da dies der global minimale Vektor ist, ist er auch ein kanonischer Repräsentant und wird gespeichert.
Als nächstes folgt $f = (0,0,0,1)$. Bei $((i,e,i,e),(24))$ ist $\tilde{f} = (1,1,1,0) > f$. Bei $((e,e,e,e),(12)(34))$ dann $\tilde{f} = (0,0,1,0) > f$ usw. Schließlich erweist sich f als neuer Repräsentant.
Nun wollen wir $f = (0,0,1,0)$ betrachten. Unter $((i,e,i,e),(24))$ erhalten wir $\tilde{f} = (0,0,0,1) < f$. Dies ist also kein Repräsentant. $(0,0,1,1)$ ist dagegen wieder einer, während $(0,1,0,0)$ unter $((e,i,e,i),(1234))$ zu $(0,0,0,1)$ wird.
Am Ende liefert uns der Algorithmus 2.26 die Repräsentanten

$$(0,0,0,0)$$
$$(0,0,0,1)$$
$$(0,0,1,1)$$
$$(0,1,1,0).$$

Es gibt also zu Tetramethylcyclobutan vier Stereoisomere.

Obwohl bei diesem Algorithmus meist nicht alle Elemente der Konfigurationssymmetriegruppe durchlaufen werden müssen, ist es bislang unvermeidlich, alle Vektoren zu durchlaufen. Unter Umständen können zwischen zwei kanonischen Repräsentanten eine ganze Reihe von nicht-kanonischen liegen. Die Abbildung, welche die Minimalität verwirft, kann aber Aufschluß darüber geben, welcher Repräsentant als nächstes getestet werden muß. Dies soll das nächste Lemma genauer fassen (nach [9]):

2.28 Lemma *Sei* $(\epsilon,\pi) \in C(\gamma,\beta)$ *und* $f \in \{0,1\}^s$ *mit* $\tilde{f} < f$, *d.h.*

$$\exists j < s:\ \epsilon(\kappa)\cdot f(\pi^{-1}(\kappa)) = f(\kappa) \text{ für } \kappa < j \text{ und } \epsilon(j)\cdot f(\pi^{-1}(j)) < f(j).$$

Sei $\theta := \max\{\pi^{-1}(\kappa) \mid \kappa < j,\ \epsilon(\kappa)f(\pi^{-1}(\kappa)) = 0\}$ *und* $\Theta := \max\{j, \pi^{-1}(j), \theta\}$. *Dann gilt für jedes* $g \in \{0,1\}^s$ *mit* $g(\kappa) = f(\kappa)\,\forall 1 \le \kappa \le \Theta$:

$$\tilde{g} < g.$$

Beweis: Nach Definition ist $j, \pi^{-1}(j) \le \Theta$. Damit gilt: $g(\pi^{-1}(j)) = f(\pi^{-1}(j))$, also auch

$$\epsilon(j)g(\pi^{-1}(j)) = \epsilon(j)f(\pi^{-1}(j)) < f(j) = g(j).$$

Für einen indirekten Beweis nehmen wir an, daß $\tilde{g} \ge g$. Dann gibt es ein $\kappa < j$ mit $\epsilon(\kappa)g(\pi^{-1}(\kappa)) > g(\kappa)$.
Falls $\epsilon(\kappa)f(\pi^{-1}(\kappa)) = 0$, so liefert die Voraussetzung gemäß der Definition von Θ:

$$\epsilon(\kappa)g(\pi^{-1}(\kappa)) = \epsilon(\kappa)f(\pi^{-1}(\kappa)) = f(\kappa) = g(\kappa),$$

im Widerspruch zu $\epsilon(\kappa)g(\pi^{-1}(\kappa)) > g(\kappa)$.
Ist andererseits $\epsilon(\kappa)f(\pi^{-1}(\kappa)) = 1$, dann folgt wegen $\kappa < j$ nach Voraussetzung

$$g(\kappa) = f(\kappa) = \epsilon(\kappa)f(\pi^{-1}(\kappa)) = 1,$$

was ebenso einen Widerspruch zu $\epsilon(\kappa)g(\pi^{-1}(\kappa)) > g(\kappa)$ darstellt.

□

Da 2.26 die Repräsentanten in lexikographisch aufsteigender Reihenfolge konstruiert, wirkt sich dies Lemma folgendermaßen aus:

2.29 Folgerung *Wurde für $f \in \{0,1\}^{\Sigma}$ ein $(\epsilon, \pi) \in C(\gamma, \beta)$ gefunden mit $\hat{f} < f$, so kann erst ein Repräsentant g mit $g(\Theta) > f(\Theta)$ wieder kanonisch sein.*

2.30 Beispiel Bei der Verbindung aus 2.17.iii) erweist sich beim Test als nichtkanonisch der Repräsentant $f = (0,0,0,0,0,0,1,0,0,0,0,0,0)$, da er unter der Abbildung

$$((e,e,e,e,e,e,e,e,i,i,e,e),(23)(45)(68)(79)(10,12)(11,13))$$

zu $\hat{f} = (0,0,0,0,0,0,0,0,1,0,0,0,0) < f$ wird. Damit ist $\theta = 8$, und wir erhalten $\Theta = \max\{7,9,8\} = 9$. Daher kann der Algorithmus die Konstruktion sofort mit $f = (0,0,0,0,0,0,1,0,1,0,0,0,0)$ fortsetzen und so 16 Tests einsparen.
Mit Hilfe von 2.28 reduziert sich bei dieser Struktur insgesamt die Zahl der Tests von 8192 auf 5307, also um etwa 35,2%.

Noch haben wir nicht erklärt, welche tatsächlichen Konfigurationen den Werten 0 und 1 zugeordnet werden sollen. Wir wollen uns dabei an [22] halten und vereinbaren:

2.31 Konvention Die Werte 0 und 1 der konstruierten Stereoisomere entsprechen folgenden Konfigurationen der Liganden:

Die Darstellung mit Hilfe von 0,1-Vektoren hat den zusätzlichen Vorteil, daß so die Repräsentanten als (u.U. lange) Ganzzahlen in ihrer Binärdarstellung aufgefaßt werden können. Auf diese Weise gelingt eine sehr kompakte Speicherung der Stereoisomere, z.B. in Datenbanken.

2.6 Zusammenfassung und Diskussion

Fassen wir alle in diesem Kapitel beschriebenen Funktionen zusammen, so erhalten wir eine Übersicht über den Nourseschen Algorithmus:

2.32 Algorithmus (Noursescher Konstruktionsalgorithmus)

i) Lies die Eingabestruktur ein.

ii) Bestimme die Atome des Moleküls, die Teil eines aromatischen Systems sind.

iii) Suche potentielle Stereozentren gemäß 2.2. Stop, falls keine gefunden werden.

iv) Wandle die Doppelbindungen zwischen potentiellen Stereozentren nach 2.4 um.

v) Rufe die Funktion zur Berechnung der Automorphismengruppe auf.

vi) Untersuche die Permutationen der Liganden der Stereozentren. Verwende 2.21, um die Stereozentren mit ungerade permutierenden Liganden zu verwerfen. Stop, falls danach keine Stereozentren mehr vorhanden sind.

vii) Bestimme die Inversionen in der Konfigurationssymmetriegruppe mit 2.14.

viii) Berechne durch Aufruf von 2.26 die kanonischen Bahnenrepräsentanten der Operation der Konfigurationssymmetriegruppe auf der Menge der Abbildungen von der Menge der Stereozentren nach $\{0,1\}$. Vermeide dabei überflüssige Tests[5] durch Zuhilfenahme von 2.28.

ix) Gib alle Stereoisomere aus.

◇

Nach dieser endgültigen Formulierung ist es nun an der Zeit, die Möglichkeiten und Grenzen dieses Verfahrens unter die Lupe zu nehmen:

2.33 Diskussion des Algorithmus
Vorteile:

- Es werden alle Stereoisomere *vollständig und redundanzfrei* berechnet, die durch Effekte an drei- und vierfach koordinierten Atomen hervorgerufen werden, die mit einem geometrischen Modell beschreibbar sind (vgl. 1.1). Dies ist auch in der Praxis von Bedeutung, da ein Großteil der Stereochemie organischer Verbindungen darauf beruht.

- Ausgangspunkt für die Berechnung ist nur die Bindungsmatrix. Keinerlei räumliche Informationen über das Molekül werden benötigt.

- Die Ausgabe ist sehr kompakt und daher gut zur Weiterverarbeitung in Datenbanken o.ä. geeignet.

[5] Dieser Punkt ist, wie bereits bemerkt, in der Originalarbeit [22] nicht so effizient gelöst.

- Die Stereoisomerie wird sowohl lokal durch Betrachtung der Umgebung der Stereozentren als auch global mit Hilfe der Konfigurationssymmetriegruppe untersucht. Eine rein lokale Untersuchung gerichte bei hochsymmetrischen Strukturen äußerst kompliziert, während die Betrachtung der Symmetrie als globale Eigenschaft nicht immer zu eindeutigen Zuordnungen führt, da viele Moleküle nur geringe Symmetrie aufweisen.

Nachteile:

- Bei der Verwendung der topologischen Automorphismengruppe wird nicht in Betracht gezogen, daß darin einige Permutationen existieren können, die aufgrund der realen Bindungsabstände und -winkel keine Automorphismen sind.

- Die Modellierung des Moleküls als Multigraphen macht es sehr kompliziert, die Stereochemie anorganischer Substanzen, die durch ionische oder Metallbindung gebunden sind, zu beschreiben.

- Die Stereoisomerie wird durch das Stützen auf den Begriff des Stereozentrums als von lokalen Verhältnissen ausgehend angesehen. Daher werden alle Formen der Stereoisomerie, die nicht auf lokalen Zentren beruhen, wie Atropisomerie oder topologische Isomerie, nicht nur nicht erfaßt; sie können vielmehr nicht einmal beschrieben werden.

- Nur höchstens vierfach koordinierte Atome können Stereozentren sein.

Es gab mehrere Ansätze, um diese Nachteile zu beheben.
In [33] wird vorgeschlagen, anstelle der topologischen Automorphismengruppe die Punktgruppe, also die räumliche Symmetriegruppe zu verwenden. Diese kann entweder experimentell oder rechnerisch ermittelt werden; beides ist mit hohem Aufwand verbunden, der zum einen die Effizienz und Geradlinigkeit des Algorithmus zunichte macht und zum anderen einen nur für bestimmte Anwender interessanten Fall betrachtet, mithin also die Universalität der Methoden der mathematischen Chemie in Frage stellt.
Die Berücksichtigung weiterer Substanzklassen über die organische Chemie hinaus ist vom mathematischen Standpunkt her wünschenswert. In der Form 2.32 ist eine Anwendung dort in der Tat nur schwer möglich. Durch eine geeignete Umformulierung zusammen mit einer Erweiterung werden allerdings viele Probleme faßbar; näheres siehe [31] und [32].
Zur Behandlung topologischer Isomerie u.ä. ist das hier verwendete mathematische Modell weitgehend untauglich. In [29], [30] und weiteren Arbeiten wird versucht, diese Form der Stereoisomerie durch die Betrachtung globaler Symmetrien in den Griff zu bekommen. Ob durch diese Vorgehensweise andere Stereoeffekte vernachlässigt werden, ist im gegenwärtigen Forschungsstand noch nicht abzusehen.

2.34 Bemerkung Ein weiterer Nachteil ist nicht algorithmischer sondern chemischer Natur. Denn nicht alle Stereoisomere, die mit 2.32 erzeugt werden, sind auch chemisch sinnvoll, d.h. physikalisch stabil[6]. Insbesondere handelt es sich dabei um *trans*-Konfigurationen von Doppelbindungen, die Teil von Ringen mit weniger als acht Atomen sind, um polyzyklische Systeme, die einen anellierten kleinen Ring in *trans*-Stellung enthalten, und um invertierte Doppelringe. Eine genauere Diskussion der durch ein Programm ausschließbaren Stereoisomere findet sich in [23]. In einem Programmsystem der Praxis könnte diese Plausibilitätslücke so geschlossen werden, daß nach der Generierung der Stereoisomere mit 2.32 ein optionaler Filter die nach den oben beschriebenen Kriterien unrealistischen Resultate herausnimmt.

2.7 Abzählen der Stereoisomere

Wie bereits bei der Definition erwähnt, kann die Konfigurationssymmetriegruppe neben der Generierung auch auf die Abzählung angewandt werden. Vom theoretischen Blickpunkt aus kommt dem Abzählen von Bahnen von Gruppenoperationen – also ohne Transversalen zu konstruieren – einige Bedeutung zu. Um das grundlegende Lemma 1.8 anwenden zu können, müssen wir noch einen Begriff einführen (nach [15]):

2.35 Definition *Sei $G \leq S_{\underline{n}}$ und $(\psi, \pi) \in H \wr_{\underline{n}} G$, wobei π in* Standardzyklennotation *gegeben sei, d.h. $\pi = \prod_{\kappa}^{c(\pi)}(j_\kappa \dots \pi^{l_\kappa - 1} j_\kappa)$ mit $j_\kappa \leq \pi^m j_\kappa$ für alle $m \in \mathbb{N}$ und $j_\kappa < j_{\kappa+1}$ für $\kappa < c(\pi)$. Dann heißt*

$$h_\kappa(\psi, \pi) := \psi(j_\kappa)\psi(\pi^{-1}j_\kappa) \cdot \ldots \cdot \psi(\pi^{-l_\kappa+1}j_\kappa)$$

das κ-te Zyklenprodukt von (ψ, π). •

Zu beachten ist, daß $h_\kappa(\psi, \pi) \in H$.

2.36 Lemma *Es operieren ${}_G\underline{n}$ und ${}_H Y$.*

i) Sei $f \in Y^{\underline{n}}$ und $(\psi, \pi) \in H \wr_{\underline{n}} G$ mit π in Standardzyklennotation. Dann bleibt f genau dann fest unter (ψ, π), wenn gilt:

(a) $f(j_\kappa) \in Y_{h_\kappa(\psi,\pi)}$

(b) Für die anderen Werte von f gilt:

$$f(j_\kappa) = \psi(j_\kappa)f(\pi^{-1}j_\kappa) = \psi(j_\kappa)\psi(\pi^{-1}j_\kappa)f(\pi^{-2}j_\kappa) = \dots$$

ii) Für die Anzahl der Bahnen gilt:

$$|H \wr_{\underline{n}} G \backslash\backslash Y^{\underline{n}}| = \frac{1}{|H \wr_{\underline{n}} G|} \sum_{(\psi,\pi) \in H \wr G} \prod_{\kappa=1}^{c(\pi)} |Y_{h_\kappa(\psi,\pi)}|$$

[6] Die Annahme einer physikalischen Instabilität ist allerdings nur mit Vorbehalten hinzunehmen. In der Vergangenheit stellten sich mehrfach Substanzen plötzlich als synthetisierbar heraus, die vorher für völlig instabil gehalten worden waren.

Zum Beweis s. [15].
Wir wollen diese Formel nun auf die Abzählung von Stereoisomeren anwenden. Ist p_2 die Projektion auf das zweite Element des Kranzproduktes, d.h. $p_2((\psi,\pi)) = \pi$, so können wir eine Transversale der Konjugiertenklassen von $p_2(C(\gamma,\beta))$ wählen; diese wollen wir $\mathcal{C}$ nennen. Falls $\overline{\pi}$ in der i-ten Konjugiertenklasse ist, sei

$$n_{i\kappa} := |\{k \mid \epsilon_\pi(k) = (01),\ k = \pi^\tau j_\kappa,\ \tau = 0,\ldots,l_\kappa - 1\}|,$$

die Zahl der Inversionen, die zum κ-ten Zykel des Repräsentanten dieser Klasse gehören. Damit können wir formulieren:

2.37 Satz *Die Anzahl der Stereoisomere zu einem Molekül mit s Stereozentren und Konfigurationssymmetriegruppe $C(\gamma,\beta)$ ist*

$$\frac{1}{|C(\gamma,\beta)|}\sum_{\overline{\pi}\in\mathcal{C}}|C^S(\overline{\pi})|\cdot 2^{c(\overline{\pi})}\prod_{\kappa=1}^{c(\overline{\pi})}((n_{i\kappa}+1)\bmod 2)$$

Beweis: Durch Anwendung von 2.36 folgt:

$$\begin{aligned}
|C(\gamma,\beta)\backslash\backslash\{0,1\}^{\underline{s}}| &= \frac{1}{|C(\gamma,\beta)|}\sum_{(\epsilon_\pi,\overline{\pi})\in C(\gamma,\beta)}\prod_{\kappa=1}^{c(\overline{\pi})}|\{0,1\}_{h_\kappa(\epsilon_\pi,\overline{\pi})}| \\
&= \frac{1}{|C(\gamma,\beta)|}\sum_{\overline{\pi}\in\mathcal{C}}|C^S(\overline{\pi})|\prod_{\kappa=1}^{c(\overline{\pi})}|\{0,1\}_{h_\kappa(\epsilon_\pi,\overline{\pi})}| \\
&= \frac{1}{|C(\gamma,\beta)|}\sum_{\overline{\pi}\in\mathcal{C}}|C^S(\overline{\pi})|\prod_{\kappa=1}^{c(\overline{\pi})}2\cdot((n_{i\kappa}+1)\bmod 2)
\end{aligned}$$

da gilt:

$$\begin{aligned}
h_\kappa((\epsilon_\pi,\overline{\pi}) &= \epsilon_\pi(j_\kappa)c_\pi(\overline{\pi}^{-1}j_\kappa)\ldots\epsilon_\pi(\overline{\pi}^{-l_\kappa+1}j_\kappa) \\
&= \begin{cases} 1 & \text{falls die Anzahl der Inversionen gerade} \\ (01) & \text{sonst} \end{cases}
\end{aligned}$$

Somit ist $|\{0,1\}_{h_\kappa(\epsilon_\pi,\overline{\pi})}| = 2\cdot((n_{i\kappa}+1)\bmod 2)$. Zieht man den Faktor 2 aus dem Produkt, folgt die Behauptung. □

2.38 Beispiel Die Konfigurationssymmetriegruppe des Tetramethylcyclobutan (s. 2.11.i) hat die Ordnung 8 und die Konjugiertenklassen

$$\{1\},\{(13),(24)\},\{(12)(34),(13)(24),(14)(23)\},\{(1234),(1432)\}.$$

Damit ist die Anzahl der Stereoisomere

$$\frac{1}{8}(1\cdot 2^4\cdot 1\cdot 1\cdot 1\cdot 1+2\cdot 2^3\cdot 1\cdot 0\cdot 0+3\cdot 2^2\cdot 1\cdot 1+2\cdot 2^1\cdot 1)=\frac{1}{8}(16+0+12+4)=4,$$

wie wir schon bei der Generierung in 2.27 festgestellt haben.

Für den Mathematiker könnte an dieser Stelle die Arbeit eigentlich beendet sein. Doch der Anwender, der Chemiker, ist mit der Ausgabe der Stereoisomere in Form von 0,1-Vektoren nicht zufrieden. Interessant ist vielmehr die tatsächliche räumliche Anordnung. Nur so können die in diesem Kapitel errechneten theoretischen Resultate auch in der Praxis Bedeutung erlangen. Die folgenden Abschnitte werden diese Fragen behandeln.

3 Visualisierung der Ergebnisse

3.1 Identifikation der Konfiguration einer Referenzplazierung

Wir wollen nun annehmen, wir hätten eine räumliche Plazierung des zu betrachtenden Moleküls gegeben. Dessen Konfiguration muß bestimmt werden, um anschließend Plazierungen aller generierten Stereoisomere errechnen zu können.
Es gibt mehrere Methoden, um eine Referenzplazierung zu erhalten. Neben den experimentellen basieren die wichtigsten mathematischen auf Distanzgeometrie (nach [6]), auf Konformationsanalyse (etwa [13]) oder auf Energiebetrachtungen, etwa nach [1]; Der letzte Weg wurde auch hier gewählt.

Eine der umfassendsten Arbeit der letzten Jahre auf dem Gebiet der mathematischen Stereochemie ist die Dissertation von L. Zlatina ([33]). Darin werden neben zweidimensionaler Plazierung, der Distanzgeometrie von Crippen und dem Algorithmus von Nourse auch die Bestimmung der Automorphismengruppe einer dreidimensionalen Plazierung, die Anpassung einer räumlichen Anordnung an gegebene Symmetrieelemente und die Konstruktion der Koordinaten aller Stereoisomere aus einer Referenzplazierung behandelt. Letzteres bildet auch die Grundlage der hier vorgestellten Methoden.

Der erste Schritt besteht, wie erwähnt, in der Analyse der gegebenen Koordinaten. Es muß ermittelt werden, ob die Konfiguration eines Stereozentrums dem Code 0 oder 1 entspricht.
Blickt man vom vierten Liganden auf die Ebene der ersten drei, so sind diese gemäß 2.31 beim Code 0 entgegen dem Uhrzeigersinn angeordnet. Sind sie dagegen im Uhrzeigersinn, so entspricht dies dem Code 1. Daher gilt es nun, eben diese Drehrichtung zu bestimmen.

Wir betrachten zunächst ein vierwertiges Stereozentrum. Die Liganden seien darum räumlich angeordnet an den Punkten X_1, X_2, X_3 und X_4. Die Ebene, die von X_1, X_2 und X_3 aufgespannt wird, heiße A. Sei X ein weiterer Punkt oberhalb der Ebene, dessen Orthogonalprojektion in A, der Punkt X_0, innerhalb des Dreiecks $X_1X_2X_3$ liege (s. Abb. 2). Ferner sei $\alpha := \angle X_1X_0X_2$ und $\vec{n} := \overrightarrow{X_1X_2} \times \overrightarrow{X_1X_3}$. Dann gilt:

3.1 Lemma *Bilden X_1, X_2 und X_3 ein gleichseitiges Dreieck, und ist X_0 der Mittelpunkt, so ist das Bild der Drehung von X_1 um die Achse XX_0 mit dem Winkel α genau dann X_2, wenn gilt: $\vec{n} \cdot \overrightarrow{XX_1} \geq 0$.*

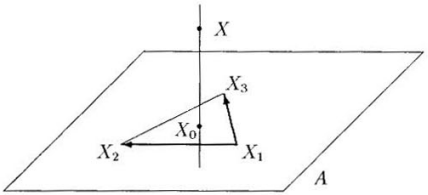


Abbildung 2: Ligandenanordnung in der Ebene

Beweis: Zuerst von links nach rechts: Ist das Bild von X_1 der Punkt X_2, so sind X_1, X_2 und X_3 von X aus gesehen im Uhrzeigersinn angeordnet. Darüber hinaus bilden $\overrightarrow{X_1X_2}$, $\overrightarrow{X_1X_3}$ und $\vec{n}$ ein Rechtssystem, d.h. die Drehung von $\overrightarrow{X_1X_2}$ in die Richtung von $\overrightarrow{X_1X_3}$ beschreibt eine Rechtsschraube. Daher bilden $\vec{n}$ und $\overrightarrow{XX_1}$ einen spitzen Winkel, woraus $\vec{n} \cdot \overrightarrow{XX_1} \geq 0$ folgt.
Von rechts nach links: Falls $\vec{n} \cdot \overrightarrow{XX_1} \geq 0$ gilt, dann stehen $\vec{n}$ und $\overrightarrow{XX_1}$ im spitzen Winkel zueinander. Angenommen, das Bild von X_1 wäre X_3. Dann würden $\overrightarrow{X_1X_2}$, $\overrightarrow{X_1X_3}$ und $\vec{n}$ ein Linkssystem bilden, was im Widerspruch zu den Eigenschaften des Vektorproduktes steht.

□

In den folgenden Algorithmen werden Unterfunktionen für Vektoroperationen aufgerufen. Da es sich nur um Anwendung von Rechenregeln handelt, wird auf eine detaillierte Darstellung verzichtet (Genaueres etwa in [2]).

3.2 Algorithmus (Identifikation der Konfiguration eines Zentrums) Seien die Positionen der Nachbarn die Punkte X_1, X_2, X_3, falls das Zentrum dreiwertig ist; sonst sei zusätzlich X_4 die Position des weiteren Nachbarn. Ferner sei X die Position des Zentrums. Berechne $\vec{n} := \overrightarrow{X_1X_2} \times \overrightarrow{X_1X_3}$. Ermittle dann $\lambda := \vec{n} \cdot (\vec{X_1} - \vec{X})$. Ist $\lambda < 0$, so gib 1 zurück, sonst gib 0 zurück. ◇

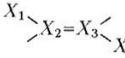
3.3 Algorithmus (Identifikation der Konfiguration einer Doppelbindung)

i) Durchlaufe alle doppelt gebundenen Stereozentren. Suche zu jedem den Bindungspartner nur unter den Atomen höherer Atomnummer.

ii) Bestimme zu beiden unter ihren restlichen Nachbarn jeweils den mit der höchsten Atomnummer.

iii) Numeriere die Punkte wie folgt[7]:

[7] Die Position von X_4 hängt natürlich von der Konfiguration ab. Die Zeichnung stellt daher nur eine Möglichkeit dar; im anderen Fall erfolge die Numerierung analog.



iv) Berechne $\vec{n}_\perp := \overrightarrow{X_2X_3} \times \overrightarrow{X_2X_1}$ und $\vec{n}_\parallel := \vec{n}_\perp \times \overrightarrow{X_2X_3}$.

v) Haben $\vec{n}_\parallel \cdot \overrightarrow{X_2X_1}$ und $\vec{n}_\parallel \cdot \overrightarrow{X_2X_4}$ dasselbe Vorzeichen, so gib 1 zurück, sonst 0.

◇

Aufgrund der Hybridisierung der doppelt gebundenen Atome liegen alle vier Punkte in einer Ebene (zumindest näherungsweise). Mit $\vec{n}_\parallel$ wird dann ein Vektor bestimmt, der ebenfalls in dieser Ebene liegt und senkrecht auf der Verbindung der Zentren steht. Sind die Liganden auf der gleichen Seite, so ergibt das Skalarprodukt ihrer Verbindungsvektoren mit einem Zentrum mit $\vec{n}_\parallel$ dasselbe Vorzeichen. Der Algorithmus liefert also 1 für die *cis*- und 0 für die *trans*-Konfiguration.

3.2 Erzeugung der räumlichen Stereoisomere

Im letzten Abschnitt haben wir gesehen, wie die Konfiguration einer Referenzplazierung bestimmt werden kann. Nun wollen wir die räumliche Gestalt der mit 2.32 berechneten Stereoisomere anhand der Referenzplazierung errechnen.

Die verschiedenen Stereoisomere unterscheiden sich von der Referenzplazierung immer nur in den Konfigurationen einzelner Stereozentren. Beim Übergang muß also an den differierenden Zentren eine Spiegelung vorgenommen werden. Wichtig ist dabei die Klärung der Frage, welche anderen Zentren bei dieser Abbildung gespiegelt werden und welche fest bleiben.
Die Spiegelung selbst geht so vor sich, daß alle Atome in zwei Liganden des Stereozentrums an der Ebene gespiegelt werden, die vom Zentrum und (bei vierfach koordinierten Zentren) den anderen beiden Nachbarn bzw. (bei dreiwertigen Zentren) einem anderen Nachbarn und dem Mittelpunkt der Verbindungsgerade der beiden ersten aufgespannt wird. Da bei dieser Abbildung die Konfigurationen aller Atome in den zu spiegelnden Liganden betroffen sind, ist es sicher von Vorteil, die Liganden als fest zu wählen, welche die meisten Stereozentren enthalten. Die Auswahl der Liganden erfolgt mit derselben Funktion, die bereits in 2.21.ii) zur Bestimmung der Anzahl der Stereozentren in einem Liganden verwendet wurde.

Entscheidend für die Durchführung der Spiegelung ist das Vorkommen von Ringen. Je nachdem, welche Position das Stereozentrum in der zyklischen Struktur hat, müssen folgende Fälle unterschieden werden (wieder nach [33]):

3.4 Spiegelung an Stereozentren

i) *Das Zentrum liegt nicht in einem Ring.* Dies ist der einfachste Fall. Die Liganden, welche die kleinste Anzahl an Stereozentren enthalten, werden gespiegelt.

ii) *Das Zentrum liegt in genau einem Ring.* Dann sind zwei Liganden Teil des Rings, und der oder die anderen nicht. Ist die Anzahl der Stereozentren in den Ring-Liganden größer als die in den anderen, so bleiben die im Ring liegenden Nachbarn fest; im umgekehrten Fall werden sie gespiegelt.

iii) *Das Zentrum liegt in zwei Ringen einer Spiro-Verbindung* (s. 1.5.iii). Hier wählt man die Liganden des Ringes als fest, der die kleinere Anzahl an Stereozentren enthält.

iv) *Das Zentrum liegt in mehr als einem Ring; die Ringe haben mehr als ein Atom gemeinsam.* In diesem Fall ist es nicht möglich, die Stereoisomere auseinander geometrisch zu konstruieren. Um doch noch ein Bild der Isomere zu erhalten, kann eine Hilfskonstruktion eine Näherung schaffen. Dabei sind zwei Fälle zu unterscheiden:

   (a) Es gibt im Molekül nur zwei derartige Stereozentren. Dann werden zwei Ring-Liganden und ein außerhalb des Doppelrings liegender an der Ebene gespiegelt, die senkrecht zu der vom dritten Liganden, dem Zentrum und dem anderen Stereozentrum aufgespannten Ebene steht.

   (b) Es gibt mehrere derartige Stereozentren. Hier kann wieder im obigen Sinne gespiegelt werden. Fest bleiben die Atome, die in einem Ring nicht am Anfang oder Ende vorkommen (also Brückenatome), oder solche, die nicht Teil eines Ringes sind.

Doch bevor wir die eigentlichen Konstruktionen diskutieren können, müssen wir uns zunächst mit der Erkennung der in 3.4 unterschiedenen Fälle beschäftigen. Diese erfolgt mit einfachen Such-Algorithmen, die Doppelringe mit einem oderen mehreren gemeinsamen Atomen erkennen.

3.5 Algorithmus (Einordnung eines Stereozentrums in 3.4)

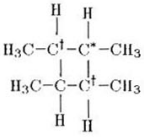
i) Durchlaufe alle Liganden des Zentrums und zähle die anderen Stereozentren in diesen Liganden sowie die Ringe, in denen das Zentrum und dieser Ligand liegen. Bestimme die Liganden, welche die meisten Stereozentren enthalten.

ii) Ist das Stereozentrum ein asymmetrisches Atom und liegt in mindestens einem Ring, so untersuche die Ringstruktur. Unterscheide nach der Zahl der Ringe, in denen das Zentrum liegt.

   - *1 Ring:* Enthalten die Liganden im Ring mehr Stereozentren als die anderen, so markiere diese als fix; sonst markiere die anderen als fix.
   - *2 Ringe:* Untersuche, ob es sich um ein Spiran handelt. Trifft dies nicht zu, so fahre in i) mit dem nächsten Zentrum fort. Sonst: Liegen im ersten Ring mehr Stereozentren als im zweiten, so wähle den ersten; sonst den zweiten. Markiere die Nachbarn des Zentrums, die im gewählten Ring liegen, als fix.

◇

3.6 Beispiel Bei der folgenden Verbindung sind die untersuchten Stereozentren mit * und die bei der Spiegelung fest bleibenden Nachbarn mit † markiert. Es handelt sich hierbei um ein Molekül mit einem Ring mit Stereozentren:



Nun wollen wir noch die geometrischen Aspekte betrachten.

3.7 Definition *Zu drei Punkten $X_0, X_1, X_2 \in \mathbb{R}^3$, die nicht auf einer Geraden liegen, heißt $\mathcal{E} = \{P \in \mathbb{R}^3 \mid \exists \lambda, \mu \in \mathbb{R} : P = X_0 + \lambda(X_1 - X_0) + \mu(X_2 - X_0)\}$* Ebene. •

Bei uns ist i.a. der Punkt[8] X_0 die Lage des Stereozentrums und X_1, X_2 die Positionen zweier Liganden. Die Forderung, daß die Punkte nicht auf einer Geraden liegen dürfen, die Vektoren also linear unabhängig sein müssen, ist dann aus chemisch-physikalischen Gründen stets erfüllt.

3.8 Hilfssatz und Definition *Sei $\mathcal{E} = \{P \in \mathbb{R}^3 \mid \exists \lambda, \mu \in \mathbb{R} : P = X_0 + \lambda\vec{v} + \mu\vec{w}\}$ eine Ebene und $\vec{n} := \vec{v} \times \vec{w}$ das Vektorprodukt der aufspannenden Vektoren. Dann heißt $\mathcal{E} = \{P \in \mathbb{R}^3 \mid \vec{n} \cdot P = \vec{n} \cdot X_0\}$ die* Normalenform *der Ebene $\mathcal{E}$.*

3.9 Lemma *Für den Abstand $d := \inf\{|A - P| \mid P \in \mathcal{E}\}$ eines Punktes $A \in \mathbb{R}^3$ von der Ebene $\mathcal{E} = \{P \in \mathbb{R}^3 \mid \vec{n} \cdot P = \vec{n} \cdot X_0\}$ gilt: $d = \frac{1}{|\vec{n}|} |\vec{n} \cdot A - \vec{n} \cdot X_0|$.*

Beweis: Fällt man das Lot von A auf $\mathcal{E}$, so erhält man den Punkt

$$A_0 := A - \frac{\vec{n} \cdot A - \vec{n} \cdot X_0}{\vec{n}^2} \vec{n},$$

denn $\vec{n} \cdot A_0 = \vec{n} \cdot A - \vec{n} \cdot A + \vec{n} \cdot X_0 = \vec{n} \cdot X_0$, also $A_0 \in \mathcal{E}$ und $\overrightarrow{AA_0}$ ist parallel zu $\vec{n}$. Für alle $P \in \mathcal{E}$ gilt dann: $|P - A|^2 = |(P - A_0) + (A_0 - A)|^2 \geq |A_0 - A|^2$, da $\vec{n} \cdot P = \vec{n} \cdot A_0$. Somit erhalten wir $d = |A_0 - A| = \frac{1}{|\vec{n}|} |\vec{n} \cdot A - \vec{n} \cdot X_0|$. □

Fällt man nicht nur das Lot, sondern geht noch darüber hinaus, so ergibt sich:

3.10 Folgerung *Sei $\mathcal{E}$ eine Ebene mit Normalenvektor $\vec{n}$ und $A \in \mathbb{R}^3$ ein Punkt mit Abstand d von der Ebene. Dann ist das Bild der Spiegelung von A an der Ebene der Punkt*

$$A' = A - 2\frac{\vec{n} \cdot A - \vec{n} \cdot X_0}{\vec{n}^2} \vec{n}.$$

Damit haben wir alle Hilfsmittel für die Algorithmen zusammen und können mit den Abbildungen beginnen:

[8] Vektoren werden in ihrer Koordinatendarstellung verstanden, Punkte als Vektoren vom Ursprung zum Punkt. Damit sind beide Elemente des $\mathbb{R}^3$. Somit ist auch das Skalarprodukt zwischen Vektor und Punkt definiert.

3.11 Algorithmus (Invertierung eines Stereozentrums)

i) Sei X der Koordinatenvektor des Stereozentrums und X_1, X_2, X_3 und (evtl.) X_4 die Koordinatenvektoren der Nachbarn. Ist das Stereozentrum vierwertig, gehe zu iii). Ist es nicht Teil eines Rings oder Teil eines Rings, der abgebildet werden soll, so gehe zu ii). Sonst gehe zu iii).

ii) Sei o.E. X_1 der Vektor des Nachbarn, der fest bleibt. Berechne $\vec{n} := (\frac{1}{2}(X_2 + X_3) - X) \times (X_1 - X)$. Gehe zu iv).

iii) Seien o.E. X_1 und X_2 die Vektoren der Nachbarn, die fest bleiben. Berechne $\vec{n} := (X_1 - X) \times (X_2 - X)$.

iv) Berechne $n := \vec{n}^2$ und $t := \vec{n} \cdot X$.

v) Durchlaufe alle Nachbarn des Stereozentrums, die nicht fest bleiben und die nicht bereits gespiegelt wurden. Spiegele sie mit Hilfe der folgenden rekursiven Funktion. Diese starte mit dem Stereozentrum als Vorgänger und dem im Durchlauf erreichten Nachbarn als aktuellem Atom in der Rekursionstiefe 0.

   (a) Ist das aktuelle Atom gleich dem Stereozentrum, so markiere den Vorgänger als gespiegelt und springe zurück. (D.h. das Ende eines Rings wurde erreicht.)

   (b) Ist das aktuelle Atom selbst ein Stereozentrum, so ändere dessen Konfigurationscode auf 1, falls dieser 0 war, und auf 0, falls er 1 war. Notiere, daß eine Änderung eines Codes vorgenommen wurde.

   (c) Das aktuelle Atom habe den Koordinatenvektor P. Ist $\vec{n} \cdot P = t$, so gehe zu (d). (Das Atom liegt dann nämlich in der Spiegelebene.) Sonst: Berechne $\lambda := \frac{2}{n}(t - \vec{n} \cdot P)$. Setze $P \leftarrow P + \lambda\vec{n}$.

   (d) Durchlaufe alle Nachbarn des aktuellen Atoms außer dem Vorgänger. War einer bereits in einer höheren Rekursionsebene berücksichtigt, so springe zurück. Anderenfalls rufe die Funktion mit dem aktuellen Atom als Vorgänger und dem Nachbarn als aktuell eine Ebene tiefer auf.

   (e) Springe zurück.

◇

Diese Art des Durchlaufs durch die Liganden eines Zentrums ist ja bereits u.a. aus 2.21 bekannt.

Auch die Änderung der Konfiguration an Doppelbindungen erfolgt mittels Spiegelungen. Es sind indessen nicht die Fälle von 3.4 zu unterscheiden, da es keine Wahlmöglichkeiten gibt, welche Liganden abgebildet werden und welche fest bleiben sollen; der eine oder die beiden einfach gebundenen Liganden des zweiten Zentrums einer Doppelbindung werden, falls eine Konfigurationsänderung nötig ist, stets gespiegelt.
Die geometrische Situation wurde bereits in 3.3 diskutiert.

3.12 Algorithmus (Invertierung einer Doppelbindung)

i) Seien X_2 und X_3 die Koordinatenvektoren der beiden Zentren. Bestimme den Nachbarn des ersten Zentrums mit der größten Atomnummer, der ungleich dem zweiten Zentrum ist. Sein Koordinatenvektor sei X_1.

ii) Berechne $\vec{n}_\perp := (X_3 - X_2) \times (X_1 - X_2)$ und $\vec{n} := \vec{n}_\perp \times (X_3 - X_2)$. Setze $n := \vec{n}^2$ und $t := \vec{n} \cdot X_3$.

iii) Durchlaufe alle Nachbarn des zweiten Zentrums außer dem ersten Zentrum und evtl. dem bereits gespiegelten Nachbarn. Spiegele sie mit Hilfe der rekursiven Funktion aus 3.11.v). Diese starte mit dem zweiten Stereozentrum als Vorgänger und Zentralatom und dem im Durchlauf erreichten Nachbarn als aktuellem Atom in der Rekursionstiefe 0.

◇

Nachdem alle Abbildungen erklärt sind, können die Koordinatendarstellungen aller Stereoisomere ermittelt werden.

3.13 Algorithmus (Konstruktion der Koordinatendarstellungen aller Stereoisomere)

i) Stelle mit 3.5 die zu spiegelnden Liganden der Stereozentren fest.

ii) Durchlaufe die Liste der Codes der Stereoisomere, die mit 2.32 erzeugt wurde.

iii) Setze den Vektor der Konfigurationscodes (kurz: den Konfigurationsvektor) auf die mit 3.2 und 3.3 ermittelten Werte der Referenzplazierung. Initialisiere alle Koordinatenvektoren mit der Referenzplazierung. Beginne mit dem ersten Stereozentrum.

iv) Stimmt der Konfigurationsvektor am aktuellen Zentrum mit dem Wert des gegenwärtigen Stereoisomers aus der Liste überein, so gehe zu v). Sonst: Invertiere das Zentrum mit Hilfe von 3.11. Kehre den Eintrag des Konfigurationsvektors um. Wurden bei der Spiegelung Konfigurationen anderer Zentren verändert, so starte wieder mit dem ersten Stereozentrum; sonst betrachte das nächste Zentrum.

v) Solange noch nicht alle Stereozentren, die auf einem asymmetrischen Atom beruhen, betrachtet sind, gehe zu iv). Sonst betrachte das erste Zentrum in einer Doppelbindung.

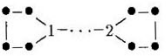
vi) Suche unter den Nachbarn des Stereozentrums das andere, mit ihm doppelt gebundene Zentrum. Stimmt dessen Konfiguration mit dem Listenwert überein, so gehe zu vii). Sonst invertiere die Liganden mittels 3.12. Kehre den Eintrag des Konfigurationsvektors des zweiten Zentrums um. Wurden bei der Spiegelung Konfigurationen anderer Zentren verändert, so starte wieder mit dem ersten asymmetrischen Stereozentrum und gehe zu iv); sonst betrachte das nächste Zentrum.

vii) Solange noch nicht alle Stereozentren, die auf einer Doppelbindung beruhen, betrachtet sind, gehe zu vi).

viii) Speichere die konstruierten Koordinaten und gehe zum nächsten Stereoisomer in der Liste. Ist dort noch eines vorhanden, so gehe zu iii).

◇

Kritisch für den Algorithmus ist das Ändern von Konfigurationen von Stereozentren, die in einem Ring liegen. Zur Untersuchung betrachten wir ein Molekül mit zwei Ringen



Im linken Ring mögen k Stereozentren liegen, im rechten m, und im Mittelteil zwischen den Atomen 1 und 2 mögen l Stereozentren liegen. Die Liganden des Zentrums 1, die im Ring liegen, enthalten $k-1$ Stereozentren, wenn man von 1 einmal absieht; in den anderen Liganden liegen $l+m$ Stereozentren. Ist $k-1 > m+l$, dann werden nach 3.4.ii) die Liganden außerhalb des Rings gespiegelt. Daraus folgt aber auch

$$k+m > k-1 > l+m > l-1,$$

also $k+m > l-1$. Dies bedeutet für das Zentrum 2, daß bei diesem es mehr Stereozentren im Ring gibt als außerhalb, und daß daher die Liganden, die nicht im Ring liegen, bei der Spiegelung festgehalten werden.
Damit haben wir gezeigt (Genaueres s. [33]):

3.14 Satz *Der iterative Algorithmus 3.13 bricht stets nach endlich vielen Schritten ab.*

Diese Aussage bedeutet insbesondere, daß keine Endlosschleifen auftreten können. Wird die Konfiguration eines Zentrums geändert, so ändern sich unter Umständen auch die Konfigurationen anderer Zentren. Um diese wieder zurückzusetzen, kann eine Spiegelung nötig sein, welche die Konfiguration des ersten Zentrums abermals ändert. So entsteht eine Endlosschleife. Mit 3.14 zeigt sich indessen, daß durch geeignete Wahl der zu spiegelnden Liganden dies vermieden werden kann.
Hätte man beispielsweise bei der dritten Struktur aus 3.6 die Nachbarn im Ring als fest gewählt, wären die Konfigurationen der Stereozentren im rechten Ring so verändert worden, daß ihre Wiederherstellung das erste Zentren erneut invertiert hätte; die Konfigurationsänderung eines Zentrums wird somit unmöglich.
Mit dem Ansatz 3.4 wird vermieden, daß bei korrespondierenden Stereozentren beidesmal die Liganden innerhalb von Ringen fix bleiben.

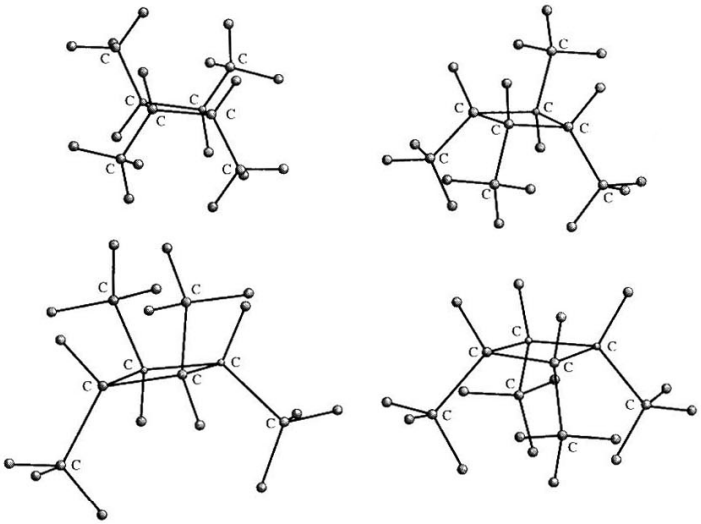


Abbildung 3: Stereoisomere des Tetramethylcyclobutan

3.15 Beispiele Wir wollen nun noch einige Ergebnisse des Algorithmus vorstellen. Sie wurden mit dem Modul *MOLVIEW* des Programmsystems **MOLGEN**$^+$ graphisch dargestellt. Sind Atome nicht beschriftet, handelt es sich dabei um Wasserstoff.

i) In 2.27 wurden die Stereoisomere des Tetramethylcyclobutan berechnet. Die unterschiedlichen räumlichen Anordnungen lassen sich am besten durch die Lage der Methyl-Substituenten im Vergleich zum Butan-Ring charakterisieren. Die Ring-Atome bilden näherungsweise eine Ebene, unter oder über der die Methyl-Gruppen liegen. Dafür gibt es vier verschiedene Möglichkeiten:

- Alle vier Gruppen liegen auf einer Seite des Rings.
- Eine Gruppe liegt entgegengesetzt den anderen.
- Je zwei benachbarte Gruppen liegen auf derselben Seite.
- Je zwei gegenüberliegende Gruppen liegen auf derselben Seite.

Alle diese Fälle wurden vom Algorithmus konstruiert; sie sind in 3 abgebildet.

ii) Das 1,5-Dichlor-1,2,4-Pentatrien

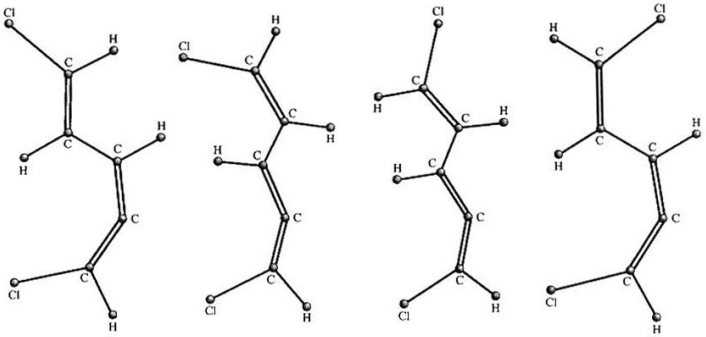
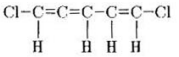


Abbildung 4: Diastereomere des 1,5-Dichlor-1,2,4-Pentatrien



kommt in vier verschiedenen Diastereomeren vor. Der Konstruktionsalgorithmus erzeugt diese korrekt aus der Referenzplazierung[9], wie Abbildung 4 zeigt.

Wie in 3.4.iv) erwähnt, lassen sich Diastereomere bestimmter Strukturen nicht auf geometrischem Wege ermitteln. In [33] wird daher vorgeschlagen, für diese eine eigene räumliche Plazierung unter Vorgabe der Konfiguration durchzuführen. Dies ist mit den hier zur Verfügung stehenden Mitteln nicht möglich, da der Energieoptimierungsalgorithmus mit einer zufälligen Plazierung der Atome startet und auch keine Vorgabe bestimmter Konfigurationen erlaubt. Um aber doch ein wenigstens ungefähres Bild von der räumlichen Struktur derartiger Verbindungen zu erhalten, wurden Modifikationen von 3.11 und 3.12 entwickelt. Die damit gewonnenen Resultate können dem Benutzer des Programms entweder einen groben Überblick vermitteln oder als Ausgangspunkt einer weiteren Energieoptimierung dienen.

Fixpunkte für die Inversion der Chiralitätszentren wurden bereits in 3.5 festgelegt. Nun kann die Spiegelung durchgeführt werden.

3.16 Algorithmus (Invertierung eines Stereozentrums in einem Doppelring)

[9] An dieser Stelle offenbart sich allerdings eine Schwäche des verwendeten Plazierungsprogramms. Das Molekül wird nämlich hier als völlig eben plaziert, mit einem Knick von 135° zwischen den beiden doppelt gebundenen Kohlenstoffen. In Wirklichkeit besteht ein solcher Knick nicht; die Ebenen der Liganden dieser beiden Atome stehen vielmehr aufeinander senkrecht (s. etwa [17]).

i) Sei X der Koordinatenvektor des Stereozentrums und X_1, X_2, X_3 und X_4 die Koordinatenvektoren der Nachbarn. Kommen im Molekül nur zwei Stereozentren in Doppelringen vor, so gehe zu ii); sonst gehe zu iii).

ii) In diesem Fall ist gem. Auswahl X_1 Fixpunkt und liegt in beiden Ringen. Setze

$$\vec{n} := ((X_1 - X) \times (X_2 - X)) \times (X_1 - X).$$

Lösche alle Fixpunktmarkierungen außer bei X_1. Gehe zu iv).

iii) Seien X_1 und X_2 die Vektoren der Nachbarn, die fest bleiben. Berechne $\vec{n} := (X_1 - X) \times (X_2 - X)$.

iv) Berechne $n := \vec{n}^2$ und $t := \vec{n} \cdot X$.

v) Durchlaufe alle Nachbarn des Stereozentrums, die nicht fest bleiben und die nicht bereits gespiegelt wurden. Spiegele sie mit Hilfe der folgenden rekursiven Funktion. Diese starte mit dem Stereozentrum als Vorgänger und dem im Durchlauf erreichten Nachbarn als aktuellem Atom in der Rekursionstiefe 0.

   (a) Ist das aktuelle Atom gleich dem Stereozentrum, so markiere den Vorgänger als gespiegelt und springe zurück. (D.h. das Ende eines Rings wurde erreicht.)

   (b) Kommt das Zentralatom in einem Doppelring vor, und liegt das aktuelle Atom in einem dieser Ringe, und ist die Rekursionstiefe bereits größer als die Hälfte der Länge des kleinsten dieser Ringe, oder ist das aktuelle Atom selbst ein Stereozentrum, so springe zurück.

   (c) Ist das aktuelle Atom selbst ein Stereozentrum, so ändere dessen Konfigurationscode auf 1, falls dieser 0 war, und auf 0, falls er 1 war. Notiere, daß eine Änderung eines Codes vorgenommen wurde.

   (d) Das aktuelle Atom habe den Koordinatenvektor P. Ist $\vec{n} \cdot P = t$, so gehe zu (e). Sonst: Berechne $\lambda := \frac{2}{n}(t - \vec{n} \cdot P)$. Setze $P \leftarrow P + \lambda\vec{n}$.

   (e) Durchlaufe alle Nachbarn des aktuellen Atoms außer dem Vorgänger. War einer bereits in einer höheren Rekursionsebene berücksichtigt, so springe zurück. Anderenfalls rufe die Funktion mit dem aktuellen Atom als Vorgänger und dem Nachbarn als aktuell eine Ebene tiefer auf.

   (f) Springe zurück.

◇

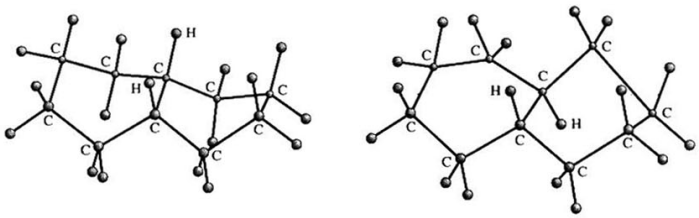
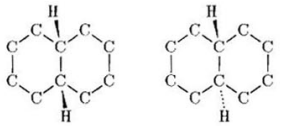


Abbildung 5: Konstruierte Stereoisomere des Decalin

3.17 Beispiel Das Standardbeispiel für den Doppelring-Fall ist das Decalin. Es kommt in zwei diastereomeren Formen vor:



Wir wollen annehmen, die Plazierung habe die linke Form ergeben. Die Abbildung in 3.16 spiegelt dann jeweils drei Atome der Ring-Liganden und das Wasserstoffatom an der Ebene, welche die beiden Stereozentren enthält und die senkrecht zur Ebene der Zentren und dem einzelnen H-Atom steht. Das Ergebnis zeigt Abb. 5.

Auch Doppelbindungen können Probleme bereiten – wenn auch in einer anderen Form. Eine kritische Situation ist dann gegeben, wenn eine Doppelbindung in einem Ring auftaucht. Die räumliche Struktur eines Rings mit einer *cis*-Bindung ist von der eines Rings mit einer *trans*-Bindung grundverschieden. Auch hier konnte eine Modifikation Abhilfe schaffen.

3.18 Algorithmus (Invertierung einer Doppelbindung in einem Ring)

i) Führe 3.12.i) und ii) aus.

ii) Durchlaufe alle Nachbarn des zweiten Zentrums außer dem ersten Zentrum. Spiegele sie wieder mit der rekursiven Funktion aus 3.11.v). Führe aber bei dieser noch einen zusätzlichen Test im Anschluß an Punkt i) durch:

i_b) Liegt die Doppelbindung in einem Ring und ist das aktuelle Atom ein Nachbar des ersten Zentrums, so springe zurück.

◇

Auf diese Weise wird verhindert, daß eine eben durchgeführte Spiegelung im nächsten Moment bereits wieder rückgängig gemacht wird.

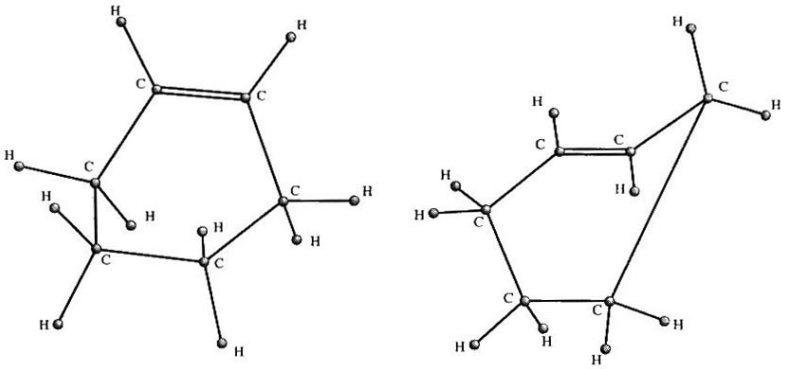


Abbildung 6: Diastereomere des Cyclohexen

3.19 Beispiel Wir betrachten das Cyclohexen, einen 6-Ring mit einer Doppelbindung. Abbildung 6 zeigt die beiden Diastereomere. Das resultierende Molekül (rechts) ist zwar nicht völlig korrekt; insbesondere die Bindungsabstände sind verzerrt. Der Konstruktionsalgorithmus hat allerdings die richtige Konfiguration der Doppelbindung erzeugt.

A Anzahlen von Stereoisomeren von Kohlenwasserstoffen

In diesem Abschnitt werden Tabellen aufgeführt, welche die Anzahlen möglicher Konstitutionsisomere und die möglicher Konfigurationsisomere zu einer Summenformel angeben. Die Strukturisomere wurden mit Hilfe des **MOLGEN**-Strukturgenerators berechnet. Die ermittelten Zahlen stimmen mit den in [22] überein, sofern die Bruttoformel dort vorkommt. Der Zeilenindex i bezeichnet jeweils die Anzahl der C-Atome, der Spaltenindex j die der H-Atome. Jeder Tabelleneintrag enthält die Zahl der Bindungsisomere und die Zahl der Stereoisomere.

A.1 Kohlenwasserstoffe C_iH_j

C_iH_j	2	4	6	8	10	12	14	16	18	20	22
1	—	1 1	—	—	—	—	—	—	—	—	—
2	1 1	1 1	1 1	—	—	—	—	—	—	—	—
3	2 3	3 4	2 2	1 1	—	—	—	—	—	—	—
4	7 15	11 22	9 13	5 6	2 2	—	—	—	—	—	—
5	21 76	40 126	40 100	26 48	10 13	3 3	—	—	—	—	—
6	85 518	185 1053	217 958	159 514	77 171	25 38	5 5	—	—	—	—
7	356 4235	920 10313	1230 10820	1031 6464	575 2447	222 620	56 101	9 11	—	—	—
8	1804 42694	5308 119777	7982 141083	7437 93365	4679 39417	2082 11350	654 2248	139 299	18 24	—	—
9	10064 498276	33860 1575402	56437 2053105	57771 1490065	40139 692420	19983 221494	7244 50270	1902 8102	338 875	35 55	—
10	64352 6576033	241297 23016067	439373 32854849	488125 25956181	369067 13133974	201578 4605287	81909 1165876	24938 216341	5568 28977	852 2640	75 136

A.2 Nitrokohlenwasserstoffe $C_iH_jN_2$ mit stereostabilen Stickstoffen

$C_iH_jN_2$	0	2	4	6	8	10	12	14	16	18
0	1 1	1 2	1 1	—	—	—	—	—	—	—
1	1 2	4 9	4 8	2 3	—	—	—	—	—	—
2	5 12	19 64	27 79	18 42	6 10	—	—	—	—	—
3	14 61	86 494	155 791	136 537	62 180	14 29	—	—	—	—
4	64 442	465 4688	1005 9143	1058 7427	633 3193	218 768	38 90	—	—	—
5	271 3761	2652 50533	6763 116231	8341 109894	5984 56126	2668 17117	716 3084	97 278	—	—
6	1448 38863	16977 617369	49516 1628764	69352 1751837	57411 1026627	30600 370221	10706 85357	2338 12133	260 870	—
7	8260 458835	117942 8364712	388019 24869431	606589 29957535	563966 19707282	344434 8098646	143857 2207282	40953 402558	7436 46451	688 2732

A.3 Oxokohlenwasserstoffe C_iH_jO

C_iH_jO	0	2	4	6	8	10	12	14	16	18
1	—	1	1	—	—	—	—	—	—	—
		1	1							
2	1	3	3	2	—	—	—	—	—	—
	1	4	3	2						
3	2	9	13	9	3	—	—	—	—	—
	3	16	19	11	3					
4	7	36	62	55	26	7	—	—	—	—
	15	97	149	105	38	8				
5	21	151	318	337	205	74	14	—	—	—
	76	702	1337	1118	503	131	18			
6	85	738	1823	2237	1623	747	211	32	—	—
	518	6252	14186	13608	7178	2308	448	47		
7	356	3971	11332	15804	13177	7166	2589	596	72	—
	4235	64742	169782	184368	110449	41256	9984	1523	123	
8	1804	24021	77431	120427	112484	69669	29797	8796	1684	171
	42694	766270	2270139	2751141	1838053	774382	218507	41781	5146	338

A.4 Oxokohlenwasserstoffe $C_iH_jO_2$

$C_iH_jO_2$	0	2	4	6	8	10	12	14	16	18
1	1	2	2	—	—	—	—	—	—	—
	1	2	2							
2	3	9	10	5	—	—	—	—	—	—
	4	12	12	5						
3	7	34	52	34	11	—	—	—	—	—
	13	67	91	49	13					
4	28	163	301	263	122	28	—	—	—	—
	71	512	861	601	216	37				
5	98	812	1821	1938	1168	400	69	—	—	—
	425	4500	9218	7777	3490	868	108			
6	459	4636	12098	15066	10893	4869	1313	179	—	—
	3468	46968	112783	110586	58559	18461	3460	325		
7	2254	28770	86246	122391	102139	54641	19154	4177	463	—
	33569	556223	1530483	1704994	1031754	383299	91068	13330	993	
8	13163	197786	666395	1055605	989647	607376	254468	72534	13190	1225
	387315	7381918	22828538	28427582	19262692	8135507	2277011	426285	50598	3070

Diese Tabellen zeigen deutlich, wie schnell die Anzahlen der Stereoisomere wachsen – auch und gerade im Verhältnis zur Zahl der Konstitutionsisomere. Für die Kohlenwasserstoffe zeigt Abb. 7 die durchschnittlichen Zahlen von Stereoisomere pro Strukturisomer. Daraus erkennt man zum einen die naheliegende Tatsache, daß es zu einer größeren Verbindung auch mehr Stereoisomere gibt, zum anderen aber auch den Einfluß der Wasserstoffe: die meisten Konfigurationsisomere existieren dort, wo es mehr als 0 und weniger als Kohlenstoffe gibt.

Literatur

[1] N.L. Allinger. MM2. A Hydrocarbon Force Field Utilizing V_1 and V_2 Torsional Terms. *J. Amer. Chem. Soc.*, **99**, S. 8127–8134, 1977.

[2] F. Barth und R. Schmid. *Analytische Geometrie.* Ehrenwirth Verlag, München, 1980.

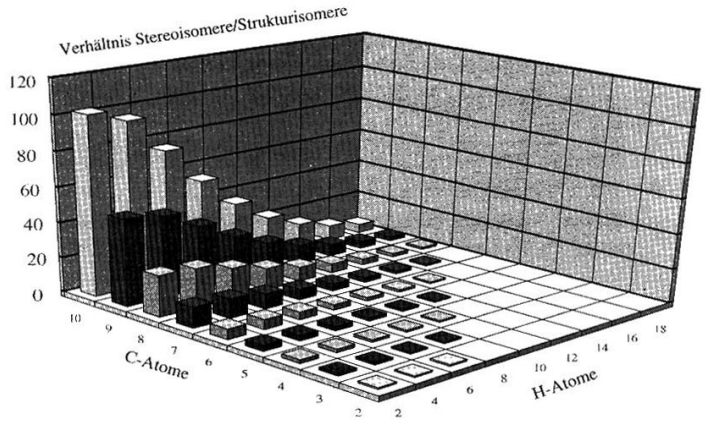


Abbildung 7: Verhältnisse Stereoisomere pro Strukturisomere bei C_iH_j

[3] H. BEYER. *Lehrbuch der organischen Chemie.* S. Hirzel Verlag, Stuttgart, 18., überarb. Aufl., 1978.

[4] E. BREITMAIER UND G. JUNG. *Organische Chemie I.* Thieme Verlag, Stuttgart, 1978.

[5] R.E. CARHART, D.H. SMITH, N.A. GRAY, J.G. NOURSE UND C. DJERASSI. GENOA: A Computer Program for Structure Elucidation Utilizing Overlapping and Alternative Substructures. *J. Org. Chem.*, **46**, S. 1708–1718, 1981.

[6] G.M. CRIPPEN. *Distance Geometry and Conformational Calculations.* New York, 1981.

[7] U. FANO UND L. FANO. *Physics of Atoms and Molecules.* The University of Chicago Press, Chicago, 1972.

[8] R. GRUND. *Computerunterstützte Berechnung von speziellen Doppelnebenklassentransversalen und deren Anwendung auf die konstruktive Kombinatorik.* Diplomarbeit, Universität Bayreuth, 1989.

[9] R. GRUND. Symmetrieklassen von Abbildungen und die Konstruktion von diskreten Strukturen. *Bayreuther Mathematische Schriften*, **31**, S. 19–54, 1990.

[10] R. GRUND, A. KERBER UND R. LAUE. MOLGEN, ein Computeralgebra-System für die Konstruktion molekularer Graphen. *MATCH*, **27**, S. 87–131, 1992.

[11] W. HÄSSELBARTH. On the interrelation between double cosets and orbits. *Theor. Chim. Acta*, **67**, S. 427–437, 1985.

[12] W. HÄSSELBARTH UND E. RUCH. Classifications of Rearrangement Mechanisms by Means of Double Cosets and Counting Formulas for the Number of Classes. *Theor. Chim. Acta*, **29**, S. 259–268, 1973.

[13] CHR. HILLER UND K. FEHNLE. *ALCOGEN - Ein algorithmischer Konformationengenerator.* Molsoft GmbH, Memmingen, 1991.

[14] B.A. KENNEDY, D.A. MCQUARRIE UND C.H. BRUBAKER, JR. Group Theory and Isomerism. *Inorg. Chem.*, **3**, S. 265–268, 1964.

[15] A. KERBER. *Algebraic Combinatorics Via Finite Group Actions.* BI-Wissenschaftsverlag, Mannheim, Wien, Zürich, 1991.

[16] A. KERBER, R. LAUE, R. GRUND, C. BENECKE UND T. WIELAND. *MOLGEN+ Handbuch.* Universität Bayreuth, 1993.

[17] H.P. LATSCHA UND H.A. KLEIN. *Organische Chemie - Basiswissen II.* Springer Verlag, Berlin, Heidelberg, 2., völlig neu bearb. Aufl., 1990.

[18] R.K. LINDSAY, B.G. BUCHANAN, E.A. FEIGENBAUM UND J. LEDERBERG. *Applications of Artificial Intelligence for Organic Chemistry: The Dendral Project.* McGraw-Hill, New York, 1980.

[19] K. MISLOW. *Einführung in die Stereochemie.* Verlag Chemie, Weinheim, 1967.

[20] J.G. NOURSE. Generalized Stereoisomerization Modes. *J. Am. Chem. Soc.*, **99**, S. 2063–2069, 1977.

[21] J.G. NOURSE. The Configuration Symmetry Group and Its Application to Stereoisomer Generation, Specification and Enumeration. *J. Am. Chem. Soc.*, **101**, S. 1210–1216, 1979.

[22] J.G. NOURSE, D.H. SMITH, R.E. CARHART UND C. DJERASSI. Exhaustive Generation of Stereoisomers for Structure Elucidation. *J. Am. Chem. Soc.*, **101**, S. 1216–1223, 1979.

[23] J.G. NOURSE, D.H. SMITH, R.E. CARHART UND C. DJERASSI. Computer-Assisted Elucidation of Molecular Structure with Stereochemistry. *J. Am. Chem. Soc.*, **102**, S. 6289–6295, 1980.

[24] G. PÓLYA. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta mathematica*, **68**, S. 145–253, 1937.

[25] E. RUCH, W. HÄSSELBARTH UND B. RICHTER. Doppelnebenklassen als Klassenbegriff und Nomenklaturprinzip für Isomere und ihre Abzählung. *Theor. Chim. Acta*, **19**, S. 288–300, 1970.

[26] E. RUCH UND D.J. KLEIN. Double Cosets in Chemistry and Physics. *Theor. Chim. Acta*, **63**, S. 447–472, 1983.

[27] C.C. SIMS. Computation with Permutation Groups. In S.R. PETRICK, Hrsg., *Proc. of the Second Symposium on Symbolic and Algebraic Manipulation*, S. 23–28. Assoc. Comput. Mach.

[28] B. TESTA. The geometry of molecules: Basic principles and nomenclatures. In CHR. TAMM, Hrsg., *Stereochemistry. New comprehensive biochemistry; vol. 3.* Elsevier Biomedical Press, Amsterdam, New York, Oxford, 1982.

[29] S. TRATCH UND Z. DEVDARIANI. *J. Org. Khimi*, S. 921–932, 1990.

[30] S. TRATCH, N. FARADEJEV ET AL. *Isled. matemat. w kompl. sistema*, S. 20–25, 1987.

[31] TH. WIELAND. Enumeration, generation and construction of stereo-isomers of high-valence stereo-centers. in Vorbereitung.

[32] TH. WIELAND. *Computerunterstützte Berechnung von Stereoisomeren.* Diplomarbeit, Universität Bayreuth, 1994.

[33] L.A. ZLATINA. *Mathematical Models of Generation of Stereoisomers and Building Space Models of Molecules.* Dissertation, All-Union Research Institute of Organic Sythesis, Moskau, 1991.