

Analyse der Einheitszelle

Um die Schwierigkeit und Komplexität des Problems der Strukturbestimmung abzuschätzen, wird die Information benötigt, wieviele Moleküle bzw. Atome in der Einheitszelle vorliegen. Diese Information ist auch entscheidend für die **Auswahl des effektivsten Verfahrens zum Lösen des Phasenproblems**.

Wieviele Moleküle oder Atome sind in der Einheitszelle ?

Für nur eine Atom- bzw. Molekülsorte gilt:

$$m = n \frac{M_s}{N}$$

Mit Hilfe der Dichte des Kristalls, des Molekulargewichts des Moleküls (Atoms) und des Volumens der Einheitszelle kann die gewünschte Größe berechnet werden:

$$\mathbf{r} = \frac{m}{V} = \frac{nM_s}{NV} \quad \longrightarrow \quad \boxed{n = 0.602 \frac{V\mathbf{r}}{M_s}}$$

Problem:

Weitere Fremdmoleküle (Salze, Wasser, etc.)
im Kristall !

m : Masse der Einheitszelle
 M_s : Molekulargewicht des
Moleküls (rel. Atommasse)
 N : Avogadrozahl
 n : Anzahl der Moleküle
in der Einheitszelle
 V : Volumen der Einheitszelle
in [\AA^3]
 \mathbf{r} : Dichte des Kristalls
in [g/cm^3]

Fouriersynthese und Bestimmung der Struktur

Das Bravaisgitter eines Kristalls wird allein durch die geometrische Anordnung der Reflexe im Beugungsbild bestimmt, während die **Anordnung der Atome in der Einheitszelle von der Intensität der Reflexe abhängt.**

Für die notwendige **Fouriersynthese zur Berechnung der Elektronendichteverteilung** bestehen unter Verwendung der gemessenen Reflexintensitäten im wesentlichen zwei Probleme:

- Aus den Beugungsdaten steht immer nur eine begrenzte Anzahl von Reflexintensitäten für die Synthese zur Verfügung. Daher treten bei der Darstellung der Elektronendichte Fehler auf, welche auch aus der Theorie der Fourierreihen bekannt sind (**Abbrucheffekte**). Dieses Problem ist bei einer genügend großen Anzahl der Reflexintensitäten **prinzipiell lösbar**.
- Die im allgemeinen unbekanntenen Phasen der Strukturfaktoren treten im Argument trigonometrischer Funktionen auf. Da sie also damit nicht linear in den Strukturfaktorgleichungen eingehen, kann prinzipiell auch bei einer sehr großen Anzahl von Daten (i.e. Gleichungen) **keine analytische Lösung für dieses Problem gefunden werden.**

Bestimmung der Phasen und Lösung der Struktur

Ansätze und Techniken zum Lösen des Phasenproblems

Bestimmung der Phasen und die absoluten Atomlagen im realen Raum von bestimmten Atomen (schwere Atome, anormale Streuer). => Bestimmung der Phasen aller anderen Atome der untersuchten Struktur durch iterativen Verfeinerungsprozess

- Isomorpher Ersatz (**MIR**: multiple isomorphous replacement)
- Anormale Streuung (**MAD**: multi-wavelength anomalous dispersion)

Molecular replacement: Eine ähnliche Struktur oder Teile der untersuchten Struktur sind bekannt => Suche nach Lage und Orientierung dieser Modelstruktur in der Einheitszelle.
=> Ausgang für weitere Strukturverfeinerung

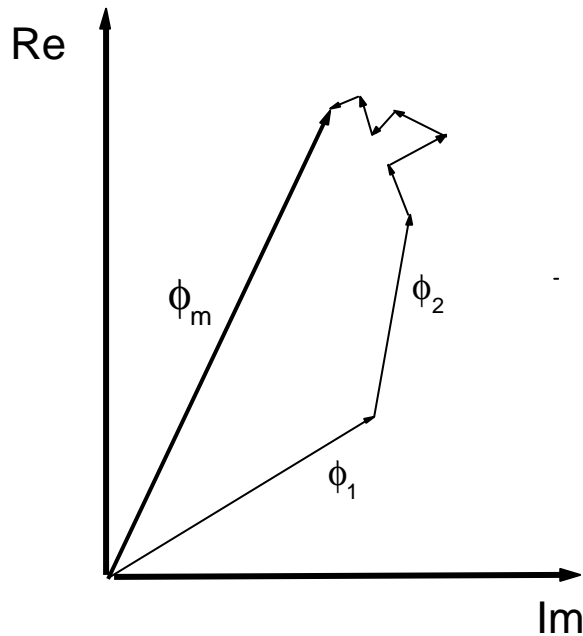
Direkte Methoden

Als direkte Methoden werden Verfahren verstanden, die zu einer unmittelbaren Festlegung der Phasenfaktoren aus beobachtbaren Intensitäten führen. Spezielle Gleichungen (**Harker-Kasper Ungleichung, Sayre-Gleichung**) schränken mögliche Beziehung zwischen Strukturfaktoramplituden und Phasenfaktoren ein. (=> nur auf kleine Strukturen anwendbar !)

Die Schweratommethode

Wenn in einer Struktur schwere Atome (d.h. mit großer Kernladungszahl Z) vorkommen oder künstlich eingefügt werden (isomorpher Ersatz, z.B. bei Proteinen), kann das ganz wesentlich die Strukturbestimmung vereinfachen. Peaks in der Pattersondichtekarte, die von Abstandsvektoren zwischen den Schweratomen resultieren, zeigen die größte Intensität. Die Koordinaten der Schweratome stellen ein gutes Anfangsmodell für die Strukturbestimmung dar.

Argand Diagramm, bei dem zwei Schweratome (mit den Phase ϕ_1, ϕ_2) und sechs weitere Atome zur Phase des gesamten Moleküls (ϕ_m) beitragen.



Für die Bestimmung der Koordinaten der „leichten“ Atome gibt es ein „optimales“ Verhältnis der Kernladungszahlen:

$$\frac{\sum Z_h^2}{\sum Z_l^2} = 1$$

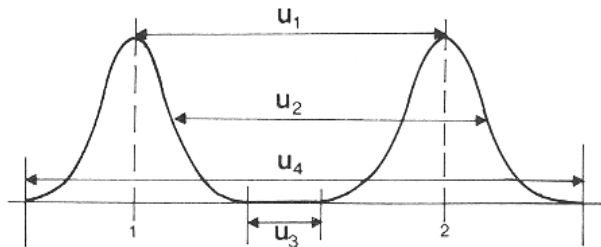
Beispiel:

hypothetisches Molekül $C_{30}H_{36}O_4X$ ($X = Br$) 2-mal in der Einheitszelle:

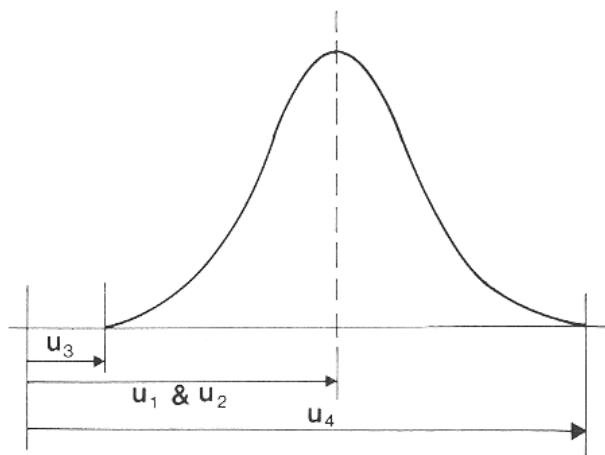
$$\sum Z_l^2 = 2708 \quad \sum Z_h^2 = 2450$$

Eigenschaften der Pattersondichte

Die Pattersonfunktion hat die gleiche Periodizität wie die Elektronendichte, und deshalb ist die Größe der Einheitszellen identisch. Auf der anderen Seite ist die Zahl der Maxima (Peaks) viel größer. Besonders bei nicht punktförmigen Elektronenverteilungen (z.B. kovalente Bindungen, siehe S. 21) kann das zu Problemen führen.



Der Überlapp verschiedener Peaks macht die Interpretation der Pattersondichtekarten unmöglich. Dieses Problem wird auch noch durch die Tatsache verstärkt, dass die Maxima der Pattersondichte breiter sind als die Maxima in der Elektronendichte (siehe Abbildung).



**Veranschaulichung der
Verarbeitung des Patterson-Peaks**

Intensitäten der Peaks

Vektoren

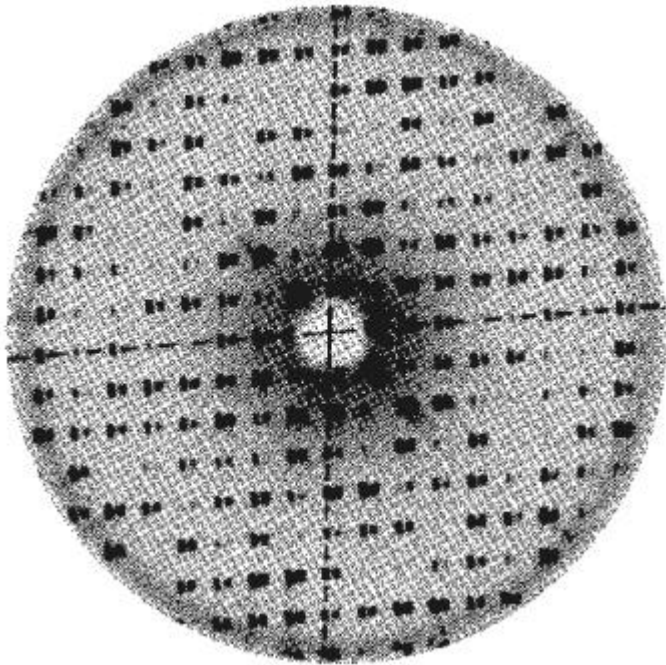
schweres Atom–schweres Atom
schweres Atom–leichtes Atom
leichtes Atom–leichtes Atom

„Peak“- Intensität

$Z_h Z_h$: sehr groß
 $Z_h Z_l$: mittelmäßig
 $Z_l Z_l$: klein

Die Schweratommethode

Isomorpher Schweratomersatz



Proteinkristall: Zwei Beugungsaufnahmen (leicht versetzte Darstellung) einmal von der nativen Struktur und einmal mit HgBr_2 in der Struktur gebunden.

Aus den Differenzen der Intensitäten können Phasen berechnet werden !

Die Elektronendichte der mit den Schweratomen gebundenen Struktur ist die Summe aus der Elektronendichte für die native Struktur und der Elektronendichte der Schweratome. Somit ergibt sich für die Strukturfaktoren:

$$F_{PH} = F_P + F_H$$

PH: Protein + Schweratom (*heavy atom*)

P: natives Protein

H: Schweratom

Die Pattersonfunktion

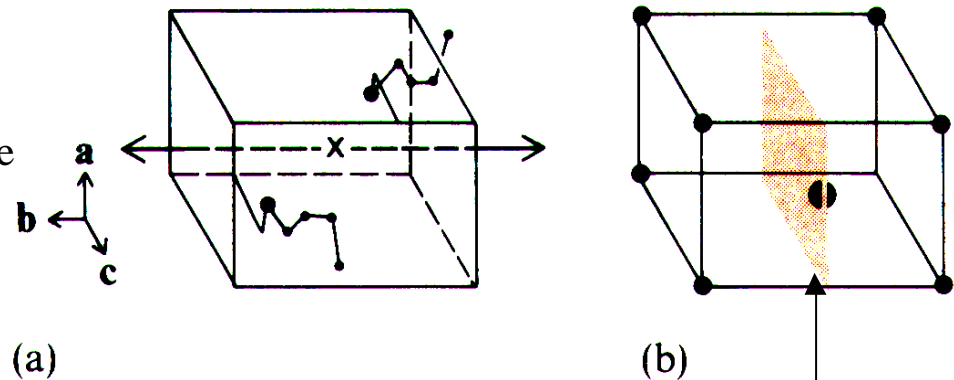
Die Pattersonfunktion zeigt die gleiche Periodizität im Kristall wie sie für die Elektronendichte beobachtet wird. Daher kann sie in der Summenschreibweise angegeben werden:

$$P(x, y, z) = \sum_h \sum_k \sum_l |F(h, k, l)|^2 e^{-2\pi i(hx+ky+lz)}$$

Ziel der Prozedur ist es nun die Lage der Schweratome mit Hilfe der Pattersondichtekarte zu bestimmen. Für Strukturen bestehend aus sehr vielen Atomen (~1000 und mehr) ist eine Zuordnung zu den entsprechenden „Interatom-Vektoren“ aufgrund der zahlreichen Peaks in der Pattersondichte sehr schwierig. Die stärksten Peaks resultieren von Abstandsvektoren zwischen Schweratomen.

Projektionen

Die meisten Raumgruppen beinhalten Moleküle die durch Symmetrierelationen innerhalb der Einheitszelle miteinander verbunden sind. Beispiel, hier $P2_1$ mit zwei Molekülen in der Einheitszelle und einer 2-zähligen Schraubenachse parallel zur Richtung von **b**. Die x,z-Koordinaten der Schweratome können in einer Pattersondichtekarte in der Ebene $(1/2)\mathbf{b}$ bestimmt werden.



Harker-Sektion: Ebenen oder Linien mit symmetrisch zueinander liegenden Patterson-Abstandsvektoren.

Bestimmung der Phasen

In der Praxis sind die Schweratom – Schweratom Vektoren häufig nicht so deutlich intensiver, als dass sie von den anderen Kombinationen (Schweratom – leichtes Atom; leichtes Atom – leichtes Atome) zu unterscheiden wären. Aus diesem Grund kann

$$P_H = (1/V) \sum_h \sum_k \sum_l |F_H(h, k, l)|^2 e^{-2\pi i(hx+ky+lz)}$$

nicht direkt aus den Daten berechnet werden. Statt dessen wird die **isomorphe Differenzpattersonfunktion** berechnet:

$$P_H = \Delta P = (1/V) \sum_h \sum_k \sum_l \left| |F_{PH}| - |F_P| \right|^2 e^{-2\pi i(hx+ky+lz)}$$

Diese Gleichung ist unter der Annahme erfüllt, dass $|F_H| = \left| |F_{PH}| - |F_P| \right|$ gilt. Dies ist für Reflexe eine gute Näherung, bei denen sich die Intensität durch das Einbringen der Schweratome stark ändert, da dann die Strukturfaktorvektoren relativ parallel zu einander liegen.

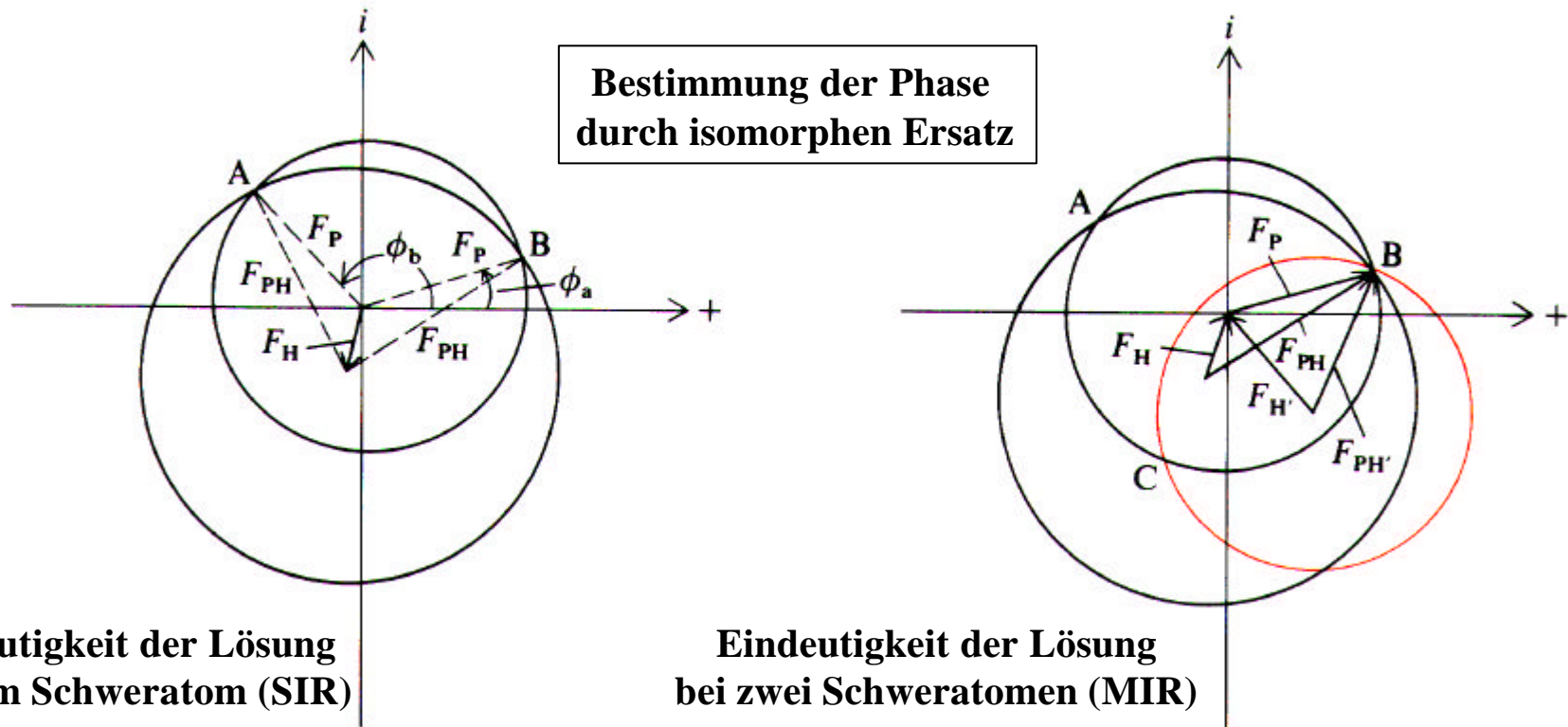


Bestimmung der Phasen

Durch die Differenzpattersonfunktion in Harker-Sektionen kann die Lage der Schweratome in der Einheitszelle bestimmt werden. Aus diesen Positionen können dann die Amplitude $|F_H|$ und die Phase ϕ_H der Strukturfaktoren der Schweratome berechnet werden. Für jeden Reflex im reziproken Gitter gilt somit:

$$I_P \sim |F_P|^2 \quad I_H \sim |F_H|^2$$

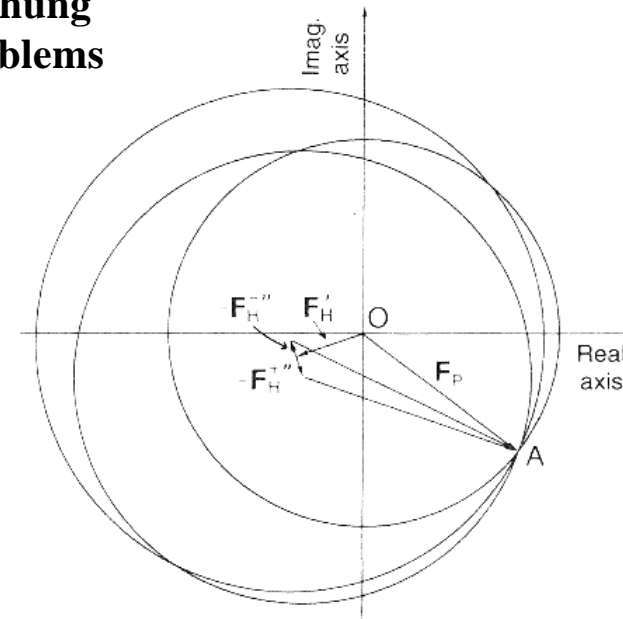
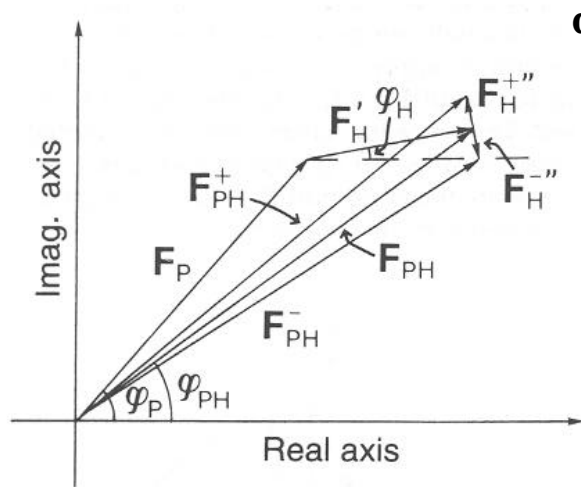
$$I_{PH} \sim |F_P + F_H|^2 = |F_P|^2 + |F_H|^2 + 2|F_H||F_P|\cos(\mathbf{j}_P - \mathbf{j}_H)$$



Anormale Streuung zur Beseitigung der Mehrdeutigkeit

Da das Herstellen mehrerer Derivate mit Schweratomen an verschiedenen Plätzen in der Einheitszelle (multipler isomorpher Ersatz) aufwendig ist, wurde auch die anormale Streuung genutzt, um das Phasenproblem zu lösen. Hierbei genügt dann **nur ein Schweratom das anormale Streuung zeigt**.

Graphische Veranschaulichung der Lösung des Phasenproblems



$$\mathbf{j}_{PH} = \mathbf{j}_H \pm \sin^{-1} \left[\left(F_{PH}^+ - F_{PH}^- \right) / 2F_H'' \right]$$

Die anormale Streuung des einen Schweratoms kann hier genutzt werden die Zweideutigkeit in der Phasenbestimmung aufzuheben. Die Genauigkeit der Phasen hängen von der Genauigkeit der sehr kleinen Bijvoet-Differenzen ab.

=> hohe Primärintensität => Synchrotronstrahlung !!

Güte der bestimmten molekularen Struktur und Strukturverfeinerung

Die bestimmten Phasen sind je nach Qualität der Daten häufig noch fehlerhaft. Die wichtigsten Fehlerquellen sind dabei:

- Ungenauigkeiten in den Reflexintensitäten und in den Phasenwinkeln
- Unvollständige Daten

Iterative Verfeinerungszyklen helfen zu einer verbesserten Übereinstimmung von Modeldaten und Messdaten. Ein Maß für die Übereinstimmung zwischen Model und den beobachteten Daten stellt der **R-Wert** dar:

$$R = \frac{\sum \left| |F_{obs}| - |F_{calc}| \right|}{\sum |F_{obs}|}$$

Für völlig zufällig verteilte Atome in der Einheitszelle einer nicht zentrosymmetrischen Struktur erhält man $R = 0.59$. Für gut definierte (kleinere) Strukturen kann man R–Werte von 0.05 – 0.1 erhalten. Bei großen, nicht überall gut definierten Strukturen (z.B. Proteine, mit Strukturbeweglichkeit und/oder Kristallwasser) erreicht man i.a. R–Werte von 0.15 – 0.25 . Daher gilt der R-Wert als die relevante Richtgröße, die angibt ob der iterative Prozess konvergiert und gegebenenfalls ob das bestmögliche Strukturmodell gefunden wurde.

Differenz-Fouriermethode

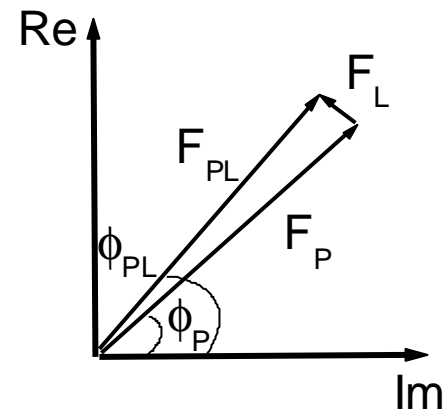
Unter der Annahme, dass die meisten Atome der Struktur schon richtig zugeordnet sind (bezüglich ihrer Kernladungszahl und ihre Position in der Einheitszelle) kann für einzelne Atome der Struktur (oder auch zusätzliche Liganden) die folgende Annahme gemacht werden:

Der Unterschied zwischen der Elektronendichte mit diesen Atomen (oder Liganden) (F_{PL}) und ohne F_P diese spiegelt sich nur in den Strukturfaktorbeträgen wider.

$$\Delta \mathbf{r}(\mathbf{r}) = \frac{1}{V} \sum_{h,k,l} |F_{PL}(h,k,l)| - |F_P(h,k,l)| \exp[-2i(hx + ky + lz) + i\mathbf{f}_p(h,k,)]$$

Es kann in diesem Ansatz also die Phase der Struktur ohne Ligand \mathbf{f}_p benutzt werden.

Diese Methode kann zur **Strukturverfeinerung** und bei der Analyse **kleiner Strukturänderungen** (Ligandenbindung, Konformationsänderungen) genutzt werden.

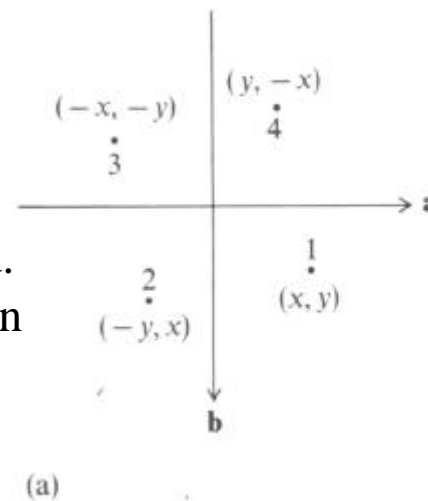


Beispiel und Interpretation einer Differenz-Pattersonprojektion

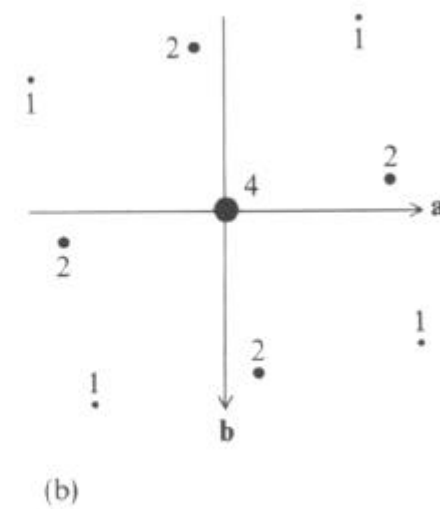
Die Lage von Schweratomen in einem Molekül: Projektion in die Ebene senkrecht zur c-Achse in einem tetragonalen Kristall (Raumgruppe $P4_1$).

In der asymmetrischen Einheit befindet sich ein Molekül mit einem Schweratom, In der Einheitszelle befinden sich vier Moleküle, die durch eine 4-zählige Schraubenachse einander zugeordnet sind. Die Schweratome liegen in den Molekülen alle an der selben Stelle.

Beschreibung der Schweratomlage:
 $\mathbf{r} = x\mathbf{a} + y\mathbf{b}$



Dazugehörige Patterson Dichte



Dazugehörige Schweratom-Schweratom Vektoren:

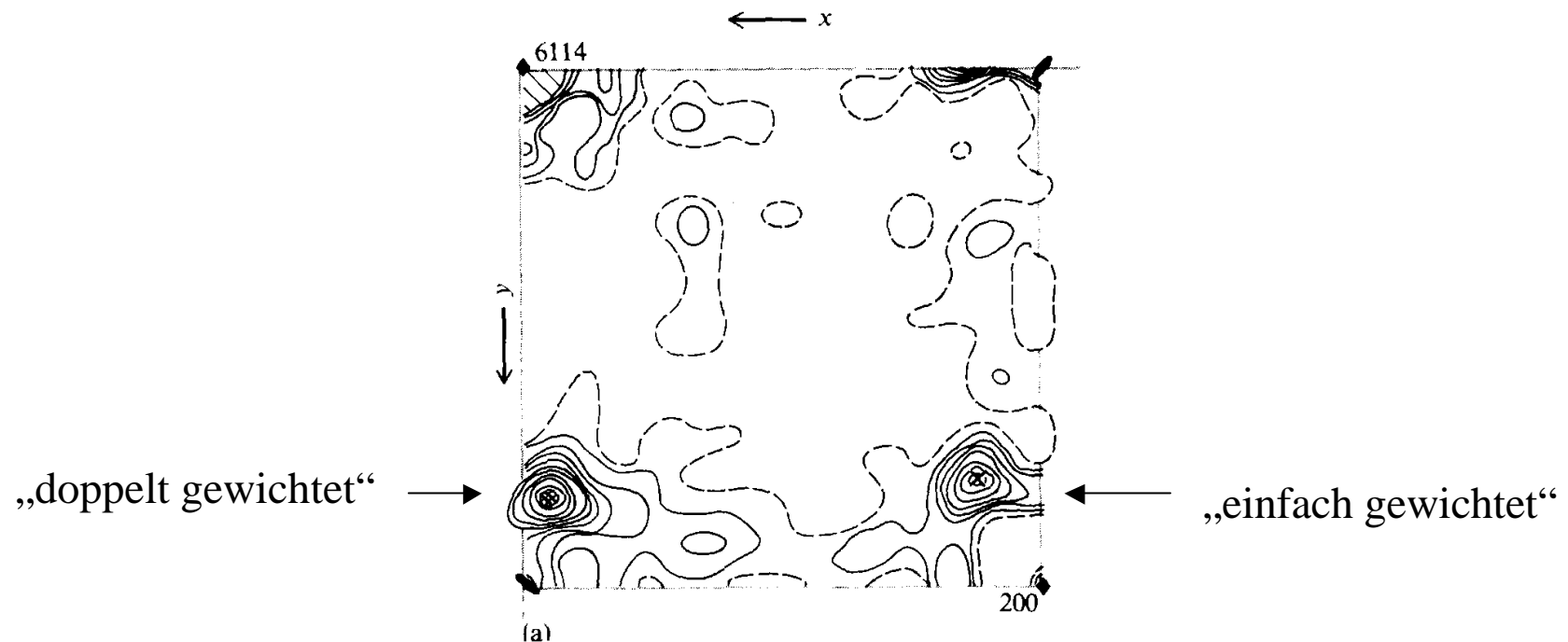
$\mathbf{r}_{12} = (x+y, y-x)$	$\mathbf{r}_{23} = (x-y, y+x)$	$\mathbf{r}_{21} = (-x-y, x-y)$	$\mathbf{r}_{32} = (x-y, y+x)$
$\mathbf{r}_{13} = (2x, 2y)$	$\mathbf{r}_{24} = (-2y, 2x)$	$\mathbf{r}_{31} = (-2x, -2y)$	$\mathbf{r}_{42} = (2y, -2x)$
$\mathbf{r}_{14} = (x-y, x+y)$	$\mathbf{r}_{34} = (-x-y, x-y)$	$\mathbf{r}_{41} = (y-x, -x-y)$	$\mathbf{r}_{43} = (x+y, y-x)$

Plus 4 Vektoren
im Ursprung.

Beispiel und Interpretation einer Differenz-Pattersonprojektion

Die gezeigte Differenz-Pattersondichtekarte wurde berechnet für eine Projektion in die Ebene senkrecht zur c-Achse eines tetragonalen Proteinkristalls (Raumgruppe $P4_1$ mit $a=b= 58.45 \text{ \AA}$; $c= 42.34 \text{ \AA}$).

$$\Delta P(x, y) = (1/A) \sum \sum \left| |F_{PH}(h, k, 0)| - |F_P(h, k, 0)| \right|^2 e^{-2\pi i(hx+ky)}$$



Vergleich S. 215 gerechnete Pattersondichte Karte unterer rechter Quadrant.

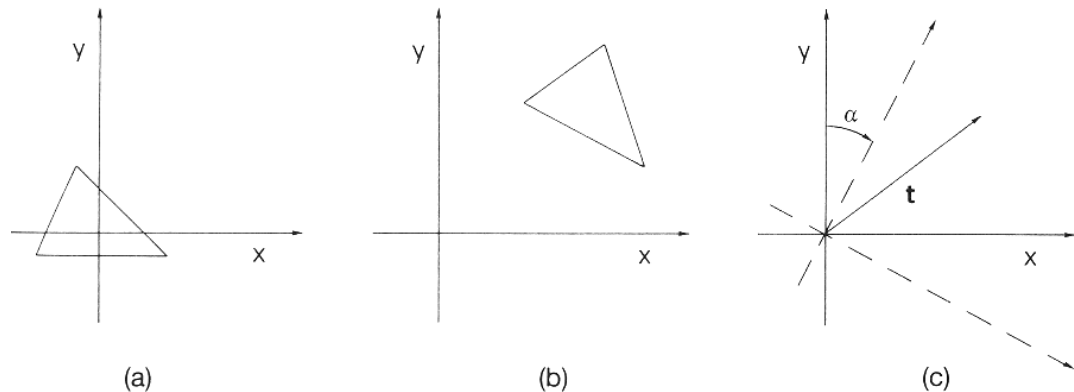
Molecular replacement

Mit dieser Technik wird für eine bekannte Modelstruktur (z.B. Molekül) oder für eine Teilstruktur, welche (große) Ähnlichkeit mit der gesuchten Struktur hat, die Position in der Einheitszelle gesucht. Im Prinzip ist es möglich simultan die 3 Translations- und 3 Rotationsfreiheitsrade zu bestimmen, die eine Minimaldifferenz zwischen F_{cal} und F_{obs} darstellen. In der Praxis wird hierfür eine erhebliche Rechnerkapazität benötigt (in Abhängigkeit von der Größe der Struktur). Das Verfahren des **molecular replacement** erfolgt in zwei Schritten:

- Zuerst wird mit Hilfe von **Rotationsfunktionen** die richtige Orientierung der Struktur gesucht.
- Dann erfolgt mit Hilfe der **Translationsfunktionen** die Suche nach der genauen Lage in der Einheitszelle.

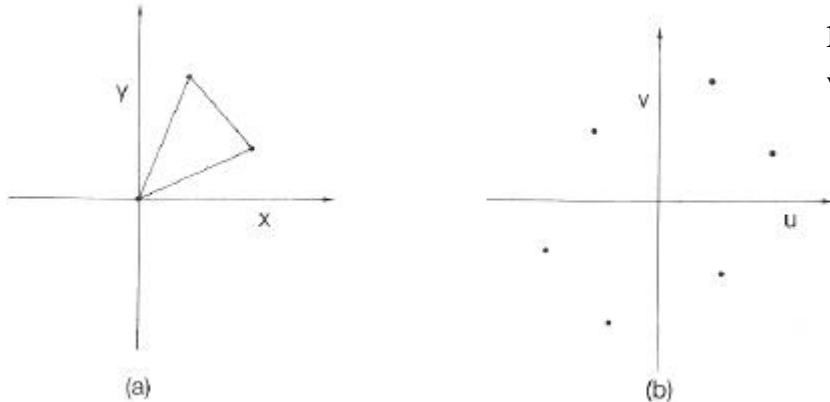
Auch wenn die benutzte Modelstruktur nur eine gewisse Ähnlichkeit mit der zu bestimmenden Struktur besitzt oder nur einen Teil der Struktur darstellt, hilft dieses Verfahren **Ausgangsphasen** zu bestimmen, mit denen die Struktur in

weiteren Verfeinerungszyklen beliebig genau **nur mit Hilfe der Reflexintensitäten zu bestimmen** ist. Es ersetzt somit in geeigneten Fällen die anderen Techniken zur Phasenbestimmung.



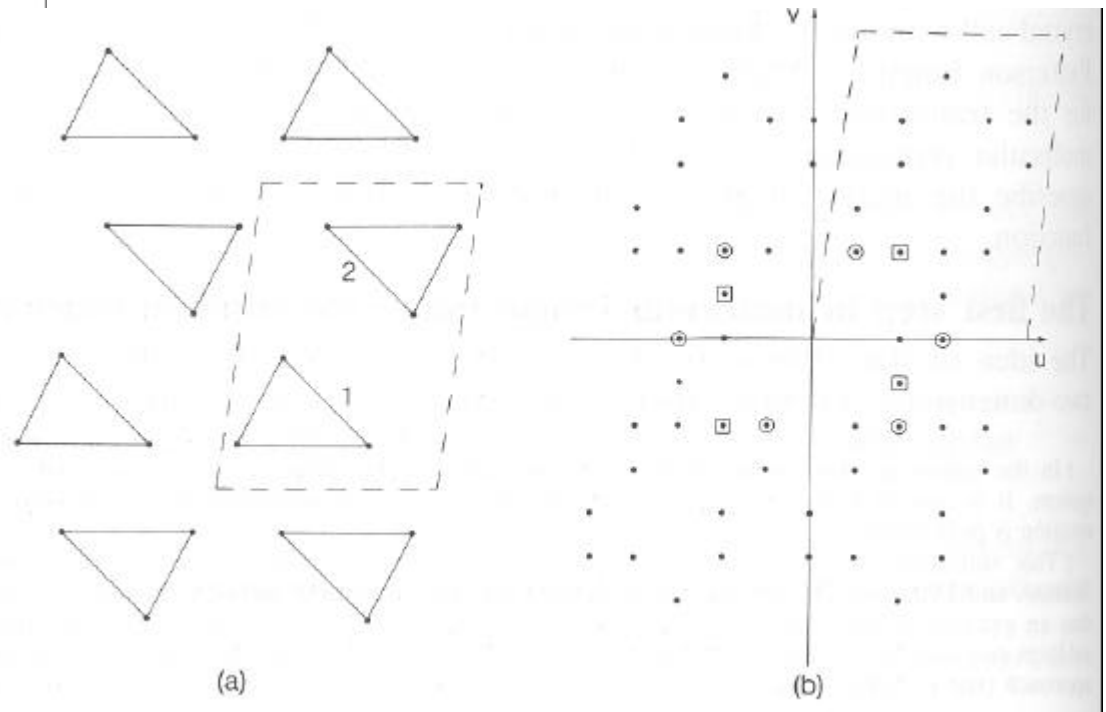
Rotations- und Translationsfunktionen

Schematische Darstellung einer Struktur (a) und deren Selbstfaltung bzw. Pattersondichte (b)



Durch Anwendung von Rotationsfunktionen auf die Pattersondichte kann zwischen intermolekularen Vektoren (**self-vectors**; Peaks mit Kreisen) und intramolekularen Vektoren (**cross-vectors**) unterschieden werden.

Periodische Anordnung der Struktur (a) und der zugehörigen Pattersondichtekarte (b) mit eingezeichneter Einheitszelle (gestrichelt)



Rotationsfunktion

$$R(\mathbf{C}) = \int_V P_{cryst}(\mathbf{u}) P_{mol}(\mathbf{Cu}) d\mathbf{u}$$

C: Rotationsmatrix

P_{cryst} : Pattersonfunktion des Kristalls

P_{mol} : Pattersonfunktion des Moleküls

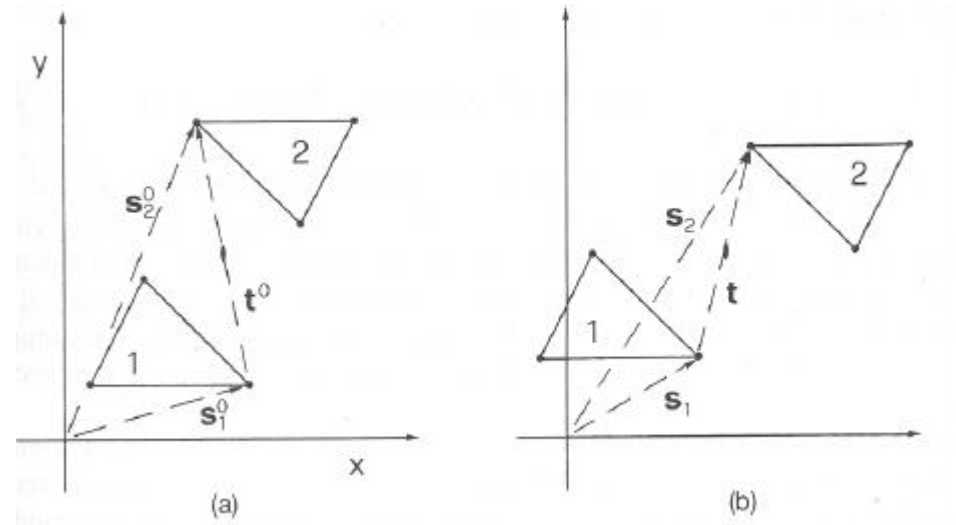
$R(\mathbf{C})$ hat ein Maximum, wenn sich Peaks von P_{cryst} und von P_{mol} überlappen.

Rotations- und Translationsfunktionen

Wenn die Orientierung des Moleküls bestimmt wurde, muss im nächsten Schritt die absolute Lage in der Einheitszelle bestimmt werden. Durch Anwendung von Translationsfunktionen ändern sich die intermolekularen Vektoren (cross-vectors) der Pattersonfunktion.

Das Ausprobieren aller möglichen Translationen führt zu den absoluten Lagen der Moleküle die z.B. durch eine Symmetrieelement vorgegeben sind (siehe Abb. (a) mit \mathbf{t}^0).

=> nur bei P_1 -Symmetrie besteht die Notwendigkeit der Anwendung von Translationsfunktionen nicht.



Anordnung zweier Strukturen mit zugehörigen Translationsvektoren mit Bezug zu einem Symmetrieelement (hier 2-zählige Achse).

Translationsfunktion

$$T(\mathbf{t}) = \int_V P_{obs}(\mathbf{u}) P_{12}(\mathbf{u}, \mathbf{t}) d\mathbf{u}$$

$$\mathbf{t} = \mathbf{s}_2 - \mathbf{s}_1$$

P_{obs} : beobachtete Pattersonfunktion

$$P_{12}(\mathbf{u}) = \int_V \mathbf{r}_1(\mathbf{x}) \mathbf{r}_2(\mathbf{x} + \mathbf{u}) d\mathbf{x}$$