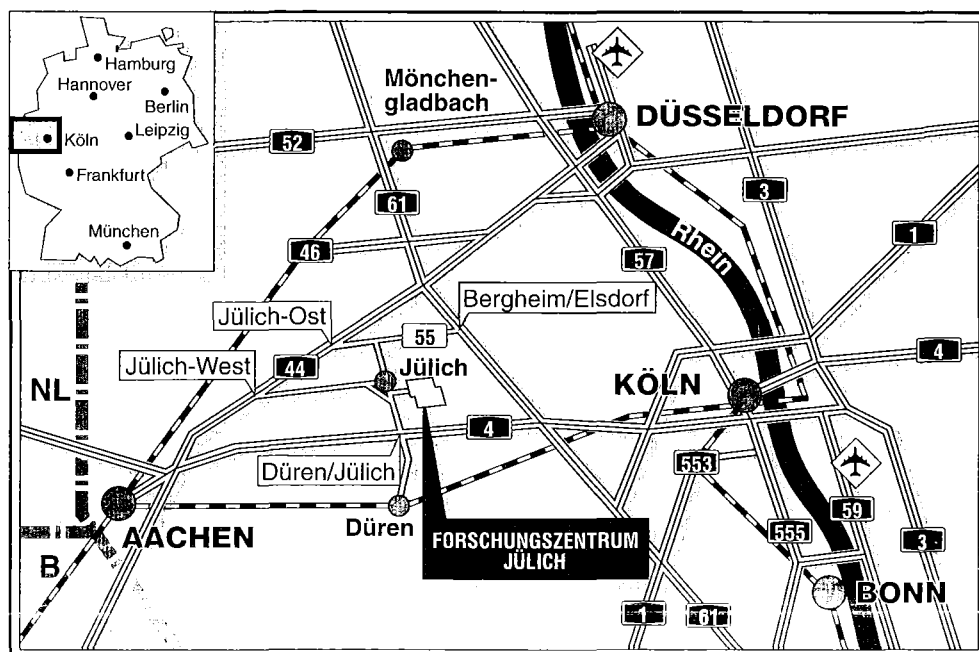


Institut für Biotechnologie

Metabolische Kohlenstoff-Markierungssysteme

**Modellierung, Simulation, Analyse,
Datenauswertung**

Wolfgang Wiechert



Berichte des Forschungszentrums Jülich ; 3301

ISSN 0944-2952

Institut für Biotechnologie Jül-3301

Zu beziehen durch : Forschungszentrum Jülich GmbH · Zentralbibliothek

D-52425 Jülich · Bundesrepublik Deutschland

Telefon : 024 61/61-61 02 · Telefax : 024 61/61-61 03 · Telex : 833 556-70 kfa d

Metabolische Kohlenstoff-Markierungssysteme

Modellierung, Simulation, Analyse,
Datenauswertung

Wolfgang Wiechert

Vorwort

Die wissenschaftliche Erforschung komplexer Systeme ist heute ohne die Zusammenarbeit verschiedener Fachdisziplinen nicht mehr denkbar. Eine der größten Herausforderungen stellen dabei biologische Systeme dar, in denen hochgradig vernetzte Strukturen eine isolierte Untersuchung einzelner Phänomene nahezu unmöglich machen. Die vorliegende Arbeit beschäftigt sich mit einem der durch Regulationsmechanismen am stärksten vernetzten biologischen Systeme, nämlich dem Zellmetabolismus. Sie verdankt ihr Entstehen einer langjährigen interdisziplinären Zusammenarbeit zwischen Biologen, Physikern, Ingenieuren, Informatikern und Mathematikern und stellt den Aspekt der Modellierung, Systemanalyse, Simulation und Datenauswertung bei metabolischen Markierungssystemen in den Vordergrund.

Die aus der Schulphysik vertraute reduktionistische Arbeitsweise ist in den biologischen Disziplinen zwangsläufig zum Scheitern verurteilt, denn die experimentelle Isolation einer Komponente — also das In-vitro-Experiment — erlaubt nur mit großen Bedenken den Rückschluß auf Funktion und Verhalten dieser Komponente innerhalb des Gesamtzusammenhangs. In dieser Situation muß versucht werden, den wohldefinierten und reproduzierbaren Bedingungen eines klassischen physikalischen Experiments möglichst nahe zu kommen, ohne das beobachtete System zu zerstören. Dieses Vorhaben ist ohne den massiven Einsatz technischer Apparaturen, hochspezialisierter Meßgeräte und spezifischer Analysemethoden meist nicht realisierbar.

Dies trifft in besonderem Maße für das Jülicher Stoffflußanalyse-Projekt zu, das sich zum Ziel gesetzt hat, stoffwechselphysiologische Studien an lebenden Mikroorganismen mit Hilfe der NMR-Technik durchzuführen. Der ohne Zweifel wichtigste Mitarbeiter des Projekts ist Herr Dr. A.A. de Graaf, der seit 1991 mit großem Elan das NMR-Labor an dem unter Leitung von Herrn Prof. Dr. H. Salm stehenden Institut für Biotechnologie 1 (kurz IBT1) aufgebaut hat. Ohne die besonderen Fähigkeiten von Herrn de Graaf wären viele der seither erzeugten In-vivo-NMR-Signale wohl bis heute uninterpretiert geblieben, denn die Anwendung der NMR-Meßtechnik erfordert eine hochspezialisierte Ausbildung. Von Herrn de Graaf habe ich gelernt, daß sich das (im Vergleich zu anderen Datenanalyseproblemen) so einfach erscheinende Problem der quantitativen Auswertung von In-vivo-NMR-Spektren wohl immer einer vollständig automatisierten Auswertung verschließen wird. Erst nach langjähriger Erfahrung vermag ein Experte einen signifikanten Peak nicht nur qualitativ zu beurteilen, sondern auch der richtigen nuklearmagnetischen Ursache zuzuordnen.

Im Falle der Stoffflußanalyse müssen diese Fähigkeiten einhergehen mit biochemischer Systemkenntnis, der Fähigkeit zur Kultivierung von Mikroorganismen und der Fertigkeit zur zuverlässigen Durchführung von Laboranalysen. Ohne die massive Mitarbeit der am IBT 1 tätigen Mikrobiologen bei der Planung, Durchführung und Interpretation der Experimente wären die vorliegenden Ergebnisse wohl kaum zustande gekommen. Ganz besonders hervorheben möchte ich die Arbeit von Herrn Dipl. Biol. Achim Marx, der im Rahmen seiner Diplomarbeit die Grundlagen für die Durchführung der in dieser Arbeit beschriebenen experimentellen Methode gelegt und die Meßdaten erhoben hat. Dankte er mir in diesem Zusammenhang für Mathematik-Nachhilfestunden, so möchte ich nun ihm ganz besonders für seine Biologie-Nachhilfe danken, vor allem für seine heftigen Einwände, wenn ich einmal wieder einen Zusammenhang zu stark vereinfacht hatte.

Mit der Fähigkeit zur Beherrschung des biologischen Systems einerseits und der NMR-Technik andererseits ist immer noch ein wichtiger Aspekt übersehen, der aus der Sicht des Naturwissenschaftlers "lediglich" technischer Natur ist: Zur Herstellung wohldefinierter physiologischer Bedingungen für das Wachstum von Mikroorganismen in einem Bioreaktor ist ein erheblicher technischer Aufwand erforderlich. Will ein biotechnologisches Forschungsvorhaben den Anfor-

derungen reproduzierbarer und quantitativ interpretierbarer Daten und Ergebnisse entsprechen, kommt ein solches Projekt nicht mehr ohne die Mitarbeit von Ingenieuren aus. Unter den erschwerten Bedingungen starker Magnetfelder, kleiner Reaktionsvolumina und hoher Zelldichten kann auch ein handwerklich geschickter Naturwissenschaftler kaum jenes systematische Arbeiten an den Tag legen, das eine reaktionstechnisch professionell aufgebaute Anlage von einer Bastelarbeit unterscheidet. Gerade hier hat sich die enge Nachbarschaft zu dem stärker bioverfahrenstechnisch ausgerichteten Institut für Biotechnologie 2 unter Leitung von Herrn Prof. Dr. C. Wandrey bewährt. So stellt der von Herrn Dipl. Ing. Andreas Hartbrich unter der Betreuung von Herrn Dr. Dirk Weuster in Jülich entwickelte Hochzelldichte-NMR-Hydrozyklonreaktor eine technische Leistung dar, der ich nicht zuletzt einige ästhetische Aspekte abgewinnen kann.

Erst unter derartigen Voraussetzungen können qualitativ hochwertige Meßwerte in großer Zahl erhoben werden. Damit steht der Experimentator vor dem Problem der Interpretation und Auswertung seiner Daten. Zwar lassen sich erste Schlüsse aufgrund einzelner Meßwerte stets leicht ziehen, jedoch ist damit die hochvernetzte Struktur des Datenmaterials noch nicht adäquat erfaßt. Insbesondere bei derartig "teuren" Daten, wie sie bei Experimenten mit ^{13}C -markierten Substanzen in Verbindung mit der NMR-Technik zustandekommen, sollte eine diesem Aufwand entsprechende Sorgfalt für deren Analyse aufgewendet werden. Hier ist schließlich eine geeignete Modellierung gefragt, denn Datenanalyse bedeutet stets die Interpretation von Meßwerten aus der Perspektive eines mathematischen Modells.

Hier zeigt sich aber auch schon die Problematik der Modellierung, denn verschiedene Modelle werden zu unterschiedlichen Interpretationen der Daten führen. Ein "Modellierer" muß daher in ständigem Dialog mit den Experimentatoren ein akzeptierbares Modell der Zusammenhänge im untersuchten System erstellen. Der Begriff "akzeptierbar" ist hier mit Bedacht gewählt, denn die Berücksichtigung aller Zusammenhänge innerhalb eines komplexen Systems führt in der Regel zu nicht mehr lösbaren mathematischen Problemen und experimentell nicht mehr validierbaren Aussagen. Die Bearbeitung der mit der Modellierung verbundenen systemtheoretischen und statistischen Fragestellungen erfordert eine breit angelegte mathematische Ausbildung, wie sich im Verlauf dieser Arbeit noch deutlich zeigen wird.

Mein Einstieg in das Stoffflußanalyse-Projekt entspringt eher dem Zufall als einem geplanten Vorhaben. Das Interesse an der NMR-Technik, persönliche Sympathie und nicht zuletzt auch einige faszinierende mathematische Probleme im Zusammenhang mit der Analyse der Experimente führten zu einer Zusammenarbeit, deren vorläufiges Ergebnis — denn der Prozeß geht weiter — hier zusammengefaßt ist. An dieser Stelle möchte ich meinem Institutsleiter Herrn Prof. Wandrey danken, der mich bei diesem interdisziplinären Vorhaben, wo immer es ihm möglich war, unterstützt hat.

Diese Arbeit stellt ein universell anwendbares Modellierungswerkzeug für die Beschreibung, Analyse, Simulation und Auswertung von Markierungsexperimenten bereit. Der Begriff "Universalität" ist dabei nicht so mißzuverstehen, daß damit ein für alle biologischen Systeme gültiges Modell formuliert wird, vielmehr wird durch die entwickelten Analysewerkzeuge ein Rahmen geschaffen, der den technischen Aufwand der Entwicklung von metabolischen Stoffflußmodellen für unterschiedliche Stoffwechselnetzwerke stark reduziert.

Die Bereitstellung von Simulations- und Datenanalysewerkzeugen für die Forschung bedeutet nicht zuletzt auch einen erheblichen Programmieraufwand. Die Programmierung wird im wissenschaftlichen Umfeld gerne als reines Handwerk behandelt und weniger als wissenschaftliche Leistung gewertet. Diese Einschätzung hat fatale Folgen für die Softwareentwicklung im akademischen Umfeld, denn Qualität kann nur dort entstehen, wo eine Arbeit als solche auch honoriert wird. Vielfach entsteht so ein Arsenal von mehr oder weniger spezialisierten Einzelprogrammen, das die ständige Neuentwicklung bereits in ähnlicher Weise vorhandener Werkzeuge

erfordert. Die objektorientierte Technologie hat hier einige Fortschritte mit sich gebracht, die zumindest innerhalb einzelner Arbeitsgruppen eine bessere langfristige Koordination ermöglicht. Jedoch stellt die starke Mitarbeiterfluktuation stets ein großes Hindernis für die kontinuierliche Entwicklung und Pflege von Software dar. An dieser Stelle möchte ich Tatjana Höner und meinen Bonner Mitarbeitern Claudia Siefke, Arndt Boland und Vera Schwingenheuer danken, die zusammen mit mir die derzeit bestehenden Module und numerischen Algorithmen des hier beschriebenen Systems erstellt haben.

Einleitung

Im Kontext eines interdisziplinären Projekts entstanden, versteht sich die vorliegende Arbeit als ein Beitrag zum Fachgebiet der Theoretischen Biologie, die sich ganz allgemein mit der Modellierung — aber auch der Modellierbarkeit — biologischer Systeme befaßt. Dabei wird die gesamte Spannweite des Modellierungsprozesses von der Diskussion der biologischen Annahmen über die Formulierung von Modellgleichungen und deren systemtheoretischer Analyse, der Simulation und Werkzeugentwicklung bis hin zur statistischen Auswertung und abschließenden Bewertung der Experimente behandelt.

Eine typische Schwierigkeit interdisziplinären Arbeitens, nämlich die Etablierung einer gemeinsamen Kommunikationsplattform, zeigt sich auch bei der Frage nach dem Adressatenkreis dieser Arbeit. Denn einerseits überfordern die durch den Gegenstand bedingten fortgeschrittenen mathematischen Hilfsmittel in aller Regel die Vorkenntnisse eines Biologen. Andererseits soll gerade der experimentelle Anwender durch die hier entwickelten theoretischen Konzepte in seiner praktischen Arbeit unterstützt werden. Es ist also mein Anliegen, diese Arbeit sowohl dem experimentellen Anwender als auch dem mathematisch vorgebildeten Leser zugänglich zu machen, auch wenn weder die biologischen Grundlagen noch die mathematischen Vorkenntnisse diesem Rahmen erschöpfend dargestellt werden können.

Es soll daher versucht werden, zumindest die Grundlagen und Ergebnisse in einer allgemein verständlichen Form darzustellen, während die Detailausführungen zum Teil umfangreiche Vorkenntnisse erfordern. So mag für den Mathematiker die umständliche Formulierung der Modellgleichungen mit vielen Beispielen und einer erst im Anschluß eingeführten abstrakten Matrixnotation einige Längen haben, jedoch kann auf diese Weise auch der interessierte Biologe dem Text weitgehend folgen und schließlich auf das genaue Verständnis des mathematischen Formalismus verzichten, der im wesentlichen beweis- und implementierungstechnische Bedeutung hat. Umgekehrt mögen mir der Biologe bzw. der NMR-Spezialist die Simplifizierungen der ersten Kapitel verzeihen, die den experimentellen Rahmen des Projekts darstellen. Konsequenterweise wurde auch auf Anhänge verzichtet, die eine Ungleichgewichtung der Inhalte suggerieren würden.

Als Leserführung für beide Leserkreise wird versucht, in der Einleitung jedes Kapitels Hinweise auf die erforderlichen Vorkenntnisse zu geben. Bei den stark mathematisch "belasteten" Kapiteln 7, 8, 10 und 11 soll zumindest versucht werden, dem nicht mathematischen ausgebildeten Leser einen Eindruck von der sachbedingten Notwendigkeit der Behandlung der jeweiligen Problematik zu verschaffen. Von besonderer Bedeutung ist in diesem Zusammenhang das didaktisch gewählte Beispiel im letzten Abschnitt des einführenden Kapitels, das die wichtigsten Probleme und Fragestellungen für den gesamten folgenden Text umreißt. Außerdem enthalten einige der stärker mathematisch orientierten Kapitel einfache Beispiele, die auch mit einem informellen Vorverständnis der vorher exakt eingeführten Begriffe lesbar sind. Der Leser möge sich hierbei von den Überschriften und den Illustrationen leiten lassen.

Ein zentraler Aspekt der Arbeit ist die Darstellung des Modellierungsprozesses. Dabei soll einem unter experimentellen Anwendern häufig anzutreffenden Fehlverständnis mathematischer Methoden entgegengewirkt werden. Denn bei den aus der Schulphysik (also dem ersten und prägenden Kontakt mit der mathematischen Modellbildung) vertrauten Systemen ist man gewohnt, daß die Anwendung mathematischer Operationen die exakte Beschreibung des Systemverhaltens gestattet. Mathematische Ausdrücke kommen hier meist durch eine direkte Umsetzung der beobachteten Elementarzusammenhänge in die Formelsprache ins Spiel. Beim Hooke'schen Gesetz ist dies z.B. der lineare Zusammenhang zwischen einwirkender Kraft und Auslenkung einer Feder. An der Gültigkeit der so formulierten Zusammenhänge besteht — zumindest im Rahmen der Alltagserfahrung — kein Zweifel, so daß der Physiker seine validierten Mo-

delle gerne als Gesetze bezeichnet.

Trotzdem ist gerade das Hooke'sche Gesetz bei einer Spiralfeder eine starke Vereinfachung für ein äußerst komplexes System, das jedoch mit gutem Grund approximativ durch eine einfache lineare Gleichung beschreibbar ist. Dementsprechend gehen statistische Verfahren zur Berechnung von Parametern aus Meßwerten bei physikalischen Systemen im Grunde immer von der Korrektheit des zugrundegelegten Modells also der mathematisch exakten Beschreibbarkeit des beobachteten Systems aus. Insbesondere in den Ingenieurwissenschaften ist diese Auffassung heute vorherrschend und meist auch sinnvoll. Ist die Korrektheit des Modells implizit vorausgesetzt, so täuscht ein statistisches Verfahren seinem Anwender ein mathematisch exaktes Ergebnis vor, das dann als der "wahre" Zustand des untersuchten Systems interpretiert wird.

Völlig anders ist dies bei der Beschreibung komplexer biologischer Systeme durch Modellgleichungen. Mathematische Modelle in der Biologie können stets nur mehr oder weniger grobe Vereinfachungen der beobachteten Zusammenhänge sein. Die Erstellung eines "akzeptierbaren" Modells (vgl. das Vorwort) ist in diesem Kontext als eine Kunst anzusehen, deren Ausübung lange Erfahrung, Systemkenntnis und einen langandauernden Diskussionsprozeß erfordert. Der aus der Schulphysik gewohnte Umgang mit Modellgleichungen und statistischen Verfahren führt in diesem Kontext zwangsläufig zu einer Fehlinterpretation der Resultate, denn ein mathematisches Modell für ein biologisches System ist im strengen Sinne immer falsch! Dies muß dennoch nicht zu einer generellen Verwerfung der Modellbildung bei biologischen Systemen führen. Ein konstruktiver Ansatz ist vielmehr, die Ergebnisse einer Datenanalyse bei biologischen Systemen stets als die Interpretation des experimentellen Datenmaterials im Lichte eines Modells zu verstehen. Dies erfordert einen ständigen Reflexionsprozeß über das Verhältnis von Modell, Meßdaten und Analyseergebnissen. Dieser muß auch vom experimentellen Anwender mitgetragen werden, der "nur" an den berechneten Parametern interessiert ist.

In diesem Zusammenhang sollte betont werden, daß die Auswertung experimenteller Daten *immer* ein mathematisches Modell voraussetzt. Im einfachsten Fall (z.B. beim Hooke'schen Gesetz) besteht dieses Modell aus einer einzigen Gleichung. Der experimentelle Anwender ist daher häufig geneigt, eine solche einfache Gleichung nicht als ein Modell mit den damit verbundenen Modellannahmen anzusehen, sondern als eine Gesetzmäßigkeit im Sinne der Schulphysik. Damit wird die Reflexion über die Beziehung von Modell und Experiment vollkommen unterdrückt, denn die Gültigkeit des zugrundegelegten Modells muß dann nicht angezweifelt werden. Als Konsequenz ergibt sich die Auffassung, daß solche Ergebnisse, die mit einfachen und unmittelbar "verständlichen" Formeln aus Meßdaten berechnet werden, stets richtig sind, wogegen aufwendige mathematische Analysen schwer nachvollziehbar und daher aus sich heraus "problematisch" sind. Die Folge ist ein Verlust an wissenschaftlicher Qualität und eine starke Tendenz zur Übereinfachung. Zu welchen Fehlinterpretationen diese Skepsis hinsichtlich der Modellbildung führen kann, zeigt das Beispiel im letzten Kapitel dieser Arbeit.

Erst wenn Modellierung als ein Beitrag zum wissenschaftlichen *Diskussionsprozeß* verstanden wird, können Meßdaten also verantwortungsvoll interpretiert werden. Insofern ist es ein gutes Zeichen, wenn das Resultat einer sorgfältigen Datenanalyse aus verschiedenen biologischen Gründen zumindest partiell angezweifelt werden kann, denn in dieser Diskrepanz ist der Antrieb für weitere Fortschritte zu suchen. Alle gewonnenen Ergebnisse sind also stets relativ zum Modell zu lesen. Der Anwender muß daher sowohl die zugrundegelegten Modellannahmen als auch die mathematische Formulierung des Modells in gewissem Umfang verstehen, um die erzielten Ergebnisse einschätzen zu können. Nicht zuletzt deshalb ist die breite Darlegung dieser Grundlagen auch für den experimentellen Anwender ein besonderes Anliegen dieser Arbeit.

Verwendete chemische und mathematische Notation

Die bei der Bezeichnung von chemischen Stoffen, Kohlenstoffatomen und Isotopomeren sowie deren Quantifizierung durch Mengen- oder Konzentrationsvariablen verwendete Notation ist in Kapitel 3 (insbesondere Definition 3.1 und Tabelle 3.1) und in Abschnitt 4.1 zusammengestellt. Grundsätzlich werden dabei chemische Stoffe mit einem serifenlosen Font gekennzeichnet, also etwa mit *Glu*, wogegen die entsprechenden Mengen- oder Konzentrationsvariablen kursiv gesetzt sind, wie in *glu*, *glu*₁ oder *glu*₁₀₀₀₀₀.

Die in dieser Arbeit durchgängig verwendeten mathematischen Symbole sind dem nachfolgenden Symbolverzeichnis zu entnehmen. In den Abschnitten 5.3, 6.3 und 8.1 finden sich zudem Zusammenstellungen aller für das hier entwickelte universelle Modell und die entsprechenden Rechenverfahren wichtigen Definitionen und Gleichungen. An dieser Stelle seien daher nur einige Eigenarten der in dieser Arbeit verwendeten mathematischen Notation — die stark durch eine objektorientierte Denkweise beeinflusst ist — dargestellt. Geprägt durch den jahrelangen Umgang mit objektorientierten Methoden ist eines meiner Hauptanliegen die Entlastung des Schriftbildes von überflüssigen Indizes und Formelsymbolen. Gerade bei der Entwicklung spezialisierter Modelle muß eine solche einfache Notation angestrebt werden, da ansonsten ein kaum mehr zu überblickender "Zoo" von Symbolen und Indizes entsteht. Bei der Umsetzung dieses Vorhabens helfen einige Parallelen zu Konventionen der objektorientierten Programmierung:

Schreibkonventionen: Grundsätzlich stehen fett gedruckte Kleinbuchstaben wie **v**, **x**, **a** für Vektoren und fette Großbuchstaben wie **P**, **Q**, **Σ** für Matrizen. Eine Ausnahme bilden die Vektoren **η**, **Θ**, **ε** im Statistikteil. Skalare Werte, aber auch Funktionssymbole sind stets ohne Fettdruck, z.B. als *c*, *α*, *f*, *φ* notiert.

Objekte und Methoden: Vektoren und Matrizen sind Objekte, denen insbesondere ihre Dimensionen zugeordnet ist. Es werden daher an keiner Stelle explizite Symbole *m*, *n*, *p*, *q* usw. zur Bezeichnung von Dimensionen verwendet. Vielmehr wird stets **dim x** für die Dimension eines Vektors und gelegentlich **rowdim M** und **coldim M** für die Zeilen- und Spaltendimension einer Matrix geschrieben. Es ist also

$$\mathbf{x} \in \mathbb{R}^{\text{dim } \mathbf{x}} \quad \text{und} \quad \mathbf{M} \in \mathbb{R}^{\text{rowdim } \mathbf{M} \times \text{coldim } \mathbf{M}}.$$

Bei expliziter Darstellung von Matrizen werden Nullen zwecks besserer Lesbarkeit durch Punkte ersetzt; etwa bei der Einheitsmatrix

$$\mathbf{1} = \begin{pmatrix} 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{pmatrix}$$

In Kapitel 4.4 wird darüber hinaus eine spezielle Notation für mehrdimensionale Matrizen eingeführt.

Symbolüberladung: Die fettgedruckte **0** kann sowohl für den Null-Vektor als auch für die Null-Matrix jeweils passender Dimension stehen. Die fettgedruckte **1** bezeichnet entweder einen mit Einsen besetzten Vektor oder die Einheitsmatrix (siehe oben).

Operatorüberladung: Eine Relation oder eine Funktion, die normalerweise auf skalaren Werten definiert ist, wird bei Anwendung auf Vektoren auf jeden der Einträge einzeln bezogen. Beispielsweise wird geschrieben:

$$\mathbf{x} > \mathbf{y} \quad \text{oder} \quad \mathbf{x} = \min(\mathbf{y}, \mathbf{z})$$

Ebenso bezeichnet $\text{diag}(\mathbf{M})$ sowohl die Diagonale einer Matrix $\mathbf{M}$ (also einen Vektor) als auch die diagonale Matrix $\text{diag} \mathbf{x}$ zu einem Vektor $\mathbf{x}$.

Symbolverzeichnis

Die nachfolgend zusammengestellten Formelsymbole werden in dieser Arbeit durchgehend verwendet. Symbole, die nur lokal innerhalb eines Kapitels benutzt werden, sind hier nicht eigens aufgeführt.

Zustandsvektoren

$\mathbf{x}$	Markierungsvektor	Definition 4.2 und 4.3
$\mathbf{x}^A$	Markierungsvektor für C-Atome	nach Definition 4.3
$\mathbf{x}^I$	Markierungsvektor für Isotopomere	nach Definition 4.3
$\mathbf{x}^{\text{inp}}$	Input-Markierungsvektor	Definition 4.2 und 4.3
$\mathbf{x}^{\text{inp},A}$	Input-Markierungsvektor für C-Atome	nach Definition 4.3
$\mathbf{x}^{\text{inp},I}$	Input-Markierungsvektor für Isotopomere	nach Definition 4.3
$\mathbf{v}$	Gesamtflußvektor	Definition 4.1
$\mathbf{v}^{\rightarrow}$	Hinflußvektor	Definition 4.1
$\mathbf{v}^{\leftarrow}$	Rückflußvektor	Definition 4.1
$\mathbf{v}^{\text{net}}$	Nettofluß-Vektor	Definition 5.5
$\mathbf{v}^{\text{xch}}$	Austauschflußvektor	Definition 5.5
$\mathbf{v}^{\text{xch}[0,1]}$	kompaktifizierter Austauschflußvektor	Definition 5.5

Strukturmatrizen

$\mathbf{Q}$	bimolekulare Transitionsmatrix für Isotopomere (bimolekular)	Definition 4.10
$\mathbf{P}$	monomolekulare Transitionsmatrix	Definition 4.9 und 4.10
$\mathbf{P}^{\text{inp}}$	Input-Transitionsmatrix	Definition 4.9 und 4.10
$\mathbf{Z}_M^A$	Zugehörigkeitsmatrix für C-Atome und Metabolite	Definition 4.11
$\mathbf{Z}_M^{A,\text{inp}}$	Zugehörigkeitsmatrix für Input-C-Atome und -Metabolite	Definition 4.11
$\mathbf{Z}_M^I$	Zugehörigkeitsmatrix für Isotopomere und Metabolite	Definition 4.11
$\mathbf{Z}_M^{I,\text{inp}}$	Zugehörigkeitsmatrix für Input-Isotopomere und -Metabolite	Definition 4.11
$\mathbf{Z}_A^I$	Zugehörigkeitsmatrix für Isotopomere und C-Atome	Definition 4.11
$\mathbf{Z}_A^{I,\text{inp}}$	Zugehörigkeitsmatrix für Input-Isotopomere und -C-Atome	Definition 4.11

Abbildungen

$\phi^{[0,1]}$	skalare Kompaktifizierungsfunktion	Definition 5.3
$\min_{\alpha}$	geglättete Minimumfunktion	Definition 5.2
$\max_{\alpha}$	geglättete Maximumfunktion	Definition 5.2
Ψ	kompaktifizierte Koordinaten als lineare Funktion der freien Flüsse	Gleichung 8.2
$\Phi^{[0,1]}$	Koordinatentransformation zwischen kompaktifizierten Flußkoordinaten und Anwendungs-Flußkoordinaten	Definition 5.5
Φ_{α}	geglättete Koordinatentransformation zwischen Anwendungs-Flußkoordinaten und natürlichen Flußkoordinaten	Definition 5.5
Γ	Markierungszustand als Funktion der natürlichen Flußkoordinaten	Satz 7.3

Lineare Nebenbedingungen

N^{net}	Nebenbedingungsmatrix für Nettoflüsse	Definition 6.1
$N^{\text{xch}[0,1]}$	Nebenbedingungsmatrix für Austauschflüsse	Definition 6.2
N^{free}	zusätzliche Nebenbedingungsmatrix für freie Flüsse	Definition 6.3
N	Gesamt-Nebenbedingungsmatrix	Definition 6.4
n^{net}	Wertevektor für Nettofluß-Nebenbedingungen	Definition 6.1
$n^{\text{xch}[0,1]}$	Wertevektor für Austauschfluß-Nebenbedingungen	Definition 6.2
n^{free}	zusätzlicher Wertevektor für freie Flüsse	Definition 6.3
n	Gesamt-Wertevektor für Flüsse	Definition 6.4
U^{net}	Ungleichungs-Nebenbedingungsmatrix für Nettoflüsse	Definition 6.5
$U^{\text{xch}[0,1]}$	Ungleichungs-Nebenbedingungsmatrix für Austauschflüsse	Definition 6.5
U	Gesamt-Ungleichungs-Nebenbedingungsmatrix	Definition 6.5
u^{net}	Ungleichungs-Wertevektor für Nettofluß-Nebenbedingungen	Definition 6.5
$u^{\text{xch}[0,1]}$	Ungleichungs-Wertevektor für Austauschfluß-Nebenbedingungen	Definition 6.5
u	Gesamt-Ungleichungs-Wertevektor für Flüsse	Definition 6.5

Meßgleichungen

w	Flußmeßvektor	Definition 6.6
y	Markierungsmeßvektor	Definition 6.6
Y	Poolgrößenmeßvektor	Definition 6.6
M_w	Fluß-Meßmatrix	Definition 6.6
M_y	Markierungs-Meßmatrix	Definition 6.6
M_Y	Poolgrößen-Meßmatrix	Definition 6.6
Σ_w	Kovarianzmatrix der Flußmessung	Definition 6.6
Σ_y	Kovarianzmatrix der Markierungsmessung	Definition 6.6
Σ_Y	Kovarianzmatrix der Poolgrößenmessung	Definition 6.6

Weitere Symbole

α	Glättungsparameter für Flußtransformationen
β	Konfidenzniveau
χ_n^2	Chi-Quadrat-Verteilung mit n Freiheitsgraden

Inhaltsverzeichnis

I	Experimentelle Grundlagen	1
1	Metabolische Markierungsexperimente	3
1.1	Stationäre Stoffflußanalyse	3
1.1.1	Stoffwechselphysiologische Forschung und Metabolic Engineering . .	3
1.1.2	Zielsetzung der stationären Stoffflußanalyse	4
1.1.3	Anwendungen im Rahmen des Metabolic Engineering	6
1.1.4	Angewendete Methoden	7
1.1.5	Markierte Substanzen	7
1.1.6	Markierungsexperimente	8
1.1.7	Positionsmarkierung und Isotopomere	10
1.1.8	Rohdaten für die Stoffflußanalyse	10
1.1.9	Typische experimentelle Anordnungen	12
1.2	Ein instruktives Beispiel	13
1.2.1	Metabolit-Pools	13
1.2.2	Flußschätzung auf der Grundlage extrazellulärer Flußmessungen	13
1.2.3	Hinzunahme von Positions-Markierungsdaten	15
1.2.4	Hinzunahme von Isotopomeren-Daten	16
1.2.5	Zyklische Wege und bidirektionale Reaktionsschritte	16
2	Experimente und Modellannahmen	18
2.1	Standard-Meßtechnik	19
2.1.1	Extrazelluläre Flüsse	19
2.1.2	Einbeziehung der Biomassezusammensetzung	19
2.2	NMR-Technik und Massenspektroskopie	21
2.2.1	Funktionsweise eines NMR-Geräts	21
2.2.2	¹³ C- und ¹ H-Spektren	23
2.2.3	Multipllettanalyse	24
2.2.4	Massenspektroskopie	26
2.2.5	Problematik von In-vivo-NMR-Messungen	27
2.2.6	Analyse von Gemischspektren	28
2.3	Stoffflußanalyse unter wohldefinierten Bedingungen	28
2.3.1	Grundannahmen der stationären Stoffflußanalyse	28
2.3.2	Homogenität	29
2.3.3	Metabolische Stationarität	30
2.3.4	Biochemische Systemkenntnis	31
2.3.5	Isotopen-Masseneffekte	31
2.3.6	Zerstörungsfreiheit des Meßvorgangs	31
2.3.7	Unverzerrtheit der Messung	32

2.4	Zusätzliche Annahmen	32
2.4.1	Energiebilanzierung	33
2.4.2	Direktionalitätsannahmen	34
2.4.3	Pool-Lumping	34
2.5	Eine neue experimentelle Methode	35
2.5.1	Entkopplung des Markierungs-Experiments von der -Messung	35
2.5.2	Verzerrungsfreiheit der Messung	36
2.5.3	Fehlerabschätzung für die Markierungsmessung	39
2.5.4	Charakteristika der neu entwickelten Methode	40
2.6	Beziehung zu metabolisch instationären Modellansätzen	41
2.6.1	Dynamische Studien	41
2.6.2	Bedenken gegen den dynamischen Ansatz	42
2.6.3	Einordnung der stationären Stoffflußanalyse	43
II	Modellbildung	45
3	Kohlenstoff-Netzwerke	47
3.1	Metabolit-, C-Atom- und Isotopomeren-Netzwerke	47
3.1.1	Formale Notation von Reaktionsschritten	47
3.1.2	Ein einfaches Beispiel	50
3.1.3	Formale Beschreibung der Netzwerke	50
3.1.4	Netzwerkvereinfachung	51
3.2	Formulierung realer Netzwerke	52
3.2.1	Einige Begriffsdefinitionen	52
3.2.2	Erweiterte Ausdrucksmöglichkeiten und einschränkende Bedingungen .	54
3.2.3	Formulierung von Sonderfällen	55
3.2.4	Scrambling-Reaktionen	57
3.3	Beispielnetzwerke	57
3.3.1	Typen von Beispielen	57
3.3.2	Reaktionen im Zentralstoffwechsel	58
4	Bilanzgleichungen	61
4.1	Zustandsvariablen	62
4.1.1	Flußvariablen	62
4.1.2	Nebenbedingungen und Annahmen über Flüsse	63
4.1.3	C-Atom-Markierungsvariablen	63
4.1.4	Isotopomeren-Markierungsvariablen	64
4.1.5	Nebenbedingungen für Markierungsvariablen	65
4.1.6	Poolgrößenvariablen	65
4.2	Formulierung von Bilanzgleichungen	65
4.2.1	Metabolit-Bilanzen	65
4.2.2	C-Atom-Bilanzen	66
4.2.3	Isotopomeren-Bilanzen	66
4.2.4	Isotopisch instationärer Fall	67
4.3	Explizite Behandlung des Beispielsystems	68
4.3.1	Explizite Lösung der Bilanzgleichungen	68
4.3.2	Redundanzbeziehungen	69

4.3.3	Flußschätzung	70
4.3.4	Realisierbare Markierungszustände	71
4.4	Mehrdimensionale Matrizen	72
4.4.1	Allgemeine Gestalt der Bilanzgleichungen	72
4.4.2	Definition mehrdimensionaler Matrizen	73
4.4.3	Matrix-Vektor-Produkte	74
4.4.4	Rechenregeln für Matrix-Vektor-Produkte	74
4.4.5	Matrix-Matrix-Produkte	75
4.5	Matrixnotation der Bilanzgleichungen	75
4.5.1	Metabolit-Bilanzen	75
4.5.2	C-Atom-Bilanzen	76
4.5.3	Isotopomeren-Bilanzen	77
4.5.4	Zugehörigkeitsmatrizen	78
4.5.5	Beziehungen zwischen den C-Atom- und Isotopomeren-Bilanzgleichungen	79
4.5.6	Bilanzgleichungen im isotopisch instationären Fall	79
5	Netto- und Austauschflüsse	81
5.1	Alternative Beschreibung von Stoffflüssen	82
5.1.1	Definition von Austauschflüssen	82
5.1.2	Eigenschaften von Austauschflüssen	83
5.1.3	Differenzierbare Approximation von Austauschflüssen	85
5.1.4	Eigenschaften des approximierten Austauschflusses	86
5.2	Behandlung großer Austauschflüsse	87
5.2.1	Charakterisierung der asymptotischen Grenzlinie des realisierbaren Gebietes	87
5.2.2	Systemeigenschaften im Bereich großer Austauschflüsse	88
5.2.3	Kompaktifizierte Definition von Austauschflüssen	89
5.3	Zusammenfassung der Flußkoordinatensysteme	90
6	Modellannahmen und Meßgleichungen	92
6.1	Spezifikation von Modellannahmen	92
6.1.1	Formulierung von Annahmen durch lineare Gleichungen	92
6.1.2	Annahmen über Nettoflüsse	93
6.1.3	Annahmen über Austauschflüsse	94
6.1.4	Freie Flüsse	95
6.1.5	Ungleichheits-Nebenbedingungen und zulässige Flüsse	96
6.2	Meßgleichungen	98
6.2.1	NMR-Multipllettanalyse	99
6.2.2	Massenspektroskopische Multipllettanalyse	100
6.2.3	Formulierung von Kovarianzmatrizen	101
6.3	Zusammenfassung der Modellgleichungen	102
III	Simulation	105
7	Stabilitätsanalyse	107
7.1	Stabilitätsanalyse von C-Atom-Markierungssystemen	107
7.1.1	Kompartimentelle Matrizen	108

7.1.2	Fallen in C-Atom-Netzwerken	108
7.1.3	Eindeutige Lösbarkeit von C-Atom-Bilanzgleichungen	109
7.1.4	Untersuchung der Fallenfreiheit	110
7.1.5	Stabilität von dynamischen Systemen	111
7.1.6	Eigenwerte kompartmenteller Matrizen	112
7.1.7	Instationäre Bilanzgleichungen	112
7.2	Dissipative Systeme	113
7.2.1	Dissipative Systeme	113
7.2.2	Dissipativität lokal dissipativer Systeme	114
7.2.3	Dissipativität lokal kompartmenteller Systeme	115
7.2.4	Ein globales Stabilitätskriterium für dissipative Systeme	116
7.3	Stabilitätsanalyse von Isotopomerensystemen	118
7.3.1	Die zulässige Markierungsmenge eines Isotopomeren-Systems	118
7.3.2	Ein Stabilitätskriterium für Isotopomeren-Systeme	120
7.4	Beispiele für Stabilitätsanalysen	120
7.4.1	Eine Rechenhilfe	120
7.4.2	Isotopomeren-Stabilität im Citratzyklus	122
7.4.3	Eine Transformationsmethode	122
8	Numerische Lösung von Bilanzgleichungen	124
8.1	Numerische Lösung der Bilanzgleichungssysteme	124
8.1.1	Rechenschritte einer Simulation	125
8.1.2	Beurteilung und Lösung des Flußsystems	126
8.1.3	Lösung der stationären C-Atom-Bilanzgleichungen	127
8.1.4	Lösung der stationären Isotopomeren-Bilanzgleichungen	127
8.1.5	Lösung der instationären Bilanzgleichungen	128
8.2	Ableitungen und große Austauschflüsse	128
8.2.1	Ableitung des C-Atom-Markierungszustands nach den Flüssen	128
8.2.2	Lösung gewichteter Gleichungssysteme	129
8.2.3	Anwendung auf Markierungssysteme	131
8.2.4	Die Grenzgleichung für schnelle Gleichgewichte	132
9	Simulation von Markierungssystemen	134
9.1	Durchführung von Simulationen	134
9.1.1	Simulationsstrategie	134
9.1.2	Anforderungen an ein Simulationswerkzeug	135
9.2	Ein Beispiel	135
9.2.1	Simulatoreingabe	137
9.2.2	Stationäre C-Atom-Markierungsanteile	139
9.2.3	Isotopomeren-Verteilungen	141
9.2.4	Instationäre Simulationsergebnisse	142
9.3	Implementierung und Architektur des Simulators	144
9.3.1	Generische Architektur eines Datenanalysesystems	144
9.3.2	Programmkomponenten für das Jülicher Stoffflußanalyse-Projekt	146
9.3.3	Netzwerksynthese	147
9.3.4	Numerische und algebraische Komponenten	148
9.3.5	Visualisierung	149

9.3.6	Objektorientierte Methoden	149
-------	--------------------------------------	-----

IV Auswertung von Experimenten 151

10 Stationäre Flußschätzung 153

10.1	Das allgemeine Schätzproblem	153
10.1.1	Verschiedene Ansätze	153
10.1.2	KQ-Flußschätzung	155
10.1.3	Vereinfachung des Schätzproblems	156
10.1.4	Das allgemeine Regressionsmodell	157
10.1.5	Numerische Problematik des Schätzproblems	158
10.2	Optimierung unter Nebenbedingungen	159
10.2.1	Mathematische Programmierung	159
10.2.2	Aktive und inaktive Nebenbedingungen	160
10.2.3	Die Kuhn-Tucker-Bedingung	161
10.2.4	Optimierungsverfahren mit hoher Konvergenzordnung	162
10.2.5	Sequentielle Quadratische Programmierung	163
10.2.6	Quadratische Programmierung	164
10.2.7	Das Verfahren von Goldfarb und Idnani	165
10.3	Numerische Lösung des Flußschätzproblems	165
10.3.1	Der Levenberg-Marquardt-Algorithmus	166
10.3.2	Adaptation des Verfahrens für restringierte KQ-Probleme	167
10.3.3	Ausschluß von Nebenbedingungen	168
10.3.4	Anmerkungen zur Konvergenztheorie	169

11 Statistische Auswertung 171

11.1	Linearisierung nichtlinearer Modelle	172
11.1.1	Zur Problematik biometrischer Methoden	172
11.1.2	Der Linearisierungsansatz	173
11.2	Lineare Regressionsanalyse	174
11.2.1	Eigenschaften linearer KQ-Schätzer	174
11.2.2	Sensitivitätsmaße	176
11.2.3	Skalierung von Sensitivitäts- und Kovarianzmatrizen	177
11.2.4	Numerische Analyse von Sensitivitäts- und Kovarianzmatrizen	179
11.2.5	Redundanzanalyse	180
11.2.6	Ausreißererkennung	181
11.2.7	Versuchsplanung	181
11.3	Ein Beispielsystem	182
11.3.1	Ein Modellsystem	183
11.3.2	Explizite Lösungen	183
11.3.3	Realisierbare Markierungszustände	184
11.3.4	Exakte Verteilung des Schätzers	185
11.3.5	Linearisierungsergebnisse	186
11.3.6	Nichtlineare Parameterreskalierung	188
11.3.7	Zur Wahl einer nichtlinearen Reskalierung	188
11.3.8	Praktische Vorgehensweise	189
11.4	Regression unter Nebenbedingungen	189

11.4.1	Restringierte KQ-Schätzung	189
11.4.2	Verteilung des restringierten KQ-Schätzers	190
11.4.3	Anwendung auf die Flußanalyse	191
12	Beispielauswertung	192
12.1	Stoffflußanalyse bei <i>Corynebacterium glutamicum</i>	192
12.1.1	Das metabolische Netzwerk	193
12.1.2	Freie Flüsse und extrazelluläre Flußdaten	194
12.1.3	Markierungsdaten	196
12.1.4	Parameterschätzung	197
12.2	Statistische Analyse	200
12.2.1	Vertrauensbereiche für Netto- und Austauschflüsse	200
12.2.2	Korrelationsanalyse	201
12.2.3	Hauptachsenanalyse	202
12.2.4	Sensitivitätsanalyse	204
12.2.5	Anpassungsgüte und Ausreißertest	204
12.2.6	Versuchsplanung	205
12.3	Ein instationäres Markierungsexperiment	206
12.3.1	Experimentelle Bedingungen	206
12.3.2	Abschätzung der Zeitkonstanten	207
12.4	Vergleich der experimentellen Ansätze	209
13	Schlußdiskussion und Ausblicke	210
13.1	Biologische Resultate	210
13.2	Austauschflüsse	211
13.3	Problematik der Modellbildung	211
13.4	Einsatz mathematischer Methoden	212
13.5	Versuchsplanung	214

Teil I

Experimentelle Grundlagen

Kapitel 1

Metabolische Kohlenstoff-Markierungsexperimente

Die ersten beiden Kapitel führen in die allgemeine Problematik der stationären Stoffflußanalyse sowie in die experimentelle Technik der Kohlenstoff-Markierungsexperimente zur Quantifizierung intrazellulärer Stoffflüsse ein. Dabei wird in diesem Kapitel auf das grundlegende Funktionsprinzip eines Markierungsexperiments eingegangen und die in dieser Arbeit verwendete Terminologie erläutert. Mit den experimentellen und meßtechnischen Details der Durchführung eines Markierungsexperiments beschäftigt sich das folgende Kapitel.

Nach einer Einführung in die Zielsetzung der Stoffflußanalyse und die Funktionsweise von Markierungsexperimenten soll in diesem Kapitel zunächst der bei einem Markierungsexperiment erhobene Datensatz charakterisiert werden. Dabei werden die für die vorliegende Arbeit zentralen Begriffe der intra- und extrazellulären Flüsse sowie der Positions- und Isotopomeren-Markierungsanteile in intrazellulären Metabolit-Pools sorgfältig definiert.

Abschließend wird anhand eines sehr einfachen Beispiels gezeigt, wie mit Hilfe von Markierungsexperimenten grundsätzlich Aufschlüsse über intrazelluläre metabolische Vorgänge gewonnen werden können. Dabei wird sich als zentrales Ergebnis zeigen, daß Markierungsdaten Informationen über intrazelluläre Flüsse enthalten, die derzeit mit keiner anderen Meßtechnik gewonnen werden können.

1.1 Stationäre Stoffflußanalyse

1.1.1 Stoffwechselphysiologische Forschung und Metabolic Engineering

Stoffwechselphysiologische Untersuchungen am lebenden Organismus stellen seit langem eine besondere Herausforderung für die biologische Grundlagenforschung dar [BS82]. Bis heute ist die quantitative Kenntnis der dynamischen Abläufe in einer lebenden Zelle eher gering im Gegensatz zum aktuellen molekularbiologischen Wissensstand. Voraussetzung für die Erforschung der zellulären Regulationsmechanismen ist u.a. die Entwicklung meßtechnischer Methoden zur In-vivo-Untersuchung lebender Zellen.

In den letzten Jahren gewinnen derartige Ergebnisse der Grundlagenforschung aber auch für die biotechnologische Anwendung an Bedeutung. Mikroorganismen werden in der Biotechnologie auf vielfältige Weise zur Produktion organischer Komponenten eingesetzt. Vor allem bei solchen Stoffen, die in großen Mengen produziert werden (z.B. Aminosäuren, organische Säuren oder Ethanol), spielen wirtschaftliche und damit auch quantitative Aspekte eine besondere Rolle [Del87, EG96]. Bei derartigen Produktionsverfahren können bereits geringe, auf fundiertem

Wissen beruhende Verbesserungen der Produktivität entscheidende Vorteile erbringen.

Mit der zunehmenden Verfügbarkeit gentechnischer Methoden zur Veränderung von Stoffwechselnetzwerken hat sich in den letzten Jahren ein Forschungszweig etabliert, der mit dem Sammelbegriff *Metabolic Engineering* [Bai91, SS93, KA93, EG96] bezeichnet wird. Darunter werden Methoden zur gezielten Verbesserung von Stoffwechselleistungen zusammengefaßt. Der Ingenieuransatz betont hierbei eine quantifizierbare und reproduzierbare Arbeitsweise, die nicht zuletzt auf mathematische Modellansätze zur Beschreibung der Regulation metabolischer Systeme zurückgreift.

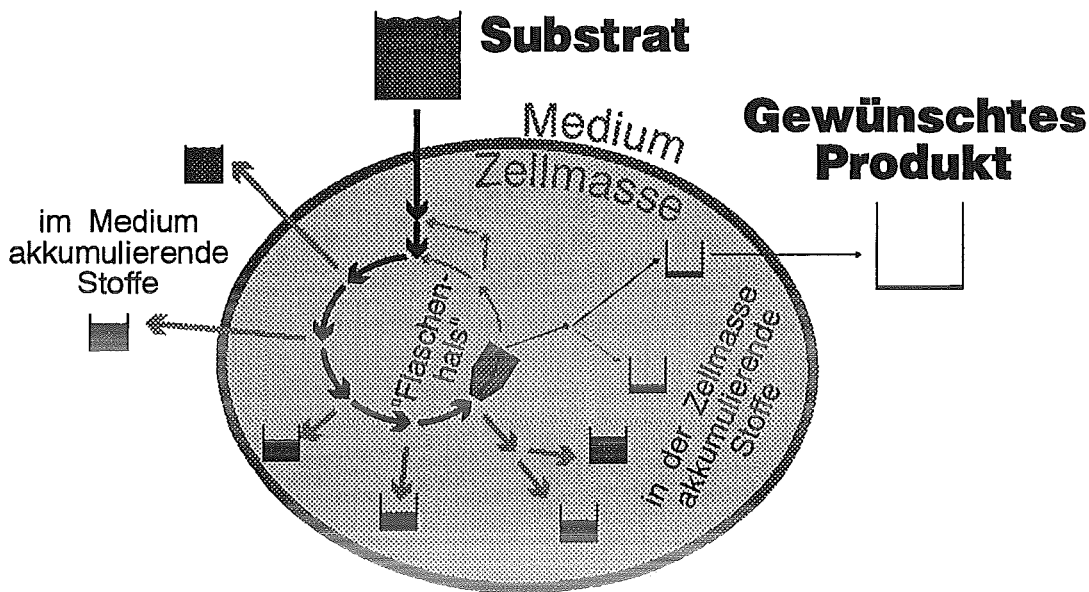


Abbildung 1.1: Schematisierte Darstellung der Zielsetzung einer stationären Stoffflußanalyse: Die Flüsse vom Substrat zum gewünschten Produkt sollen quantifiziert werden. Mit gentechnischen Methoden kann versucht werden, die Ausbeute zu verbessern.

1.1.2 Zielsetzung der stationären Stoffflußanalyse

In diesem Abschnitt sollen vorweg einige wichtige, im Zusammenhang mit der Stoffflußanalyse verwendete Begriffe definiert werden (vgl. auch [Val91]).

Definition 1.1 (metabolische Stationarität, stationäre Stoffflußanalyse) *Ein Stoffwechselsystem befindet sich in einem metabolisch stationären Zustand, wenn alle intrazellulären Stoffflüsse — und damit auch die intrazellulären Metabolit-Konzentrationen — zeitlich konstant sind. Eine stationäre Stoffflußanalyse ist eine auf Meßdaten und Bilanzgleichungen¹ gestützte Methode zur Quantifizierung von Stoffflüssen in einem Stoffwechselabschnitt einer Zelle unter metabolisch stationären Bedingungen.*

Mit den heute verfügbaren experimentellen und meßtechnischen Methoden kann vor allem der Zentralstoffwechsel einer Zelle als ganzer oder in seinen Teilen analysiert werden.

¹Vgl. dazu Kapitel 4.

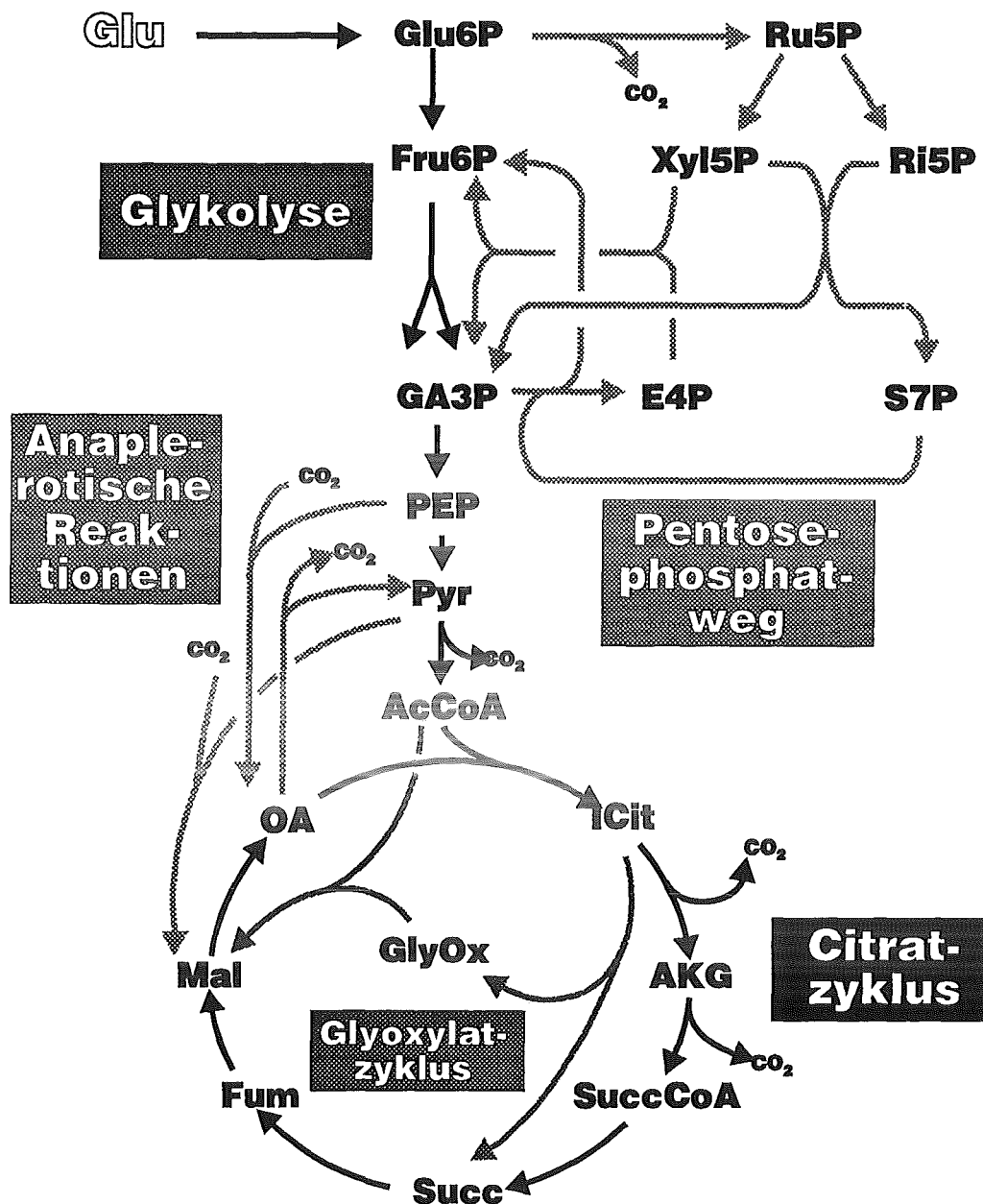


Abbildung 1.2: Repräsentatives metabolisches Netzwerk mit den wichtigsten Metaboliten und Reaktionsschritten des Zentralstoffwechsels. Einige Produktionsschritte sind verkürzt wiedergegeben (vgl. dazu Tabelle 3.2 und 3.3), andere spezifisch für bestimmte Organismen. Zu den Stoffabkürzungen vgl. Tabelle 3.1.

Definition 1.2 (Zentralstoffwechsel) *Im Rahmen der vorliegenden Arbeit sollen unter dem Zentralstoffwechsel einer Zelle die katabolischen Stoffwechselwege der Glykolyse, des Pentosephosphatwegs und des Citrat- und Glyoxylatzyklus zusammengefaßt werden sowie die im wesentlichen anabolischen Wege der anaplerotischen Reaktionen, der Aminosäureproduktion und der Produktion organischer Säuren und Alkohole (vgl. auch Abb. 1.2).*

Eine über den Zentralstoffwechsel hinausgehende quantitative Analyse von Stoffflüssen ist dagegen selten aussichtsreich, weil Flüsse außerhalb des Zentralstoffwechsels meist zu gering sind, um sie mit derzeit verfügbaren Meßinstrumenten exakt zu erfassen [Val91, Sch91, Fie84].

Die nachfolgende Differenzierung zwischen intrazellulären und extrazellulären Flüssen ist insofern von besonderer Bedeutung, als extrazelluläre Flüsse meist direkt gemessen werden können (vgl. Abschnitt 2.1), während die intrazellulären Flüsse die zu bestimmenden Größen einer Stoffflußanalyse sind (vgl. auch Abb. 1.6).

Definition 1.3 (intra- und extrazelluläre Metabolite, intra- und extrazelluläre Flüsse) *Als extrazellulärer Metabolit in bezug auf einen betrachteten Stoffwechselabschnitt werden solche Stoffe bezeichnet, die jeweils von außen zugeführt oder nach außen abgeführt werden, also insbesondere Substrate, Produkte und Komponenten für die Biomassesynthese². Die entsprechenden Flüsse (also z.B. Substrataufnahme und Produktausscheidung) heißen extrazelluläre Flüsse. Die Intermediate des betrachteten Abschnitts heißen intrazelluläre Metabolite, die sie verbindenden Flüsse intrazelluläre Flüsse.*

1.1.3 Anwendungen im Rahmen des Metabolic Engineering

Die stationäre Stoffflußanalyse sollte als ein Beitrag zum Metabolic Engineering verstanden werden. Denn die Verfügbarkeit einer quantitativen Methode zur intrazellulären Stoffflußanalyse ermöglicht die Beantwortung zentraler Fragen im Prozeß der Veränderung von Stoffwechselleistungen (vgl. auch Abb. 1.1):

- Nachweis und Quantifizierung von Enzymaktivitäten und in beide Richtungen ablaufenden Reaktionsschritten in vivo [RDC⁺95].
- Charakterisierung verschiedener physiologischer Zustände mit dem Ziel einer empirischen quantitativen Beschreibung von regulatorischen Mechanismen [VS93, Jor95, SSGM⁺95].
- Detaillierte quantitative Untersuchung von Hybridoma-Zellen oder anderen gentechnisch manipulierten Mikroorganismen [MST⁺94, ZS95].
- Systematische Kontrollanalyse basierend auf einer metabolischen Kontrolltheorie [Kac88, KW86, AW92].

Damit steht ein Werkzeug zur Verfügung, welches eine detaillierte quantitative Analyse der Auswirkung gentechnischer Manipulationen ermöglicht. Dies stellt einen erheblichen Fortschritt im Vergleich zu einer auf Mutmaßungen und globalen Befunden beruhenden "Trial-and-Error"-Methode dar. Zwar erlaubt die Stoffflußanalyse nach wie vor keine *Vorhersage* der Auswirkung von Veränderungen, jedoch können auf der Basis des gemessenen Datenmaterials gezielter solche Experimente vorgeschlagen werden, die in Bezug auf die jeweilige Fragestellung aussichtsreich sind. Genau dies ist die Zielsetzung des Metabolic Engineering.

²Man beachte, daß die Systemgrenze hier *nicht* durch die Zellwand definiert ist, sondern durch die Unterscheidung zwischen betrachteten und nicht betrachteten Stoffwechselkomponenten.

1.1.4 Angewendete Methoden

Grundsätzlich lassen sich zwei experimentelle Ansätze zur stationären Stoffflußanalyse unterscheiden, die in den folgenden Abschnitten noch detailliert beschrieben werden (vgl. auch Abb. 1.6):

1. In der Bioverfahrenstechnik hat in den letzten Jahren die sogenannte *Metabolit-Bilanzierungstechnik* große Aufmerksamkeit erregt. Diese Methode basiert auf der alleinigen Messung aller extrazellulären Flüsse mit Standard-Meßinstrumenten [VS93, GFJA93] [HHHR⁺94, VP94a, Jor95].
2. Im Bereich der physiologisch/biochemischen Forschung werden dagegen seit langer Zeit *Markierungsexperimente* mit isopenangereicherten Substraten für den Nachweis und die Quantifizierung intrazellulärer Flüsse verwendet. Dabei spielt heute die NMR-Meßtechnik eine entscheidende Rolle [WHK⁺82, CSKW83, WK84, MSJ88].

Beide Methoden weisen – jeweils allein eingesetzt – einige Unzulänglichkeiten auf, die nur mit ergänzendem Datenmaterial überwunden werden können. So hat sich z.B. herausgestellt, daß die Anzahl der meßtechnisch zugänglichen extrazellulären Flüsse nicht ausreicht, um eine komplette Flußquantifizierung zu erreichen. Daher werden in der Regel weitere Annahmen über das metabolische System getroffen (vgl. Abschnitt 2.4). Andererseits kann mit Markierungsexperimenten stets nur eine relative Flußquantifizierung erreicht werden, d.h. die absolute Größe von Stoffflüssen ist allein aus Markierungsdaten nicht berechenbar (vgl. Satz 7.3).

Die in der Literatur beschriebenen Experimente sind in der Regel Mischformen aus beiden Ansätzen, wobei meist eine der beiden Methoden deutlich dominiert. Erst in jüngster Zeit kam es zu einer engeren Integration — allerdings ist dazu ein erheblicher technischer Aufwand erforderlich, der sowohl den Einsatz verfahrenstechnischer Methoden als auch der NMR-Technik umfaßt [STM⁺94, ZS95, MGWE⁺96, Szy95, WG96a].

Derzeit stellt diese Kombination aus der Markierungs- und der Metabolit-Bilanzierungsmethode die leistungstärkste Technik zur Gewinnung stationärer intrazellulärer Flußdaten dar. Sie liefert sowohl einen vollständigen Satz extrazellulärer Flüsse als auch eine große Zahl von Markierungsdaten unter definierten Bedingungen und von hoher meßtechnischer Qualität (vgl. Abb. 1.6). Ausgehend von diesen Daten gelingt — wie in dieser Arbeit gezeigt wird (vgl. auch [MGWE⁺96]) — eine vollständige Quantifizierung aller intrazellulären Flüsse im Zentralstoffwechsel einer Zelle, ohne daß umfangreichere Modellannahmen über das lebende System getroffen werden müssen.

1.1.5 Markierte Substanzen

Während extrazelluläre Flüsse mit chemisch-analytischen Methoden leicht bestimmt werden können (vgl. Abschnitt 2.1), erfordert die Gewinnung von Isotopen-Anreicherungsdaten die Durchführung eines *Markierungsexperiments* [And83]. Voraussetzung hierfür ist die Verfügbarkeit von ¹³C- oder ¹⁴C-markierten chemischen Stoffen.³

³Isotope anderer Atome, wie z.B. ²H oder ¹⁵N spielen für die Stoffflußanalyse eine untergeordnete Rolle (vgl. dazu auch Abschnitt 1.1 und [KWR88, RCRL90, Lon92, MBW94, RKB94]). Diese Arbeit konzentriert sich daher auf Kohlenstoff-Markierungsexperimente. Außerdem eignen sich natürlich vorhandene Isotope hervorragend für die direkte Messung intrazellulärer Konzentrationen, jedoch nicht für Markierungsexperimente, da diese auf der Unterscheidung markierter und unmarkierter Atome beruhen. Beispielsweise tritt Phosphor in der Natur zu 100 % in der im NMR-Gerät beobachtbaren ³¹P-Form auf (vgl. auch [Coh83, MTBK89, GSB90, AAC⁺91, GSPW⁺91, SGSW⁺93, LVHHV95]).



Für die wichtigsten metabolischen Substrate sind nahezu beliebig markierte Varianten verfügbar. Man spricht beispielsweise von $[1-^{13}\text{C}]$ -Glucose oder $[5,6-^{13}\text{C}_2]$ -Glucose, wenn das Glucosemolekül auf der ersten oder auf der fünften und sechsten C-Atom-Position markiert ist (Abb. 1.3). Der Preis eines Gramms markierter Glucose bewegt sich derzeit zwischen 160 DM (für $[1-^{13}\text{C}]$ -Glucose) und 7500 DM (für $[5,6-^{13}\text{C}_2]$ -Glucose).

Bei einem *Markierungsexperiment* [And83] wird ein Substrat des Stoffwechsels (in der Regel Glucose) durch eine markierte Variante ausgetauscht. Vom Moment der Zugabe des markierten Substrats an verteilt sich die Markierung über den gesamten Stoffwechsel. Besonders Kohlenstoff-Markierungen lassen sich aufgrund der hochgradigen Vernetzung des Systems nach kurzer Zeit in den meisten intrazellulären Metaboliten nachweisen.⁴ Die Isotopenanreicherung wird wie folgt durch relative Anteile quantifiziert (vgl. auch Abb. 1.5)⁵:

⁶Prozentsätze werden in der modellmäßigen Beschreibung als Zahlen im Intervall $[0, 1]$ aufgefaßt.

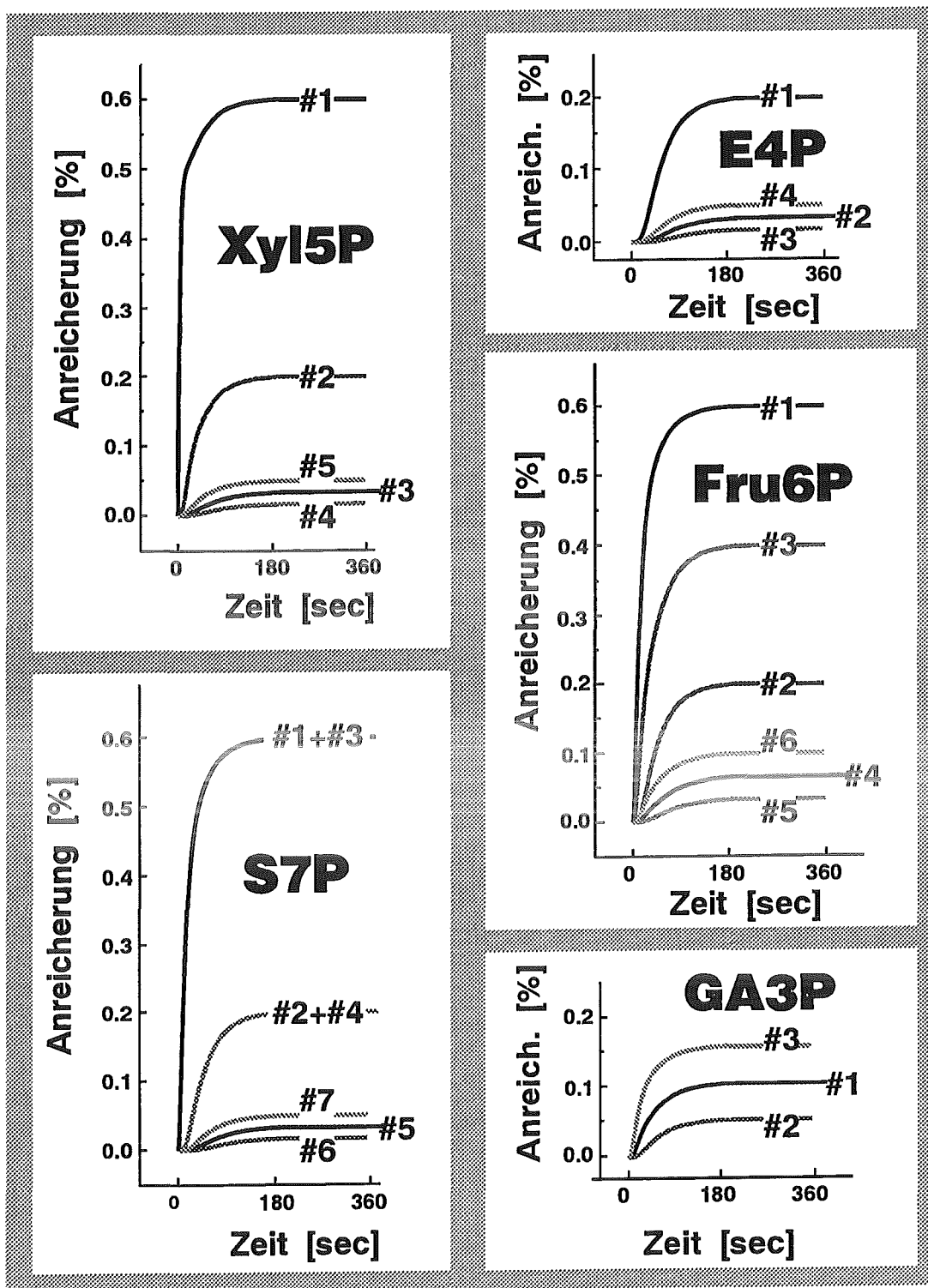


Abbildung 1.4: Simulierter zeitlicher Verlauf des isotopisch instationären Markierungszustands in den C-Atomen der Glykolyse und des Pentosephosphatwegs während eines Markierungsexperiments. Das zugrundeliegende Stoffwechselnetzwerk in Abbildung 9.1 beschreibt die xyloseverwertende Hefe *Pichia stipitis* [LPdPB88]. Als markiertes Substrat wurde $[1-^{13}\text{C}]$ -Xylose eingesetzt. Die Simulationsparameter sind in Abschnitt 9.2 angegeben.

Ist eine Verwechslung mit dem in Definition 1.7 eingeführten Isotopomeren-Anteil ausgeschlossen, so wird kurz von einem Markierungsanteil an einer C-Atom-Position gesprochen.

Im folgenden wird bei allen Experimenten ein metabolisch stationärer Zustand vorausgesetzt. Im Gegensatz dazu sind die intrazellulären Markierungsanteile während des Experiments jedoch zeitabhängige Größen. Abbildung 1.4 zeigt die Simulation eines solchen instationären Verlaufs im Zuckerstoffwechsel⁷. Wie in Abschnitt 7.1 ganz allgemein bewiesen wird, stabilisieren sich die Markierungsanteile nach einiger Zeit, so daß das System sich asymptotisch einem isotopisch stationären Zustand annähert.

Definition 1.6 (isotopische Stationarität) *In einem isotopisch stationären Zustand ist für jede intrazelluläre C-Atom-Position die Menge der pro Zeiteinheit zu- und abfließenden Markierung gleich. Insbesondere sind alle Markierungsanteile stabil.*

Während des Experiments befindet sich die Zelle also in einem metabolisch stationären Zustand, durchläuft aber einen isotopisch instationären Zustand, der schließlich in einen isotopisch stationären Zustand einmündet.

1.1.7 Positionsmarkierung und Isotopomere

Bisher wurde die Quantifizierung der Markierungsanteile im metabolischen System auf die C-Atom-Positionen intrazellulärer Metabolite bezogen. Wie nachfolgend gezeigt wird, kann mit Hilfe moderner Meßtechnik jedoch noch zusätzliche Information über die Isotopenverteilung gewonnen werden. Dies ist dann der Fall, wenn im Stoffwechsel mehrfach markierte Moleküle entstehen.

Definition 1.7 (Isotopomer, Isotopomeren-Anteil) *Ein Molekül mit n Kohlenstoff-Atomen kann in insgesamt 2^n verschiedenen Markierungszuständen bzgl. der Isotope ^{12}C und ^{13}C angetroffen werden (vgl. Abb. 1.5a). Diese Zustände werden als Isotopomere des Moleküls bezeichnet. Der Prozentsatz eines bestimmten Isotopomers bezogen auf die Menge aller Moleküle des entsprechenden Stoffs heißt Isotopomeren-Anteil.*

Isotopomeren mit Mehrfachmarkierungen können im Stoffwechsel dann auftreten, wenn markierte Metabolite, die einen Zyklus durchlaufen haben, wieder mit dem markierten Substrat reagieren [JRMS91, CKS90]. Es ist daher zu erwarten, daß die Kenntnis der Isotopomeren-Verteilung weiteren Aufschluß über die Flußverteilung in metabolischen Zyklen gibt. Isotopomeren-Verteilungen können bis zu einem gewissen Grad mit Hilfe der NMR-Meßtechnik oder der Massenspektroskopie gemessen werden (siehe hierzu Definition 2.2 im folgenden Kapitel).

Aus der Kenntnis der Isotopomeren-Verteilung können unmittelbar die Positions-Markierungsanteile berechnet werden (vgl. Abb. 1.5b). Die Gesamtmarkierungsanteile auf der Kohlenstoff-Atom-Position i ergibt sich aus der Summe aller Anteile der verschiedenen Isotopomere, die auf dieser Position markiert sind.

1.1.8 Rohdaten für die Stoffflußanalyse

Zuverlässiges und aussagekräftiges physiologisches Datenmaterial kann nur auf der Basis von Meßmethoden gewonnen werden, die das lebende System nicht beeinflussen ("zerstörungsfreie" Methoden). Grundsätzlich können drei verschiedene Typen von Meßmethoden unterschieden werden, deren jeweilige Problematik an anderer Stelle noch detaillierter diskutiert wird:

⁷Mit dem Begriff Zuckerstoffwechsel wird die Glykolyse und der Pentosephosphatweg zusammengefaßt, also die Reaktionen zum Hexose- und Pentoseabbau.

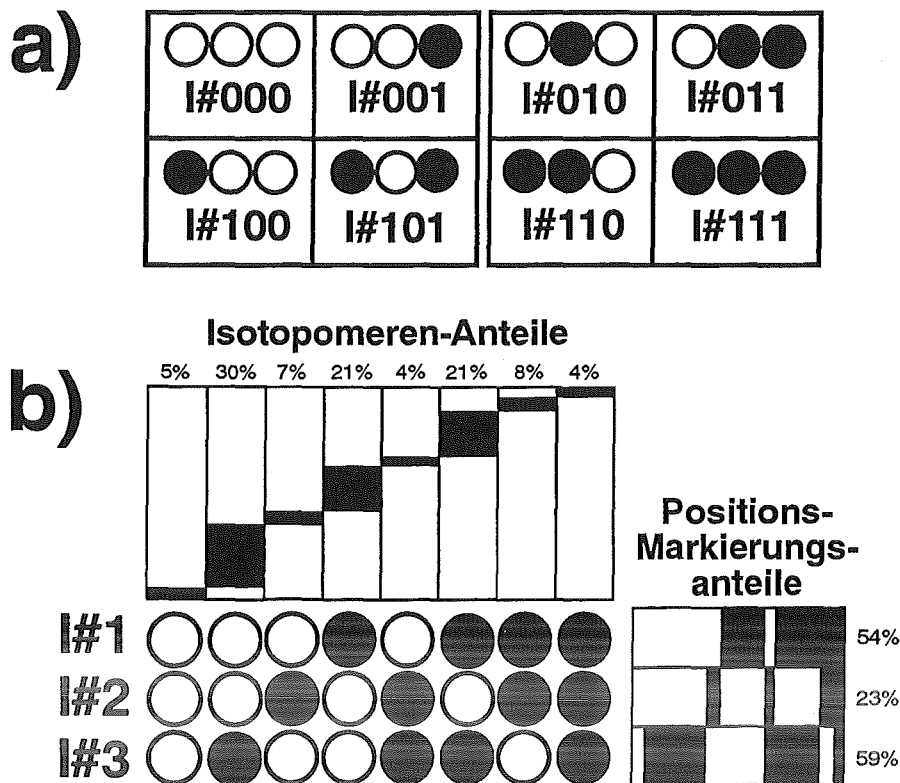


Abbildung 1.5: a) Die $2^3 = 8$ verschiedenen ^{13}C -Isotopomeren eines Moleküls I mit 3 Kohlenstoff-Atomen. Weiße Kreise entsprechen ^{12}C , schwarze Kreise ^{13}C . b) Eine denkbare Isotopomeren-Verteilung, quantifiziert durch die Isotopomeren-Anteile an der Gesamtstoffmenge.

1. Alle mengenmäßig bedeutsamen extrazellulären Stoffflüsse können heute mit Standardinstrumenten der Bioverfahrenstechnik aus Konzentrationsdaten bestimmt werden [Fie84, BO86, Sch91] (vgl. Abschnitt 2.1).
2. Mit Hilfe der NMR-Meßtechnik sowie der Massenspektroskopie lassen sich ^{13}C -Atome in einer bestimmten Position sowie Isotopomere ohne chemische Vorbehandlung der Metabolite nachweisen⁸ (vgl. dazu Abschnitt 2.2). Darüber hinaus kann die NMR-Technik auch auf Stoffgemische angewendet werden, im Extremfall sogar auf ein lebendes System [GWPS⁺92, CB93, WG96a]. Man spricht dann im engeren Sinne von einer *In-vivo-NMR-Methode*. Es hängt jedoch vom experimentellen Ansatz ab, ob die isotopisch instationären Zustände des Systems im Verlauf des Markierungsexperimentes beobachtet werden können oder nur der abschließende isotopisch stationäre Zustand. Isotopomeren-Anteile können gewöhnlich nur bei längeren Meßzeiten im isotopisch stationären Zustand beobachtet werden.
3. Mit Hilfe einer chemischen Analyse können intrazelluläre Stoffkonzentrationen gemessen werden (vgl. dazu Abschnitt 2.1). Diese Konzentrationsdaten werden für die Auswer-

⁸Die klassische Methode der ^{14}C -Bestimmung auf der Basis von Radioaktivitäts-Messungen ist mit einem extrem hohen Aufwand verbunden, da alle Substanzen chemisch degradiert werden müssen, um die Kohlenstoff-Atome zu isolieren [SB79, CRSK81, DRM⁺93]. Isotopomeren-Messungen sind praktisch undurchführbar. Die ^{14}C -Methode wird daher zur Stoffflußanalyse heute kaum noch angewendet.

tung isotopisch instationärer Experimente benötigt [WG96a]. In vielen Fällen sind intrazelluläre Konzentrationen aber auch aufgrund von zusätzlichen NMR-Messungen direkt bestimmbar. Darüberhinaus verwenden einige Autoren thermodynamische Gleichgewichte zur Abschätzung der Konzentrationen [PSMC93].

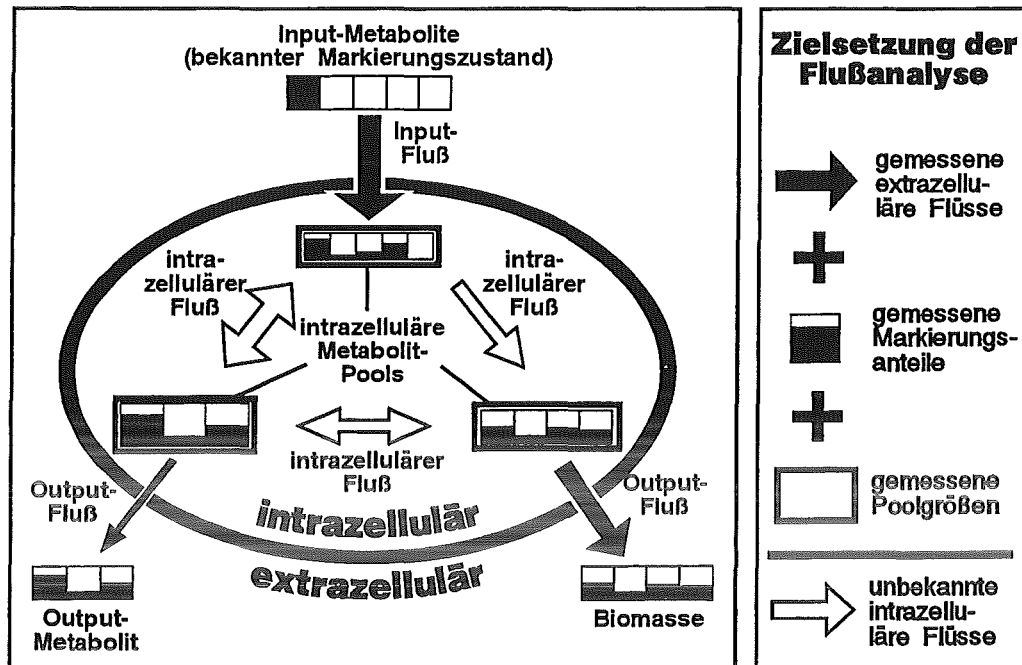


Abbildung 1.6: Schematische Darstellung der verfügbaren experimentellen Daten für die Auswertung eines Markierungsexperiments. Zu den extrazellulären Kompartimenten zählt auch die Zellmasse. Im isotopisch instationären Fall liegen die Markierungsdaten zeitlich aufgelöst vor.

1.1.9 Typische experimentelle Anordnungen

Der aus der Anwendung der beschriebenen Meßtechniken resultierende Datensatz für die Bestimmung stationärer intrazellulärer Stoffflüsse ist in Abb. 1.6 visualisiert. Ist dieser Datensatz hinreichend ergiebig, können alle gesuchten Flüsse unter Rückgriff auf nur wenige Annahmen über das lebende System quantifiziert werden.

Einige typische experimentelle Anordnungen zur Flußquantifizierung können unterschieden werden:

1. Lediglich extrazelluläre Flüsse werden gemessen (reines Metabolit-Bilanzierungsexperiment) [VS93, GFJA93, Jor95].
2. Lediglich die Markierungsanreicherungen in einem isotopisch stationären Zustand werden gemessen (reines Markierungsexperiment) [WHK⁺82, WK84, SEGS93].
3. Extrazelluläre Flußmessungen werden ergänzt durch die Messung von Markierungsanteilen im isotopisch stationären Zustand [SB79, JWBL89, STM⁺94, MGWE⁺96, ZS95].
4. Sowohl extrazelluläre Flüsse als auch der Zeitverlauf der Markierungsanteile werden in einem isotopisch instationären Zustand beobachtet [KKW79, FHBS90, WGKF⁺92] [CFG95].

5. Sowohl extrazelluläre Flüsse als auch die Isotopomeren-Verteilungen werden in einem isotopisch stationären Zustand quantifiziert [MSJ88, KWL93, KCS93, DRM⁺93].

In einigen Fällen wurde darüber hinaus das Experiment in demselben physiologischen Zustand mit unterschiedlich markierten Substraten wiederholt, wodurch sich die verfügbare Information erheblich erhöht [SB79, RB85, CB83].

1.2 Ein instruktives Beispiel

Dieser Abschnitt führt anhand eines elementaren Beispiels aus [WG96c] in die Auswertung von Markierungsexperimenten mit Hilfe von Modellgleichungen ein.

1.2.1 Metabolit-Pools

Die mathematische Modellierung von Flußsystemen beruht auf der Einteilung in sogenannte Kompartimente [And83, Hol85], die in dieser Arbeit auch als Pools bezeichnet werden. Ein Pool stellt eine modelltechnische Vereinfachung eines räumlich verteilten Systemzustands dar ⁹.

Definition 1.8 (Metabolit-, C-Atom-, Isotopomeren-Pool) *Ein Metabolit-Pool ist die Gesamtheit aller Moleküle einer bestimmten Stoffgruppe innerhalb eines räumlichen Abschnitts des metabolischen Systems. Analog wird die Gesamtheit von C-Atomen an derselben Position eines Metaboliten oder von in derselben Weise markierten Isotopomeren als C-Atom-Pool oder Isotopomeren-Pool bezeichnet. Es wird angenommen, daß innerhalb eines Pools eine räumlich homogene Stoffverteilung vorliegt (vgl. dazu auch Abschnitt 2.3).*

In dieser Arbeit ist der räumliche Abschnitt durch das gesamte Cytosol der Zellen definiert, da Prokaryonten untersucht werden. Ein Beispiel für ein enger eingegrenztes Kompartiment ist z.B. der intramitochondriale Raum bei Eukaryonten [BM88a].

1.2.2 Flußschätzung auf der Grundlage extrazellulärer Flußmessungen

Das in Abb. 1.7 dargestellte elementare Beispiel soll aufzeigen, wie intrazelluläre Stoffflüsse auf der Basis von extrazellulären Flußmessungen und der Messung von Markierungsanteilen oder Isotopomeren-Verteilungen bestimmt werden können. Das Beispiel geht von einem Stoffwechselabschnitt mit einem Substrat S, einem Intermediat A und einem (intrazellulären) Produkt P aus, daß im nächsten Schritt ausgeschieden wird. Der einfachste Fall eines linearen Stoffwechselweges von S nach P ist in Abb. 1.7a dargestellt. Hier bezeichnet v_S die (extrazelluläre) Substrataufnahmerate, v_P die (extrazelluläre) Produktbildungsrate und v_A den einzigen intrazellulären Fluß.

Nimmt man an, daß sich das System in einem metabolisch stationären Zustand befindet und die Produktbildung v_P meßbar ist, so fließt die gleiche Stoffmenge in den Pool A hinein und hinaus. Für den intrazellulären Fluß v_A gilt folglich

$$v_A = v_P \quad .$$

⁹Auf die Problematik der Kompartimentierung und der Homogenitätsannahme wird in Abschnitt 2.3 noch eingehender eingegangen.

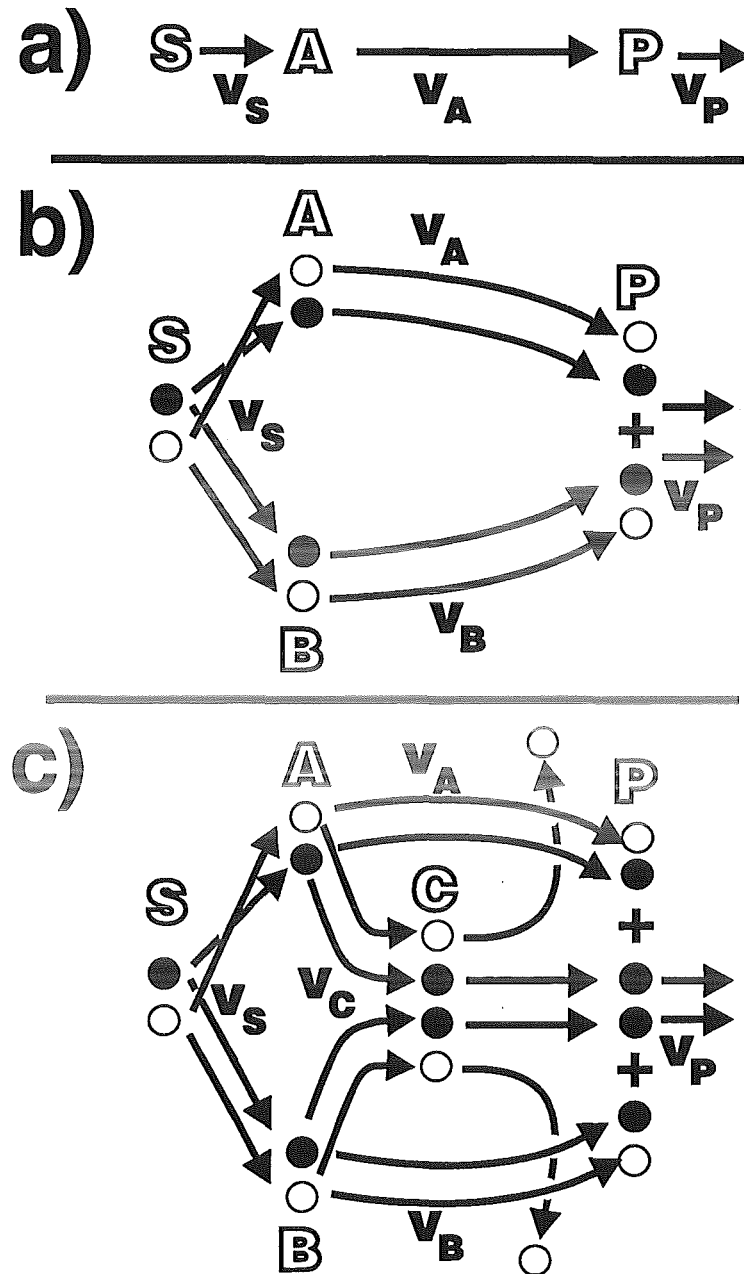


Abbildung 1.7: Elementares Beispiel zur Verdeutlichung des Informationsgehaltes von Meßwerten für extrazelluläre Flüsse, Positions-Markierungsanteile und Isotopomeren-Anteile. Im Übergang von a) nach b) nach c) wird je eine intrazelluläre Reaktion hinzugenommen.

Damit wird ein intrazellulärer Fluß aufgrund einer extrazellulären Flußmessung zugänglich. Dieser Bilanzierungsansatz wird auch als *Metabolit-Bilanzierungstechnik* bezeichnet (vgl. [Hof86, VS93, VP94a]).

Liegt zusätzlich ein Meßwert für die Substrataufnahme v_S vor, so gilt offenbar $v_A = v_S = v_P$, d.h. die beiden Messungen von v_S und v_P sind redundant. Für die Praxis kann dies verschiedene Konsequenzen haben:

1. Der experimentelle Aufwand kann durch Verzicht auf eine Messung reduziert werden.
2. Die statistische Qualität der Flußschätzung ist durch Regressionsrechnung (hier einer einfachen Mittelwertbildung) verbesserbar [BW88, SW89].
3. Grobe Meßfehler können durch Vergleich der Meßdaten erkannt werden [NS83] [HRHH⁺94].
4. Eine Verletzung der Redundanzbeziehung kann aber auch auf Modellfehler (z.B. unberücksichtigte Flüsse) hinweisen [Him78, AC74]:

Auf die Redundanzanalyse wird in den Abschnitten 4.3 und 11.2 noch eingegangen.

1.2.3 Hinzunahme von Positions-Markierungsdaten

In Abb. 1.7b wurde das Beispielnetzwerk um ein zweites intrazelluläres Intermediat B erweitert, das einen alternativen Stoffwechselweg parallel zu dem bereits vorhandenen eröffnet. Offenbar ist es dann nicht mehr möglich, aus den gemessenen extrazellulären Flüssen die beiden intrazellulären Flüsse v_A , v_B zu bestimmen. Hier hilft die Kenntnis von Positions-Markierungsanteilen weiter. Voraussetzung hierfür ist, daß die C-Atome in den beiden parallelen Stoffwechselwegen auf verschiedene Weise in das Produkt überführt werden. Offenbar ist dies der Fall, denn ihre Position wird im zweiten Weg vertauscht, während sie im ersten Weg unverändert bleibt.

Die prozentualen Anreicherungen der ^{13}C -Markierung auf den beiden Kohlenstoff-Positionen von P seien mit p_1, p_2 bezeichnet. Das Verhältnis dieser beiden Markierungsanteile entspricht genau dem Verhältnis der beiden Flüsse v_A und v_B :

$$\frac{v_A}{v_B} = \frac{p_1}{p_2}$$

Dies zeigt bereits, wie mit Hilfe von Markierungsdaten Informationen über das Verhältnis intrazellulärer Stoffflüsse gewonnen werden können.

Die Kenntnis der Produktbildungsrate v_P kann nun dazu verwendet werden, auch die absoluten Flüsse durch die beiden Wege zu quantifizieren, denn aus der Metabolit-Bilanz $v_A + v_B = v_P$ für den Pool P folgt:

$$v_A = p_1 \cdot v_P \quad \text{und} \quad v_B = p_2 \cdot v_P$$

Darüber hinaus wird deutlich, daß ohne Rückgriff auf extrazelluläre Flußdaten lediglich Flußverhältnisse bestimmbar sind.

Die beschriebene Verzweigungssituation tritt beispielsweise in *Corynebacterium glutamicum* auf, wo Lysin über zwei verschiedene Stoffwechselwege ausgehend von Pyruvat produziert wird [SSK⁺91]. Ein entsprechendes Argument wurde in [SEGS93] verwendet, um die Nutzung der beiden Stoffwechselwege zu quantifizieren (vgl. auch ein verwandtes Beispiel in [WK84]).

Offenbar erhielte man keine verwertbare Information, wenn die Markierung in beiden Ästen der Verzweigung auf derselben Position bliebe. Prinzipiell hängt die mittels Markierungsexperimenten zugängliche Flußinformation stark von der Häufigkeit von Positionswechseln in den

Kohlenstoff-Gerüsten der Metabolite ab. Von Vorteil ist hier die starke Vernetzung des Zentralstoffwechsels einer Zelle. Vielfach wird eine in die Glykolyse eingebrachte C-Atom-Markierung über das gesamte Netzwerk verteilt und ist schließlich in fast jeder Position auffindbar. Daraus resultiert die besondere Bedeutung von Kohlenstoff-Markierungsexperimenten für die Flußquantifizierung. Im Gegensatz dazu kann z.B. ^{15}N nur für ganz bestimmte Untersuchungen verwendet werden.

1.2.4 Hinzunahme von Isotopomeren-Daten

Das Beispiel wird nun nochmals erweitert, indem ein drittes intrazelluläres Intermediat C hinzugenommen wird (Abb. 1.7c). Dieses Intermediat wird in einem bimolekularen Reaktionsschritt aus den bereits eingeführten Intermediaten A und B gebildet. Aus C wiederum entsteht nach Abspaltung der äußeren Kohlenstoff-Atome das auf beiden Positionen markierte Isotopomer des Produkts P. Es soll angenommen werden, daß die von C abgespaltenen Atome in anderen Stoffwechselwegen Verwendung finden und deshalb nicht direkt nachgewiesen werden können. In dieser Situation reichen offenbar die bisher gemessenen Daten nicht aus, um alle drei intrazellulären Flüsse v_A , v_B , v_C zu bestimmen.

Charakteristisch für das erweiterte Netzwerk ist, daß im Produktpool ein zusätzliches Isotopomer entsteht. Die jeweiligen Anteile der drei entstehenden Isotopomere im Produktpool werden entsprechend der Abbildung 1.7 sinnfälligerweise mit p_{01} , p_{11} , p_{10} bezeichnet. Wegen $v_P = v_A + v_B + v_C$ gilt dann:

$$v_B = p_{01} \cdot v_P \quad v_C = p_{11} \cdot v_P \quad v_A = p_{10} \cdot v_P$$

Damit wird deutlich, daß Isotopomeren-Anteile einen höheren Informationsgehalt aufweisen können als Positions-Markierungsanteile. Umgekehrt lassen sich die Positions-Markierungsanteile stets aus der Isotopomeren-Verteilung rekonstruieren, da

$$p_1 = p_{10} + p_{11} \quad p_2 = p_{01} + p_{11}$$

Die Isotopomeren-Analyse ist also grundsätzlich mindestens so aussagekräftig wie die positionale Auswertung von Kohlenstoff-Markierungen. Voraussetzung für ihre Überlegenheit im Beispiel ist die dort auftretende bimolekulare Reaktion, die ein mehrfach markiertes Molekül erzeugt. Eine entsprechende Situation findet sich z.B. im Citratzyklus, wenn markiertes Acetyl-Coenzym A in das System eingebracht wird. Dies wurde für die Quantifizierung von Flüssen im Citratzyklus wiederholt verwendet [MSJ88, Lee93].

Der Informationsgewinn von Isotopomeren-Messungen gegenüber positionalen Messungen wird im Laufe dieser Arbeit noch an anderer Stelle thematisiert (vgl. Abschnitt 4.3).

1.2.5 Zyklische Wege und bidirektionale Reaktionsschritte

Aus den obigen Ausführungen wird deutlich, daß Markierungs- bzw. Isotopomeren-Messungen insbesondere für die Quantifizierung von Stoffflüssen in solchen metabolischen Zyklen von Bedeutung sind, die bei mehrfachem Durchlaufen verschiedene Isotopomere produzieren [MSJ88, KWL93, DRM⁺93]. Der kleinste in einem Stoffwechselnetz auftretende Zyklus ist ein bidirektionaler Reaktionsschritt, d.h. eine Reaktion, die in der Hin- und Rückrichtung zugleich abläuft (vgl. Definition 2.4). Solche bidirektionalen Reaktionen stellen keine thermodynamische Unmöglichkeit dar, vielmehr sind sie bei einem vergleichbaren Energieniveau von Substrat und

Produkt und geringer Aktivierungsenergie zu erwarten (Abb. 1.8) und tatsächlich in vivo beobachtet worden [RDC⁺95, WGM95]. Enzymkinetische Studien berücksichtigen diese Möglichkeit nur selten (vgl. z.B. [Seg75, DLC⁺84, MTBK89, WBA92]). Insbesondere wird meist nur die Netto-Reaktionsgeschwindigkeit des enzymatischen Schrittes beschrieben.

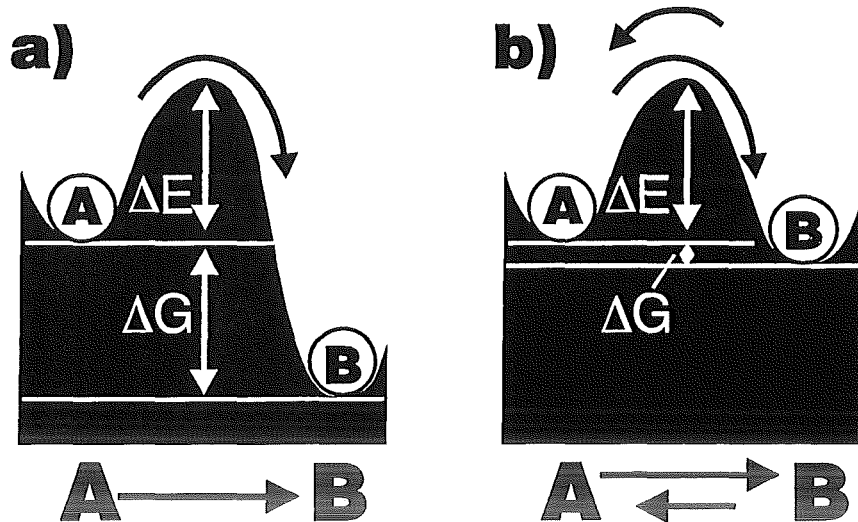


Abbildung 1.8: Unidirektional und bidirektional ablaufender Reaktionsschritt. Es hängt von der In-vivo-Energiedifferenz ΔG und der Aktivierungsenergie ΔE ab, ob und wie schnell die Reaktion in beiden Richtungen abläuft.

Viele bidirektionale Reaktionsschritte laufen in beiden Richtungen so schnell ab, daß die durch diese Schritte verbundenen Pools im Modell zu einem einzigen zusammengefaßt werden können (*Pool-Lumping*) (vgl. dazu auch Definition 2.5). Damit wird eine flexible Umverteilung von Vorläufermetaboliten für die Verwendung in den abzweigenden Stoffwechselwegen ermöglicht.

Allerdings ist auch der Fall denkbar, daß eine Reaktion nur in geringerem Umfang auch in der Rückrichtung abläuft. Über diese Zwischenstufen zwischen Bidirektionalität und Unidirektionalität ist in quantitativer Hinsicht wenig bekannt. Wie im weiteren Verlauf dieser Arbeit gezeigt wird, liegt eine besondere Stärke der Markierungstechnik darin, daß unter bestimmten Umständen beide Reaktionsgeschwindigkeiten eines bidirektionalen Schrittes quantifiziert werden können. Dies zeigt erneut, daß die Markierungstechnik der Metabolit-Bilanzierungstechnik grundsätzlich überlegen ist.

Kapitel 2

Experimentelle Grundlagen und Modellannahmen

Nachdem in Kapitel 1 die für eine Stoffflußanalyse benötigte Datenbasis charakterisiert wurde, muß nun geklärt werden, wie die entsprechenden Messungen konkret durchgeführt werden. Allerdings würde die Schilderung der praktischen Details des experimentellen Aufbaus, der NMR-Meßprozeduren und der chemischen Analysen den Rahmen dieser Arbeit sprengen. Hier wird jeweils auf die entsprechenden Originalpublikationen verwiesen. In diesem Kapitel werden nur diejenigen Details des Meßvorgangs diskutiert, die für Verständnis der nachfolgend gewählten modellmäßigen Beschreibung erforderlich sind. Ein Schwerpunkt liegt dabei auf der Beurteilung der statistischen Qualität der Meßwerte.

Die Konzeption von Experimenten steht in einer wechselseitigen Beziehung zu den grundlegenden Modellannahmen, die für die quantitative Beschreibung des Systems getroffen werden. Grundsätzlich ist die Auswertung von Daten aus einem biologischen Experiment nur unter mehr oder weniger stark vereinfachenden Annahmen möglich. Die sorgfältige Dokumentation dieser Annahmen ist Voraussetzung für ihre wissenschaftliche Diskussion und die Einschätzung der erzielten Ergebnisse. Ein derartiger Diskussionsprozeß kann zu einer Verfeinerung der Modellansätze führen, aber auch zu einer Veränderung des experimentellen Ansatzes, um die Gültigkeit der gemachten Annahmen zu gewährleisten. Es wird sich in Abschnitt 2.3 zeigen, daß bei vielen der in der Literatur beschriebenen Experimente die Gültigkeit der einen oder anderen dort getroffenen Modellannahme weder experimentell kontrolliert noch meßtechnisch überprüft werden kann.

Es ist ein besonderes Anliegen des Jülicher Stoffflußanalyse-Projekts, durch den interdisziplinären Zusammenschluß von Biologen, Ingenieuren und "Modellierern" zu einer neuen Qualität hinsichtlich der Gewährleistung wohldefinierter experimenteller Bedingungen bei gleichzeitiger Universalität der modellmäßigen Beschreibung vorzustoßen. Vor diesem Hintergrund ist nicht nur der in dieser Arbeit vorgestellte allgemeine Rahmen für die Modellierung metabolischer Kohlenstoff-Markierungssysteme, sondern auch der in Abschnitt 2.5 vorgestellte neuartige experimentelle Ansatz das Ergebnis eines langjährigen Diskussionsprozesses zwischen experimentellen und theoretischen Mitarbeitern des Jülicher Projekts. Die neue experimentelle Methode gewährleistet erstmals die Einhaltung und Überprüfung fast aller für die Kompartiment-Modellierung erforderlichen Bedingungen. Infolgedessen ist von den hiermit erzielten Ergebnissen eine bisher nicht erreichte Zuverlässigkeit zu erwarten.

Abschließend wird in diesem Kapitel die Beziehung zwischen einer metabolisch stationären und einer metabolisch instationären Analyse von Stoffwechselsystemen diskutiert. Es wird sich dabei zeigen, daß beide Ansätze komplementär sind, so daß zukünftige Fortschritte am ehesten

von einer Kombination der entsprechenden Techniken zu erwarten sind.

2.1 Standard-Meßtechnik

Unter dem Begriff *Standard-Meßtechnik* werden hier die in den meisten Fermenterlaboren verfügbaren Meßinstrumente [BO86, Sch91] und chemischen Analysemethoden zur Bestimmung von Stoffkonzentrationen, zum Aufschluß von Zellen und zur quantitativen Auftrennung von Stoffgemischen zusammengefaßt. Nicht zur Standardausrüstung gehören dagegen ein NMR-Instrument oder ein Massenspektrometer, die im folgenden Abschnitt behandelt werden.

2.1.1 Extrazelluläre Flüsse

Zur Konzentrationsmessung in flüssiger Phase werden verschiedene analytische Instrumente wie HPLC (High Performance Liquid Chromatography), GC (Gaschromatographie) oder FIA (enzymatische Flow Injektion Analysis) sowie Sonden zur Messung der Gelöstgaskonzentrationen verwendet. Die Biomassekonzentration kann aus einer Probe durch Trocknen und Wägen bestimmt werden. Schließlich erfolgt die Messung der Abgaszusammensetzung meist auf der Grundlage einer optischen Messung im Infrarotbereich. Auf dieser Datenbasis können die entsprechenden extrazellulären Flüsse wie nachfolgend beschrieben aus einfachen Massenbilanzen errechnet werden.

Die benötigten Bilanzgleichungen werden hier exemplarisch für einen kontinuierlich betriebenen Reaktor mit einer Biomassekonzentration X , der Konzentration S eines metabolischen Substrats und der Konzentration P eines Produkts angegeben (vgl. [Wie90] für eine allgemeinere Darstellung). Diese Konzentrationen hängen über die Durchflußrate D des Reaktors und die Konzentration S_{inp} des zulaufenden Substrats wie folgt mit den gesuchten biomassespezifischen Raten μ für das Wachstum, σ für die Substrataufnahme und π für die Produktbildung zusammen:

$$\begin{array}{rcl} \text{Änderung} & = & \begin{array}{c} \text{Produktion} \\ -\text{Verbrauch} \end{array} \quad -\text{Abfluß} \quad +\text{Zufluß} \\ \dot{X} & = & \mu \cdot X \quad -X \cdot D \\ \dot{S} & = & -\sigma \cdot X \quad -S \cdot D \quad +S_{\text{inp}} \cdot D \\ \dot{P} & = & \pi \cdot X \quad -P \cdot D \end{array} \quad (2.1)$$

Im metabolisch stationären Zustand des Systems erhält man hieraus unmittelbar:

$$\begin{pmatrix} \mu \\ \sigma \\ \pi \end{pmatrix} = \frac{D}{X} \cdot \begin{pmatrix} X \\ S_{\text{inp}} - S \\ P \end{pmatrix} \quad (2.2)$$

Folglich können die gesuchten Flüsse aus den gemessenen Konzentrationen im Reaktor direkt berechnet werden. Analoge Gleichungen gelten für die Gasbildung.

2.1.2 Einbeziehung der Biomassezusammensetzung

Nach Definition 1.3 gehören auch die anabolischen Abflüsse in die Biomasse zu den extrazellulären Flüssen. Ein wichtiger Beitrag zur Bestimmung intrazellulärer Stoffflüsse war die detaillierte Ermittlung auch dieser Flüsse auf der Basis einer Analyse der Zellmassenzusammensetzung [Hol86, Val91]. Die biologische Grundlage hierfür ist, daß der Biosyntheseweg jeder

einzelnen Zellkomponente eindeutig auf einen oder mehrere von zwölf Vorläufermetaboliten im Zentralstoffwechsel zurückverfolgt werden kann [NIS90]. Diese sind Glucose-6-phosphat, Fructose-6-phosphat, Ribose-5-phosphat, Erythrose-4-phosphat, Glycerinaldehyd-3-phosphat, 3-Phosphoglycerat, Phosphoenolpyruvat, Pyruvat, Acetyl-Coenzym A, α -Ketoglutarat, Succinyl-Coenzym A und Oxalacetat (vgl. Abb. 2.1).

Bei bekannter Zellzusammensetzung kann aus der bereits bestimmten Wachstumsrate μ direkt auf die zwölf von den Vorläufer-Pools ausgehenden anabolischen Stoffflüsse zurückgerechnet werden (Abb. 2.1). Ist der prozentuale Anteil des Vorläufers i an der Biomasse durch c_i gegeben, so gilt für den entsprechenden Biomasseabfluß:

$$\mu_i = c_i \cdot \mu, \quad i = 1, \dots, 12 \quad (2.3)$$

Dies führt zu einer erheblichen Verbesserung der verfügbaren Information.

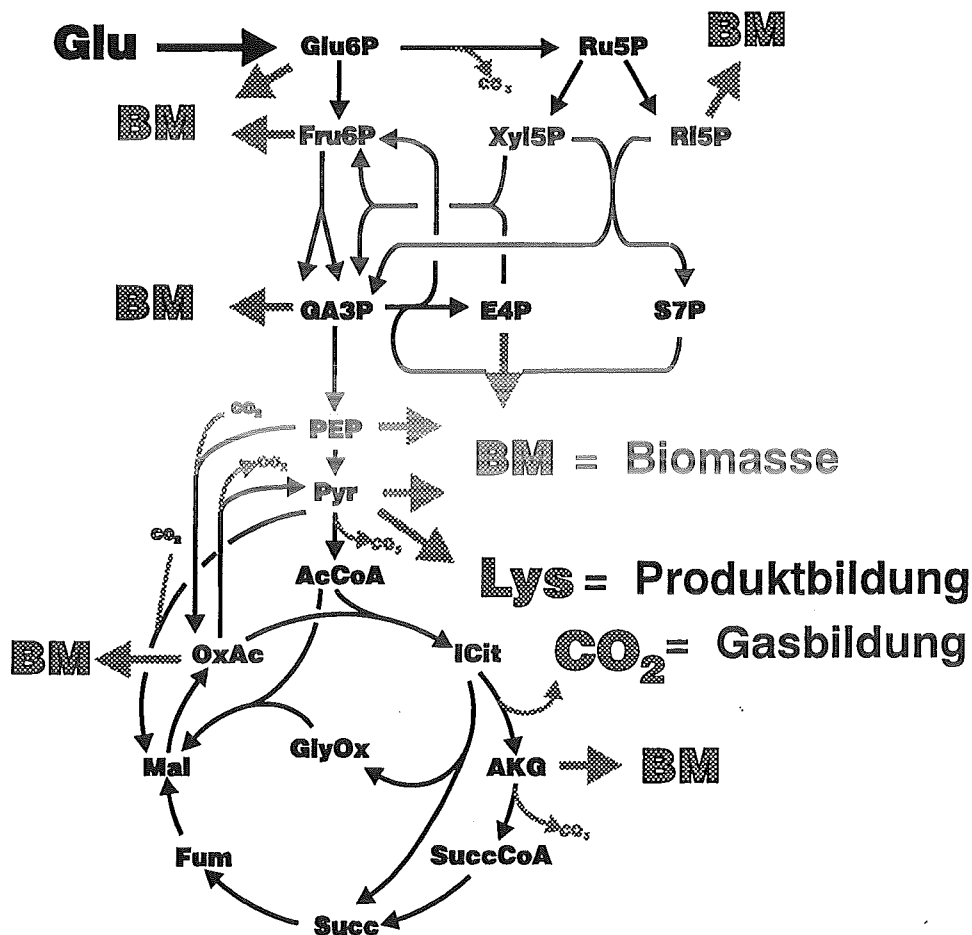


Abbildung 2.1: Meßbare extrazelluläre Flüsse in einem Bioreaktor. Aus dem Fluß in die Biomasse können bei Kenntnis der Biomassezusammensetzung anabolische Flüsse, die von Vorläufern im Zentralstoffwechsel ausgehen, detailliert quantifiziert werden.

Die meisten Anwendungen dieser Technik beruhen auf der Annahme, daß die Biomassezusammensetzung in Mikroorganismen über eine große Bandbreite von metabolischen Zuständen hinweg konstant ist und sich sogar zwischen verschiedenen Mikroorganismen kaum unterscheidet [Roe83, Val91]. Diese Annahme wurde gelegentlich experimentell bestätigt, jedoch sollte sie stets mit Vorsicht gehandhabt werden. Genaugenommen ist für jedes einzelne Experiment erneut

die Zellzusammensetzung zu bestimmen, was allerdings einen unzumutbaren Arbeitsaufwand bedeutet.

Die Gesamtanzahl der mit den beschriebenen Meßmethoden bestimmbaren extrazellulären Stoffflüsse liegt — je nach Anzahl der gemessenen extrazellulären Produkte — bei etwa 15-18.

2.2 NMR-Technik und Massenspektroskopie

Das Phänomen der kernmagnetischen Resonanz ist bereits seit den 40er Jahren bekannt. Heute ist die NMR-Technik eine der wichtigsten Meßmethoden für die chemische Strukturaufklärung und die Computertomographie. Auf biologische Systeme wird die Methode seit etwa 1970 angewandt, v.a. für den Nachweis intrazellulärer Metabolite, für die Messung von intrazellulären Konzentrationen und schließlich die Durchführung von Markierungsexperimenten [Gad82] [Mat82, Lon88, LHHV90]. Die Massenspektroskopie hat ein ähnliches Einsatzfeld, ist jedoch bei bestimmten Fragestellungen — wie nachfolgend erläutert — im Vergleich zur NMR-Technik wesentlich weniger leistungsfähig.

Seit den Anfängen hat die NMR-Technik ständig Fortschritte gemacht, denen heute zwei Zeitschriften gewidmet sind, nämlich das "Journal of Magnetic Resonance" und "NMR in Biomedicine". Da eine Einführung in die moderne NMR-Technik ein Buch füllt [Fri86], können hier nur einige Grundzüge dargestellt werden. Es soll daher — mit dem Risiko einer zu starken Vereinfachung — nur eine elementare Vorstellung von den Möglichkeiten der NMR-Technik vermittelt werden. Im einzelnen verfügt nur ein geschulter Experte über das erforderliche Detailwissen zur Konzeption und Auswertung von NMR-Experimenten. Alle dieser Arbeit zugrundeliegenden NMR-Messungen wurden unter Anleitung von Herrn Dr. A.A. de Graaf im Jülicher NMR-Labor durchgeführt.

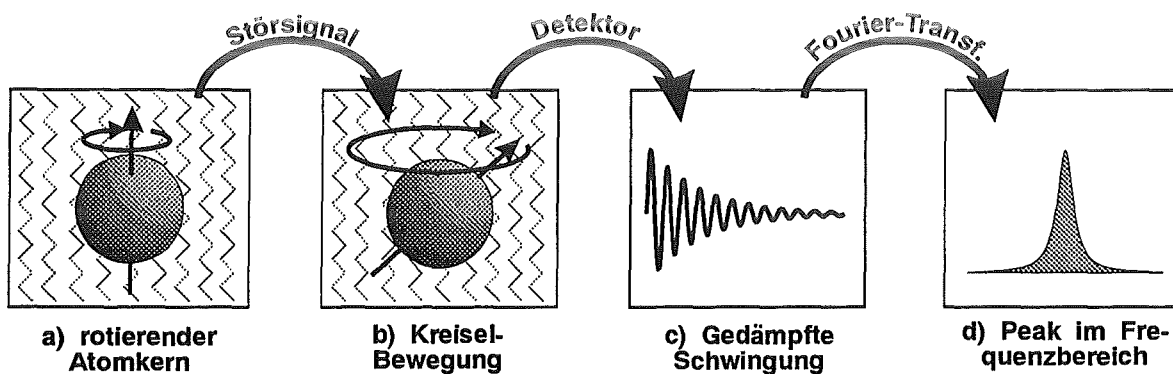


Abbildung 2.2: Funktionsprinzip der NMR-Technik. a) Die Drehachsen rotierender geladener Teilchen richten sich im Magnetfeld aus. b) Ein Störimpuls löst eine Kreisbewegung mit atomspezifischer Frequenz aus. c) Die erzeugten exponentiell gedämpften Signale werden detektiert. d) Nach Summation aller Einzelsignale entsteht mit Hilfe von Fourier-Transformationsmethoden das NMR-Spektrum.

2.2.1 Funktionsweise eines NMR-Geräts

Die NMR-Technik beruht darauf, daß Atomkerne aufgrund ihres Kernspins rotierende Ladungsverteilungen im Raum darstellen. Somit können solche Kerne mit einem Spin ungleich Null durch ein Magnetfeld beeinflusst werden. Hierzu sind allerdings hohe Feldstärken erforderlich

(etwa 0.5-20 Tesla). Die für die biologische Anwendung wichtigsten detektierbaren Kerne sind ^1H , ^{13}C , ^2H , ^{15}N , ^{23}Na und ^{31}P .

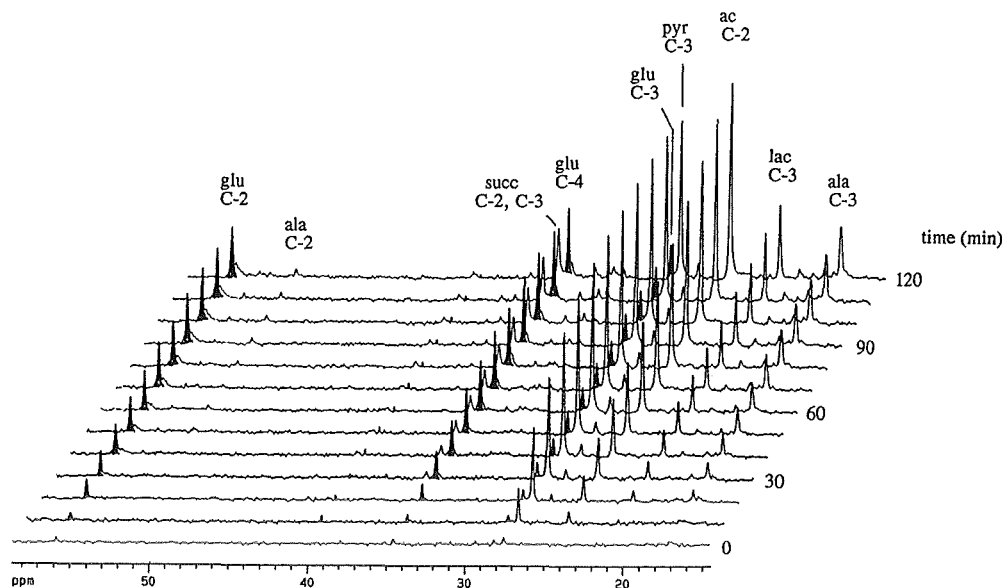


Abbildung 2.3: Beispiel einer in zeitlicher Reihenfolge aufgenommenen Sequenz von niedrig aufgelösten ^{13}C -NMR-Gemischspektren während der isotopisch instationären Phase eines Markierungsexperiments mit *Corynebacterium glutamicum* mit $[1-^{13}\text{C}]$ -Glucose als markiertem Substrat (vgl. [WG96a]). Jedes C-Atom in der Probe erzeugt einen charakteristischen Resonanzpeak. Der Markierungsaufbau in einzelnen C-Atom-Positionen ist deutlich zu erkennen (Quelle: A.A. de Graaf, Jülich).

Ein NMR-Meßgerät besteht im wesentlichen aus einer supraleitenden Spule, die ein starkes homogenes Magnetfeld erzeugt und eine zentrale Öffnung für die Einführung der Probe aufweist. Innerhalb des Magnetfeldes richten sich die Drehachsen der magnetisierbaren Isotope der Probe in einer Richtung aus (Abb. 2.2a). Diese wird von einer zweiten, wesentlich kleineren Anregungsspule umschlossen. Die Anregung erfolgt durch Abstrahlung eines zum Hauptmagnetfeld transversalen magnetischen Störimpulses in die Probe. Die dadurch "getroffenen" Atome reagieren durch eine Spinauslenkung und in der Folge durch eine kreiselnde Spinbewegung (Abb. 2.2b). Diese Bewegung relaxiert sich schließlich innerhalb des Hauptmagnetfeldes wieder in die alte Richtung.

Während ihrer kreiselnden Bewegung senden die rotierenden Spins einer Substanz die eingestrahlte Energie in Form einer elektromagnetischen Welle auf einer für die Position des Kerns innerhalb des Moleküls charakteristischen Frequenz wieder ab (Abb. 2.2c). Diese Signale können mit Hilfe von Detektorspulen empfangen werden.

Einzelne Resonanzsignale sind jedoch schwach. Um quantitativ auswertbare Signale zu erhalten, müssen in der Regel sehr viele solcher Resonanzsignale in Folge erzeugt und aufsummiert werden. Das Summensignal wird anschließend mit Hilfe einer Fourier-Transformation in den Frequenzbereich übersetzt (Abb. 2.2d), man erhält das eigentliche NMR-Spektrum. Eine Einzelmessung dauert etwa 1-30 Sekunden, zur Erzeugung eines Spektrums sind meistens 4-4000 solcher Signale erforderlich.

Die Auswertung solcher NMR-Spektren beruht darauf, daß die Resonanzfrequenz eines

Kerns einzig von seiner Stellung innerhalb eines Moleküls abhängt. Damit lassen sich die Peaks eines NMR-Spektrums exakt den einzelnen in der Probe enthaltenen Stoffen zuordnen (Abb. 2.3). Die Fläche eines Peaks im Spektrum ist direkt proportional zur Menge der detektierbaren Atome an einer speziellen Position.

2.2.2 ^{13}C - und ^1H -Spektren

Die Resonanzfrequenzen verschiedener Arten von Atomen, z.B. Kohlenstoff, Stickstoff oder Wasserstoff, liegen in sehr unterschiedlichen Frequenzbereichen, so daß jeweils spezielle Meßeinrichtungen erforderlich sind. Man spricht z.B. von ^{13}C - oder ^1H -NMR-Spektroskopie.

Abbildung 2.3 zeigt ein niedrigauflösendes ^{13}C -NMR-Spektrum aus einem In-vivo-Experiment. Es wird deutlich, daß das Auftreten vieler Resonanzpeaks eine möglichst hohe Auflösung erfordert, um diese voneinander trennen zu können.

In der Praxis werden zum Nachweis von ^{13}C -Atomen häufig nicht direkt die Resonanzen der markierten Kohlenstoff-Atome gemessen, sondern vielmehr die der umliegenden ^1H -Protonen. Denn ein markiertes Kohlenstoff-Atom verändert seine elektrochemische Umgebung, so daß sich die Frequenzen der umliegenden Protonen verschieben. Auf diese Weise sind Protonen, die an ein ^{12}C -Atom gebunden sind, von solchen zu unterscheiden, die an ein ^{13}C -Atom gebunden sind. Mit Hilfe einer speziellen Resonanzentkopplungstechnik ist es dann möglich, aus den Daten mehrerer Spektrenmessungen direkt auf den Prozentsatz der markierten Kohlenstoff-Atome zurückzuschließen (Abb. 2.4). Die Kenntnis der absoluten Konzentration der vermessenen Substanz ist also nicht erforderlich, um den Markierungsanteil zu berechnen.

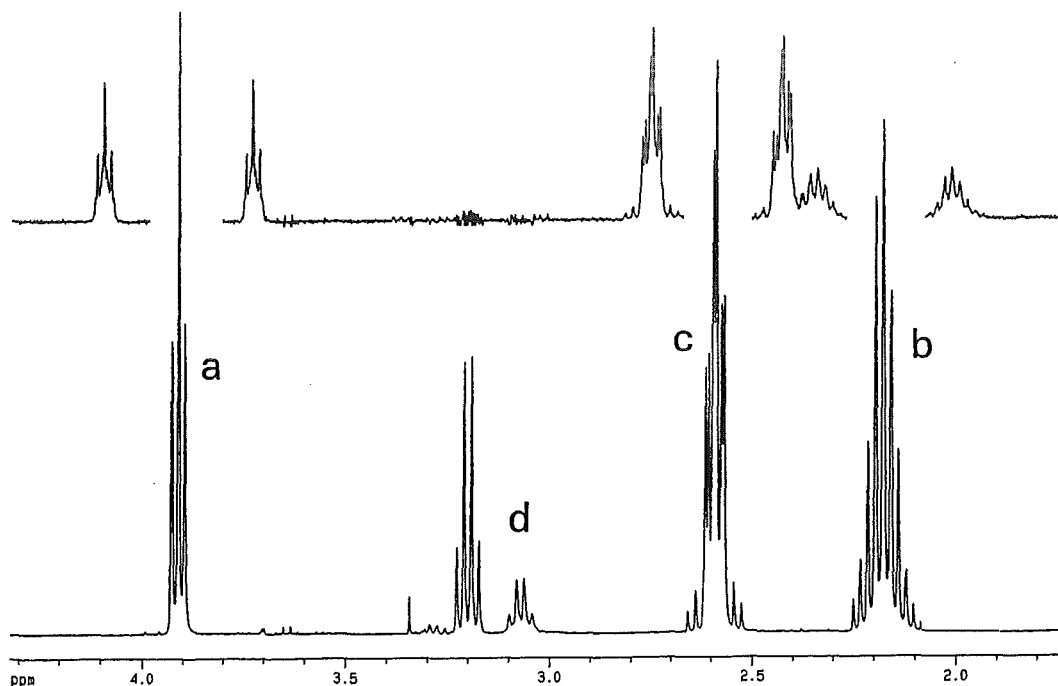


Abbildung 2.4: ^1H -NMR-Spektren von ^{13}C -markiertem Glutamat, das aus dem Zellprotein von *C. glutamicum* nach Inkubation mit $[1-^{13}\text{C}]$ -Glucose gewonnen wurde. Untere Reihe: ^{13}C -entkoppeltes Spektrum mit den a) H-2-, b) H-3- und c) H-4-Protonen des Glutamats sowie d) Signalen aus Unreinheiten. Obere Reihe: entsprechende ^{13}C -Satellit-Signale nach Subtraktion des ^{13}C -entkoppelten Spektrums vom "normalen" Protonen-Spektrum wie in [SEGS93] (aus: [WG96c]).

Protonenspektren weisen extrem komplexe Peakstrukturen auf und sind damit wesentlich schwerer zu interpretieren als ^{13}C -Spektren (vgl. Abb. 2.4). Daher kann die ^1H -Spektroskopie nicht direkt auf Stoffgemische angewendet werden. Vielmehr müssen die darin enthaltenen Stoffe zunächst isoliert und aufkonzentriert werden [SEGS93].

2.2.3 Multipliettanalyse

Wenn die Meßzeit hinreichend lang ist, können hochauflösende ^{13}C -NMR-Spektren gemessen werden. Diese zeigen eine Struktur, die für die Stoffflußanalyse von großer Bedeutung ist, nämlich die Aufspaltung von Peaks (vgl. [Lon88, JRMS91, KCS93]). Diese resultiert daraus, daß die elektrochemische Umgebung eines C-Atoms bei verschiedenen Isotopomeren eines Stoffes differiert. Ist auch ein Nachbar eines C-Atoms in einem Molekül markiert, so besitzt das Atom zwei verschiedene Resonanzfrequenzen infolge sogenannter skalarer Spin-Spin-Kopplung und der Resonanzpeak spaltet sich in ein sogenanntes Doublett auf. Analoges gilt für zwei markierte Nachbarn, es kommt zu insgesamt vier verschiedenen Resonanzen, von denen jedoch die beiden mittleren zusammenfallen können, so daß ein Triplet entsteht (vgl. Abb. 2.5, 2.6 und 2.7).

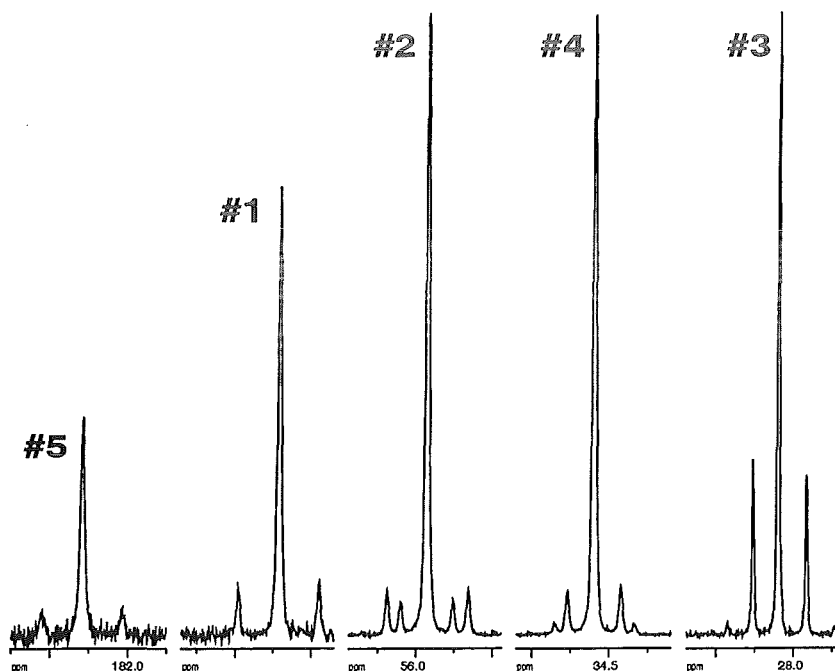


Abbildung 2.5: ^{13}C -Isotopomeren-Spektren von ^{13}C -markiertem Glutamat entsprechend den in Abbildung 2.4 gezeigten ^1H -Spektren. Multipletts treten aufgrund von skalaren Spin-Spin-Kopplungen mit benachbarten ^{13}C -Atomen auf (aus: [WG96c], vgl. auch [KCS93]).

Definition 2.1 (Multipliett) Ein Multipliett ist ein Mehrfachpeak in einem NMR-Spektrum, der durch die Aufspaltung eines einzelnen Peaks aufgrund von skalaren Spin-Spin-Kopplungen innerhalb verschiedener Isotopomeren entsteht. Zu unterscheiden sind Doubletts, Tripletts, Quartetts und noch komplexere Strukturen (Abb. 2.6).

Generell hängt die Anzahl der Peaks in einem Multipliett von der Anzahl der markierten Nachbarn ab. Allerdings müssen die genauen elektrochemischen Bedingungen innerhalb eines

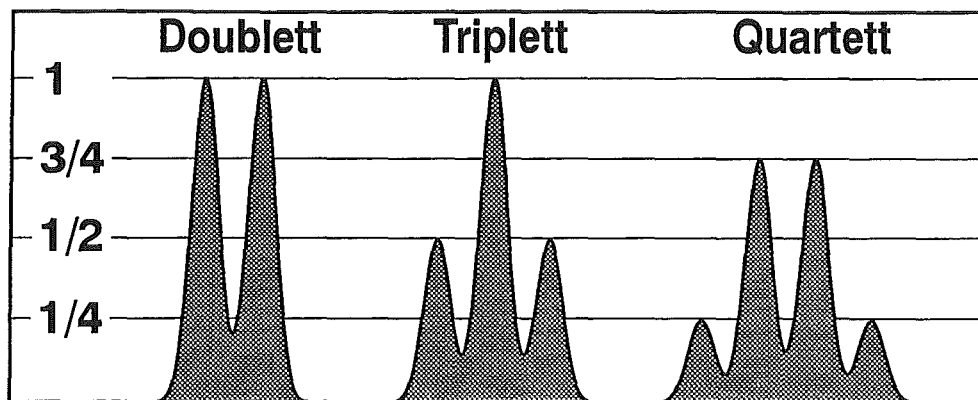


Abbildung 2.6: Ideale Multiplett-Peaks einer C-Atom-Resonanz. Bei einem Triplet überlagern sich zwei Resonanzen auf der mittleren Frequenz, bei einem Quartett je drei Resonanzen auf den beiden mittleren Frequenzen (Binomialverteilung). Die Gesamtfläche der Multipletts ist jeweils gleich.

Moleküls bekannt sein, um vorhersagen zu können, welche Markierungsform zu welchem Spektrum führt. Abbildung 2.7 zeigt schematisch die möglichen Multiplettspektren eines idealen Moleküls mit drei C-Atomen, während Abb. 2.5 einige Multiplettspektren eines realen Metaboliten wiedergibt.

Entscheidend für die Datenauswertung ist, daß die exakte Korrespondenz zwischen Markierungsform und Spektrum in der Regel aufgrund von publizierten NMR-Daten oder Kalibrierungsmessungen im Vorhinein bekannt ist (vgl. z.B. [BD85]). Auf eine detaillierte Darstellung ergänzender Meßmethoden (z.B. durch Resonanz-Entkopplungstechniken oder die Hinzunahme von ^1H -Spektren), die weitere aussagekräftige Daten liefern, muß an dieser Stelle verzichtet werden.

In der Praxis überlagern sich die Multiplettspektren aller im Gemisch vorliegenden Isotopomere. Aus Abb. 2.7a wird deutlich, daß eine vollständige Trennung aller Isotopomere eines Moleküls selbst auf der Basis von hochauflösenden ^{13}C -NMR-Spektren nicht möglich ist [KCS93, JSJ⁺93, DRM⁺93]. Beispielsweise produzieren die Isotopomere I#100 und I#001 in Abb. 2.7a Singulett-Peaks auf der ersten und dritten Position. An denselben Stellen entstehen aber auch dann Singulett-Peaks, wenn das Isotopomer I#101 in der Mischung vorhanden ist. Singuletts an der ersten und dritten Position können also nicht eindeutig einer Kombination von I#100, I#001 und I#101 zugeordnet werden.¹ Dagegen entspricht ein Triplet-Peak an Position 2 eindeutig dem Isotopomer I#111. Dies motiviert die folgende Begriffsbildung (vgl. auch Abb. 2.7):

Definition 2.2 (Multiplettanalyse) Die (teilweise) Quantifizierung von Isotopomeren-Anteilen auf der Basis eines Multiplettspektrums wird als Multiplettanalyse bezeichnet.

Notwendige Voraussetzung für die Multiplettanalyse sind hochauflösende ^{13}C -NMR-Spektren. Ist diese Bedingung erfüllt, so bietet sie allerdings eine Informationsfülle, die heute mit keiner anderen Meßtechnik erreicht werden kann.

¹Eine präzise quantitative Analyse dieser Zusammenhänge findet sich in Abschnitt 6.2, insbesondere Gleichung (6.13).

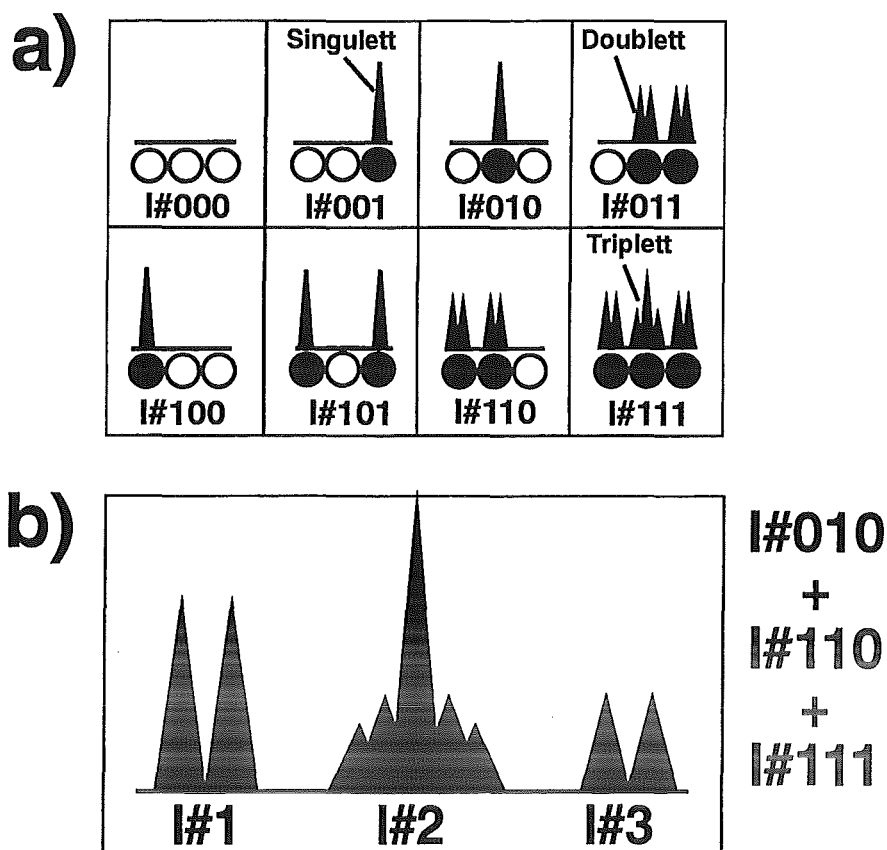


Abbildung 2.7: a) Die acht verschiedenen ^{13}C -Isotopomeren eines idealen Moleküls I mit drei Kohlenstoff-Atomen und ihre NMR-Multiplettspektren unter idealen Bedingungen. b) Ein überlagertes Spektrum des Isotopomeren-Gemischs aus I#010, I#110 und I#111.

2.2.4 Massenspektroskopie

Eine weitere Meßtechnik, die — allerdings nur in beschränktem Umfang — Aufschluß über Isotopomeren-Verteilungen geben kann, ist die Massenspektroskopie (MS) [IL87, DRM⁺93] [KWL93]. Sie wird heute vielfach in direkter Verbindung mit einem Gaschromatographen zur Stofftrennung verwendet (GC-MS). In einem Massenspektrometer können Isotopomere eines Moleküls aufgrund ihrer unterschiedlichen Massen voneinander getrennt werden. In der Situation aus Abb. 2.7 können beispielsweise das unmarkierte Isotopomer I#000, die drei einfach markierten Isotopomere I#100, I#010, I#001, die drei doppelt markierten Isotopomere I#110, I#101, I#011 und schließlich das dreifach markierte Isotopomer I#111 voneinander getrennt werden. Wie bei der NMR-Technik ist dabei eine Bestimmung der Prozentsätze dieser Isotopomeren-Gruppen ohne Kenntnis der absoluten Konzentration möglich.

Die Massenspektroskopie wird heute vielfach zusätzlich mit einer chemischen Degradierung der Moleküle verknüpft [DRM⁺93]. Dabei werden bestimmte Kohlenstoff-Atome der Kette abgespalten und die Restmoleküle erneut im Massenspektrometer vermessen. Daraus ergeben sich weitere Informationen über die Isotopomeren-Verteilung im ursprünglichen Molekül. Schließlich erlaubt eine Kombination von Massenspektroskopie und NMR-Technik — wenn auch verbunden mit einigem Aufwand — die detaillierte Quantifizierung aller Isotopomere bei Molekülen mit bis zu etwa fünf Kohlenstoff-Atomen (vgl. dazu Abschnitt 6.2).

2.2.5 Problematik von In-vivo-NMR-Messungen

Das Hauptproblem bei der Messung von In-vivo-NMR-Spektren ist die geringe Sensitivität dieser Technik im Vergleich zu anderen Meßmethoden, z.B. der Massenspektroskopie [Gad82] [Mat82, Lon88, LHHV90]. Die Qualität eines Spektrums ist abhängig von der Anzahl der aufsummierten Einzelmessungen (also der Gesamtdauer der Messung) und der Anzahl der in der Probe vorhandenen markierten Atome. Für eine Isotopomeren-Quantifizierung sind dabei wesentlich höhere Anforderungen an die Auflösung zu richten, während für die Bestimmung von Positions-Markierungsanteilen niedrigere Auflösungen ausreichend sein können. Soll der isotopisch instationäre Zustand des Systems beobachtet werden, so ist jedoch die Zeitdauer für die Aufnahme eines Einzelspektrums stark limitiert (etwa 0.1-10 Minuten), was generell zu einem schlechten Signal-Rausch-Verhältnis führt.

Teilweise konnten diese Probleme in den letzten Jahren durch stärkere NMR-Instrumente, erhöhte Verfügbarkeit von ^{13}C -markierten Substraten und spezielle experimentelle Techniken überwunden werden. Einige typische Anordnungen zur In-vivo-NMR-Spektroskopie seien hier kurz vorgestellt:

- In der Vergangenheit wurde ein Markierungsexperiment meist mit aufkonzentrierten Zellen in einem Reagenzglas durchgeführt [WHK⁺82, WK84, HUBB⁺86]. Dabei können zwar sehr hohe Zelldichten erreicht werden, jedoch ist die Meßzeit streng limitiert, da die Stationarität der Stoffwechselvorgänge nur über sehr kurze Zeiträume gewährleistet ist.
- Die stärksten Signale werden derzeit mit tierischen Organen (Herz, Leber, Gehirn) erzielt, die als ganzes in einer Perfusionslösung für längere Zeit in Funktion gehalten werden [MSJ88, DRM⁺93]. Hiermit gelingt sogar die Messung von isotopisch instationären Zuständen [CFG95] und Multiplettspektren [MSJ88, KCS93]. Allerdings stellt sich auch hier das Problem der Sicherung eines wohldefinierten stationären Zustands und darüber hinaus das der biologischen Homogenität der Probe.
- In einigen Fällen wurde versucht, einen miniaturisierten Bioreaktor [GWPS⁺92, CB93] im Magnetkern des NMR-Instruments zu betreiben. Inzwischen gelingt dies sogar in kontinuierlicher Kultur, was jedoch einen erheblichen technischen Aufwand erfordert [Har95]. Die erzielbaren Signale sind jedoch sehr schwach verglichen mit den beiden anderen Techniken, weil die erreichbare Zellkonzentration deutlich geringer ist.
- Eine wesentliche Verbesserung dieser Anordnung wurde mit einem in Jülich entwickelten speziellen NMR-Bioreaktor erreicht (Abb. 2.8). Dabei wird durch eine Biomasserückhaltung über einen im Bypass liegenden Filtrationsmodul die Verweilzeit der Biomasse im System von der Verweilzeit des Mediums entkoppelt (vgl. Abb.12.5). In einem solchen Reaktor kann eine Zelldichte von 30-50g/l Trockenmasse erreicht und bei gleichbleibend guter Substratversorgung online vermessen werden. Das vorgestellte Experiment ist äußerst kostenaufwendig, da die aufkonzentrierte Biomasse einen extrem hohen Nährstoffbedarf hat.

Für Metaboliten im Konzentrationsbereich von 10 mmol/l kann ein Signal mit einem akzeptierbaren Signal-Rausch-Verhältnis in weniger als einer Minute erzeugt werden. Auf diese Weise konnte der in Abb. 12.6 dargestellte instationäre Markierungsverlauf innerhalb des Reaktors gemessen werden. Die so erreichbare Qualität der Spektren ist jedoch nicht ausreichend für Isotopomeren-Bestimmungen.

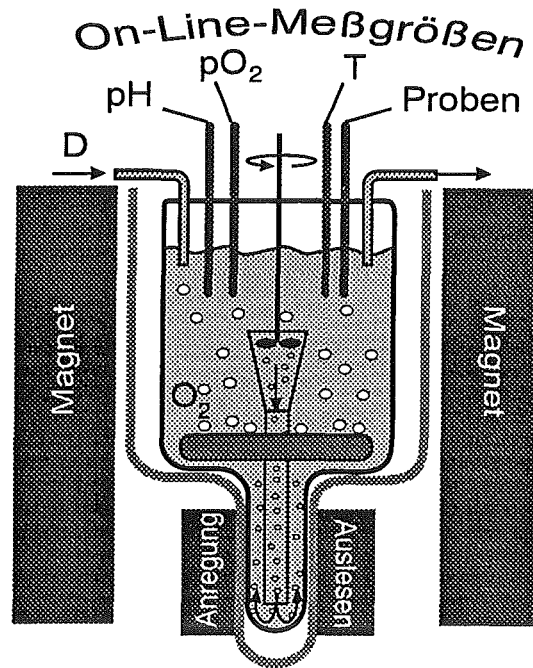


Abbildung 2.8: Schemazeichnung des in Jülich entwickelten NMR-Bioreaktors für den Betrieb einer kontinuierlichen Kultur im NMR-Gerät (vgl. [GWPS⁺92]).

2.2.6 Analyse von Gemischspektren

Bei Gemischspektren, insbesondere also bei In-vivo-Experimenten, stellen sich bezüglich der Spektralanalyse besondere Probleme (vgl. Abb. 2.3):

- Das Spektrum weist eine große Menge verschieden starker Signale unterschiedlicher chemischer Stoffe aufgrund der stark variierenden Konzentrationen dieser Stoffe auf.
- Bei schwachauflösenden Spektren können sich die einzelnen Peaks stark überlappen [SB88].
- Bei Messung in einem In-vivo-NMR-Bioreaktor kommt es aufgrund von konstruktionsbedingter Störungen zu einer weiteren Peakverbreiterung.
- In einer nicht standardisierten Probe hängt die Resonanzfrequenz der Atome in nicht genau bekannter Weise von weiteren Parametern, z.B. dem intrazellulären pH-Wert, ab.

Aus diesen experimentell bedingten Schwierigkeiten resultiert das Problem der rechnerischen Zerlegung von NMR-Gemischspektren zur Berechnung der Peakflächen. Eine detaillierte Darstellung der Verfahren der NMR-Gemischspektralanalyse sprengt jedoch den Rahmen dieser Arbeit. Der Leser sei hierzu verwiesen auf [DAM88, NS89, MTG94, WMWG95, Möl95].

2.3 Stoffflußanalyse unter wohldefinierten Bedingungen

2.3.1 Grundannahmen der stationären Stoffflußanalyse

Einige — nachfolgend aufgeführte — Grundannahmen werden bei der Durchführung und Auswertung von Markierungsexperimenten vorausgesetzt. Nur wenn diese Annahmen weitgehend

zutreffen, kann eine In-vivo-Stoffflußanalyse unter wohldefinierten physiologischen Bedingungen durchgeführt werden.

Definition 2.3 (wohldefinierte physiologische Bedingungen, Grundannahmen der stationären Stoffflußanalyse) *Es wird von einer stationären In-vivo-Stoffflußanalyse unter wohldefinierten physiologischen Bedingungen gesprochen, wenn die nachfolgenden Grundannahmen zutreffen:*

Homogenität : *Die Metabolite sind innerhalb eines für die Systembeschreibung gewählten Kompartiments homogen verteilt.*

Metabolische Stationarität : *Das System befindet sich während der gesamten Dauer des Meßvorgangs in einem metabolisch stationären Zustand.*

Biochemische Systemkenntnis : *Für alle betrachteten Stoffwechselwege sind die Reaktionsschritte des zugrundeliegenden biochemischen Netzwerks und insbesondere die Positionswechsel von Kohlenstoff-Atomen bekannt.*

Keine Isotopen-Masseneffekte : *Es treten keine Isotopen-Masseneffekte auf, d.h. Enzyme unterscheiden hinsichtlich der Reaktionsgeschwindigkeit nicht zwischen verschiedenartig markierten Metaboliten.*

Zerstörungsfreie Messung : *Der Meßvorgang beeinflusst den Zustand des Systems nicht.*

Verzerrungsfreie Messung : *Der Meßvorgang ist verzerrungsfrei, d.h. die beobachtete Grösse wird durch nachfolgende chemische Behandlungen des Probenmaterials nicht verändert.*

Im folgenden wird die Gültigkeit dieser Annahmen in Hinblick auf die in Kapitel 1.1 beschriebenen Experimente eingehender diskutiert.

2.3.2 Homogenität

Die zentrale Voraussetzung für die kompartmentelle Modellierung ist die Homogenitätsannahme, da sie eine Systembeschreibung mit Hilfe von gewöhnlichen Differentialgleichungen bzw. stationären Bilanzgleichungen ermöglicht. Ohne diese Annahme ist eine überprüfbare Beschreibung von metabolischen Vorgängen nicht möglich, weil die experimentelle Untersuchung von Konzentrationsgradienten innerhalb einzelner Bakterienzellen bis heute weit außerhalb der meßtechnischen Möglichkeiten liegt. Es ist daher u.a. aus pragmatischen Gründen notwendig, von einer räumlich homogenen Verteilung aller beteiligten Stoffe innerhalb einzelner Kompartimente auszugehen. Die folgende Diskussion der wichtigsten Einwände gegen die Homogenitätsannahme zeigt, daß diese in Bezug auf niedere Zellen (also Bakterien) vergleichsweise unproblematisch ist.

- Extrazelluläre Stoffkonzentrationen sind bei Reagenzglasexperimenten mit hoher Zelldichte sicher nicht konstant, da der Stofftransport in dem die Zelle umgebenden Medium limitiert ist. Dieses Problem tritt nicht auf in einem gerührten Bioreaktor, in dem das Medium bis zu einer gewissen Gefäßgröße im Rahmen der meßtechnischen Genauigkeit vollkommen homogen ist.

- Die einzelnen Zellen einer Population können sich in unterschiedlichen physiologischen Zuständen befinden. Beispielsweise ist eine solche Inhomogenität durch den Zellzyklus gegeben. Die ermittelten Meßwerte sind im Grunde stets Mittelungen über die Population. Allerdings kann bei den schnell wachsenden Bakterien kaum noch von einem Zellzyklus im engeren Sinne gesprochen werden. Es ist z.B. bekannt, daß sich die DNA bei *E. coli* bei hohen Wachstumsraten in ununterbrochener Teilung befindet. Bereits während der Bildung der Tochterzellen beginnt die erneute Teilung der DNA-Stränge für die Enkelzellen [BM88b]. Dies spricht gegen die These eines zeitlich stark oszillierenden Stoffflusses innerhalb des einzelnen Organismus.
- Bei Eukaryonten muß ein kompartimentelles Modell auch das Volumen von Zellorganellen (z.B. Mitochondrien) berücksichtigen. Dies kann zu einer nicht mehr handhabbaren Zahl von Kompartimenten führen. Im Gegensatz dazu ist eine räumliche Unterteilung bei Prokaryonten nicht ausgeprägt.
- Intrazelluläre Diffusionsvorgänge können in höheren (und damit erheblich größeren) Zellen durchaus von Bedeutung sein. Innerhalb der Kompartimente kleiner (insbesondere bakterieller) Zellen kann jedoch weitgehend von einer räumlichen Gleichverteilung von Substraten und Enzymen bzw. Enzymkomplexen ausgegangen werden.

2.3.3 Metabolische Stationarität

Ein metabolisch stationärer Zustand (oft auch als balanciertes Wachstum bezeichnet [BM88b]) kann praktisch nur in einem kontinuierlich betriebenen Bioreaktor eingestellt und überwacht werden. Mit einer solchen Anordnung sind sogar unterschiedliche Zustände des Systems — ggfs. mit Hilfe regelungstechnischer Methoden [Wie90] — stabilisierbar. Allerdings wurden nur sehr wenige der in der Literatur beschriebenen Experimente unter den Bedingungen einer kontinuierlichen Kultur durchgeführt, am seltensten ist dies bei reinen Markierungsexperimenten der Fall.

Stattdessen werden in der Literatur meist verschiedene Approximationen an einen stationären Zustand gewählt. Dabei muß die Stationarität jeweils in Relation zur Dauer des Meßvorgangs beurteilt werden. Bei Experimenten, die lediglich extrazelluläre Flüsse messen, nimmt dieser kaum Zeit in Anspruch. Dagegen erfordern Markierungsexperimente Meßzeiten zwischen wenigen Minuten und einigen Stunden (vgl. Abb. 12.6). Folgende — zum Teil mit Nachteilen verknüpfte — Annäherungen an einen stationären Zustand werden typischerweise zugrundegelegt:

- In einer im Bioreaktor betriebenen Batch- oder Fed-Batch-Kultur kann während einer exponentiellen Wachstumsphase und bei geringer Meßdauer von einem annähernd stationären Verhalten des Systems ausgegangen werden.
- Bei Versuchen im Schüttelkolben können prozeßrelevante Größen, wie Sauerstoffkonzentration und pH-Wert, bereits nicht mehr konstant gehalten und nur schwer überprüft werden.
- Es ist zweifelhaft, ob Versuche mit aufkonzentrierten Zellsuspensionen im Reagenzglas noch Aussagen über den "natürlichen" Zustand des Systems zulassen.
- Bei Experimenten mit ganzen Tierorganen ist der physiologische Zustand des Systems ebenfalls schwer einstellbar.

2.3.4 Biochemische Systemkenntnis

Die biochemischen Reaktionsschritte im Zentralstoffwechsel einer Zelle (Abb. 1.2) können seit langen Jahren als nahezu vollständig aufgeklärt gelten. Lehrbücher geben detailliert Auskunft über die Positionswechsel von Kohlenstoff-Atomen in jedem einzelnen Reaktionsschritt. Allerdings können im einzelnen große Unterschiede zwischen dem Zentralstoffwechsel von Mikroorganismen bzw. höheren Zellen auftreten. In der Regel können Fachleute jedoch genaue Auskunft geben, welche Reaktionen in dem gerade betrachteten Organismus auftreten.

Natürlich werden auch heute noch im Zentralstoffwechsel bestimmter Mikroorganismen neue Reaktionsschritte entdeckt. Gerade die Markierungstechnik hat zu ihrer Aufklärung einen entscheidenden Beitrag geleistet. Für die quantitative Stoffflußanalyse stellen die wenigen in Frage stehenden Reaktionsschritte jedoch kein großes Problem dar, da das gemessene Datenmaterial in der Regel eine Diskriminierung zwischen einer überschaubaren Anzahl von Alternativmodellen erlauben wird. Damit ermöglicht die Stoffflußanalyse zugleich den Nachweis von Reaktionsschritten.

2.3.5 Isotopen-Masseneffekte

Isotopen-Masseneffekte sind bei chemischen Reaktionen tatsächlich beobachtet worden, denn Metabolite mit verschiedenen Massen haben auch eine unterschiedliche Beweglichkeit [Wol82, OLea82, WKKS82]. Damit verändert sich auch die Bindungswahrscheinlichkeit in bezug auf das aktive Zentrum eines Enzyms.

Am häufigsten sind biologisch relevante Masseneffekte im Zusammenhang mit CO_2 festgestellt worden.² Allerdings liegen keine Ergebnisse über in wässriger Phase gelöstes CO_2 vor, denn in wässriger Phase ist die Molekülbeweglichkeit im Vergleich zur Gasphase um Größenordnungen kleiner. Schon bei Metaboliten mit zwei Kohlenstoff-Atomen ist die Atommasse jedoch so groß, daß der Isotopenmasseneffekt verglichen mit der Genauigkeit der verwendeten Meßtechnik vernachlässigt werden kann.

Im Rahmen dieser Arbeit (wie auch in allen bisher vorliegenden Markierungsstudien) werden Masseneffekte daher vernachlässigt. Prinzipiell könnten jedoch in den hier beschriebenen Modellen auch Korrekturfaktoren für Isotopen-Masseneffekte Eingang finden. Voraussetzung dafür ist allerdings eine quantitative Charakterisierung des Masseneffektes.

2.3.6 Zerstörungsfreiheit des Meßvorgangs

Die Unabhängigkeit des Systemzustands vom Meßvorgang ist ein klassisches Problem experimentellen Arbeitens. Bei In-vivo-Studien kommt dieser Problematik eine besondere Bedeutung zu. Bezogen auf einzelne hier verwendete Meßtechniken ist im Hinblick auf diese Fragestellung festzuhalten:

- Die Analyse des Kulturabgases ist offenbar unbedenklich, da sie außerhalb des Reaktors stattfindet.
- Die Messung extrazellulärer wie intrazellulärer Konzentrationen beruht auf einer Probenahme [WG96a]. Solange das Volumen der Probe im Verhältnis zum Gesamtvolumen hinreichend klein ist, wird der Zustand der im System verbleibenden Biomasse durch eine derartige Operation nicht beeinträchtigt. Wird die Messung kostenbedingt in einem sehr

²Natürlich spielen Masseneffekte in der Atmosphärenchemie eine zentrale Rolle.

kleinen Bioreaktor durchgeführt, muß allerdings eine Minimierung des Probevolumens angestrebt werden.

- Viele Markierungsexperimente werden im Magnetfeld des NMR-Gerätes durchgeführt. Der Einfluß extrem starker Magnetfelder auf lebende Organismen wird seit einiger Zeit heftig diskutiert. Auch bei mikrobiellen Systemen sind magnetische Effekte festgestellt worden [Oku94]. Bei den in Jülich durchgeführten Studien zeigte sich im Hinblick auf die gemessenen extrazellulären Flußdaten jedoch keine Veränderung gegenüber einem außerhalb des Magnetfeldes betriebenen Kontrollsystem. Dies kann zumindest als Indiz für einen unverändert ablaufenden Zentralstoffwechsel gewertet werden.

Im übrigen finden die im nachfolgenden Abschnitt 2.5 beschriebenen neuartigen Experimente überhaupt nicht mehr im Magnetfeld des NMR-Gerätes statt und sind daher im Hinblick auf magnetische Effekte völlig unbedenklich.

2.3.7 Unverzerrtheit der Messung

Die Entwicklung zerstörungs- und verzerrungsfreier Meßmethoden ist nach wie vor eine der größten Herausforderungen in der biologischen Forschung. Bezogen auf die Stoffflußanalyse muß vor allem die Problematik der Probenahme mit lebenden Zellen diskutiert werden. Mit in neuerer Zeit entwickelten Methoden kann der Stoffwechsel einer lebenden Zelle innerhalb einer Sekunde vollständig gestoppt werden [WG96a]. Diese Inaktivierung erfolgt in der Regel durch einen kombinierten Einsatz von Perchlorsäure, die außerdem die Zellwände aufschließt, und Kühlung auf tiefe Temperaturen. Bei Mikroorganismen, die einen sehr schnellen Stoffumsatz haben, kann diese Zeitkonstante durchaus kritisch sein, da gering konzentrierte Stoffe (etwa unter 1 mmol/l) noch während der Inaktivierungsphase vollständig umgesetzt werden können.

Wie bereits in Abschnitt 1.1 erläutert, erfordern Markierungsexperimente, die sich auf den isotopisch stationären Zustand des Systems beschränken, nicht die Messung von Stoffkonzentrationen von Intermediaten im Zentralstoffwechsel. Dies trifft auch auf die in Abschnitt 2.5 beschriebene experimentelle Technik zu.

2.4 Zusätzliche Annahmen

Die Gültigkeit der bisher vorgestellten Annahmen wird in der experimentellen Literatur zumeist nicht thematisiert. In der Tat stellen sie als Grundlage für die datengestützte Modellierung eines biologischen Systems äußerst schwache Annahmen dar (vgl. Abschnitt 2.6). Bis auf die methodenbedingte Homogenitätsannahme sind alle anderen Prämissen mit heutiger Meßtechnik zumindest global überprüfbar.

Im Gegensatz dazu stellen die nachfolgend aufgeführten optionalen Annahmen Vor-Urteile über das metabolische System dar, die in die Modellierung einfließen. Die Hinzunahme solcher Annahmen erfordert also einige Sorgfalt, da das verfügbare Datenmaterial im Lichte dieser Prämissen interpretiert wird und daraus die Gefahr einer verfälschten Darstellung resultiert. Die sorgfältige Dokumentation solcher Modellannahmen ist daher ein besonderes Anliegen der vorliegenden Arbeit.

2.4.1 Energiebilanzierung

Gründet sich die Stoffflußanalyse nur auf extrazelluläre Flußmessungen, so reichen die gemessenen Daten allein nicht aus, um auf der Basis von Massenbilanzen die unbekannten intrazellulären Stoffflüsse zu schätzen. In diesem Fall wird häufig eine Energiebilanzierung durchgeführt. Folgende Typen von Bilanzen sind dabei in der Literatur belegt [Roe83, VS93, GFJA93, Jor95]:

- Entscheidend für den Energiehaushalt der Zelle sind die energieübertragenden Metabolite, also ATP, NADH, FADH und einige andere energiereiche Substanzen. Ihre Bilanzierung erweist sich deshalb als problematisch, weil nicht alle erzeugenden und verbrauchenden Reaktionsschritte bekannt sind.
- Ebenso problematisch ist die Bilanzierung spezieller Kofaktoren, die zur Katalyse bestimmter Reaktionen erforderlich sind [Val91]. So ist beispielsweise NADPH für die Lysin-Synthese von großer Bedeutung. Trotzdem besteht hinsichtlich der Bilanzierung von NADPH in *Corynebacterium glutamicum* bis heute ein großer Differenzbetrag zwischen den bekannten Erzeugern und Verbrauchern [MGWE⁺96, SSGM⁺95].
- Darüber hinaus kann sogar Energie aufgrund von Futile Cycles [CF80, HBS81, RB85, Pat92, CL94, CMCM⁺94] und Lecks in Membranen verloren gehen, wodurch die Bilanzen nicht mehr schlüssig sind.
- Die oxidative Phosphorylierung kann nicht als Reaktionsschritt mit konstanten stöchiometrischen Koeffizienten betrachtet werden (vgl. dazu [Roe83, WD87, NM94]). Zwar wird hier Energie durch die Übertragung von Protonen auf Sauerstoff gewonnen, jedoch ist der Wirkungsgrad dieses Vorgangs vom Organismus und den äußeren Bedingungen abhängig. Der sogenannte P/O-Quotient, also das gebildete ATP bezogen auf ein mol Sauerstoff ist nur aufgrund aufwendiger Reihenexperimente mit verschiedenen Substraten zugänglich [Roe83, WD87].
- Genauso problematisch ist die Bestimmung der benötigten Energiemenge für die Synthese eines Gramms der Zellmasse. Hier können Schätzwerte aufgrund einer Zellmassenanalyse und der Kenntnis aller Biosynthesewege angegeben werden [NIS90]. Dies setzt allerdings voraus, daß der Erhaltungsstoffwechsel vernachlässigbar ist.
- Als noch problematischer erweist sich die Bilanzierung von freien Reaktionsenergien (ΔG) [Roe83, Mav93]. Bekannt ist im Grunde nur die jeweilige freie Energie ΔG^0 einer Reaktion unter standardisierten Bedingungen. Unter In-vivo-Bedingungen können ΔG -Werte hiervon stark abweichen; im Extremfall laufen Reaktionen in der thermodynamisch nicht bevorteilten Richtung ab. Freie Reaktionsenergien können daher im Grunde nur für qualitative Argumentationen herangezogen werden [Mav93].

Annahmen über den Energiehaushalt einer Zelle sind aus diesen Gründen äußerst problematisch [WD87]. Darüber hinaus entziehen sie sich einer meßtechnischen Überprüfung und machen damit Vorurteile des Experimentators zur Grundlage einer Modellierung. In dieser Arbeit wird daher konsequent auf eine Energiebilanzierung verzichtet, obwohl diese in dem vorgestellten Rahmen ohne weiteres hinzugenommen werden können.

2.4.2 Direktionalitätsannahmen

Bereits in Abschnitt 1.2 (vgl. auch Abb. 1.8) wurde auf die Möglichkeit bidirektional ablaufender Reaktionsschritte eingegangen. In der vorliegenden Arbeit wird dabei sorgfältig unterschieden zwischen den Begriffspaaren bidirektional/unidirektional und reversibel/irreversibel [ALS77]:

Definition 2.4 (reversibler und irreversibler Reaktionsschritt, bidirektionaler und unidirektionaler Reaktionsschritt) *Eine Reaktion heißt irreversibel unter gegebenen In-vivo-Bedingungen, wenn sie in der intrazellulären Umgebung aus thermodynamischen Gründen ausschließlich in einer Richtung ablaufen kann, andernfalls heißt sie reversibel. Eine Reaktion heißt unidirektional unter gegebenen In-vivo-Bedingungen, wenn sie im untersuchten Zustand des Systems (fast) ausschließlich in einer Richtung abläuft, andernfalls heißt sie bidirektional.*

Während die ersten beiden Begriffe also die grundsätzliche Möglichkeit des Ablaufs in beiden Richtungen charakterisieren, beschreiben die letzteren den tatsächlichen Befund bei einem gegebenen Zustand des Stoffwechsels. Eine Reaktion kann also (thermodynamisch) reversibel sein, aber trotzdem in einer bestimmten Situation ausschließlich in der einen Richtung ablaufen, weil die Nachfolgereaktionen einen Rückfluß verhindern. Lediglich die (thermodynamische) Irreversibilität einer Reaktion garantiert auch deren Unidirektionalität. Offenbar sind alle Unterscheidungen graduell, d.h. eine Reaktion kann praktisch irreversibel sein, während im strengen thermodynamischen Sinn stets eine positive Wahrscheinlichkeit für beide Reaktionsrichtungen gegeben ist.

Nachfolgend wird gelegentlich die Annahme getroffen, daß eine große freie Standard-Reaktionsenergie ΔG^0 in vitro auf eine Unidirektionalität der Reaktion in vivo schließen läßt. Auch diese Annahme muß im einzelnen diskutiert werden, denn es kann vorkommen, daß eine Reaktion in vivo sogar in der thermodynamisch nicht begünstigten Richtung abläuft [Str88]. In strittigen Fällen sollte eine Reaktion daher stets als bidirektional angenommen werden.

2.4.3 Pool-Lumping

Eine weitere Annahme, die bereits in Abschnitt 1.2 angeschnitten wurde, ist das sogenannte Pool-Lumping:

Definition 2.5 (Pool-Lumping) *Unter einem Pool-Lumping versteht man die Zusammenfassung mehrerer Pools eines kompartmentellen Modells zu einem neuen Pool. Dabei werden alle Zu- und Abflüsse der zusammengefaßten Pools auf den neuen Pool übertragen.*

Diese Standardvorgehensweise der kompartmentellen Modellierung ist eine Notwendigkeit, da ausnahmslos alle biochemischen Reaktionsschritte über einige mehr oder weniger stabile Zwischenstufen ablaufen [Syk88]. Ihre detaillierte Berücksichtigung würde zu einer nicht mehr handhabbaren Modellkomplexität führen. Insofern stellt bereits die Abbildung 1.2 eine auf die wichtigsten Metabolit-Pools reduzierte Sicht des Zentralstoffwechsels dar. Beispielsweise wurden mehrere Intermediate der Glykolyse kombiniert und der Citrat-Pool mit dem Isocitrat-Pool zusammengefaßt.

Generell ist die Zusammenfassung einer linearen Sequenz von Reaktionsschritten unproblematisch, wenn keine Abzweigungen von den Zwischenprodukten ausgehen. Darüber hinaus werden üblicherweise solche Pools, die in sehr schnellem wechselseitigen Austausch stehen, zusammengefaßt. Ein Beispiel sind die drei verschiedenen Pentose-5-phosphat-Pools (Ribose-,

Ribulose- und Xylulose-5-phosphat) am Eingang des Pentosephosphatwegs (vgl. Abb. 9.1). Von diesen drei Pools zweigt sowohl die DNA- und RNA-Synthese als auch der eigentliche Pentosephosphatweg ab. Sie werden durch äußerst aktive Isomerasen wechselseitig ineinander umgewandelt [Woo85]. Auch die übrigen bimolekularen Reaktionsschritte des Pentosephosphatwegs werden häufig als schnelle Gleichgewichte angesehen [SSH92] (vgl. Abb. 9.5). Die Auswirkung dieser Annahme auf die Modellierung wird in Abschnitt 9.2 diskutiert.

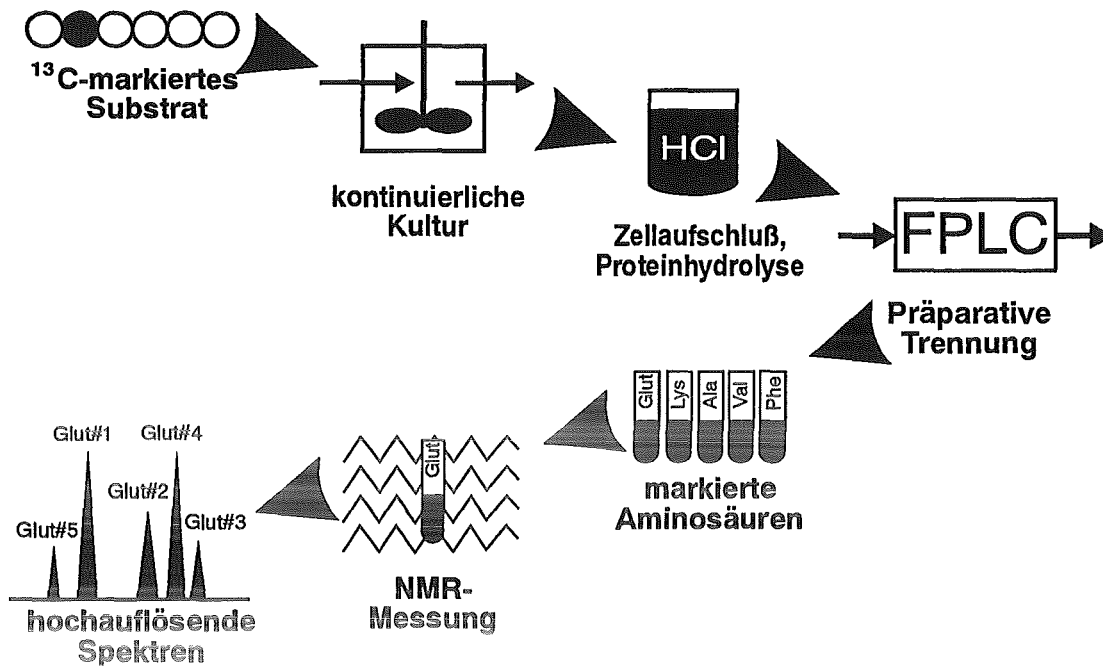


Abbildung 2.9: Entkopplung der NMR-Messung von der Produktion markierten Materials. Markierte Aminosäuren werden durch Hydrolyse des Zellproteins gewonnen und einzeln vermessen.

2.5 Eine neue experimentelle Methode

Aufgrund der Diskussion der Modellannahmen verschiedener gängiger experimenteller Ansätze wird deutlich, daß keine der bestehenden Methoden unproblematisch hinsichtlich einer modellmäßigen Interpretation der Daten ist. Aus diesem Grunde wurde in der Jülicher Arbeitsgruppe die nachfolgend beschriebene neue experimentelle Technik, die in [Mar94, MGWE⁺96] publiziert ist, entwickelt. Obwohl einzelne Bestandteile der Methode (wie etwa der Zellaufschluß) bereits früher in anderen Arbeitsgruppen angewendet wurden, führt erst die Integration von kontinuierlicher Fermentation, Zellaufschluß, hochauflösender NMR-Vermessung und modellgestützter Auswertung zu einer neuen Qualität der stationären Stoffflußanalyse. Insbesondere wird sich in dieser Arbeit zeigen, daß die neue Technik in mehrfacher Hinsicht zu einer wesentlichen Verbesserung der Qualität als auch der Quantität des Datenmaterials führt. Eine vergleichbare Entwicklung hat in jüngster Zeit an der ETH Zürich stattgefunden [Szy95].

2.5.1 Entkopplung des Markierungs-Experiments von der -Messung

Der Grundgedanke der neu entwickelten Methode ist die Entkopplung des Markierungsexperiments von der NMR-Messung.

Definition 2.6 (zeitliche Entkopplung von Markierungs-Experiment und -Messung) *Eine Methode zur Durchführung eines Markierungsexperiments, bei dem die NMR-Vermessung markierter intrazellulärer Metabolite zu einem anderen Zeitpunkt stattfindet als das unter wohldefinierten physiologischen Bedingungen durchgeführte Experiment und zusätzlich die Verzerrungsfreiheit der Messung gewährleistet ist, wird als zeitliche Entkopplung von Markierungs-Experiment und -Messung bezeichnet.*

Ein solches Experiment gelingt nicht, wenn lediglich Proben aus einem Bioreaktor gezogen werden, und diese dann im NMR-Gerät zu vermessen. Hierzu sind die meisten intrazellulären Metaboliten bei weitem zu niedrig konzentriert, so daß sie im NMR-Spektrum nicht direkt beobachtbar sind. Die praktische Umsetzung des Entkopplungsgedankens gelingt unter Ausnutzung der anabolischen Stoffwechselwege. Die Zelle akkumuliert das von ihr produzierte markierte Material in verschiedenen Zellkomponenten der Zellmasse. Die nachfolgend beschriebene Umsetzung dieses Vorgehens zeichnet die in Jülich neu entwickelte Methode aus.

Die meßtechnisch am einfachsten zugänglichen Zellkomponenten sind Polymere wie Proteine, RNA oder DNA. In ihrer polymerisierten Form sind die Einzelmoleküle einer NMR-Messung nicht zugänglich, jedoch können sie durch Hydrolyse in Kombination mit einer präparativen analytischen Auftrennung leicht isoliert werden (Abb. 2.9). Schon in der Vergangenheit haben verschiedene Autoren die Extraktion von Zellkomponenten wie Glycerol aus Lipiden [ESS83, CB83, RB85], Ribonukleinsäuren aus RNA und DNA [ESS83, ESW⁺93] oder Aminosäuren aus Proteinen [ESS83, SEBF92, PWK94] beschrieben. Jedoch wurde erstmals in Jülich [Mar94] ein solches Experiment in einer kontinuierlichen Kultur unter wohldefinierten physiologischen Bedingungen durchgeführt, zugleich eine quantitative Analyse der Markierungen vorgenommen und schließlich auch eine Messung aller relevanten extrazellulären Flüsse mit hoher Präzision durchgeführt.

Das Markierungsexperiment wird in einem Bioreaktor durchgeführt, wobei einige Biomasseverweilzeiten abgewartet werden, so daß sich das markierte Material im Zellprotein anreichert. Sodann wird der gesamte Inhalt des Bioreaktors aufgearbeitet und die einzelnen Zellkomponenten extrahiert. Weil Aminosäuren aus Vorläufern an vielen verschiedenen Stellen des Zentralstoffwechsels entstehen (Abb. 2.10), ist offenbar die ergiebigste quantitative Informationsquelle durch die Aminosäurefraktionen gegeben. Aus ihrem Markierungszustand kann direkt auf die Markierungsanteile in ihren Vorläufern zurückgeschlossen werden. Die meisten dieser Vorläufermetabolite sind mit anderen, direkten Methoden völlig unzugänglich, weil sie in vivo nur in extrem geringen Konzentrationen vorliegen. Damit ist es erstmals möglich, intrazelluläre Markierungsdaten von einem kontinuierlich kultivierten Mikroorganismus zu erhalten.

Die gewonnenen Aminosäurefraktionen können beliebig lange in einem NMR-Instrument vermessen werden, so daß sogar hochauflösende Spektren zur Bestimmung von Multipletts erzeugt werden [SEBF92]. Dabei ist es ein besonderer Vorteil des isotopisch stationären Ansatzes, daß keine absoluten Poolgrößen benötigt werden, so daß durch den Aufbereitungsvorgang bedingte Massenverluste nicht ins Gewicht fallen. Dagegen ist der isotopisch instationäre Zustand mit dieser Technik nicht zugänglich.

2.5.2 Verzerrungsfreiheit der Messung

Auch die neue Technik muß im Hinblick auf die in Definition 2.3 für die stationäre Stoffflußanalyse gemachten Modellannahmen kritisch untersucht werden. Dabei bleibt offenbar nur die Frage nach der Verzerrungsfreiheit des Meßvorgangs zu beantworten (vgl. Abschnitt 2.3). Diese ist in der Tat nicht vollständig gewährleistet, denn der zelluläre Anabolismus, also die Bildung

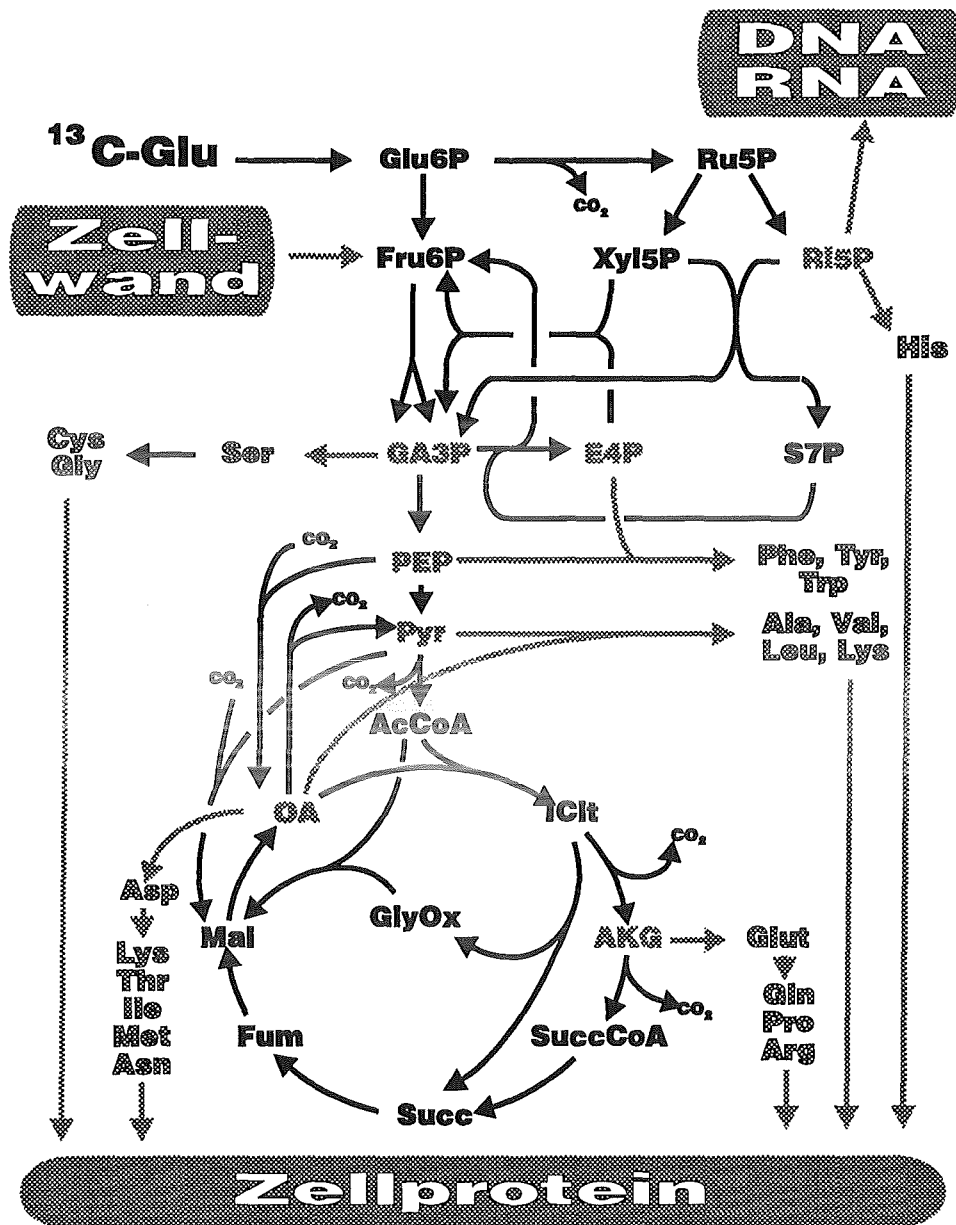


Abbildung 2.10: Produktion von Aminosäuren, Nukleinsäuren und Zellwandbestandteilen im Zentralstoffwechsel. Mit der Kenntnis der Markierungen liegen breit verteilte Stützpunkte für die quantitative Flußbestimmung vor.

von Zellmaterial aus den Vorläufern des Zentralstoffwechsels, ist ein Teil des Meßvorgangs. Der Vorgang der Einlagerung von markiertem Material in die Zellmasse muß also bei der Datenauswertung mit berücksichtigt werden [BS82, ESW⁺93].

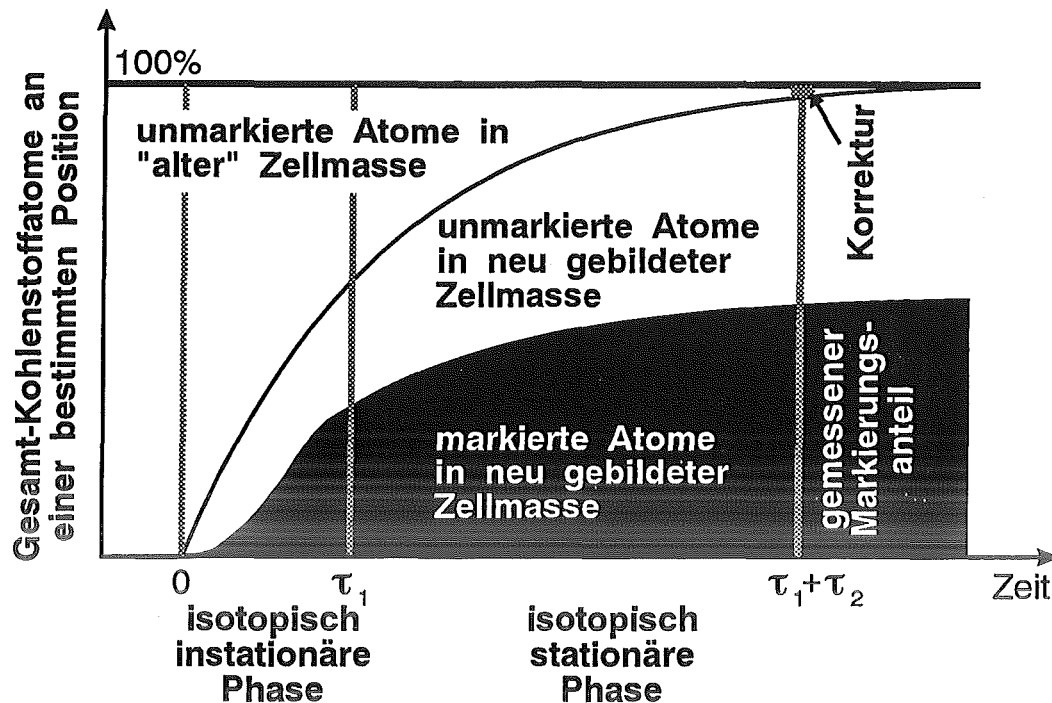


Abbildung 2.11: Phasen eines Markierungsexperiments mit zeitlich entkoppelter Vermessung von Zellkomponenten. Dargestellt ist die durchschnittliche ^{13}C -Anreicherung eines C-Atom-Pools in der Zellmasse. Während der Zentralstoffwechsel nach einer Zeit τ_1 annähernd isotopisch stationär ist (vgl. auch Abb. 1.4), nehmen die Markierungsanteile in der Zellmasse einen exponentiell sättigenden Verlauf, der durch die langsamere Auswaschung der im Bioreaktor verbleibenden unmarkierten Biomasse verursacht wird. Das Experiment wird nach der Zeit $\tau_1 + \tau_2$ beendet.

Nachfolgende Überlegungen führen auf eine quantitative Beschreibung der Markierungsanreicherung in der Zellmasse. Es wird davon ausgegangen, daß bis zum Einsatz des markierten Substrates zum Zeitpunkt Null alle intrazellulären Intermediate sowie die gesamte Zellmasse unmarkiert sind. Die zu diesem Zeitpunkt im Reaktor befindliche Zellmasse wird nachfolgend als die "alte" Zellmasse bezeichnet. Die nun einsetzende Produktion markierter Intermediate im Zentralstoffwechsel führt dazu, daß markierte Substanzen in die neu gebildete Zellmasse eingelagert werden. Hinsichtlich der Markierungsanteile der eingelagerten C-Atome sind nun zwei Phasen zu unterscheiden (Abb. 2.11):

Isotopisch instationäre Phase ($0 \leq t \leq \tau_1$): In diesem Zeitraum befinden sich die Metabolite im Zentralstoffwechsel noch im instationären Übergang zu der sich schließlich einstellenden zeitlich konstanten Markierungsverteilung (vgl. Abb. 1.4). Dementsprechend ist auch der Markierungszufluß in die Biomasse während dieser Phase nicht konstant. Über die Markierungsanteile der C-Atome im Zentralstoffwechsel und in der Biomasse kann während dieses Zeitraums — mangels Meßdaten — nichts gesagt werden. Die instationäre Phase ist (praktisch) beendet, wenn sich die Markierungsanteile im Zentralstoffwechsel

nicht mehr meßbar verändern.³

Isotopisch stationäre Phase ($0 \leq t \leq \tau_1 + \tau_2$): Erst vom Zeitpunkt τ_1 ab ist der Markierungszufluß in jede Komponente der Biomasse konstant. Die von nun an neu gebildeten Zellkomponenten sind also ebenfalls konstant und identisch mit den jeweiligen Vorläufern im Zentralstoffwechsel markiert. Aufgrund der Zellauswaschung in der kontinuierlichen Kultur nimmt zugleich der Anteil des noch unmarkierten alten Materials und des in der isotopisch instationären Phase produzierten Zellmaterials ab. Mit der Zeit wird sich also die durchschnittliche Markierung in der Zellmasse der Markierung in den Intermediaten angleichen. Jedoch wird das Experiment bereits zum Zeitpunkt $\tau_1 + \tau_2$ aus Kostengründen abgebrochen.

Nach Beendigung der Fermentation wird die durchschnittliche Markierung in der Gesamtzellmasse gemessen. Dies ist offenbar eine Verzerrung der zum Zeitpunkt $\tau_1 + \tau_2$ vorliegenden isotopisch stationären Markierungsanteile im Zentralstoffwechsel, da damit außerdem die noch im Reaktor befindliche alte und die während der instationären Phase neu gebildete Zellmasse mit vermessen wird. Gesucht ist nun nach einem Faktor, der diese Verzerrung korrigiert.

2.5.3 Fehlerabschätzung für die Markierungsmessung

Sei nun $x_{ZS}(t)$ der in einem C-Atom-Pool im Zentralstoffwechsel vorliegende Markierungsanteil. Für diesen gilt also

$$x_{ZS}(\tau_1) = x_{ZS}(t) = x_{ZS}(\tau_1 + \tau_2), \quad \text{wenn} \quad \tau_1 \leq t \leq \tau_2.$$

Ferner sei dieser C-Atom-Pool ein direkter Vorläufer eines C-Atom-Pools in der Zellmasse mit dem entsprechenden Markierungsanteil x_{BM} . Gesucht sind nun zwei Korrekturfaktoren $\underline{c}$ und $\bar{c}$, die eine möglichst enge Abschätzung der Form

$$\underline{c} \leq \frac{x_{BM}(\tau_1 + \tau_2)}{x_{ZS}(\tau_1 + \tau_2)} \leq \bar{c}$$

zulassen. Das Problem liegt offenbar darin, daß der genaue dynamische Verlauf von $x_{ZS}(t)$ für $0 \leq t \leq \tau_1$ unbekannt und darüber hinaus nicht konstant ist.

Sei nun τ die (volumenbezogene) Verweilzeit im Bioreaktor (also die reziproke Durchflußrate [Ber83]). Bekanntlich hat dann jede Konzentration $y(t)$ eines von außen nicht zugeführten Stoffs die exponentielle Auswaschungscharakteristik

$$y(t) = y(0) \cdot (1 - \exp(-\frac{t}{\tau})) \quad (2.4)$$

Dies kann wie folgt ausgenutzt werden:

Untere Abschätzung: Es wird (pessimistisch) angenommen, daß alle Metabolite im Zentralstoffwechsel während der instationären Phase völlig unmarkiert bleiben, d.h. $x_{ZS}(t) = 0$ für $0 \leq t \leq \tau_1$. Es folgt daraus $x_{BM}(t) = 0$ im entsprechenden Zeitraum. Dann gilt aufgrund von Gleichung (2.4), angewandt auf die alte Biomasse:

$$\frac{x_{BM}(\tau_1 + \tau_2)}{x_{ZS}(\tau_1 + \tau_2)} = 1 - \exp(-\frac{\tau_2}{\tau}) \quad (2.5)$$

³Eine mögliche präzise Definition für τ_1 wäre z.B. das Erreichen des 99-prozentigen stationären Markierungsanteils.

Obere Abschätzung: Es wird (optimistisch) angenommen, daß die Metabolite im Zentralstoffwechsel zum Zeitpunkt 0 bereits ihre Äquilibriumsmarkierung annehmen, d.h. $x_{ZS}(t) = x_{ZS}(\tau_1 + \tau_2)$ für $t \geq 0$. Dann gilt entsprechend:

$$\frac{x_{BM}(\tau_1 + \tau_2)}{x_{ZS}(\tau_1 + \tau_2)} = 1 - \exp\left(-\frac{\tau_1 + \tau_2}{\tau}\right) \quad (2.6)$$

Aus Gleichung (2.5) und (2.6) folgt der⁴

Satz 2.1 (Auswaschkorrektur bei zeitlicher Entkopplung von Markierungs-Experiment und -Messung) Mit den oben eingeführten Konstanten τ, τ_1, τ_2 gilt:

$$1 - \exp\left(-\frac{\tau_2}{\tau}\right) \leq \frac{x_{BM}(\tau_1 + \tau_2)}{x_{ZS}(\tau_1 + \tau_2)} \leq 1 - \exp\left(-\frac{\tau_1 + \tau_2}{\tau}\right)$$

Der relative Fehler bei Anwendung des Korrekturfaktors $\delta = [1 - \exp(-\frac{\tau_1 + \tau_2}{\tau})]^{-1}$ ist beschränkt durch:

$$0 \leq \frac{x_{ZS} - \delta \cdot x_{BM}}{x_{ZS}} \leq \frac{1 - \exp(-\frac{\tau_1 + \tau_2}{\tau})}{1 - \exp(-\frac{\tau_2}{\tau})} - 1$$

Ausgehend von Simulationsstudien (vgl. Abb. 1.4) liegt der wahre Wert des Quotienten erfahrungsgemäß deutlich näher an der oberen Abschätzung. Die Zeitkonstante τ_1 ist im konkreten Fall größenordnungsmäßig bekannt, da die intrazellulären Poolgrößen ungefähr bekannt sind (vgl. auch die On-line-Messung in Abb. 12.3). In der Praxis muß stets gelten $\tau_1 < \tau$, denn bekanntlich ist τ in kontinuierlicher Kultur bis auf den Faktor $\ln 2$ gleich der Verdopplungszeit des Organismus. Realistische relative Werte sind also $\tau = 1$, $\tau_1 = 0.25$ und $\tau_2 = 2.75$. Es folgt

$$0.936 \leq \frac{x_{BM}(\tau_1 + \tau_2)}{x_{ZS}(\tau_1 + \tau_2)} \leq 0.950 \quad ,$$

also ein maximaler relativer Meßfehler von etwa 1%. Bedenkt man, daß Markierungsanteile meist im Bereich unter 0.2 liegen, denn es wird meist nur eines der Substrat-C-Atome markiert, so folgt ein absoluter Fehler unterhalb von 0.002. Dies liegt an der unteren Grenze der im NMR-Instrument erzielbaren Meßgenauigkeit. Die Tabelle 2.1 zeigt einige maximal zu erwartende Fehler für verschiedene Werte der Zeitkonstanten.

2.5.4 Charakteristika der neu entwickelten Methode

Die neu entwickelte Meßtechnik unterscheidet sich von bisher verfügbaren Methoden dadurch, daß sie alle nachfolgenden Kriterien *zugleich* erfüllt:

1. Das Experiment wird in einer kontinuierlichen Kultur unter definierten und kontrollierbaren physiologischen Bedingungen durchgeführt.
2. Über die Protein- und RNA/DNA-Hydrolyse ist eine mit direkten Methoden bisher nicht erreichte Anzahl von aussagekräftigen intrazellulären Meßdaten verfügbar.

⁴Für einen strengen Beweis der Abschätzung müßte zusätzlich angenommen werden, daß $x_{ZS}(t)$ nicht über- oder unterschwingt. Dies wird bei Markierungsexperimenten nur in vernachlässigbarem Ausmaß beobachtet (vgl. dazu auch Abschnitt 7.2).

		τ_1				
		0.2	0.4	0.6	0.8	1.0
τ_2	1.0	10.5	19.0	26.2	32.0	36.8
	2.0	2.8	5.1	7.0	8.6	9.9
	3.0	0.9	1.7	2.4	2.9	3.3
	4.0	0.3	0.6	0.8	1.0	1.2
	5.0	0.1	0.2	0.3	0.4	0.4
	6.0	0.0	0.1	0.1	0.1	0.2

Tabelle 2.1: Maximaler zu erwartender prozentualer Meßfehler bei Auswaschkorrektur gemäß Satz 2.1 für verschiedene Werte der Zeitkonstanten τ_1, τ_2 bezogen auf die normierte Verweilzeit $\tau = 1$.

3. Außerdem sind alle relevanten extrazellulären Flüsse mit hoher Präzision bestimmbar.
4. Der Meßvorgang ist vollkommen zerstörungsfrei, da das Experiment nicht innerhalb eines NMR-Gerätes stattfindet.
5. Der Meßvorgang ist im Rahmen der erforderlichen Genauigkeit verzerrungsfrei.
6. Es können beliebig hoch auflösende NMR-Spektren gemessen werden, so daß stationäre Isotopomeren-Verteilungen ermittelbar sind.

Damit wird es möglich, eine Stoffflußanalyse unter Verzicht auf diverse sonst in der Literatur häufig anzutreffende Annahmen durchzuführen. Die Entwicklung einer Methodik zur Auswertung der entstehenden, komplex strukturierten Datensätze ist das wichtigste Ziel der vorliegenden Arbeit. Während die erforderlichen experimentellen und meßtechnischen Voraussetzungen inzwischen bei mehreren Arbeitsgruppen gegeben sind, ist ein zu der vorliegenden Arbeit vergleichbarer allgemeiner methodischer Ansatz für die Modellierung, Auswertung und Analyse von stationären Markierungssystemen in der Literatur nicht annähernd dokumentiert.

2.6 Beziehung zu metabolisch instationären Modellansätzen

2.6.1 Dynamische Studien

Obwohl sich diese Arbeit auf die Analyse metabolisch stationärer Stoffflüsse beschränkt, soll zumindest kurz die Beziehung der stationären Stoffflußanalyse zu physiologischen Studien unter metabolisch instationären Bedingungen diskutiert werden. Solche Studien beruhen auf der Messung intra- und extrazellulärer Stoffkonzentrationen zu verschiedenen Zeitpunkten eines Experimentes bei einem sich dynamisch verändernden metabolischen Zustand.⁵ In der Regel werden sogenannte Shift-Experimente durchgeführt, bei denen eine Eingangsgröße (z.B. eine Stoffkonzentration) sprunghaft verändert wird. Aufgrund der Sprungantwort des Systems wird dann eine quantitative Charakterisierung des dynamischen Systems vorgenommen [RTQ⁺96, WG96a].

Zielsetzung derartiger Studien ist die Entwicklung eines dynamischen Modells für den Stoffwechsel in Form eines Differentialgleichungssystems [DLC⁺84, WBA92, RTQ⁺96]. Das Problem "klassischer" grobstrukturierter Modelle aus dem Bereich der Bioverfahrenstechnik bildet die Zusammenfassung ganzer Stoffwechselwege, z.B. Katabolismus, Anabolismus und Produktbildung, im Modell und die daraus resultierende, unzureichende Wiedergabe des Stoffwechsels

⁵Der Einsatz markierten Materials stellt bei solchen Studien eher eine Ausnahme dar.

[Roe83, Baz83, Wie90]. Dagegen wird schon seit langer Zeit versucht, detailliertere Stoffwechselmodelle auf der Basis einer enzymkinetischen Beschreibung der einzelnen Reaktionsschritte aufzustellen, die auch kurzfristige dynamische Prozesse abbilden können [DLC⁺84, RTQ⁺96]. Derartige Modelle verwenden jedoch bis auf den heutigen Tag in relativ großem Umfang In-vitro-Parameter, die mehr oder weniger unreflektiert auf die In-vivo-Situation übertragen werden. Erst in jüngster Zeit wurden experimentelle Techniken zur Messung intrazellulärer Stoffkonzentrationen in Mikroorganismen innerhalb sehr kurzer Zeitintervalle entwickelt [RTQ⁺96, WG96a]. Sie erlauben erstmals eine zunehmend detaillierte und validierbare dynamische Vermessung des metabolischen Systems.

2.6.2 Bedenken gegen den dynamischen Ansatz

Eine Reihe von Bedenken stellt den kinetischen Zugang zur dynamischen Modellierung grundsätzlich in Frage:

- Eine detaillierte Abbildung des Zentralstoffwechsels mit Hilfe enzymkinetischer Terme führt zu einem äußerst komplexen Differentialgleichungssystem. Wird jeder Reaktionsschritt abgebildet, so gelangt man ohne weiteres zu einer Größenordnung von hundert Gleichungen mit mehreren hundert Parametern! Die Validierung eines derartigen Modells auf der Basis vorgegebener Meßdaten im Sinne strenger statistischer Gütemaße scheint aussichtslos, wenn nicht grundsätzlich unmöglich.
- Generell lassen sich enzymkinetische Parameter, die in In-vitro-Experimenten gewonnen wurden, nicht auf die In-vivo-Situation übertragen. Der Grund liegt darin, daß das Cytosol einer Zelle keineswegs als eine Ansammlung von in einem wässrigen Medium gelösten Proteinen betrachtet werden kann. Vielmehr ist die Beweglichkeit eines Teilchens im Inneren einer mit Proteinen angefüllten Zelle — im Vergleich zu einem Reagenzglas — stark eingeschränkt.
- Nicht einmal die kinetischen Mechanismen einer enzymkatalysierten Reaktion können ohne weiteres von der In-vitro- auf die In-vivo-Situation übertragen werden. In jüngster Zeit häufen sich Indizien dafür, daß Enzyme in der lebenden Zelle anders arbeiten können als in wässriger Phase. Hierbei spielt ihre räumliche Anordnung in der Zelle eine wesentliche Rolle [Sre90, Mat93]. In der Literatur wird von sogenannten Enzymkomplexen, Metabolons oder Kanälen gesprochen [Mat93]. In einem Enzymkomplex wird die Verfügbarkeit eines Zwischenprodukts für außerhalb liegende Enzyme stark herabgesetzt. Es kommt zu einem sogenannten Channelling-Effekt.
- Zusätzlich zur Enzymkinetik ist bei der Beschreibung lebender Zellen auch die Kinetik von Transportschritten über Membranen zu berücksichtigen. Bei höheren Zellen ist insbesondere auch die intrazelluläre Kompartimentierung zu beachten, was wiederum die Komplexität der zu entwickelnden Modelle erhöht.
- Viele enzymatische Schritte laufen in vivo in beiden Richtungen gleichzeitig ab (vgl. Abb. 1.8). In vielen Fällen findet dieser Ausgleich extrem schnell statt, so daß die betroffenen Pools ohne weiteres zusammengefaßt werden können (Lumping). Jedoch kann auch ein

langsamer Austausch auftreten. In diesem Fall müssen enzymkinetische Terme für bidirektionale Reaktionen eingesetzt werden [Seg75, MTBK89, WBA92].⁶

Die genannten Bedenken sprechen nicht grundsätzlich gegen einen dynamischen Ansatz zur Beschreibung metabolischer Systeme. In vielen Fällen hat sich herausgestellt, daß auf eine detaillierte Beschreibung vieler enzymatischer Mechanismen weitgehend verzichtet werden kann, weil das System insensitiv gegenüber den auftretenden Parametern ist. Biologische Systeme müssen unter sehr stark wechselnden äußeren Bedingungen ein stabiles Verhalten zeigen. Es erstaunt daher nicht, wenn alle Regulationsmechanismen extrem robust sind, d.h. auf Schwankungen in Eingangsgrößen und Systemparametern innerhalb großer Bereiche unempfindlich reagieren. Das Ziel einer Modellierung muß sein, die regulatorisch relevanten Schritte zu identifizieren.

2.6.3 Einordnung der stationären Stoffflußanalyse

Die genannten Einwände zeigen generell, daß die derzeit verfügbare quantitative Information über die Dynamik von Stoffwechselsystemen *in vivo* immer noch zu gering für eine detaillierte Modellierung ist. Gerade hier setzt die vorliegende Arbeit an. Denn es wird der Versuch gemacht, eine Quantifizierung aller relevanten intrazellulären Vorgänge unter stationären Bedingungen zu erreichen, die vergleichsweise wenige Annahmen über das lebende System erfordern. Das gewonnene Datenmaterial kann dann eine weitergehende kinetische Modellierung stützen und zur Validierung dynamischer Modelle beitragen.

Die nachfolgend zusammengestellten Beziehungen zwischen dem stationären und dem dynamischen Ansatz machen deutlich, daß beide Methoden im wesentlichen komplementär zueinander sind:

- Es ist gängige Praxis, kinetische Modelle für die Stoffwechseldynamik "um einen stationären Zustand herum" zu entwickeln [SV91, WBA92]. D.h. das Modell soll zunächst die stationären Zustände des Systems in Abhängigkeit von der Substratkonzentration reproduzieren. Für den Übergang zwischen zwei stationären Zuständen wird dann ein quasistationäres Verhalten angenommen.⁷ Auf diese Weise werden aus den stationären kinetischen Gleichungen unmittelbar die entsprechenden instationären Gleichungen gewonnen. Stationäre Stoffflußdaten können also wichtige Ecksteine für die dynamische Modellierung liefern.
- Bei der stationären Stoffflußanalyse kann auf viele Annahmen, insbesondere über den Energiehaushalt der Zelle und kinetische Mechanismen, verzichtet werden. Aus diesem Grunde ist den geschätzten stationären intrazellulären Flüssen eine hohe Zuverlässigkeit und Genauigkeit zuzuschreiben. Baut die dynamische Modellierung auf einem derartigen Datensatz auf, so wird eine zweistufige Vorgehensweise erreicht, was die Modellvalidierung erleichtert.

⁶Als einfachstes Beispiel sei die reversible Michaelis-Menten-Kinetik genannt [Seg75], die einen bidirektionalen Michaelis-Menten-Mechanismus beschreibt (S Substrat-, P Produktkonzentration):

$$v(S, P) = \frac{v_s/k_S \cdot S - v_P/k_P \cdot P}{1 + 1/k_S \cdot S + 1/k_P \cdot P}$$

Die Zahl der Parameter erhöht sich gegenüber dem einfachen Michaelis-Menten-Mechanismus um zwei.

⁷Über die Korrektheit von Quasi-Stationaritätsannahmen soll hier nicht diskutiert werden (vgl. [Roe83, Wie90]).

- Die stationäre Stoffflußanalyse mit Hilfe von Markierungsexperimenten erlaubt den Nachweis bidirektionaler Reaktionsschritte *in vivo* [WGM95]. Dies ist eine wichtige Hilfe bei der Formulierung der enzymkinetischen Mechanismen.
- Die Resultate der stationären Flußanalyse sind offenbar nur in einem sehr engen physiologischen Zustandsbereich gültig [VP94b]. Vorhersagen über das regulatorische Verhalten unter geänderten physiologischen Bedingungen sind grundsätzlich nicht möglich. Gerade hierin liegt die Stärke von kinetischen Modellen. Im Extremfall könnten — validierte — kinetische Modelle des Stoffwechsels sogar die Vorhersage der Auswirkung von genetischen Modifikationen ermöglichen.

Sowohl für den stationären als auch für den dynamischen Ansatz ist eine Serie aufeinanderfolgender Experimente unter verschiedenen physiologischen Bedingungen erforderlich, um aussagekräftiges Datenmaterial zu produzieren.

Nach dieser Einordnung der stationären Stoffflußanalyse in den Gesamtzusammenhang der Modellierung von Stoffwechselsystemen, konzentriert sich diese Arbeit im folgenden auf die Auswertung und Analyse von metabolisch stationären Experimenten.

Teil II

Modellbildung

Kapitel 3

Kohlenstoff-Netzwerke

Dieses Kapitel leitet in die modellmäßige Beschreibung von Markierungsexperimenten ein. Aus einer im Vergleich zu den vorangehenden Kapiteln stärker systemtheoretischen Perspektive wird ein metabolisches System und sein dynamisches Verhalten bei ^{13}C -Markierungsexperimenten als ein *Metabolisches Kohlenstoff-Markierungssystem* bezeichnet. Grundlage für die Beschreibung eines Markierungssystems ist die Festlegung der diskreten Netzwerkstruktur, also des Metabolit-, C-Atom- und Isotopomeren-Netzwerks. Diese erfolgt in vorliegendem Kapitel — mit Rücksicht auf den experimentellen Anwender — zunächst auf einem informalen Niveau. In [Wie94b] wird weiterführend eine abstrakte formale Beschreibung metabolischer Netzwerke mit Hilfe graphentheoretischer Methoden vorgestellt.

3.1 Metabolit-, C-Atom- und Isotopomeren-Netzwerke

Dieser Abschnitt befaßt sich mit der Spezifikation der Struktur eines metabolischen Netzwerkes im Hinblick auf Kohlenstoff-Markierungsexperimente. Die hierzu verwendete formale Notation soll den folgenden Anforderungen genügen:

1. Unmittelbare Verständlichkeit für den experimentellen Anwender.
2. Dokumentierbarkeit getroffener Annahmen.
3. Überschaubarkeit und Kompaktheit.
4. Maschinenlesbarkeit als Grundlage für eine automatische Generierung entsprechender Datenstrukturen.

Es genügt bereits, wenn das Metabolit- und Kohlenstoff-Netzwerk durch die Notation eindeutig festgelegt wird, da alle Isotopomeren-Übergänge daraus rekonstruiert werden können [Wie94b].

3.1.1 Formale Notation von Reaktionsschritten

Die nachfolgende erstmals in [WG93] vorgestellte Repräsentation von Reaktionsschritten lehnt sich an die bekannte Summennotation für biochemische Reaktionen an [Val91]. Sie wird erweitert durch die Angabe der Atom-Transitionen aller an einer Reaktion beteiligten Kohlenstoff-

Atome. Ergänzend werden die Bezeichner für C-Atom- und Isotopomeren-Pools aus denen der entsprechenden Metabolit-Pools abgeleitet.¹

Definition 3.1 (formale Notation von Metaboliten, C-Atomen, Isotopomeren und Reaktionsschritten)

1. *Metabolite werden mit frei wählbaren alphanumerischen Bezeichnern in einer serifenlosen Schrift bezeichnet, also beispielsweise für den Glycerinaldehyd-3-phosphat-Pool:*

GA3P

2. *Die zu einem Metaboliten gehörenden Kohlenstoff-Atome werden unter Verwendung des #-Symbols und der Angabe der Nummer des Kohlenstoff-Atoms notiert, also beispielsweise als*

GA3P#1, GA3P#2, GA3P#3 .

3. *Zur Bezeichnung von Isotopomeren wird eine Binärcode-Notation unter Verwendung der Zeichen 0 und 1 für unmarkierte und markierte C-Atome verwendet (vgl. auch [FR95]). Beispielsweise werden die acht verschiedenen Isotopomere von GA3P notiert als²*

GA3P#000, GA3P#001, GA3P#010, GA3P#011,
GA3P#100, GA3P#101, GA3P#110, GA3P#111.

4. *Reaktionsschritte werden ebenfalls mit alphanumerischen Bezeichnern versehen. Beispielsweise kann die Transaldolase-Reaktion im Pentosephosphatweg bezeichnet werden mit (vgl. Tabelle 3.2 zur verwendeten Numerierung von Reaktionsschritten)*

PPW5 .

5. *Schließlich wird ein Reaktionsschritt mit einer durch die C-Atom-Transitionen erweiterten chemischen Summennotation spezifiziert. Die Gleichung*

PPW5: GA3P + S7P > E4P + Fru6P
#ABC + #abcdefg > #defg + #abcABC

besagt, daß das erste Kohlenstoff-Atom von GA3P (hier mit dem Buchstaben A hinter dem #-Zeichen bezeichnet) zum vierten Kohlenstoff-Atom von Fru6P wird usw. Die Buchstaben in der zweiten Zeile sind frei wählbar, dürfen auf jeder Seite der Reaktionsgleichung jedoch nur einmal benutzt werden.

¹Die Einführung der Notation erfolgt hier auf einem allgemein verständlichen informellen Niveau. Ein formales System, das eine Netzwerkspezifikation im mathematisch strengen Sinne erlaubt, und den Entwurf von Algorithmen zur Netzwerksynthese ermöglicht, wird in [Wie94b] vorgestellt.

²Bei einatomigen Metaboliten kann diese Notation mehrdeutig sein. Es wird jedoch darauf verzichtet, diesen Sonderfall eigens zu berücksichtigen.

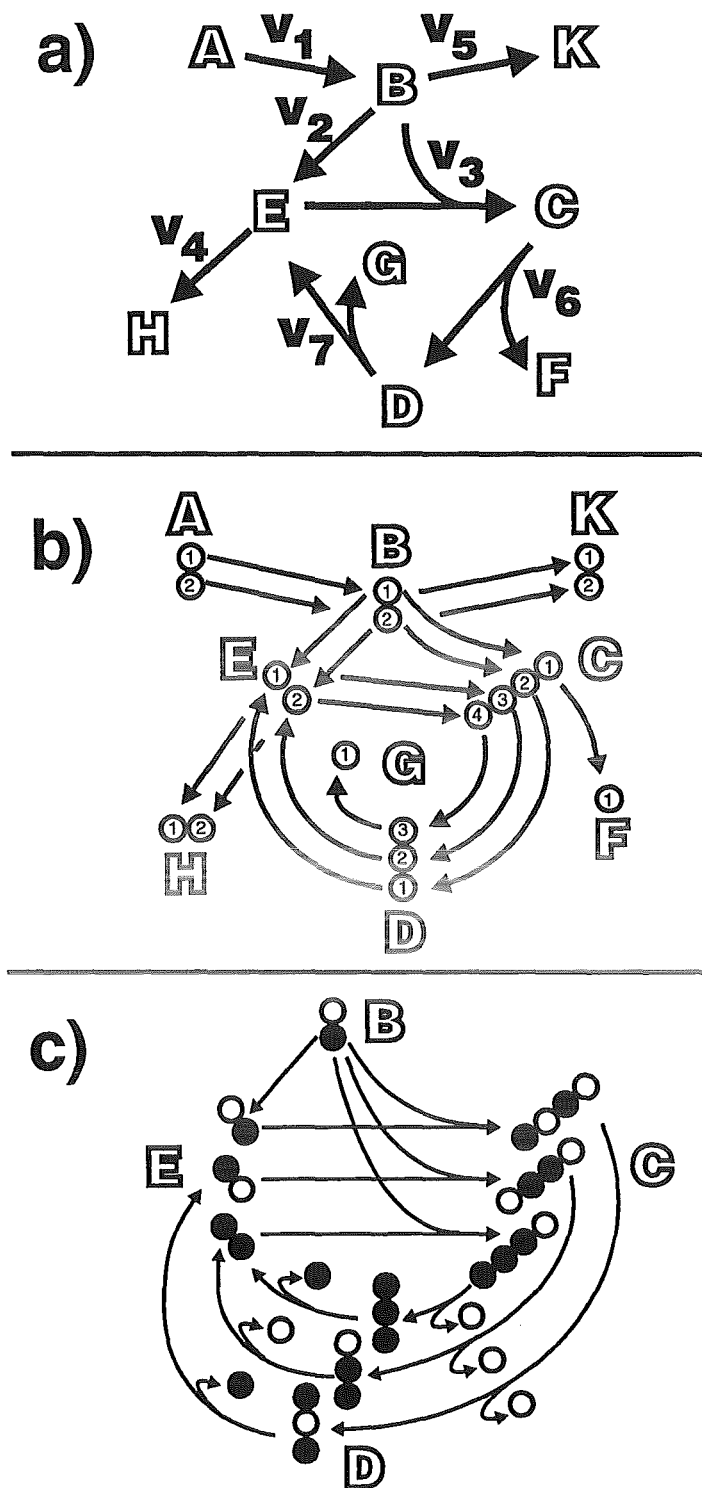


Abbildung 3.1: Elementares Beispiel zur Veranschaulichung des a) Metabolit-, b) C-Atom- und c) Isotopomeren-Netzwerks eines metabolischen Kohlenstoff-Markierungssystems. Das Isotopomeren-Netz zeigt nicht alle möglichen Isotopomere, sondern wurde auf diejenigen reduziert, die bei einem Markierungs-Input mit A#01 — unter der Voraussetzung, daß alle Schritte nur in der angezeigten Richtung ablaufen — tatsächlich entstehen. Die schwarzen und weißen Kreise stehen für markierte und unmarkierte C-Atome.

3.1.2 Ein einfaches Beispiel

Ein einfaches Beispiel, das im Verlauf dieser Arbeit noch häufiger herangezogen wird, soll die eingeführten Notationen verdeutlichen (vgl. Abb. 3.1). Es stellt einen einfachen metabolischen Zyklus dar und ist in der formalen Notation gegeben durch:



Es gibt drei verschiedene Sichtweisen dieses Netzwerks, deren wechselseitige Beziehung im weiteren Verlauf dieser Arbeit noch analysiert wird (vgl. Definition 4.11).

Definition 3.2 (Metabolit-, C-Atom-, Isotopomeren-Netzwerk) Das Metabolit-Netzwerk besteht aus allen Metaboliten des metabolischen Systems und den zwischen ihnen stattfindenden Reaktionen (Abb. 3.1a). Das C-Atom-Netzwerk besteht aus allen zu den Metaboliten gehörigen Kohlenstoff-Atomen und deren Transitionen (Abb. 3.1b). Das Isotopomeren-Netzwerk besteht aus allen möglichen Isotopomeren der Metabolite und den zwischen ihnen stattfindenden Reaktionen (Abb. 3.1c).

3.1.3 Formale Beschreibung der Netzwerke

Unter Verwendung der formalen Notation lassen sich die einzelnen Netzwerke wie folgt notieren:

Metabolit-Netzwerk: Dieses setzt sich unmittelbar aus den ersten Zeilen der formalen Festlegung des Netzwerks (Definition 3.1) zusammen:



C-Atom-Netzwerk: Auf der Ebene der C-Atome stellt sich das Netz als eine Reihe von einzelnen Atom-Transitionen dar:

v1: A#1 > B#1	v5: B#1 > K#1
v1: A#2 > B#2	v5: B#2 > K#2
v2: B#1 > E#1	v6: C#1 > F#1
v2: B#2 > E#2	v6: C#2 > D#1
v3: B#1 > C#1	v6: C#3 > D#2
v3: B#2 > C#2	v6: C#4 > D#3
v3: E#1 > C#3	v7: D#1 > E#1
v3: E#2 > C#4	v7: D#2 > E#2
v4: E#1 > H#1	v7: D#3 > G#1
v4: E#2 > H#2	

Isotopomeren-Netzwerk: Auf der Ebene der Isotopomeren müssen schließlich alle möglichen Reaktionen zwischen Isotopomeren aufgeführt werden. Dies führt bei der bimolekularen Reaktion v3 zu insgesamt $2^2 \cdot 2^2 = 16$ Einzelreaktionen. Von diesem Netzwerk wird daher aus Platzgründen nur ein Teil wiedergegeben:

v1: A#00 > B#00	v3: B#00 + E#00 > C#0000
v1: A#01 > B#01	v3: B#00 + E#01 > C#0001
v1: A#10 > B#10	v3: B#00 + E#10 > C#0010
v1: A#11 > B#11	v3: B#00 + E#11 > C#0011
v2: B#00 > E#00	v3: B#01 + E#00 > C#0100
v2: B#01 > E#01	v3: B#01 + E#01 > C#0101
v2: B#10 > E#10	v3: B#01 + E#10 > C#0110
v2: B#11 > E#11	v3: B#01 + E#11 > C#0111
v4: E#00 > H#00	v3: B#10 + E#00 > C#1000
v4: E#01 > H#01	v3: B#10 + E#01 > C#1001
v4: E#10 > H#10	v3: B#10 + E#10 > C#1010
v4: E#11 > H#11	v3: B#10 + E#11 > C#1011
	v3: B#11 + E#00 > C#1100
	v3: B#11 + E#01 > C#1101
	v3: B#11 + E#10 > C#1110
	v3: B#11 + E#11 > C#1111

...

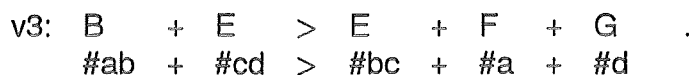
Bereits bei einem derart einfachen Beispiel ist die Anzahl der möglicherweise auftretenden Isotopomere mit insgesamt 48 so groß, daß eine graphische Darstellung nicht mehr überschaubar ist. Die Abb. 3.1c zeigt daher nur einen Ausschnitt des vollständigen Isotopomeren-Netzwerks unter der Annahme, daß sich das Substrat A nur aus einem einzigen Isotopomer A#01 zusammensetzt. In dieser Situation entsteht nur ein Bruchteil der möglichen Isotopomeren im System.

3.1.4 Netzwerkvereinfachung

Bereits angesichts des einfachen Beispiels ist ersichtlich, daß bei hinsichtlich der experimentellen Anwendung realistisch komplexen Netzwerken derart viele Isotopomere zu erwarten sind, daß sich eine manuelle Eingabe von Isotopomeren-Netzwerken in der oben verwendeten Notation aufgrund ihrer Fehleranfälligkeit verbietet. Analoges gilt für die im nächsten Kapitel eingeführten Bilanzgleichungen, denn jedem Metaboliten, C-Atom oder Isotopomer entspricht eine solche Gleichung und jeder Reaktion, C-Atom-Transition oder Isotopomeren-Reaktion ein

Term in einer der Gleichungen. Die automatische Erzeugung von metabolischen Kohlenstoff-Netzwerken wird in Abschnitt 9.3 behandelt. Andererseits kann eine Netzwerkvereinfachung, bei der die Markierungsinformation nicht verringert wird, die Zahl der Gleichungen erheblich reduzieren. Eine solche Vereinfachung wird nun für das oben beschriebene System aus Abb. 3.1 durchgeführt, um seine spätere algebraische Behandlung zu erleichtern.

Ein Ad-hoc-Ansatz zur Netzwerkvereinfachung bei isotopisch stationären Systemen beruht wesentlich auf der Kenntnis unidirektionaler Reaktionsschritte. Im Beispiel sei angenommen, daß die Reaktionsschritte v6 und v7 nur in einer Richtung ablaufen. Unter diesen Umständen können unverzweigte Reaktionssequenzen zusammengezogen werden. Dies ist auch dann noch möglich, wenn dabei Produkte abgespalten werden. Dementsprechend ist der Grundgedanke zur Vereinfachung des vorliegenden Beispiels das Zurückverfolgen der Produkt-C-Atome F#1 und G#1 durch das System. Offenbar macht es keinen Unterschied, ob diese Atome bereits auf der Stufe B#1 und E#2 oder erst von C#1 und D#3 abgespalten werden. Dies erlaubt, die Kohlenstoff-Atome C#1, D#3 und C#4 ganz aus dem System zu entfernen. Nach der Elimination gehen von den Metaboliten C und D keine Verzweigungen mehr ab, folglich können auch sie weggelassen werden. Das Netzwerk reduziert sich dann auf die in Abbildung 3.2 gezeigten Reaktionsschritte v1, v2, v4 und v5 mit dem veränderten Reaktionsschritt



Obwohl dieses Netzwerk im Vergleich zu realen metabolischen Netzwerken ungewöhnlich aussieht, können alle quantitativen Beziehungen, die für das isotopisch stationäre Originalnetzwerk gelten, aus der Analyse des reduzierten Netzwerks rekonstruiert werden (vgl. auch Abschnitt 7.4). Der tiefere Grund hierfür liegt darin, daß jeder der durchgeführten Netzwerk-Vereinfachungsschritte eine Äquivalenzumformung der entsprechenden in Kapitel 4 eingeführten Bilanzgleichungssysteme entspricht. Auf die Ausführung der technischen Details wird hier verzichtet. Das reduzierte Netzwerk wird in den nachfolgenden Kapiteln stets anstelle des ursprünglichen Netzwerks als Beispiel herangezogen.

Die oben vorgestellten informellen Überlegungen geben einen Eindruck davon, welche typischen Vereinfachungsoperationen üblicherweise angewendet werden, um Netzwerke zu reduzieren (vgl. auch [WK84, JRMS91]). Allgemeine Ansätze zur Netzwerkvereinfachung werden in [RML93, GS91] behandelt.

3.2 Formulierung realer Netzwerke

3.2.1 Einige Begriffsdefinitionen

Das Beispielnetzwerk kann herangezogen werden, um einige Begriffe einzuführen, die nachfolgend durchgängig verwendet werden.

Definition 3.3 (intra- und extrazelluläre Metabolite, Input- und Output-Metabolite) *Die Menge der Metabolite eines Netzwerks läßt sich wie folgt in zwei Teilmengen aufteilen:*

1. *Alle Metabolite, die sowohl auf der linken als auch auf der rechten Seite von Reaktionsgleichungen auftreten, heißen intrazelluläre Metabolite oder auch Intermediate.*
2. *Alle übrigen Metabolite werden als extrazellulär bezeichnet. Diese Menge zerfällt weiter in zwei Untermengen:*

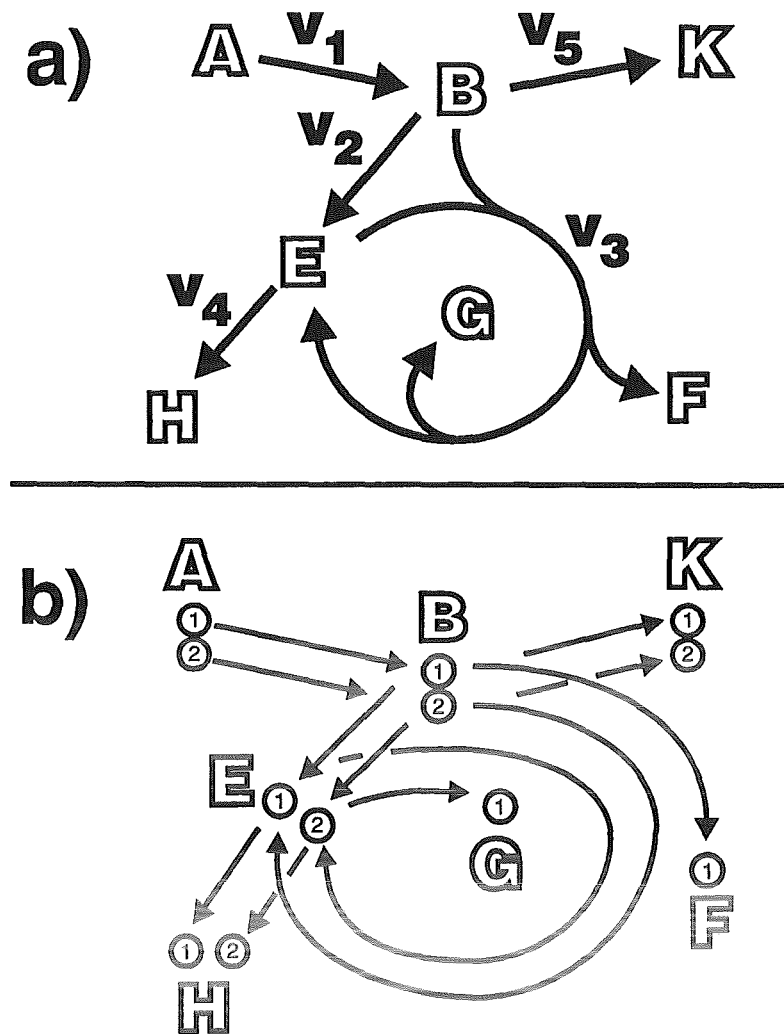


Abbildung 3.2: Aus Abb. 3.1 durch Netzwerkvereinfachung gewonnenes Netzwerk.

a) Metabolit-, b) C-Atom-Netz.

(a) Metabolite, die ausschließlich auf der linken Seite von Reaktionsgleichungen auftreten, heißen Input-Metabolite oder auch Substrate.

(b) Metabolite, die ausschließlich auf der rechten Seite von Reaktionsgleichungen auftreten, heißen Output-Metabolite oder auch Produkte.

Analog wird zwischen *intrazellulären*, *Input-* und *Output-Kohlenstoff-Atomen* sowie *intrazellulären*, *Input-* und *Output-Isotopomeren* unterschieden.

Definition 3.4 (intra- und extrazelluläre Reaktionsschritte, Input- und Output-Reaktionsschritte) Entsprechend dieser Terminologie werden auch die Reaktionsschritte in Teilmengen unterteilt:

1. Intrazelluläre Reaktionsschritte verbinden ausschließlich intrazelluläre Metabolite miteinander.
2. Alle übrigen Reaktionsschritte heißen extrazellulär. Weiter wird unterschieden zwischen:

- (a) Input-Reaktionen oder Substrataufnahmeschritten, die einen Input-Metaboliten auf der linken Seite der zugehörigen Reaktionsgleichung aufweisen.
- (b) Output-Reaktionen oder Produktbildungsschritten, die einen Output-Metaboliten auf der rechten Seite der zugehörigen Reaktionsgleichung aufweisen.

Hiermit wird nicht ausgeschlossen, daß ein Reaktionsschritt sowohl Input- als auch Output-Metabolite verbindet. Im laufenden Beispiel aus Abbildung 3.1 ergibt sich die nachfolgende Klassifikation der Metabolite und Reaktionen:

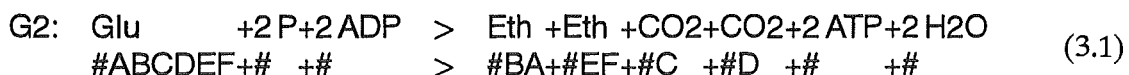
	Metabolite	Flüsse
intrazellulär	B, C, D, E	v2, v3
Input	A	v1
Output	F, G, H, K	v4, v5, v6, v7

3.2.2 Erweiterte Ausdrucksmöglichkeiten und einschränkende Bedingungen

Die Netzwerknnotation soll nun um eine Ausdrucksmöglichkeit zur Bilanzierung von Stoffen mit am Markierungsexperiment unbeteiligten C-Atomen erweitert werden. Beispielsweise enthält ATP zwar C-Atome, diese nehmen aber nicht an den Reaktionen des Zentralstoffwechsels teil. So wird etwa die alkoholische Gärung notiert in der Form



Darüber hinaus dürfen unbeteiligte Metabolite — im Gegensatz zu solchen mit bilanzierten C-Atomen — auch stöchiometrische Koeffizienten führen, die sogar nicht-ganzzahlig sein können (vgl. dazu den nachfolgenden Abschnitt):



Bei den bisherigen Ausführungen nicht berücksichtigt wurde die Uni- bzw. Bidirektionalität einer Reaktion. Nachfolgend soll ein formal notierter Reaktionsschritt grundsätzlich sowohl für die Hin- als auch für die mögliche Rückreaktion dieses Schrittes stehen. Die Annahme, ob dieser Schritt bidirektional oder unidirektional verläuft, wird gesondert formuliert. Detaillierte Ausführungen dazu folgen in Kapitel 6.

Abschließend werden einige einschränkende Bedingungen für die formale Notation von Reaktionsschritten formuliert. Sie sollen eine spätere automatische Behandlung sowie die Modellbildung erheblich vereinfachen, indem Sonderfälle ausgeschlossen werden. Diese Ausnahmesituationen können jedoch bei Bedarf leicht in einer alternativen, zulässigen Form formuliert werden. Der nachfolgende Abschnitt bietet dazu einige Beispiele.

Definition 3.5 (zulässige Netzwerkspezifikation) Eine formale Netzwerkspezifikation mit Hilfe der formalen Notation aus Definition 3.1 heißt zulässig, wenn sie die nachfolgenden Bedingungen erfüllt:

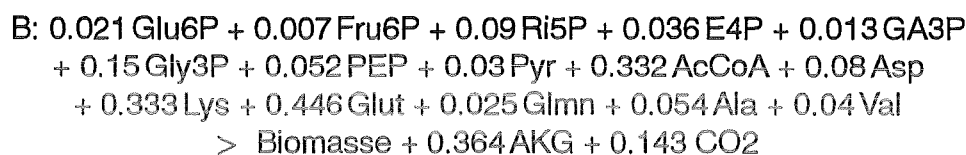
1. Ein Metabolit, dessen Kohlenstoff-Transitionen spezifiziert sind, tritt immer ohne Koeffizient auf.

2. Extrazelluläre Flüsse haben stets genau einen Metaboliten auf der linken und rechten Seite.³
3. Output-Metabolite stehen nur in genau einer Gleichung.
4. Für Input- und Output-Reaktionsschritte wird grundsätzlich die Unidirektionalität angenommen.
5. Falls Isotopomeren-Netzwerke betrachtet werden, wird jede Reaktion auf maximal zwei Edukte und Produkte mit berücksichtigten C-Atome beschränkt.⁴

3.2.3 Formulierung von Sonderfällen

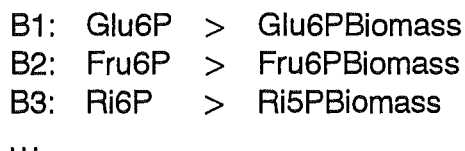
Trotz der vorgenommenen Einschränkungen lassen sich viele spezielle Strukturen in metabolischen Netzwerken leicht so notieren, daß sie den Forderungen der Definition 3.5 genügen. Die nachfolgenden Beispiele verdeutlichen zugleich die Gründe für die formulierten Einschränkungen.

Nicht-ganzzahlige Koeffizienten : In der Literatur über die Metabolit-Bilanzierungstechnik ist es üblich, eine summarische "Biomassebildungsreaktion" anzugeben, die alle Abflüsse in die Biomasse zusammenfaßt (nach [Roe83, VS93]):



Die Koeffizienten geben hierbei die Biomassezusammensetzung an — sie entsprechen also den c_i in Gleichung (2.3) — während die "Reaktionsgeschwindigkeit" gerade die Wachstumsrate μ ist.

Bei Markierungsexperimenten ist eine derartige Zusammenfassung nicht mehr sinnvoll, da dabei alle Metabolite samt ihren C-Atomen bis in die Biomasse hinein verfolgt werden müssen. Dazu kann die summarische Reaktion in einzelne Biomasseabflüsse zerlegt werden, also:

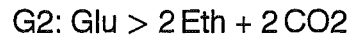


Der jeweils entsprechende "stöchiometrische Koeffizient" wird als Meßwert des Biomasseabflusses bereitgestellt (entsprechend $c_i \cdot \mu$ in Gleichung (2.3)).

Ganzzahlige Koeffizienten : Müssen einzelne Kohlenstoff-Atome durch das System verfolgt werden, so sind ganzzahlige Koeffizienten ungleich eins nicht sinnvoll interpretierbar. Wie die Reaktion

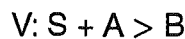
³Für das Beispielnetzwerk aus Abbildung 3.1 trifft dies offenbar nicht zu. Es wird jedoch im nachfolgenden Abschnitt gezeigt, wie die Bedingung durch leichte Modifikation stets erfüllt werden kann.

⁴Diese Einschränkung hat lediglich technische Gründe, da sie den in den nachfolgenden Kapiteln entwickelten mathematischen Formalismus zugunsten einer größeren Klarheit entlastet.



unter Verdopplung der Metabolit-Symbole Eth und CO₂ ohne Koeffizienten formuliert wird, zeigt bereits Gleichung (3.1).

Reaktionsgekoppelte extrazelluläre Flüsse : Substrataufnahme- und Produktbildungsschritte werden grundsätzlich als unidirektional angesehen. Dies ist eine notwendige Voraussetzung für die Modellbildung, da bei Umkehrung eines solchen Schrittes aus einem Substrat ein Produkt würde und umgekehrt. Dieser Fall kann jedoch modelltechnisch nicht erfaßt werden (vgl. Definition 4.2 bzw. 4.3).⁵ Reaktionsgekoppelte extrazelluläre Flüsse, wie z.B.



mit einem extrazellulären Substrat S und einem intrazellulären Metaboliten A, können unter Einführung eines Hilfsintermediats wie folgt umformuliert werden zu

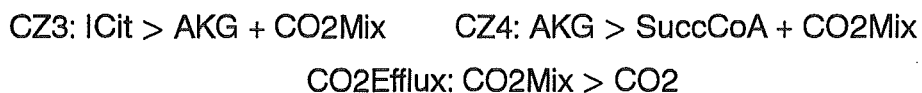


Die Unidirektionalitätsannahmen für den extrazellulären Substrataufnahmeschritt V1 hat dann keine Konsequenzen für den intrazelluläre Fluß V2. Analoges gilt für Produktbildungsschritte.

Produktabfluß an mehreren Stellen : Produkte sollen das System stets an genau einer Stelle verlassen. Dies entspricht der meßtechnischen Realität, da verschiedene Stoffströme eines Stoffes nicht voneinander getrennt werden können. Beispielsweise entsteht CO₂ an zwei Stellen im Citratzyklus:



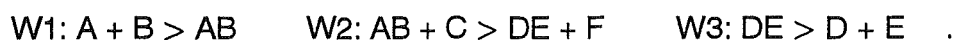
Die Gasabflüsse werden aber im Medium zunächst vermischt, bevor sie das Meßgerät erreichen. Diese Durchmischung läßt sich auch innerhalb der formalen Notation unter Zuhilfenahme eines Misch-Pools realistisch formulieren:



Multimolekulare Reaktionsschritte : Enzymatische Umsetzungen (insbesondere solche mit Cosubstraten) können durchaus zwei und mehr Partner haben. Jedoch sind stets nur C-Atome von höchstens zwei Partnern an der Reaktion beteiligt. Aus diesem Grunde genügt es hier, sich auf monomolekulare und bimolekulare Reaktionen zu beschränken. Trotzdem ergeben sich bei Zusammenfassung von Reaktionsschritten oft formal Reaktionen mit mehr als zwei Partnern. Diese können jedoch unter Zuhilfenahme von Hilfsmetaboliten umformuliert werden. Aus



wird beispielsweise



⁵Muß ein solcher Fall berücksichtigt werden (z.B. bei einem wechselseitigen Austausch zwischen einer Medienkomponente und dem Cytosol), so muß ein entsprechender Pool (z.B. für die Medienkomponente) mit zum Stoffwechselnetzwerk hinzugenommen werden.

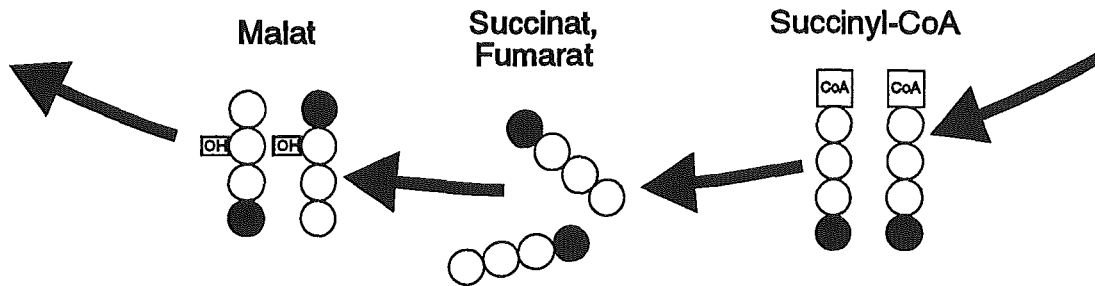
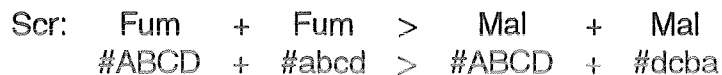


Abbildung 3.3: Markierungs-Scrambling im Citratzyklus.

3.2.4 Scrambling-Reaktionen

Das sogenannte *Markierungs-Scrambling* im Citratzyklus ist ein in der gesamten Literatur zu Markierungsexperimenten bekanntes Phänomen [MSJ88, SSMS93]. Dabei handelt es sich um die beiden Reaktionsschritte vom Succinyl-Coenzym A zum Malat (Abb. 3.3). Die beiden Intermediate Succinat und Fumarat sind jeweils rotationssymmetrische Moleküle, d.h. das Molekül ist nicht von dem um 180 Grad gedrehten Teilchen unterscheidbar. Die C-Atom-Positionen 1-4 kehren also bei einer Drehung im Raum ihre Reihenfolge um. Eine Markierung wechselt damit ihre Position, ohne daß eine chemische Reaktion stattfindet. Dieser Effekt verringert offenbar die aus den Markierungsdaten zu gewinnende Information.

In der Regel wird die Wahrscheinlichkeit eines Markierungs-Scrambling in einem der beiden Schritte mit 50% angenommen. Dies läßt sich z.B. wie folgt formal notieren:



Allerdings mehren sich die Anzeichen, daß der Citratzyklus durch einen Enzymkomplex katalysiert wird [Sre90, ESS+93]. Die dabei entstehenden Zwischenprodukte verbleiben dann in der Umgebung des Komplexes, d.h. sie sind im Raum nicht frei beweglich. In diesem Fall ist die 50%-Scrambling-Annahme nicht korrekt. Ein Nachweis dieser Hypothese wird derzeit versucht (vgl. aber auch [SSMS93]).

Eine weitere Scrambling-Reaktion tritt speziell bei der Lysinproduktion in *Corynebacterium glutamicum* auf. Hier wird Lysin auf zwei parallelen Wegen aus Pyruvat und Oxalacetat gebildet, nämlich dem Diaminopimelat-Dehydrogenase-Weg (ddh) und dem Succinyl-Diaminopimelat-Dehydrogenase-Weg (succ-dap) [SEGS93]. Der zweite Weg enthält einen Scrambling-Schritt, so daß genaugenommen drei Reaktionswege vorliegen. Jedoch liefern diese drei Wege nur zwei verschiedene Vertauschungsmuster für die Kohlenstoff-Atome, so daß nachfolgend die beiden Wege mit gleichen Atom-Transitionen zu einem zusammengefaßt werden. Die 50%-Annahme trifft für die verbleibenden beiden Wege natürlich nicht mehr zu.

3.3 Beispielnetzwerke

3.3.1 Typen von Beispielen

Im Verlaufe dieser Arbeit werden diverse Beispiele herangezogen, die Eigenschaften von Kohlenstoff-Markierungsnetzwerken illustrieren oder konkrete Experimente beschreiben. Die entsprechenden graphischen Darstellungen geben jeweils das Metabolit-Netzwerk und bei Beispielen geringen Umfangs ggfs. auch das C-Atom-Netzwerk wieder. Es kann zwischen drei Kategorien unterschieden werden:

Ac	Acetat	Gly3P	Glycerin-3-phosphat
AcCoA	Acetyl-Coenzym A	ICit	Isocitrat
AKG	α -Ketoglutarat	Lact	Lactat
Ala	Alanin	Lys	Lysin
Asp	Aspartat	Mal	Malat
CO ₂	Kohlendioxid	OA	Oxalacetat
E4P	Erythrose-4-phosphat	PEP	Phosphoenolpyruvat
Eth	Ethanol	Pyr	Pyruvat
Fru6P	Fructose-6-phosphat	Ri5P	Ribose-5-phosphat
Fru16BP	Fructose-1,6-bisphosphat	Ru5P	Ribulose-5-phosphat
Fum	Fumarat	Succ	Succinat
GA3P	Glycerinaldehyd-3-phosphat	SuccCoA	Succinyl-Coenzym A
Glu	Glucose	S7P	Sedoheptulose-7-phosphat
Glut	Glutamat	Xyl5P	Xylulose-5-phosphat
Glmn	Glutamin	Xy	Xylose
Glu6P	Glucose-6-phosphat	Val	Valin
GlyOx	Glyoxylat		

Tabelle 3.1: In diesem Text verwendete Stoffabkürzungen.

Elementare Beispiele: Sehr einfach konstruierte Beispiele dienen dazu, grundsätzliche systemtheoretische Eigenschaften von Kohlenstoff-Markierungsnetzwerken aufzuzeigen. Sie sind möglichst elementar gewählt, um einen bestimmten Effekt herauszustellen oder als didaktische Hilfe zu dienen. Zu dieser Kategorie gehören die Beispiele aus Abbildung 1.7, 3.1 und 3.2.

Stoffwechsell Ausschnitte: Reale Netzwerke in ihrer vollen Komplexität entziehen sich meist einer expliziten Analyse mit algebraischen Methoden. Jedoch können Ausschnitte aus komplexen Netzwerken oder Vereinfachungen solcher Ausschnitte herangezogen werden, um bestimmte bei einer numerischen Auswertung eines konkreten Experiments aufgetretene Effekte zu analysieren und besser zu verstehen. Ein solches Beispiel zeigt Abbildung 9.5 (vgl. auch [Wie96a]).

Komplexe Stoffwechselmodelle: Komplette Abschnitte des Zentralstoffwechsels können ohne einschränkende Annahmen nur mit Hilfe numerischer Methoden auf dem Wege der Simulation untersucht werden. Insbesondere erfordert die Auswertung des Datenmaterials aus gegebenen Experimenten die Behandlung des ganzen Netzes. Beispiele zeigen die Abbildungen 9.1 und 12.3.

3.3.2 Reaktionen im Zentralstoffwechsel

Ein Metabolit-Netzwerk mit den wichtigsten Reaktionsschritten des Zentralstoffwechsels wurde bereits in Abb. 1.2 dargestellt. Die Tabellen 3.2 und 3.3 geben die entsprechenden C-Atom-Transitionen wieder. Sie sollen als Referenz für alle Netzwerkformulierungen im weiteren Text dienen.

Glykolyse					
Gly1:	Glu	+	GA3P	>	Glu6P + Pyr
	#ABCDE	+	#abc	>	#ABCDE + #abc
Gly2:	Glu6P			>	Fru6P
	#ABCDE			>	#ABCDE
Gly3:	Fru6P			>	Fru16BP
	#ABCDE			>	#ABCDE
Gly4:	Fru16BP			>	GA3P + GA3P
	#ABCDE			>	#CBA + #DEF
Gly5:	GA3P			>	PEP
	#ABC			>	#ABC
Gly6:	PEP			>	Pyr
	#ABC			>	#ABC

Pentosephosphatweg					
PPW0:	Xyl			>	Xy5P
	#ABCDE			>	#ABCDE
PPW1:	Glu6P			>	CO2 + Ru5P
	#ABCDE			>	#A + #BCDEF
PPW2:	Ru5P			>	Xyl5P
	#ABCDE			>	#ABCDE
PPW3:	Ru5P			>	Ri5P
	#ABCDE			>	#ABCDE
PPW4:	Xyl5P	+	Ri5P	>	S7P + GA3P
	#ABCDE	+	#abcde	>	#ABabcde + #CDE
PPW5:	GA3P	+	S7P	>	E4P + Fru6P
	#ABC	+	#abcdefg	>	#defg + #abcABC
PPW6:	Xyl5P	+	E4P	>	GA3P + Fru6P
	#ABCDE	+	#abcd	>	#CDE + #ABabcd

Produktbildung aus Pyruvat					
PB1:	Pyr			>	Eth + CO2
	#ABC			>	#BC + #A
PB2:	Pyr			>	Lact
	#ABC			>	#ABC

Tabelle 3.2: C-Atom-Transitionen für die wichtigsten Reaktionsschritte in der Glykolyse und dem Pentosephosphatweg sowie angrenzenden Produktbildungsschritten. Zu den Stoffabkürzungen vgl. Tabelle 3.1.

Citratzyklus					
CZ1:	Pyr		>	AcCoA	+ CO2
	#abc		>	#bc	+ #a
CZ2:	AcCoA	+ OA	>	ICit	
	#ab	+ #ABCD	>	#DCBAba	
CZ3:	ICit		>	AKG	+ CO2
	#ABCDEF		>	#ABCEF	+ #D
CZ4:	AKG		>	SuccCoA	+ CO2
	#ABCDE		>	#BCDE	+ #A
CZ5:	SuccCoA		>	Succ	
	#ABCD		>	#ABCD	
CZ6a:	Succ		>	Fum	
	#ABCD		>	#ABCD	
CZ6b:	Succ		>	Fum	
	#ABCD		>	#DCBA	
CZ7a:	Fum		>	Mal	
	#ABCD		>	#ABCD	
CZ7b:	Fum		>	Mal	
	#ABCD		>	#DCBA	
CZ8:	Mal		>	OA	
	#ABCD		>	#ABCD	

Glyoxylatzyklus					
GOZ1:	ICit		>	GlyOx	+ Succ
	#ABCDEF		>	#AB	+ #CDEF
GOZ2:	GlyOx	+ AcCoA	>	Mal	
	#AB	+ #ab	>	#ABba	

Anaplerotische Reaktionen					
AR1:	PEP	+ CO2	>	OA	
	#ABC	+ #a	>	#ABCa	
AR2:	OA		>	Pyr	+ CO2
	#ABCD		>	#ABC	+ #D
AR3:	Pyr	+ CO2	>	Mal	
	#ABC	+ #a	>	#ABCa	

Aminosäureproduktion					
AS1a:	OA	+ Pyr	>	Lys	+ CO2
	#ABCD	+ #abc	>	#ABCDcb	+ #a
AS1b:	OA	+ Pyr	>	Lys	+ CO2
	#ABCD	+ #abc	>	#abcDCB	+ #A
AS2:	AKG		>	Glut	
	#ABCDE		>	#ABCDE	

Tabelle 3.3: C-Atom-Transitionen für die wichtigsten Reaktionsschritte im Citratzyklus, im Glyoxylatzyklus und dem anaplerotischen Abschnitt sowie einige Aminosäureproduktionen. Zu den Stoffabkürzungen vgl. Tabelle 3.1. Die Buchstaben a und b in einem Flußnamen kennzeichnen die beiden Varianten einer Markierungs-Scrambling-Reaktion (vgl. Abb. 3.3).

Kapitel 4

Bilanzgleichungen

Bilanzgleichungen stellen das Kernstück der mathematischen Modellierung von Markierungssystemen dar. Kohlenstoff-Markierungsexperimente gehören zur allgemeinen Kategorie der sogenannten Tracer-Experimente, die in engem Zusammenhang zu Kompartimentmodellen stehen [And83, Hol85]. Allgemeine Modellansätze und mathematische Werkzeuge für die Analyse von Kompartimentmodellen sind schon in den 70er Jahren entwickelt worden. Eine exzellente Einführung gibt [And83]. Trotzdem sind einige spezielle Aspekte metabolischer Kohlenstoff-Markierungssysteme bei der Modellierung besonders zu berücksichtigen:

1. Bei klassischen — vor allem medizinischen — Anwendungen der Tracer-Technik findet man selten Systeme mit mehr als 10 Kompartimenten, von denen meistens nur ein oder zwei direkt beobachtet werden konnten [LR83]. Dieser Mangel an Meßwerten wurde dadurch kompensiert, daß zeitlich aufgelöste Datensätze für den instationären Verlauf des Markierungsexperimentes verfügbar waren. Im Falle komplexer metabolischer Systeme — insbesondere bei Anwendung der neuen Meßzeit-Entkopplungstechnik aus Definition 2.6 — ist die Datenlage gerade entgegengesetzt. Hier muß eine sehr große Anzahl von Kompartimenten in das System aufgenommen werden, nämlich eines für jedes Kohlenstoff-Atom im System. Jedoch sind nun sehr viele dieser Pools einer Quantifizierung zugänglich, wenn auch nur im isotopisch stationären Zustand. Im Gegensatz zum instationären Fall wurde die systematische Analyse derart komplexer stationärer Markierungssysteme, die im Mittelpunkt dieser Arbeit steht, bisher kaum thematisiert.
2. Ein weiteres Charakteristikum metabolischer Markierungsexperimente im Gegensatz zu vielen medizinischen Tracer-Experimenten besteht darin, daß jeder Metabolit-Fluß mehrere C-Atom-Transitionen zugleich impliziert. Offenbar finden alle diese Transitionen mit derselben Reaktionsgeschwindigkeit statt, wodurch die Anzahl der Freiheitsgrade für die C-Atom-Flüsse im System erheblich reduziert wird. In der algebraischen Beschreibung des Systems resultieren daraus vermehrt Redundanzbeziehungen, die in [Wie95a] erstmals systematisch analysiert wurden.
3. Wie in Abschnitt 4.3 gezeigt wird, kann die intrazelluläre Markierung sowohl von der Hin- als auch von der Rückreaktion bidirektionaler Reaktionsschritte beeinflusst werden. Dieser Effekt ist von allgemeinen Tracer-Experimenten wohlbekannt [And83, LR83], wurde jedoch bei metabolischen Markierungssystemen bislang selten berücksichtigt [SB79, CB83, KC83, CFGC95]. In dem hier vorgestellten Ansatz werden bidirektionale Schritte erstmals als Regelfall in das Modell aufgenommen und systematisch behandelt.

4. Isotopomeren-Systeme sind mit klassischen linearen Kompartimentmodellen nicht beschreibbar, weil zur Modellierung bimolekularer Reaktionsschritte Gleichungen zweiter Ordnung herangezogen werden müssen. Jedoch zeigen auch Isotopomeren-Bilanzgleichungen eine allgemeine Struktur, die mit Hilfe von mehrdimensionalen Matrizen dargestellt werden kann. Eine solche allgemeine Matrix-Repräsentation wird hier erstmals eingeführt. Zudem werden die entsprechenden Systemeigenschaften analysiert.
5. Zur systemtheoretischen Analyse von Isotopomeren-Systemen gibt es bis heute keine Ansätze, die über das explizite oder numerische Lösen der Isotopomeren-Bilanzen hinausgehen. Hierzu müssen die strukturellen Eigenschaften von Isotopomeren-Netzwerken mit mathematischen Hilfsmitteln herausgearbeitet werden, so daß systemanalytische Werkzeuge angesetzt werden können. In dieser Hinsicht werden in dieser Arbeit einige grundsätzliche Fortschritte erzielt (vgl. auch [Wie94b]).

Die für allgemeine Kompartimentsysteme formulierbaren Bilanzgleichungen haben im Falle metabolischer Markierungsexperimente also eine spezielle Struktur. Diese Gleichungen werden im vorliegenden Kapitel zunächst anhand eines Beispiels eingeführt. Sie beschreiben quantitativ die Beziehung zwischen den Zustandsgrößen der intrazellulären Markierungsanteile und Flüsse.

Im Hinblick auf systemanalytische Fragestellungen und die numerische Behandlung der Gleichungen ist es ratsam, eine abstrakte Formulierung von C-Atom- und Isotopomeren-Bilanzen mit Hilfe einer Matrixnotation einzuführen. Die in Abschnitt 4.5 definierten Matrizen repräsentieren die Struktur des metabolischen Netzwerks, während die Variablen dessen Fluß- und Markierungszustand quantifizieren. Derartige Beschreibungsansätze sind als *strukturelle Modelle* bekannt [Red88]. Ausgehend von der Matrixnotation können dann in den folgenden Kapiteln leistungsfähige mathematische Werkzeuge zur Simulation, Datenauswertung und Systemanalyse entwickelt werden.

Derartige universelle Beschreibungsansätze für metabolische Markierungssysteme sind in der Literatur nur in wenigen Fällen beschrieben [HS93, ZS94, MGWE⁺96]. Der nachfolgend entwickelte allgemeine Ansatz zeichnet sich durch die besondere Berücksichtigung von Eigenarten metabolischer Markierungsexperimente aus. Insbesondere wird in Abschnitt 4.5 erstmals eine abstrakte strukturelle Beschreibung von Isotopomeren-Markierungssystemen vorgestellt.

4.1 Zustandsvariablen

4.1.1 Flußvariablen

Als Zustandsgrößen zur Beschreibung von Markierungssystemen werden die Werte von Stoffflüssen und die prozentualen Markierungsanteile von C-Atomen oder Isotopomeren verwendet. Die Einführung in die Formulierung von Modellgleichungen erfolgt hier anhand des schon in Kapitel 3 verwendeten Beispielnetzwerks aus Abbildung 3.2. Zu seiner quantitativen Beschreibung müssen die jeweiligen Hin- und Rückrichtungen aller Reaktionsschritte explizit berücksichtigt werden. Grundsätzlich werden dabei alle extrazellulären Flüsse als unidirektional angesehen, während die übrigen Flüsse als bidirektional angenommen werden. Ergänzende Annahmen über die Direktionalität intrazellulärer Flüsse müssen dann explizit getroffen werden.

Für jeden Reaktionsschritt werden zwei Variablen eingeführt, die den jeweiligen Fluß in der Hin- bzw. Rückrichtung bezeichnen. Die Flußvariablen tragen dieselbe Bezeichnung wie die entsprechenden Reaktionsschritte in der formalen Notation aus Definition 3.1, sind jedoch in Absetzung davon in Kursiven gesetzt. Dem Reaktionsschritt PPW2 entspricht also die Flußvariable

*ppw*₂. Der Hin- bzw. Rückfluß jeder Reaktion wird durch einen vorwärts oder rückwärts gerichteten Pfeil am Flußsymbol angedeutet. Im laufenden Beispiel ergeben sich dann zwei Vektoren von Flußvariablen mit

$$\mathbf{v}^{\rightarrow} = (v_1^{\rightarrow}, v_2^{\rightarrow}, v_3^{\rightarrow}, v_4^{\rightarrow}, v_5^{\rightarrow})^T \quad \text{und} \quad \mathbf{v}^{\leftarrow} = (v_1^{\leftarrow}, v_2^{\leftarrow}, v_3^{\leftarrow}, v_4^{\leftarrow}, v_5^{\leftarrow})^T.$$

Definition 4.1 (Hin- und Rückflußvektor, Gesamtflußvektor, natürliche Flußkoordinaten)

Die den Hin- und Rückflüssen aller Reaktionsschritte zugeordneten Flußvariablen werden zum Hinflußvektor $\mathbf{v}^{\rightarrow}$ und zum Rückflußvektor $\mathbf{v}^{\leftarrow}$ zusammengefaßt, die gemeinsam den Flußzustand des Systems beschreiben. Dazu wird eine Reihenfolge der Flüsse fixiert, die für $\mathbf{v}^{\rightarrow}$ und $\mathbf{v}^{\leftarrow}$ jeweils gleich ist. Die beiden Vektoren $\mathbf{v}^{\rightarrow}$ und $\mathbf{v}^{\leftarrow}$ werden auch als die natürlichen Flußkoordinaten zur Beschreibung von Markierungssystemen bezeichnet¹. Wenn die Unterscheidung zwischen Hin- und Rückflüssen nicht von Bedeutung ist, so wird der Gesamtflußvektor

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}^{\rightarrow} \\ \mathbf{v}^{\leftarrow} \end{pmatrix}$$

zur simultanen Beschreibung aller Flüsse verwendet.

4.1.2 Nebenbedingungen und Annahmen über Flüsse

Aufgrund der Definition gilt stets $\dim \mathbf{v} = 2 \cdot \dim \mathbf{v}^{\rightarrow} = 2 \cdot \dim \mathbf{v}^{\leftarrow}$. Weiterhin ist wichtig, daß alle Flüsse nicht negativ sind, was durch die komponentenweise zu lesende Ungleichung

$$\mathbf{v} \geq 0 \tag{4.1}$$

ausgedrückt wird.²

Als physikalische Einheit für Flüsse ist im Rahmen der Stoffflußanalyse eine molare Einheit zweckmäßig, also mol/(l·h) (konzentrationsbezogen) oder auch mol/h (mengenbezogen). Der Grund für diese Einheitenwahl liegt darin, daß strenggenommen drei verschiedene Flußvariablen für Metabolit-Flüsse, C-Atom-Flüsse und Isotopomeren-Flüsse eingeführt werden müßten. Bei Verwendung der in der Bioverfahrenstechnik üblichen Konzentrationseinheit g/l hätten diese Variablen auch unterschiedliche Werte. Wählt man jedoch eine molare Einheit, so kann ein und dieselbe Variable alle drei Arten von Flüssen zugleich beschreiben, da der molare Wert aller drei Flußarten gleich ist.

Die grundsätzlich getroffene Unidirektionalitätsannahme für extrazelluläre Flüsse wird im Beispiel durch die Gleichungen

$$v_1^{\leftarrow} = 0 \quad v_3^{\leftarrow} = 0 \quad v_4^{\leftarrow} = 0 \quad v_5^{\leftarrow} = 0 \tag{4.2}$$

ausgedrückt.

4.1.3 C-Atom-Markierungsvariablen

Ein zweiter Satz von Variablen beschreibt die prozentualen Anteile markierten Materials in den einzelnen Metabolit-Pools. Dabei muß zwischen positionalen C-Atom-Markierungsanteilen und

¹Die Wortwahl wird durch die nachfolgenden Ausführungen zur Formulierung der Bilanzgleichungen für den Kohlenstoff-Fluß motiviert.

²Diese Ungleichung wird in Definition 6.5 noch erweitert.

Isotopomeren-Anteilen differenziert werden. Bei den Kohlenstoff-Atomen ist zunächst eine Variable für jede intrazelluläre Atom-Position einzuführen, die den Prozentsatz markierter Atome an dieser Stelle beschreibt. Die Variablennamen werden erneut von den in der formalen Notation aus Definition 3.1 verwendeten Namen der intrazellulären Metabolite abgeleitet. Im laufenden Beispiel ergibt sich aus B#1, B#2, E#1, E#2 der Gesamtmarkierungsvektor

$$\mathbf{x} = (b_1, b_2, e_1, e_2)^T.$$

Die durch das gewählte Substrat gegebenen Input-C-Atom-Markierungen von A#1 und A#2 werden in einem zweiten Vektor zusammengefaßt, dem Input-Markierungsvektor

$$\mathbf{x}^{\text{inp}} = (a_1, a_2)^T.$$

Definition 4.2 (C-Atom-Markierungsvektor, C-Atom-Input-Markierungsvektor) Die Markierungsanteile aller intrazellulären C-Atome werden in einer vorher fixierten Reihenfolge zum C-Atom-Markierungsvektor $\mathbf{x}$, die Markierungsanteile der Input-C-Atome zum C-Atom-Input-Markierungsvektor $\mathbf{x}^{\text{inp}}$ zusammengefaßt.

Natürlich muß für Markierungsvariablen stets gelten, daß $x_i \in [0, 1]$.

4.1.4 Isotopomeren-Markierungsvariablen

Wenn Isotopomere betrachtet werden, wird die Beschreibung wesentlich aufwendiger. In diesem Fall muß pro Metabolit-Pool je eine Variable für jeden Isotopomeren-Anteil innerhalb dieses Pools eingeführt werden. Für den Metaboliten B erhält man beispielsweise die Variablen $b_{00}, b_{01}, b_{10}, b_{11}$, die analog zur Notation aus Definition 3.1 binär indiziert sind. Die intrazellulären Isotopomeren-Anteile werden zu dem Isotopomeren-Markierungsvektor

$$\mathbf{x} = (b_{00}, b_{01}, b_{10}, b_{11}, e_{00}, e_{01}, e_{10}, e_{11})^T$$

zusammengefaßt. Der entsprechende Input-Vektor heißt:

$$\mathbf{x}^{\text{inp}} = (a_{00}, a_{01}, a_{10}, a_{11})^T$$

Definition 4.3 (Isotopomeren-Markierungsvektor, Isotopomeren-Input-Markierungsvektor) Die Anteile aller intrazellulären Isotopomere werden in einer vorher fixierten Reihenfolge zum Isotopomeren-Markierungsvektor $\mathbf{x}$, die Anteile der Input-Isotopomere zum Isotopomeren-Input-Markierungsvektor $\mathbf{x}^{\text{inp}}$ zusammengefaßt.

Für Kohlenstoff-Atom- und Isotopomeren-Markierungsvektoren werden nachfolgend also dieselben Bezeichner $\mathbf{x}, \mathbf{x}^{\text{inp}}$ verwendet, um unnötige Indizierungen zu vermeiden. Im Verlauf dieser Arbeit ist jeweils aus dem Kontext ersichtlich, ob über C-Atome oder Isotopomere gesprochen wird. Werden in seltenen Fällen die Isotopomeren-Vektoren zugleich mit den entsprechenden C-Atom-Markierungsvektoren benötigt, so sind die Variablen eindeutig gekennzeichnet als

$$\mathbf{x}^A, \quad \mathbf{x}^I \quad \text{bzw.} \quad \mathbf{x}^{A,\text{inp}}, \quad \mathbf{x}^{I,\text{inp}}.$$

4.1.5 Nebenbedingungen für Markierungsvariablen

Es ist zu beachten, daß die Isotopomeren-Anteile stets zusätzlich die Nebenbedingungen

$$\begin{aligned} a_{00} + a_{01} + a_{10} + a_{11} &= 1 \\ b_{00} + b_{01} + b_{10} + b_{11} &= 1 \\ e_{00} + e_{01} + e_{10} + e_{11} &= 1 \end{aligned} \quad (4.3)$$

erfüllen müssen, was für positionale Anreicherungen nicht gelten muß (vgl. Abb. 1.5). Außerdem gelten offensichtlich die folgenden Beziehungen zwischen Positions-Markierungsanteilen und Isotopomeren-Anteilen:

$$\begin{aligned} a_{10} + a_{11} &= a_1 & b_{10} + b_{11} &= b_1 & e_{10} + e_{11} &= e_1 \\ a_{01} + a_{11} &= a_2 & b_{01} + b_{11} &= b_2 & e_{01} + e_{11} &= e_2 \end{aligned} \quad (4.4)$$

4.1.6 Poolgrößenvariablen

Für die Modellierung eines isotopisch instationären Markierungsexperiments werden zusätzlich zu den bisher eingeführten Variablen auch die absoluten Poolgrößen der im System beteiligten Metabolite benötigt. Sie werden mit groß geschriebenen Kursiven bezeichnet, wie z.B. *B* für die Poolgröße des Metaboliten B. Die Poolgrößenvariablen werden zusammengefaßt zu

$$\mathbf{X} = \begin{pmatrix} B \\ E \end{pmatrix}.$$

Definition 4.4 (Poolgrößenvektor) Die Größen aller intrazellulären Pools werden in einer vorher fixierten Reihenfolge zum Poolgrößenvektor $\mathbf{X}$ zusammengefaßt.

Die Einheit von $\mathbf{X}_i$ ist naheliegenderweise mol oder mol/l. Man beachte ferner, daß in der Regel $\dim \mathbf{X} < \dim \mathbf{x}$ gilt.

4.2 Formulierung von Bilanzgleichungen

Zur Beschreibung und Auswertung von ^{13}C -Markierungsexperimenten werden nun die Bilanzgleichungen für Metabolit-, C-Atom- und Isotopomeren-Pools formuliert, welche die Beziehungen zwischen Flüssen und Markierungsanteilen quantitativ ausdrücken. Dabei wird zunächst mit dem isotopisch stationären Fall begonnen.

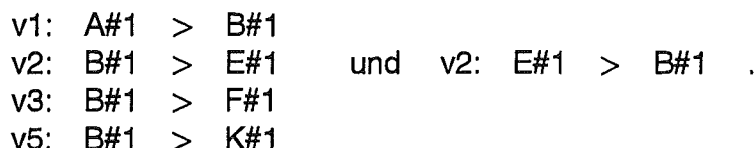
4.2.1 Metabolit-Bilanzen

Das erste Gleichungssystem ist durch die wohlbekannten stöchiometrischen Bilanzgleichungen gegeben. Für jeden Metabolit-Pool werden die hieran partizipierenden Flüsse bilanziert. Da sich die absoluten Poolgrößen im metabolisch stationären Zustand nicht verändern, müssen die Summen der in einen Metabolit-Pool hinein- bzw. aus ihm hinausfließenden Flüsse jeweils gleich sein. Im laufenden Beispiel erhält man aufgrund der getroffenen Annahmen $v_1^{\leftarrow} = 0$, $v_3^{\leftarrow} = 0$, $v_4^{\leftarrow} = 0$, $v_5^{\leftarrow} = 0$ also zwei stöchiometrische Gleichungen:

$$\begin{aligned} \text{B: } v_1^{\rightarrow} + v_2^{\leftarrow} &= v_2^{\rightarrow} + v_3^{\rightarrow} + v_5^{\rightarrow} \\ \text{E: } v_2^{\rightarrow} + v_3^{\rightarrow} &= v_2^{\leftarrow} + v_3^{\leftarrow} + v_4^{\rightarrow} \end{aligned} \quad (4.5)$$

4.2.2 C-Atom-Bilanzen

Zur Formulierung der Bilanzgleichungen für Kohlenstoff-Atome wird ein bestimmter C-Atom-Pool im System betrachtet und die Menge aller C-Atom-Transitionen, die in diesen Pool hinein- bzw. hinausführen. Für den Pool B#1 sind dies (aufgrund der getroffenen Direktionalitätsannahmen) beispielsweise die Hin- und Rückreaktionen

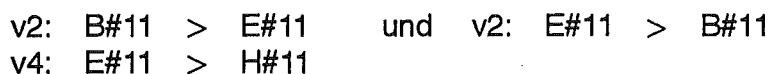


Die Menge der markierten Kohlenstoff-Atome, die in den Pool B#1 hineintransportiert wird, ergibt sich wie folgt. Die zuführenden Reaktionen, also die Hinreaktion von v1 und die Rückreaktion von v2, transportieren die absolute Menge $v_1^{\rightarrow}$ und $v_2^{\leftarrow}$ an markierten und unmarkierten Teilchen pro Zeiteinheit. Da der Anteil der zugeführten markierter Atome am Gesamtfluß durch a_1 bzw. e_1 gegeben ist, ergibt sich als Gesamtfluß von markierten Kohlenstoff-Atomen in den Pool B#1 folglich $v_1^{\rightarrow} \cdot a_1 + v_2^{\leftarrow} \cdot e_1$. Entsprechend ist die Menge des durch die Flüsse v2, v3 und v5 aus dem Pool heraustransportierten markierten Materials gleich $(v_2^{\leftarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_1$. Im stationären Zustand ändert sich die Markierungsverteilung nicht, so daß die zu- und abgeführten Teilchenmengen übereinstimmen. Daraus ergibt sich der folgende Satz von Markierungs-Bilanzgleichungen:

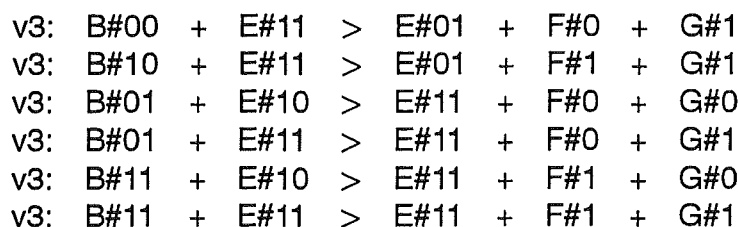
$$\begin{array}{ll} B\#1: & v_1^{\rightarrow} \cdot a_1 + v_2^{\leftarrow} \cdot e_1 = (v_2^{\leftarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_1 \\ B\#2: & v_1^{\rightarrow} \cdot a_2 + v_2^{\leftarrow} \cdot e_2 = (v_2^{\leftarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_2 \\ E\#1: & v_3^{\rightarrow} \cdot b_2 + v_2^{\leftarrow} \cdot b_1 = (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_1 \\ E\#2: & v_3^{\rightarrow} \cdot e_1 + v_2^{\leftarrow} \cdot b_2 = (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_2 \end{array} \quad (4.6)$$

4.2.3 Isotopomeren-Bilanzen

Die Formulierung von Isotopomeren-Bilanzen ist aufwendiger als die von C-Atom-Bilanzen, da bei Reaktionen mit mehreren Partnern auch Terme höherer Ordnung auftreten. Exemplarisch sei dies anhand des Isotopomeren-Pools E#11 des Beispielsystems dargestellt. Zu diesem Pool tragen die monomolekularen Hin- und Rückreaktionen



sowie die bimolekularen Umsetzungen



bei.³ Die monomolekularen Schritte werden analog zu den Kohlenstoff-Übergängen durch bzgl. Flüssen und Markierungen lineare Terme beschrieben. Der entsprechende Zuflußterm in der Bi-

³Interessanterweise wurden Isotopomeren-Bilanzgleichungen bisher stets für den Citratzyklus formuliert, in dem im wesentlichen nur ein bimolekularer Schritt vorkommt [CSKW83, KWL93, KCS93, FR95]. Da in diesen Studien zudem einer der beiden Reaktionspartner dieses Schritts (nämlich Acetyl-Coenzym A) als Input-Metabolit vorgegeben war, reduziert sich das entsprechende System sogar auf ein lineares Modell bzgl. x.

lanzgleichung für E#11 ist $v_2^{\rightarrow} \cdot b_{11}$ und der Abflußterm $(v_2^{\leftarrow} + v_4^{\rightarrow}) \cdot e_{11}$. Die bimolekulare Reaktion trägt insgesamt vier weitere Zuflußterme zur Bilanzgleichung bei, entsprechend den vier möglichen Isotopomeren-Kombinationen, die E#11 bilden. Die Wahrscheinlichkeit, daß die Reaktion v3 die beiden Isotopomere B#01 und E#11 zusammenführt, ist durch das Produkt ihrer Anteile, also $b_{01} \cdot e_{11}$, gegeben [KCS93]. Der entsprechende Gesamtzufluß ist also

$$v_3^{\rightarrow} \cdot b_{01} \cdot e_{10} + v_3^{\rightarrow} \cdot b_{01} \cdot e_{11} + v_3^{\rightarrow} \cdot b_{11} \cdot e_{10} + v_3^{\rightarrow} \cdot b_{11} \cdot e_{11} = v_3^{\rightarrow} \cdot (b_{01} + b_{11}) \cdot (e_{10} + e_{11}) \quad .$$

Der Gesamtabfluß aus dem Pool ist nach derselben Gesetzmäßigkeit und Anwendung von Gleichung (4.3) gegeben durch

$$v_3^{\rightarrow} \cdot b_{00} \cdot e_{11} + v_3^{\rightarrow} \cdot b_{01} \cdot e_{11} + v_3^{\rightarrow} \cdot b_{10} \cdot e_{11} + v_3^{\rightarrow} \cdot b_{11} \cdot e_{11} = v_3^{\rightarrow} \cdot (b_{00} + b_{01} + b_{10} + b_{11}) \cdot e_{11} = v_3^{\rightarrow} \cdot e_{11} \quad .$$

Ganz allgemein kombiniert eine bimolekulare Reaktion ein Isotopomer mit jeder Form seines Partners. Die Summe der Markierungsanteile der Partner ist also Eins, so daß für die aus dem Pool herausführenden Reaktionsschritte die Molekularität nicht von Bedeutung ist. Die entsprechenden Terme sind dann von erster Ordnung. Zusammenfassend ergeben sich die Bilanzgleichungen:

$$\begin{aligned} \text{B\#00} : v_1^{\rightarrow} \cdot a_{00} + v_2^{\leftarrow} \cdot e_{00} &= (v_2^{\rightarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_{00} \\ \text{B\#01} : v_1^{\rightarrow} \cdot a_{01} + v_2^{\leftarrow} \cdot e_{01} &= (v_2^{\rightarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_{01} \\ \text{B\#10} : v_1^{\rightarrow} \cdot a_{10} + v_2^{\leftarrow} \cdot e_{10} &= (v_2^{\rightarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_{10} \\ \text{B\#11} : v_1^{\rightarrow} \cdot a_{11} + v_2^{\leftarrow} \cdot e_{11} &= (v_2^{\rightarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_{11} \\ \text{E\#00} : v_2^{\rightarrow} \cdot b_{00} + v_3^{\rightarrow} \cdot (b_{00} + b_{10}) \cdot (e_{00} + e_{01}) &= (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_{00} \\ \text{E\#01} : v_2^{\rightarrow} \cdot b_{01} + v_3^{\rightarrow} \cdot (b_{00} + b_{10}) \cdot (e_{10} + e_{11}) &= (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_{01} \\ \text{E\#10} : v_2^{\rightarrow} \cdot b_{10} + v_3^{\rightarrow} \cdot (b_{01} + b_{11}) \cdot (e_{00} + e_{01}) &= (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_{10} \\ \text{E\#11} : v_2^{\rightarrow} \cdot b_{11} + v_3^{\rightarrow} \cdot (b_{01} + b_{11}) \cdot (e_{10} + e_{11}) &= (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_{11} \end{aligned} \quad (4.7)$$

4.2.4 Isotopisch instationärer Fall

In der isotopisch instationären Situation sind die oben hergeleiteten Bilanzgleichungen nicht mehr korrekt, da sich die Poolgrößen noch verändern. In diesem Fall müssen die absoluten Poolgrößen einbezogen werden, weil sie die Kapazität des jeweiligen Metabolit-Pools für markiertes Material festlegen. Die absolute Menge an markierten Kohlenstoff-Atomen an der ersten Position von B ist gegeben durch $B \cdot b_1$. Da sich B zeitlich nicht verändert (metabolisch stationärer Zustand), folgt für die zeitliche Veränderung der Markierungsmenge:

$$d/dt (B \cdot b_1) = B \cdot d/dt b_1$$

Durch Ergänzung der stationären Markierungs-Bilanzgleichungen (4.6) um den differentiellen Term ergibt sich:

$$\begin{aligned} \text{B\#1} : B \cdot d/dt b_1 &= v_1^{\rightarrow} \cdot a_1 + v_2^{\leftarrow} \cdot e_1 - (v_2^{\rightarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_1 \\ \text{B\#2} : B \cdot d/dt b_2 &= v_1^{\rightarrow} \cdot a_2 + v_2^{\leftarrow} \cdot e_2 - (v_2^{\rightarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b_2 \\ \text{E\#1} : E \cdot d/dt e_1 &= v_3^{\rightarrow} \cdot b_2 + v_2^{\rightarrow} \cdot b_1 - (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_1 \\ \text{E\#2} : E \cdot d/dt e_2 &= v_3^{\rightarrow} \cdot e_1 + v_2^{\rightarrow} \cdot b_2 - (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_2 \end{aligned} \quad (4.8)$$

Die instationären Gleichungen für die Isotopomeren-Anteile können analog gebildet werden.

4.3 Explizite Behandlung des Beispielsystems

An dieser Stelle können anhand des laufenden Beispiels bereits einige allgemeine Prinzipien und zentrale Probleme der Lösung von stationären Bilanzgleichungen sowie der Flußschätzung aus Meßdaten aufgezeigt werden, die für den weiteren Verlauf dieser Arbeit von zentraler Bedeutung sind. Ganz allgemein sind diese Fragestellungen als Redundanz- und Identifizierbarkeitsprobleme bekannt [Wal87].

Die im folgenden vorgestellten Ergebnisse wurden mit Hilfe des Computeralgebra-Systems Maple [Kof94] erzielt. Der Einsatz solcher Systeme [Hea91, Kof92] hat in den letzten Jahren in allen wissenschaftlichen Disziplinen an Bedeutung gewonnen [LR87, MMN89, PT95]. Insbesondere bei der Behandlung polynomialer Gleichungssysteme — wie z.B. Markierungs-Bilanzgleichungen — sind in den letzten Jahren einige Fortschritte erzielt worden [CLO92, BW93]. Die ausführliche Behandlung algebraischer Methoden würde jedoch den Rahmen dieser Arbeit sprengen. Der Leser sei hier verwiesen auf [Wie95a, Wie96a].

Nachfolgend wird der Einfachheit halber angenommen, daß

$$a_{00} = 0 \quad a_{01} = 1 \quad a_{10} = 0 \quad a_{11} = 0 \quad (4.9)$$

gilt, insbesondere also nach Gleichung (4.4)

$$a_1 = 0 \quad a_2 = 1 \quad . \quad (4.10)$$

4.3.1 Explizite Lösung der Bilanzgleichungen

Als erstes sollen die Markierungsanteile als Funktion der Flüsse ausgedrückt werden. Zu diesem Zweck können die Metabolit-Bilanzen (4.5) verwendet werden, um die Flußvariablen $v_3^{\rightarrow}$ und $v_4^{\rightarrow}$ zu eliminieren:

$$\begin{aligned} v_3^{\rightarrow} &= v_1^{\rightarrow} - v_2^{\rightarrow} + v_2^{\leftarrow} - v_5^{\rightarrow} \\ v_4^{\rightarrow} &= v_2^{\rightarrow} - v_2^{\leftarrow} \end{aligned} \quad (4.11)$$

Die linear unabhängigen Flüsse $v_1^{\rightarrow}, v_2^{\rightarrow}, v_2^{\leftarrow}, v_5^{\rightarrow}$ auf der rechten Seite der Gleichungen werden als *freien Flüsse* bezeichnet (vgl. Definition 6.4). Einsetzen der rechten Seiten in die stationären C-Atom-Bilanzen (4.6) bzw. die Isotopomeren-Bilanzen (4.7) führt auf ein lineares bzw. quadratisches Gleichungssystem zur Berechnung aller Markierungsanteile in Abhängigkeit von den freien Flüssen. Während dessen Lösung im linearen Fall direkt durch eine Gauß-Elimination berechnet werden kann (vgl. Satz 7.3), gestaltet sich die explizite Lösung der Isotopomeren-Bilanzen wesentlich aufwendiger (vgl. dazu die Diskussion sogenannter Gröbner-Basen in [CLO92, Wie95a]).

Hier sei nur die Lösung der stationären C-Atom-Bilanzen

$$\begin{aligned} b_1 &= v_2^{\leftarrow} \cdot \alpha_1 \alpha_2 / \alpha_0 & e_1 &= \alpha_1 \alpha_2 \alpha_4 / \alpha_0 \\ b_2 &= \alpha_3 / \alpha_0 & e_2 &= \alpha_5 / \alpha_0 \end{aligned} \quad (4.12)$$

und der stationären Isotopomeren-Bilanzen

$$\begin{aligned} b_{00} &= 1 - b_{10} - b_{01} - b_{11} & e_{00} &= 1 - e_{10} - e_{01} - e_{11} \\ b_{10} &= v_2^{\rightarrow} v_2^{\leftarrow} \cdot \alpha_1 \alpha_2 \alpha_6 / \alpha_0^2 & e_{10} &= v_2^{\rightarrow} \cdot \alpha_1 \alpha_2 \alpha_4 \alpha_6 / \alpha_0^2 \\ b_{01} &= \alpha_2 \alpha_8 / \alpha_0^2 & e_{01} &= v_2^{\rightarrow} \cdot \alpha_7 / \alpha_0^2 \\ b_{11} &= v_2^{\leftarrow} \cdot \alpha_1^2 \alpha_2^2 \alpha_4 / \alpha_0^2 & e_{11} &= \alpha_1^2 \alpha_2^2 \alpha_4^2 / \alpha_0^2 \end{aligned} \quad (4.13)$$

wiedergegeben, wobei die Hilfskonstanten α_i gegeben sind durch:

$$\begin{aligned}
\alpha_0 &= v_1^{\rightarrow 3} + v_1^{\rightarrow 2} \cdot (3v_2^{\leftarrow} - 2v_5^{\rightarrow}) + v_1^{\rightarrow} \cdot (3v_2^{\leftarrow 2} - 4v_2^{\leftarrow} v_5^{\rightarrow} + v_5^{\rightarrow 2}) \\
&\quad + v_2^{\leftarrow} \cdot (v_2^{\leftarrow 2} - 2v_2^{\leftarrow} v_5^{\rightarrow} - v_2^{\rightarrow 2} + v_5^{\rightarrow 2}) \\
\alpha_1 &= v_1^{\rightarrow} - v_2^{\rightarrow} + v_2^{\leftarrow} - v_5^{\rightarrow} \\
\alpha_2 &= v_1^{\rightarrow} + v_2^{\leftarrow} - v_5^{\rightarrow} \\
\alpha_3 &= \alpha_0 - v_2^{\rightarrow} v_2^{\leftarrow} \alpha_1 \\
\alpha_4 &= v_1^{\rightarrow} + v_2^{\leftarrow} \\
\alpha_5 &= \alpha_0 - v_2^{\rightarrow} \alpha_1 \alpha_4 \\
\alpha_6 &= v_1^{\rightarrow 2} + v_1^{\rightarrow} \cdot (2v_2^{\leftarrow} - v_5^{\rightarrow}) + v_2^{\leftarrow} \cdot (v_2^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \\
\alpha_7 &= \alpha_0 \alpha_6 + v_2^{\rightarrow} v_2^{\leftarrow} \cdot (2\alpha_6 v_2^{\rightarrow} - \alpha_0 - v_1^{\rightarrow} v_2^{\rightarrow 2}) \\
\alpha_8 &= \alpha_7 - v_2^{\leftarrow} \alpha_1^3 \alpha_4
\end{aligned}$$

Ganz allgemein kann für das C-Atom-System und unter gewissen Bedingungen auch für das Isotopomeren-System gezeigt werden, daß der Markierungszustand eine Funktion des Flußvektors ist (siehe dazu Abschnitt 7.1 und 7.3):

$$\mathbf{x} = \mathbf{x}(\mathbf{v}) \quad (4.14)$$

Allerdings kann diese Beziehung nur für C-Atom-Bilanzen explizit unter Verwendung einer inversen Matrix angegeben werden (Satz 7.3). Für Isotopomeren-Systeme ist dies im allgemeinen nicht mehr möglich. Entsprechende algebraische Lösungsstrategien werden in [Wie95a] vorgestellt. Die oben angegebene explizite Beziehung für das Beispielsystem verdeutlicht, welchen Grad an Komplexität die Lösungen schon bei kleinen Systemen annehmen können.

4.3.2 Redundanzbeziehungen

Für die Beurteilung des Informationsgehaltes von Meßwerten spielen *Redundanzbeziehungen* eine zentrale Rolle. Kann ein Meßwert ohne Kenntnis des Systemzustands bereits aus anderen Meßwerten rekonstruiert werden, so enthält er keine zusätzliche Information über die intrazellulären Flüsse. Algorithmische Methoden zur Aufdeckung von Redundanzbeziehungen werden in [Wie96a] vorgestellt. Hier sei nur das Ergebnis der Anwendung auf das laufende Beispiel präsentiert. Die Redundanzbeziehungen sind hier:

$$\begin{aligned}
0 &= b_1 e_1 - b_2^2 - b_1 + b_2 & 0 &= b_2 - b_2^2 - b_1 + b_{11} \\
0 &= b_1 e_2 - b_2 e_1 - b_1 + e_1 & 0 &= b_1 e_{11} - e_1 b_{11}
\end{aligned} \quad (4.15)$$

Dieses Resultat überrascht zunächst, denn es folgt:

$$\begin{aligned}
e_1 &= (b_2^2 + b_1 - b_2)/b_1 \\
e_2 &= (b_2 e_1 + b_1 - e_1)/b_1
\end{aligned} \quad (4.16)$$

und sodann

$$\begin{aligned}
b_{11} &= b_2^2 + b_1 - b_2 & e_{11} &= b_{11} e_1 / b_1 \\
b_{10} &= b_{11} - b_1 & e_{10} &= e_{11} - e_1 \\
b_{01} &= b_{11} - b_2 & e_{01} &= e_{11} - e_2 \\
b_{00} &= 1 - b_{01} - b_{10} - b_{11} & e_{00} &= 1 - e_{01} - e_{10} - e_{11}
\end{aligned} \quad (4.17)$$

Es zeigt sich also, daß das gesamte System bereits durch die Kenntnis von b_1 und b_2 bestimmt ist. Alle anderen C-Atom- und Isotopomeren-Anteile können daraus berechnet werden. Folglich enthalten Isotopomeren-Meßwerte im Beispiel keine zusätzliche Information über die intrazellulären Flüsse! Es zeigt sich also, daß Isotopomeren-Anteile nicht grundsätzlich mehr Information über die gesuchten Flüsse enthalten als Positions-Markierungsanteile. Der Nachweis, ob dies der Fall ist, kann sich im Einzelfall äußerst schwierig gestalten (vgl. [Wie95a]).

4.3.3 Flußschätzung

Das Problem der Flußschätzung besteht in der Berechnung der intrazellulären Flüsse aus vorgegebenen Meßdaten. Es soll nun für das Beispiel explizit gelöst werden. Zunächst läßt sich aus den ermittelten Redundanzbeziehungen ablesen, welche Größen gemessen werden müssen, um die verbleibenden unbekannten Flüsse zu berechnen. Wie gezeigt, kann man sich dabei auf die C-Atom-Markierungen b_1, b_2 beschränken. Nimmt man die extrazellulären Flüsse $v_1^{\rightarrow}$ und $v_5^{\rightarrow}$ als meßbar an, so ergeben sich die verbleibenden intrazellulären Flüsse aus

$$\begin{aligned} v_2^{\leftarrow} &= v_1^{\rightarrow} \cdot \frac{b_2^2}{(b_2 - b_1) \cdot (b_2 + b_1 - 1)} \\ v_2^{\rightarrow} &= (v_1^{\rightarrow} + v_2^{\leftarrow} - v_5^{\rightarrow}) \cdot \frac{1 - b_2}{b_1} \end{aligned} \quad (4.18)$$

Auch die übrigen Flüsse $v_3^{\rightarrow}$ und $v_4^{\rightarrow}$ lassen sich dann nach Gleichung (4.11) berechnen. Gleichung (4.18) macht darüber hinaus deutlich, daß nicht auf die Messung von v_5 verzichtet werden kann. Ganz allgemein wird die Frage, ob sich intrazelluläre Flüsse auf der Basis von Markierungsdaten und extrazellulären Flußmessungen bestimmen lassen, als *Identifizierbarkeitsproblem* bezeichnet (vgl. [Wie95a]).

Ausdrücklich sei an dieser Stelle auf das wichtige Resultat hingewiesen, daß — wie das Beispiel mit $v_2^{\rightarrow}$ und $v_2^{\leftarrow}$ zeigt — unter bestimmten Bedingungen aus Markierungsdaten beide Flußraten von bidirektionalen Reaktionsschritten berechnet werden können.

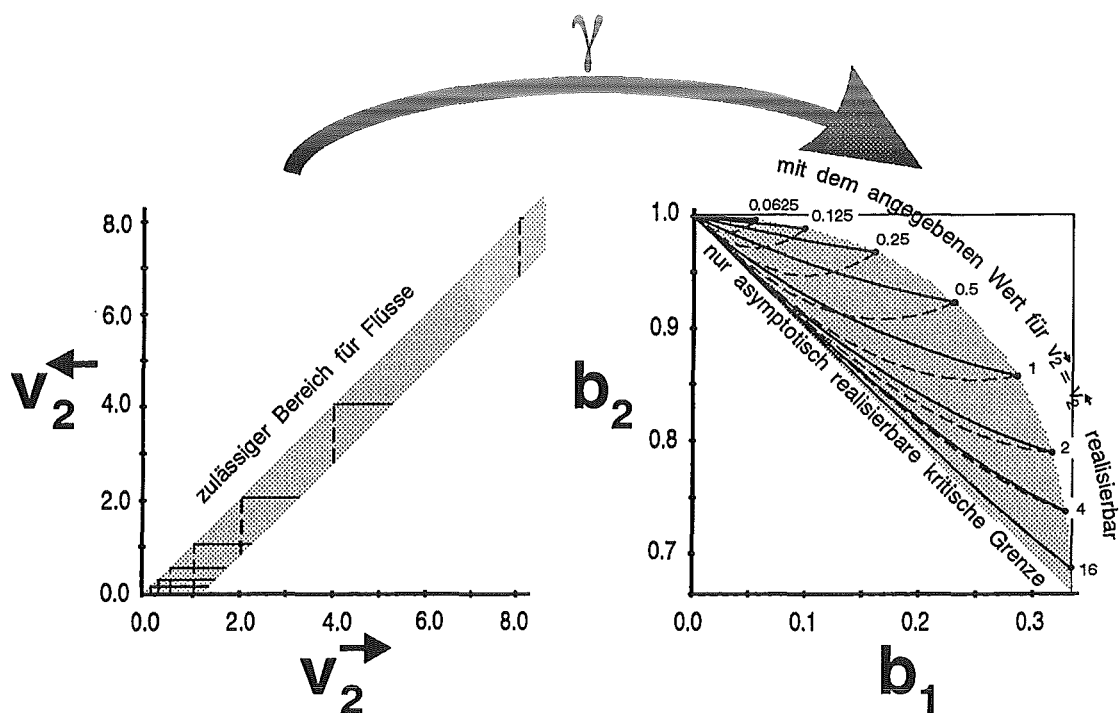


Abbildung 4.1: Realisierbare Markierungszustände für das Beispielsystem aus Abbildung 3.2 mit den Werten $v_1^{\rightarrow} = 1$ und $v_5^{\rightarrow} = 0$. Die eingezeichneten Niveaulinien zeigen, wo die freien Flüsse $v_2^{\rightarrow}$ und $v_2^{\leftarrow}$ den jeweils angegebenen Wert annehmen.

4.3.4 Realisierbare Markierungszustände

Wird davon ausgegangen, daß Meßwerte für die Flüsse $v_1^{\rightarrow}$ und $v_5^{\rightarrow}$ vorliegen, so beschreiben die Gleichungen (4.12) und (4.18) eine bijektive Beziehung

$$\gamma: (v_2^{\rightarrow}, v_2^{\leftarrow}) \xleftrightarrow{1-1} (b_1, b_2) \quad (4.19)$$

Der Bildbereich dieser Transformation wird als Menge der (physikalisch) *realisierbaren* (b_1, b_2) -Markierungszustände bezeichnet. Diese Menge soll nun exemplarisch für $v_1^{\rightarrow} = 1$ und $v_5^{\rightarrow} = 0$ berechnet werden. Dabei ist zu Bedenken, daß die möglichen Werte für $v_2^{\rightarrow}$ und $v_2^{\leftarrow}$ eingeschränkt sind durch die aus den stöchiometrischen Gleichungen (4.5) und der Bedingung $v \geq 0$ resultierenden Nebenbedingungen (Abb. 4.1)

$$\begin{aligned} 0 &\leq v_2^{\rightarrow} \\ 0 &\leq v_2^{\leftarrow} \\ 0 &\leq v_3^{\rightarrow} = v_1^{\rightarrow} - v_2^{\rightarrow} + v_2^{\leftarrow} - v_5^{\rightarrow} \\ 0 &\leq v_4^{\rightarrow} = v_2^{\rightarrow} - v_2^{\leftarrow} \end{aligned}$$

Hieraus folgt

$$0 \leq v_2^{\leftarrow} \leq v_2^{\rightarrow} \leq 1 + v_2^{\leftarrow} \quad \text{bzw.} \quad -1 \leq v_2^{\rightarrow} - 1 \leq v_2^{\leftarrow} \leq v_2^{\rightarrow} \quad (4.20)$$

Nun ist γ bijektiv und stetig, d.h. es genügt, das Bild der Randlinien des durch Gleichung (4.20) gegebenen Bereichs zu bestimmen. Diese sind gegeben durch (Abb. 4.1):

$$\begin{aligned} v_2^{\leftarrow} &= 0 && \wedge && 0 \leq v_2^{\rightarrow} \leq 1 \\ v_2^{\rightarrow} &= v_2^{\leftarrow} && \wedge && 0 \leq v_2^{\rightarrow} \\ v_2^{\rightarrow} &= 1 + v_2^{\leftarrow} && \wedge && 1 \leq v_2^{\rightarrow} \end{aligned} \quad (4.21)$$

Ferner ist das asymptotische Verhalten von b_1, b_2 für

$$v_2^{\rightarrow}, v_2^{\leftarrow} \rightarrow \infty$$

zu untersuchen. Dies kann mit Hilfe der Gleichungen (4.12) durchgeführt werden, so daß Abbildung 4.1 entsteht. Interessanterweise schrumpfen dabei zwei der in Gleichung (4.21) beschriebenen Grenzlinien auf je einen Punkt zusammen.

Aus Abbildung 4.1 ist ersichtlich, daß die Menge der realisierbaren Zustände nicht den ganzen Raum $[0, 1] \times [0, 1]$ ausfüllt. Dies hat weitreichende Konsequenzen für die Auswertung von Experimenten, da die gemessenen Markierungsanteile aufgrund von Meßfehlern durchaus auch außerhalb des realisierbaren Gebietes liegen können. In diesem Falle sind die Gleichungen (4.18) nicht mit positiven Flüssen lösbar. Die Analyse dieser Ausnahmesituation ist ein zentraler Gegenstand der statistischen Auswertung in Kapitel 11.

Während die halbkreisförmige obere Grenzlinie in Abbildung 4.1 im Falle $v_2^{\rightarrow} = v_2^{\leftarrow}$ genau erreicht wird (hier wird von einer *realisierbaren Grenzlinie* gesprochen), repräsentiert die untere Grenze das Ergebnis des Grenzübergangs $v_2^{\rightarrow}, v_2^{\leftarrow} \rightarrow \infty$. Sie selbst gehört also nicht mit zur Menge der physikalisch realisierbaren Markierungszustände und wird nachfolgend als *asymptotische Grenzlinie* bezeichnet.

Besonders problematisch ist die Flußschätzung dann, wenn der Meßwert in der Nähe der asymptotischen Grenzlinie liegt. Dann führt eine kleine Veränderung des gemessenen Markierungswertes zu beliebig großen Veränderungen der zugehörigen Flußschätzung aus Gleichung (4.18). Im Extremfall kann die Messung aufgrund des Meßfehlers auch unterhalb der gezeigten asymptotischen Grenze liegen, so daß die Flußschätzung im Unendlichen liegt. Aus diesem Grunde ist es für die praktische Anwendung wichtig, das Problem "unendlich großer Flüsse" sowohl in numerischer als auch in statistischer Hinsicht zu behandeln. Diese Thematik durchzieht die folgenden Abschnitte.

4.4 Mehrdimensionale Matrizen

In den beiden nun folgenden Abschnitten wird eine abstrakte Notation für die bisher am Beispiel eingeführten C-Atom- und Isotopomeren-Bilanzgleichungssysteme entwickelt. Das Ziel ist eine kompakte Notation der Gleichungssysteme unter Zuhilfenahme eines Matrizenkalküls. Die dabei zur Beschreibung von Markierungssystemen eingeführten mehrdimensionalen Matrizen repräsentieren die Struktur des zugrundeliegenden Netzwerks, während die bereits vorgestellten Vektoren $\mathbf{v}$ und $\mathbf{x}$ seinen Zustand quantifizieren.

4.4.1 Allgemeine Gestalt der Bilanzgleichungen

Offenbar haben alle Bilanzgleichungen die Gestalt von Multilinearformen bzgl. der Variablen $\mathbf{x}_i$, $\mathbf{x}_j^{\text{inp}}$ und $\mathbf{v}_k$. Konkret gilt

1. für den i -ten Metabolit-Pool eine Gleichung der allgemeinen Gestalt

$$0 = \sum_{j=1}^{\dim \mathbf{v}} \mathbf{N}_{i,j} \cdot \mathbf{v}_j \quad (4.22)$$

mit gewissen stöchiometrischen Koeffizienten $\mathbf{N}_{i,j}$.

2. für den i -ten C-Atom-Pool eine Gleichung der allgemeinen Gestalt

$$0 = \sum_{j=1}^{\dim \mathbf{v}} \sum_{k=1}^{\dim \mathbf{x}} \mathbf{P}_{j,i,k} \cdot \mathbf{x}_k \cdot \mathbf{v}_j + \sum_{j=1}^{\dim \mathbf{v}} \sum_{k=1}^{\dim \mathbf{x}^{\text{inp}}} \mathbf{P}_{j,i,k}^{\text{inp}} \cdot \mathbf{x}_k^{\text{inp}} \cdot \mathbf{v}_j \quad (4.23)$$

mit gewissen Koeffizienten $\mathbf{P}_{j,i,k}$, $\mathbf{P}_{j,i,k}^{\text{inp}}$.

3. für den i -ten Isotopomeren-Pool eine Gleichung der allgemeinen Gestalt⁴

$$\begin{aligned} 0 = & \sum_{j=1}^{\dim \mathbf{v}} \sum_{k=1}^{\dim \mathbf{x}} \sum_{l=1}^{\dim \mathbf{x}} \mathbf{Q}_{j,i,k,l} \cdot \mathbf{x}_k \cdot \mathbf{x}_l \cdot \mathbf{v}_j \\ & + \sum_{j=1}^{\dim \mathbf{v}} \sum_{k=1}^{\dim \mathbf{x}} \mathbf{P}_{j,i,k} \cdot \mathbf{x}_k \cdot \mathbf{v}_j + \sum_{j=1}^{\dim \mathbf{v}} \sum_{k=1}^{\dim \mathbf{x}^{\text{inp}}} \mathbf{P}_{j,i,k}^{\text{inp}} \cdot \mathbf{x}_k^{\text{inp}} \cdot \mathbf{v}_j \end{aligned} \quad (4.24)$$

mit gewissen Koeffizienten $\mathbf{P}_{j,i,k}$, $\mathbf{P}_{j,i,k}^{\text{inp}}$, $\mathbf{Q}_{j,i,k,l}$.

Ähnlich wie bei den Poolvariablen werden für die Koeffizienten von C-Atom- und Isotopomeren-Bilanzen dieselben Bezeichner $\mathbf{P}_{j,i,k}$ und $\mathbf{P}_{j,i,k}^{\text{inp}}$ verwendet, obwohl diese Koeffizienten in beiden Fällen ganz verschiedene Werte haben. Mit der genauen Definition der Koeffizienten beschäftigt sich Abschnitt 4.5.

Die oben verwendete allgemeine Summennotation legt die Einführung einer Matrixnotation nahe. Um völlig auf die umständliche Indexnotation verzichten zu können, werden nun drei- und vierdimensionale Matrizen zusammen mit entsprechenden Rechenregeln eingeführt.⁵

⁴Nach den einschränkenden Annahmen aus Definition 3.5 darf eine bimolekulare Reaktion mit einem extrazellulären Partner nicht auftreten. Daher werden keine Koeffizienten $\mathbf{Q}_{j,i,k,l}^{\text{inp}}$ benötigt.

⁵Der häufig verwendete Begriff des Tensors für mehrdimensionale Matrixstrukturen wird in dieser Arbeit vermieden, da Tensoren meist im Kontext geometrischer Transformationsinvarianzen verwendet werden. Daher sind die Indexdimensionen eines Tensors gleich, was bei 3D- und 4D-Matrizen nicht unbedingt der Fall ist [BG68].

4.4.2 Definition mehrdimensionaler Matrizen

Definition 4.5 (3D-Matrix, 4D-Matrix) 3D-Matrizen bzw. 4D-Matrizen sind Vektoren bzw. Matrizen, deren Einträge selbst wieder Matrizen sind:

$$\mathbf{M}^{3D} = (M_1, \dots, M_n) \quad \mathbf{M}^{4D} = \begin{pmatrix} M_{11} & \dots & M_{1m} \\ \vdots & \ddots & \vdots \\ M_n & \dots & M_{nm} \end{pmatrix}$$

Dabei müssen alle Einträge M_{ij} die gleichen Abmessungen haben, d.h.:

$$\forall_{i,j} \text{rowdim } M_{ij} = \text{rowdim } M_{11} \quad \wedge \quad \forall_{i,j} \text{coldim } M_{ij} = \text{coldim } M_{11}$$

Die Mengen der 3D- bzw. 4D-Matrizen werden mit $\mathbb{R}^{m \times (k \times l)}$ bzw. $\mathbb{R}^{(m \times n) \times (k \times l)}$ bezeichnet, wobei die Indizierung in der Form

$$(\mathbf{M}_i^{3D})_{k,l} \quad \text{bzw.} \quad (\mathbf{M}_{i,j}^{4D})_{k,l}$$

mit der Reihenfolge Zeile-Spalte-Subzeile-Subspalte erfolgt.

Genauso wie sich Vektoren als einspaltige Matrizen auffassen lassen, werden auch 3D-Matrizen als spezielle $(m \times 1) \times (k \times l)$ -4D-Matrizen (also einspaltige 4d-Matrizen) betrachtet. Es wird somit identifiziert:

$$\mathbb{R}^{m \times (k \times l)} \cong \mathbb{R}^{(m \times 1) \times (k \times l)}$$

Ausgedrückt in der Terminologie der Programmiersprachen werden also bei Bedarf implizite Typkonvertierungen vorgenommen [EGL89]. Aufgrund dieser Konvention können alle weiteren Definitionen, die in diesem Abschnitt für 4D-Matrizen vorgenommen werden, unmittelbar auch auf 3D-Matrizen angewendet werden.

Weiterhin treten als Ergebnis der nachfolgend eingeführten Matrix-Vektor-Produkte häufig "entartete" $(1 \times 1) \times (k \times l)$ -4D-Matrizen

$$\mathbf{M}^{4D} = \left(\begin{pmatrix} M_{11,11} & \dots & M_{11,1l} \\ \vdots & \ddots & \vdots \\ M_{11,k1} & \dots & M_{11,kl} \end{pmatrix} \right)$$

mit nur einer einzigen inneren Matrix oder $(m \times n) \times (1 \times 1)$ -4D-Matrizen

$$\mathbf{M}^{4D} = \begin{pmatrix} (M_{11,11}) & \dots & (M_{1n,11}) \\ \vdots & \ddots & \vdots \\ (M_{m1,11}) & \dots & (M_{mn,11}) \end{pmatrix}$$

bestehend aus (1×1) -Matrizen auf. Für diese entarteten Fälle werden zusätzlich die folgenden Identifikationen vorgenommen, bei denen die inneren Matrizenklammern jeweils weggelassen werden:

$$\begin{aligned} \mathbb{R}^{(1 \times 1) \times (k \times l)} &\cong \mathbb{R}^{k \times l} \\ \mathbb{R}^{(m \times n) \times (1 \times 1)} &\cong \mathbb{R}^{m \times n} \end{aligned} \quad (4.25)$$

Insbesondere kann eine entartete 4D-Matrix in der gewohnten Weise unmittelbar mit einem Vektor multipliziert werden.

4.4.3 Matrix-Vektor-Produkte

Angestrebt wird die Festlegung von Matrix-Vektor-Produkten, um die Indexsummen aus den Gleichungen (4.22), (4.23) und (4.24) kompakt ausdrücken zu können.

Definition 4.6 (inneres Produkt, äußeres Produkt) Das äußere Produkt einer 4D-Matrix M mit einem Vektor w von links bzw. v von rechts ist gegeben durch:

$$w^T \odot M \odot v = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}^T \cdot \begin{pmatrix} M_{11} & \dots & M_{1m} \\ \vdots & \ddots & \vdots \\ M_n & \dots & M_{nm} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \stackrel{def}{=} \left(\sum_{i=1}^n \sum_{j=1}^m w_i M_{ij} v_j \right)$$

Das innere Produkt einer 4D-Matrix M mit einem Vektor x von links bzw. y von rechts ist gegeben durch:

$$x^T \cdot M \cdot y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} M_{11} & \dots & M_{1m} \\ \vdots & \ddots & \vdots \\ M_n & \dots & M_{nm} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \stackrel{def}{=} \begin{pmatrix} x^T M_{11} y & \dots & x^T M_{1m} y \\ \vdots & \ddots & \vdots \\ x^T M_{nm} y & \dots & x^T M_{nm} y \end{pmatrix}$$

Beim äußeren Produkt werden Vektoren von außen so "heranmultipliziert", als ob die Einträge von M einfache Zahlen wären. Das Ergebnis ist eine entartete (1×1) -4D-Matrix, und die einzige innere Matrix hat die Abmessungen der inneren Matrizen M_{ij} .

Beim inneren Produkt werden dagegen die Vektoren an die einzelnen Einträge von M "heranmultipliziert", das Ergebnis ist eine entartete 4D-Matrix aus (1×1) -Matrizen mit den Abmessungen der äußeren Matrix M .

4.4.4 Rechenregeln für Matrix-Vektor-Produkte

Es kann leicht gezeigt werden, daß sich alle Produkte linear verhalten bezüglich Addition und Skalarmultiplikation der in ihnen auftretenden Vektoren und Matrizen. Ferner sind die eingeführten Vektor-Matrix-Produkte assoziativ und darüber hinaus auch kommutativ auf jeder Seite der Matrix.

Ein Beispiel kann die Anwendung der Rechenregeln und der Typkonvertierungsregel (4.25) illustrieren. Da mit dem " $\cdot$ "-Symbol sowohl das innere 4D-Matrix-Vektor-Produkt als auch das gewöhnliche Matrix-Vektor-Produkt gemeint sein kann, soll hier ausnahmsweise die innere Multiplikation mit " $\otimes$ " bezeichnet werden. Dann gilt:

$$x^T \cdot v^T \odot M \odot w \cdot y = x^T \otimes v^T \odot M \odot w \otimes y = v^T \odot x^T \otimes M \otimes y \odot w = v^T \cdot x^T \otimes M \otimes y \cdot w$$

Aufgrund der Einführung der Matrix-Vektor-Produkte lassen sich Multilinearformen kompakt formulieren. Z.B. ist:⁶

$$\begin{aligned} x^T \cdot v^T \odot M \odot w \cdot y &= \sum_i \sum_j \sum_k \sum_l (M_{i,j})_{k,l} \cdot x_k \cdot y_l \cdot v_i \cdot w_j \\ x^T \odot M \odot w \cdot y &= \sum_j \sum_k \sum_l (M_{i,j})_{k,l} \cdot x_k \cdot y_l \cdot w_j \end{aligned}$$

⁶Der mit der Tensorrechnung vertraute Leser wird sich hier an die Einsteinsche Summenkonvention für indizierte Mehrfachsummen erinnern [Bet87], die ebenfalls eine äußerst kompakte Schreibweise ermöglicht. Der Autor bevorzugt jedoch eine indexfreie Notation analog zur gewöhnlichen Matrizenrechnung.

4.4.5 Matrix-Matrix-Produkte

Schließlich kann jede der oben eingeführten Matrix-Vektor-Multiplikationen kanonisch zu einer Matrix-Matrix-Multiplikation erweitert werden:

Definition 4.7 (Matrix-Matrix-Produkt) Ist ein Matrix-Vektor-Produkt $\otimes$ gegeben, so definiert

$$M \otimes (K_1, \dots, K_m) \stackrel{\text{def}}{=} (M \otimes K_1, \dots, M \otimes K_m)$$

ein dazu assoziiertes Matrix-Matrix-Produkt.

Für die Matrix-Matrix-Produkte gelten die von den Matrix-Vektor-Produkten ererbten Rechenregeln.

4.5 Matrixnotation der Bilanzgleichungen

In diesem Abschnitt werden die eingeführten mehrdimensionalen Matrizen angewendet, um zu einer kompakten Notation der Bilanzgleichungen zu gelangen. Alle Definitionen werden am laufenden Beispiel aus Abbildung 3.2 illustriert. Eine formal strengere Festlegung, die auch für eine automatische Generierung der entsprechenden Matrizen geeignet ist, wird in [Wie94b] vorgestellt.

4.5.1 Metabolit-Bilanzen

Definition 4.8 (stöchiometrische Matrix bzgl. der Hin- und Rückflüsse) Die stöchiometrische Matrix bzgl. der Hin- und Rückflüsse $N \in \mathbb{R}^{\dim \mathbb{X} \times \dim \mathbb{V}}$ zu einem System von Metabolit-Reaktionsschritten ist gegeben durch die Festlegung:

$$N_{i,j} = \begin{cases} \lambda & \text{falls der } i\text{-te Metabolit mit Koeffizient} \\ & \lambda \text{ auf der rechten Seite der } j\text{-ten Reaktion auftritt} \\ -\lambda & \text{falls der } i\text{-te Metabolit mit Koeffizient} \\ & \lambda \text{ auf der linken Seite der } j\text{-ten Reaktion auftritt} \\ 0 & \text{sonst} \end{cases}$$

Hierbei ist sowohl die Hin- als auch die Rückrichtung jeder Reaktion zu betrachten. Ferner sind die Koeffizienten mehrfach auf einer Seite auftretender Metabolite zu addieren.

Im laufenden Beispiel ist⁷

$$N = \begin{pmatrix} 1 & -1 & -1 & . & -1 & -1 & 1 & 1 & . & 1 \\ . & 1 & . & -1 & . & . & -1 & . & 1 & . \end{pmatrix} \quad (4.26)$$

und die stöchiometrischen Gleichungen (4.5) lassen sich formulieren in der kompakten Form

$$0 = N \cdot v \quad (4.27)$$

Diese Gleichung hat zunächst noch einen vorläufigen Charakter, denn die Matrix N wird später in Definition 6.4 auf Nettoflüsse bezogen und zudem um weitere Zeilen ergänzt, so daß weitere lineare Beziehungen dargestellt werden können.

⁷Aufgrund der Unidirektionalitätsannahmen könnten viele Einträge in der rechten Matrixhälfte auch entfallen, da die entsprechenden Flüsse ohnehin den Wert 0 haben, jedoch werden sie aus Symmetriegründen stets mitgeführt.

4.5.2 C-Atom-Bilanzen

Definition 4.9 (C-Atom-Transitionsmatrix, Input-C-Atom-Transitionsmatrix) Die C-Atom-Transitionsmatrix $\mathbf{P} \in \mathbb{R}^{\dim \mathbf{v} \times (\dim \mathbf{x} \times \dim \mathbf{x})}$ zu einem System von C-Atom-Transitionen ist gegeben durch die Festlegung:

$$(\mathbf{P}_i)_{j,k} = \begin{cases} \lambda & \text{falls die } i\text{-te Reaktion } \lambda \text{ C-Atome vom } k\text{-ten intrazellulären} \\ & \text{C-Atom-Pool zum } j\text{-ten intrazellulären C-Atom-Pool transportiert} \\ -\lambda & \text{falls } j = k \text{ ist und die } i\text{-te Reaktion } \lambda \text{ C-Atome aus dem} \\ & \text{} j\text{-ten intrazellulären C-Atom-Pool heraus transportiert} \\ 0 & \text{sonst} \end{cases}$$

Die Input-C-Atom-Transitionsmatrix $\mathbf{P}^{\text{inp}} \in \mathbb{R}^{\dim \mathbf{v} \times (\dim \mathbf{x} \times \dim \mathbf{x}^{\text{inp}})}$ zu einem System von C-Atom-Transitionen ist gegeben durch die Festlegung:⁸

$$(\mathbf{P}_i^{\text{inp}})_{j,k} = \begin{cases} \lambda & \text{falls die } i\text{-te Reaktion } \lambda \text{ C-Atome vom } k\text{-ten Input-} \\ & \text{C-Atom-Pool zum } j\text{-ten intrazellulären C-Atom-Pool transportiert} \\ 0 & \text{sonst} \end{cases}$$

Zur Verdeutlichung der Definition am laufenden Beispiel lassen sich die Gleichungen (4.6) mit Hilfe zweier zustandsabhängiger Matrizen formulieren als:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1^{\rightarrow} & \cdot \\ \cdot & v_1^{\rightarrow} \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} -v_2^{\rightarrow} - v_3^{\rightarrow} - v_5^{\rightarrow} & \cdot & v_2^{\leftarrow} & \cdot \\ \cdot & -v_2^{\rightarrow} - v_3^{\rightarrow} - v_5^{\rightarrow} & \cdot & v_2^{\leftarrow} \\ v_2^{\rightarrow} & v_3^{\rightarrow} & -v_2^{\leftarrow} - v_3^{\rightarrow} - v_4^{\rightarrow} & \cdot \\ \cdot & v_2^{\rightarrow} & v_3^{\rightarrow} & -v_2^{\leftarrow} - v_3^{\rightarrow} - v_4^{\rightarrow} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ e_1 \\ e_2 \end{pmatrix}$$

Unter Verwendung einer naheliegenden Schreibweise ergeben sich daraus die den Flüssen zugeordneten Matrizen (nur die nichtverschwindenden Matrizen sind aufgeführt)⁹

$$\begin{aligned} \mathbf{P}_1^{\text{inp}} &= \begin{pmatrix} 1 & \cdot \\ \cdot & 1 \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} & \mathbf{P}_3^{\rightarrow} &= \begin{pmatrix} -1 & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot \\ \cdot & 1 & -1 & \cdot \\ \cdot & \cdot & 1 & -1 \end{pmatrix} \\ \mathbf{P}_2^{\rightarrow} &= \begin{pmatrix} -1 & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \end{pmatrix} & \mathbf{P}_4^{\rightarrow} &= \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix} \\ \mathbf{P}_2^{\leftarrow} &= \begin{pmatrix} \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix} & \mathbf{P}_5^{\rightarrow} &= \begin{pmatrix} -1 & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \end{aligned} \quad (4.28)$$

⁸Im Grunde müssen hier nur die Hinflüsse $\mathbf{v}^{\rightarrow}$ der Substrataufnahmeschritte berücksichtigt werden, da extrazelluläre Flüsse nach Definition 3.5 unidirektional sind. Daher ist $(\mathbf{P}_i^{\text{inp}})_{j,k}$ für jeden Rückfluß $\mathbf{v}_i$ gleich Null zu setzen.

⁹Vgl. auch die in [ZS94] eingeführten "atom mapping matrices". Sie entsprechen genau denjenigen Teilmatrizen der Atom-Transitionsmatrizen, die durch Streichung aller Zeilen und Spalten zu unbeteiligten Edukten und Produkten entstehen. Diese Teilmatrizen können in einem Computeralgebra-System sehr elegant zur Formulierung der Bilanzgleichungen verwendet werden, erschweren jedoch eine allgemeine Systemanalyse, da alle diese Matrizen unterschiedliche Abmessungen haben. Im übrigen ist bei einer Implementierung in Form von schwach besetzten Matrizen der Speicherplatzaufwand beider Varianten identisch.

und daraus die 3D-Atom-Transitionsmatrizen:

$$\mathbf{P} = (\mathbf{P}_1^{\rightarrow}, \mathbf{P}_2^{\rightarrow}, \mathbf{P}_3^{\rightarrow}, \mathbf{P}_4^{\rightarrow}, \mathbf{P}_5^{\rightarrow}, \mathbf{P}_1^{\leftarrow}, \mathbf{P}_2^{\leftarrow}, \mathbf{P}_3^{\leftarrow}, \mathbf{P}_4^{\leftarrow}, \mathbf{P}_5^{\leftarrow})$$

$$\mathbf{P}^{\text{inp}} = (\mathbf{P}_1^{\text{inp}}, \mathbf{P}_2^{\text{inp}}, \mathbf{P}_3^{\text{inp}}, \mathbf{P}_4^{\text{inp}}, \mathbf{P}_5^{\text{inp}}, 0, 0, 0, 0, 0)$$

Offenbar lassen sich die C-Atom-Bilanzgleichungen (4.6) jetzt schreiben in der Form

$$\mathbf{0} = \sum_{i=1}^{\dim \mathbf{v}^{\rightarrow}} (\mathbf{P}_i^{\text{inp}} \cdot \mathbf{v}_i) \cdot \mathbf{x}^{\text{inp}} + \sum_{i=1}^{\dim \mathbf{v}^{\rightarrow}} (\mathbf{P}_i^{\rightarrow} \cdot \mathbf{v}_i^{\rightarrow}) \cdot \mathbf{x} + \sum_{i=1}^{\dim \mathbf{v}^{\leftarrow}} (\mathbf{P}_i^{\leftarrow} \cdot \mathbf{v}_i^{\leftarrow}) \cdot \mathbf{x}$$

oder unter Zuhilfenahme der 3D-Matrizen und des Gesamtflußvektors $\mathbf{v}$ kompakt als:

$$\mathbf{0} = \mathbf{P}^{\text{inp}} \odot \mathbf{v} \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v} \cdot \mathbf{x} \quad (4.29)$$

4.5.3 Isotopomeren-Bilanzen

Die in Gleichung (4.7) auftretenden bilinearen Terme können am Beispiel des E#11-Pools mit Hilfe einer symmetrischen Matrix notiert werden als:

$$e_{11} : 0 = \frac{1}{2} \cdot v_3^{\rightarrow} \cdot \begin{pmatrix} b_{00} \\ b_{01} \\ b_{10} \\ b_{11} \\ e_{00} \\ e_{01} \\ e_{10} \\ e_{11} \end{pmatrix}^T \cdot \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \cdot \begin{pmatrix} b_{00} \\ b_{01} \\ b_{10} \\ b_{11} \\ e_{00} \\ e_{01} \\ e_{10} \\ e_{11} \end{pmatrix} + v_2^{\rightarrow} \cdot b_{11} - (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e_{11}$$

Daher muß je eine symmetrische Matrix $\mathbf{Q}_{i,j}$ pro (bimolekularem) Reaktionsschritt $\mathbf{v}_i$ und pro Isotopomer $\mathbf{x}_j$ aufgebaut werden, so daß eine 4D-Matrix

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{1,1} & \cdots & \mathbf{Q}_{\dim \mathbf{v},1} \\ \vdots & \ddots & \vdots \\ \mathbf{Q}_{1,\dim \mathbf{x}} & \cdots & \mathbf{Q}_{\dim \mathbf{v},\dim \mathbf{x}} \end{pmatrix}$$

entsteht. Im Beispiel sind jedoch nur die Matrizen zum Reaktionsschritt $\mathbf{v}_3^{\rightarrow}$ und den als Produkte entstehenden Isotopomeren $e_{00}, e_{01}, e_{10}, e_{11}$ von Null verschieden. Allgemein führt dies zu der

Definition 4.10 (4D-Isotopomeren-Transitionsmatrix, 3D-Isotopomeren-Transitionsmatrix, 3D-Input-Isotopomeren-Transitionsmatrix) Die 4D-Isotopomeren-Transitionsmatrix $\mathbf{Q} \in \mathbb{R}^{(\dim \mathbf{v} \times \dim \mathbf{x}) \times (\dim \mathbf{x} \times \dim \mathbf{x})}$ zu einem System von bimolekularen Isotopomeren-Transitionen ist gegeben durch die Festlegung:

$$(\mathbf{Q}_{i,j})_{k,l} = \begin{cases} \frac{1}{2} & \text{falls die } i\text{-te Reaktion das} \\ & k\text{-te und das } l\text{-te Isotopomer zum } j\text{-ten Isotopomer kombiniert} \\ 0 & \text{sonst} \end{cases}$$

Die 3D-Isotopomeren-Transitionsmatrix $\mathbf{P} \in \mathbb{R}^{\dim \mathbf{v} \times (\dim \mathbf{x} \times \dim \mathbf{x})}$ und die 3D-Input-Isotopomeren-Transitionsmatrix $\mathbf{P}^{\text{inp}} \in \mathbb{R}^{\dim \mathbf{v} \times (\dim \mathbf{x} \times \dim \mathbf{x}^{\text{inp}})}$ beschreiben die monomolekularen Isotopomeren-Transitionen. Sie sind vollkommen analog zur Definition 4.9 für C-Atom-Transitionsmatrizen festgelegt, haben jedoch —trotz gleicher Bezeichnung— andere Dimensionen und Einträge.

Eine zu den C-Atom-Bilanzen (4.29) völlig analoge Kompaktnotation der Isotopomeren-Bilanzen ist nun gegeben durch:

$$\mathbf{0} = \mathbf{P}^{\text{inp}} \odot \mathbf{v}^T \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v}^T \cdot \mathbf{x} + \mathbf{x}^T \cdot \mathbf{Q} \odot \mathbf{v} \cdot \mathbf{x} \quad (4.30)$$

4.5.4 Zugehörigkeitsmatrizen

Die Bilanzgleichung (4.30) enthält noch nicht die Forderung, daß die Summe aller Isotopomeren-Anteile eines Metaboliten stets eins ist (vgl. Gleichung (4.3) und Abb. 1.5). Dies kann mit der nachfolgend eingeführten Matrixnotation leicht ausgedrückt werden. Ebenso fehlt bisher eine Matrixnotation, um die Beziehung zwischen Positions-Markierungs- und Isotopomeren-Anteilen (vgl. Gleichung (4.4) und Abbildung 1.5) auszudrücken. Dazu muß in der folgenden Definition ausnahmsweise notationell zwischen dem C-Atom-Markierungsvektor $\mathbf{x}^A$ und dem Isotopomeren-Markierungsvektor $\mathbf{x}^I$ unterschieden werden:

Definition 4.11 (C-Atom-Metabolit-Zugehörigkeitsmatrix, Isotopomeren-Metabolit-Zugehörigkeitsmatrix, Isotopomeren-C-Atom-Zugehörigkeitsmatrix, Input-Zugehörigkeitsmatrizen) Die C-Atom-Metabolit-Zugehörigkeitsmatrix

$\mathbf{Z}_M^A \in \mathbb{R}^{\dim \mathbf{X} \times \dim \mathbf{x}^A}$ zu einem System von C-Atomen und Metaboliten ist festgelegt durch:

$$(\mathbf{Z}_M^A)_{i,j} = \begin{cases} 1 & \text{falls das } j\text{-te C-Atom zum } i\text{-ten Metaboliten gehört} \\ 0 & \text{sonst} \end{cases}$$

Die Isotopomeren-Metabolit-Zugehörigkeitsmatrix $\mathbf{Z}_M^I \in \mathbb{R}^{\dim \mathbf{X} \times \dim \mathbf{x}^I}$ ist analog definiert. Die Isotopomeren-C-Atom-Zugehörigkeitsmatrix $\mathbf{Z}_A^I \in \mathbb{R}^{\dim \mathbf{x}^A \times \dim \mathbf{x}^I}$ ist gegeben durch:

$$(\mathbf{Z}_A^I)_{i,j} = \begin{cases} 1 & \text{falls das } i\text{-te C-Atom im } j\text{-ten Isotopomer markiert ist} \\ 0 & \text{sonst} \end{cases}$$

Nach dem gleichen Schema werden Zugehörigkeitsmatrizen $\mathbf{Z}_M^{A,\text{inp}}, \mathbf{Z}_M^{I,\text{inp}}, \mathbf{Z}_A^{I,\text{inp}}$ für die Input-Markierungsvektoren $\mathbf{x}^{A,\text{inp}}$ und $\mathbf{x}^{I,\text{inp}}$ festgelegt. Die Indizierung mit A und I wird nachfolgend weggelassen, wenn der Bezug aus dem Kontext klar ist.

Mit diesen Hilfsmitteln können die Beziehungen (4.3) und (4.4) nun sehr einfach ausgedrückt werden durch die Matrizen-Gleichungen

$$\mathbf{Z}_M^I \cdot \mathbf{x}^I = \mathbf{1} \quad \mathbf{Z}_M^{I,\text{inp}} \cdot \mathbf{x}^{I,\text{inp}} = \mathbf{1} \quad (4.31)$$

(wobei $\mathbf{1}$ den mit Eins-Einträgen gefüllten Vektor passender Dimension meint) und

$$\mathbf{Z}_A^I \cdot \mathbf{x}^I = \mathbf{x}^A \quad \mathbf{Z}_A^{I,\text{inp}} \cdot \mathbf{x}^{I,\text{inp}} = \mathbf{x}^{A,\text{inp}} \quad (4.32)$$

Beispielsweise werden die Gleichungen (4.3) und (4.4) nun geschrieben als:

$$\begin{aligned} (1 \ 1 \ 1 \ 1) \cdot (a_{00} \ a_{01} \ a_{10} \ a_{11})^T &= (1) \\ \begin{pmatrix} 1 & 1 & 1 & 1 & . & . & . & . \\ . & . & . & . & 1 & 1 & 1 & 1 \end{pmatrix} \cdot (b_{00} \ b_{01} \ b_{10} \ b_{11} \ e_{00} \ e_{01} \ e_{10} \ e_{11})^T &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} . & . & 1 & 1 \\ . & 1 & . & 1 \end{pmatrix} \cdot (a_{00} \ a_{01} \ a_{10} \ a_{11})^T &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\ \begin{pmatrix} . & . & 1 & 1 & . & . & . & . \\ . & 1 & . & 1 & . & . & . & . \\ . & . & . & . & . & . & 1 & 1 \\ . & . & . & . & . & 1 & . & 1 \end{pmatrix} \cdot (b_{00} \ b_{01} \ b_{10} \ b_{11} \ e_{00} \ e_{01} \ e_{10} \ e_{11})^T &= \begin{pmatrix} b_1 \\ b_2 \\ e_1 \\ e_2 \end{pmatrix} \end{aligned}$$

4.5.5 Beziehungen zwischen den C-Atom- und Isotopomeren-Bilanzgleichungen

Da die C-Atom-Positions-Markierungsanteile aus den Isotopomeren-Anteilen berechnet werden können, liegt es nahe, daß auch die C-Atom-Bilanzgleichungen aus den Isotopomeren-Bilanzgleichungen rekonstruierbar sind. In der Tat müssen nur die Bilanzen derjenigen Isotopomeren-Pools aufaddiert werden, die ein bestimmtes C-Atom erhalten, um zu der entsprechenden C-Atom-Bilanz zu gelangen. Beispielsweise ist in Gleichung (4.7):

$$\begin{array}{rcl}
 \text{B\#01 : } v_1^{\rightarrow} \cdot a_{01} & + & v_2^{\leftarrow} \cdot e_{01} = (v_3^{\rightarrow} + v_2^{\rightarrow} + v_5^{\rightarrow}) \cdot b_{01} \\
 + \text{ B\#11 : } v_1^{\rightarrow} \cdot a_{11} & + & v_2^{\leftarrow} \cdot e_{11} = (v_3^{\rightarrow} + v_2^{\rightarrow} + v_5^{\rightarrow}) \cdot b_{11} \\
 \hline
 \text{B\#2 : } v_1^{\rightarrow} \cdot (a_{01} + a_{11}) & + & v_2^{\leftarrow} \cdot (e_{01} + e_{11}) = (v_3^{\rightarrow} + v_2^{\rightarrow} + v_5^{\rightarrow}) \cdot (b_{01} + b_{11})
 \end{array}$$

Die untere Zeile ist aufgrund von Gleichung (4.4) gerade die C-Atom-Bilanz zu B#2 (Gleichung (4.6)). Ganz analog weist man nach, daß die Summe aller Isotopomeren-Bilanzen zu einem Metabolit-Pool gerade die stöchiometrische Gleichung dieses Pools, also gleich Null ist. Zusammenfassend gilt der

Satz 4.1 (Beziehungen zwischen C-Atom- und Isotopomeren-Bilanzen) Die rechte Seite der C-Atom-Bilanzgleichungen eines Markierungssystems sei gegeben durch

$$f^A(\mathbf{x}^{A,\text{inp}}, \mathbf{x}^A) = \mathbf{P}^{A,\text{inp}} \odot \mathbf{v}^T \cdot \mathbf{x}^{A,\text{inp}} + \mathbf{P}^A \odot \mathbf{v}^T \cdot \mathbf{x}^A$$

und die rechte Seite der entsprechenden Isotopomeren-Bilanzgleichung durch

$$f^I(\mathbf{x}^{I,\text{inp}}, \mathbf{x}^I) = \mathbf{P}^{I,\text{inp}} \odot \mathbf{v}^T \cdot \mathbf{x}^{I,\text{inp}} + \mathbf{P}^I \odot \mathbf{v}^T \cdot \mathbf{x}^I + (\mathbf{x}^I)^T \cdot \mathbf{Q} \odot \mathbf{v} \cdot \mathbf{x}^I.$$

Es gelte die stöchiometrische Gleichung $\mathbf{N} \cdot \mathbf{v} = \mathbf{0}$ mit der stöchiometrischen Matrix $\mathbf{N}$ aus Definition 4.8. Dann gilt:

$$\begin{array}{lcl}
 \mathbf{Z}_A^I \cdot \mathbf{x}^I = \mathbf{x}^A & \wedge & \mathbf{Z}_A^I \cdot \mathbf{x}^{I,\text{inp}} = \mathbf{x}^{A,\text{inp}} \implies \mathbf{Z}_A^I \cdot f^I(\mathbf{x}^{I,\text{inp}}, \mathbf{x}^I) = f^A(\mathbf{x}^{A,\text{inp}}, \mathbf{x}^A) \\
 \mathbf{Z}_M^I \cdot \mathbf{x}^I = \mathbf{1} & \wedge & \mathbf{Z}_A^I \cdot \mathbf{x}^{I,\text{inp}} = \mathbf{1} \implies \mathbf{Z}_M^I \cdot f^I(\mathbf{x}^{I,\text{inp}}, \mathbf{x}^I) = \mathbf{0}
 \end{array}$$

Auf den rein formalen Beweis dieser Beziehungen wird hier verzichtet.

4.5.6 Bilanzgleichungen im isotopisch instationären Fall

Die Zugehörigkeitsmatrizen lassen sich auch dazu benutzen, um die Differentialgleichungen für das instationäre System zu formulieren. Beispielsweise können die instationären C-Atom-Bilanzgleichungen (4.8) in Matrizenschreibweise geschrieben werden in der Art

$$\begin{pmatrix} B \cdot d/dt b_1 \\ B \cdot d/dt b_2 \\ E \cdot d/dt e_1 \\ E \cdot d/dt e_2 \end{pmatrix} = \begin{pmatrix} B & . & . & . \\ . & B & . & . \\ . & . & E & . \\ . & . & . & E \end{pmatrix} \cdot \begin{pmatrix} d/dt b_1 \\ d/dt b_2 \\ d/dt e_1 \\ d/dt e_2 \end{pmatrix}.$$

Offenbar läßt sich die hierin auftretende Strukturmatrix mit Hilfe der transponierten C-Atom-Metabolit-Zugehörigkeitsmatrix $\mathbf{Z}_M^A$ aus Definition 4.11 und des Poolgrößen-Vektors $\mathbf{X}$ darstellen als

$$\begin{pmatrix} B & \cdot & \cdot & \cdot \\ \cdot & B & \cdot & \cdot \\ \cdot & \cdot & E & \cdot \\ \cdot & \cdot & \cdot & E \end{pmatrix} = \text{diag} \left[\begin{pmatrix} 1 & \cdot \\ 1 & \cdot \\ \cdot & 1 \\ \cdot & 1 \end{pmatrix} \cdot \begin{pmatrix} B \\ E \end{pmatrix} \right] ,$$

wobei der Operator diag aus einem Vektor die entsprechende Diagonalmatrix bildet. Folglich können die instationären C-Atom-Bilanzen in Erweiterung von Gleichung (4.29) allgemein geschrieben werden als

$$\text{diag} \left[(\mathbf{Z}_M^A)^T \cdot \mathbf{X} \right] \cdot \dot{\mathbf{x}} = \mathbf{P}^{\text{inp}} \odot \mathbf{v} \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v} \cdot \mathbf{x} \quad . \quad (4.33)$$

Analog lassen sich die instationären Isotopomeren-Bilanzgleichungen aus Gleichung (4.30) kompakt formulieren als

$$\text{diag} \left[(\mathbf{Z}_M^A)^T \cdot \mathbf{X} \right] \cdot \dot{\mathbf{x}} = \mathbf{P}^{\text{inp}} \odot \mathbf{v} \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v} \cdot \mathbf{x} + \mathbf{x}^T \cdot \mathbf{Q} \odot \mathbf{v} \cdot \mathbf{x} \quad . \quad (4.34)$$

Kapitel 5

Netto- und Austauschflüsse

In Kapitel 4 wurde die Beschreibung eines bidirektionalen Stoffflusses durch einen Hin- und Rückfluß $v^{\rightarrow} = v_i^{\rightarrow}$ und $v^{\leftarrow} = v_i^{\leftarrow}$ eingeführt. Wie sich gezeigt hat, sind solche Hin- und Rückflüsse hervorragend geeignete Variablen zur Formulierung der Markierungs-Bilanzgleichungen, so daß ihre Bezeichnung als “natürliche Flußkoordinaten” sinnfällig ist. Andererseits sind diese Koordinaten eher ungeeignet, um das System im Sinne der experimentellen Anwendung zu charakterisieren, und zwar aus mehreren Gründen:

1. Gilt das Interesse der Stoffproduktion so ist in erster Linie der Nettofluß $v^{\text{net}} = v^{\rightarrow} - v^{\leftarrow}$ relevant [Val91]. Andererseits bietet jedoch die Quantifizierung des in beide Richtungen stattfindenden Stoffaustauschs eine wichtige Zusatzinformation. Sie erlaubt ggfs. die Auf-
findung von Reaktionsschritten mit großem Kontrolleinfluß auf die Regulation des Systems.
2. Standardmeßinstrumente (vgl. Abschnitt 2.1) erlauben stets nur die direkte Messung von Nettoflüssen. Dementsprechend werden die stöchiometrischen Gleichungen (4.27) üblicherweise mit Nettoflüssen formuliert. Im Hinblick auf den Vergleich mit bestehenden Methoden sollte daher auch der hier entwickelte allgemeine Ansatz zur Bilanzierung von Markierungssystemen Nettoflüsse zur Systembeschreibung verwenden. Insbesondere sollte sich die etablierte Metabolit-Bilanzierungsmethode [VS93, VP94a] als Spezialfall der allgemeinen Methode ergeben.
3. Im Rahmen von Simulationsstudien, die den Einfluß bidirektionaler Reaktionsschritte untersuchen, ist eine Variation des bidirektionalen Austauschs bei konstantem Nettofluß sinnvoll. Dazu sind in der bisherigen Formulierung stets zwei Größen zugleich — nämlich der entsprechende Hin- und Rückfluß — zu variieren.
4. Modellannahmen, die den bidirektionalen Stoffaustausch betreffen, also insbesondere die Annahme einer unidirektionalen oder schnell austauschenden Reaktion, sollten sich mit Hilfe eines einzigen, vom Nettofluß unabhängigen Parameters ausdrücken lassen.
5. Bzgl. der statistischen Güte von Flußschätzungen wird sich in Abschnitt 11.3 zeigen, daß Nettoflüsse in der Regel gut bestimmt werden können, weil bereits viele direkte Nettoflußmessungen vorliegen. Andererseits ist eine Schätzung des bidirektionalen Stoffaustauschs ausschließlich auf der Basis von Markierungsdaten möglich (vgl. Gleichung (4.12)). Es ist also zu erwarten, daß die Einführung eines alternativen Koordinatensystems zur Flußbeschreibung die Einflüsse der beiden Arten von Meßdaten stärker entkoppelt.

Zusammenfassend ist festzustellen, daß Hin- und Rückflüsse ein ungeeignetes Koordinatensystem für die Simulation, Interpretation und Auswertung von Markierungsexperimenten darstellen. Daher wird in diesem Kapitel eine äquivalente Beschreibung bidirektionaler Stoffflüsse mit Hilfe des Nettoflusses v^{net} und eines Austauschflusses v^{xch} eingeführt, die zusammen das sogenannte Anwendungs-Koordinatensystem bilden. Letzterer quantifiziert den in beiden Richtungen zugleich ablaufenden Fluß eines Reaktionsschritts.

Während der Austauschfluß v^{xch} hervorragend geeignet ist, um Direktionalitätsannahmen des experimentellen Anwenders zu formulieren, bereitet er hinsichtlich der numerischen und statistischen Behandlung einige Schwierigkeiten. Dies führt zu einigen Varianten der Definition des Austauschflusses, die bei bestimmten Problemlösungen besser geeignet sind, ohne dabei ihre Handhabbarkeit seitens des Anwenders aufzugeben. Dies führt zur Einführung sogenannter geglätteter und kompaktifizierter Flußkoordinaten. Abschließend werden dann anhand eines Beispiels die Eigenschaften von Markierungssystemen bezüglich der verschiedenen Varianten von Austauschflüssen diskutiert. Hierbei wird sich zeigen, daß Austauschflüsse ein hervorragend geeignetes Hilfsmittel für die Beschreibung schnell austauschender Reaktionsschritte darstellen.

5.1 Alternative Beschreibung von Stoffflüssen

5.1.1 Definition von Austauschflüssen

Zur Vermeidung überflüssiger Indizierungen wird in diesem Abschnitt weiterhin von einem einzelnen Reaktionsschritt mit Flüssen $v^{\rightarrow}$ und $v^{\leftarrow}$ ausgegangen. Der Nettofluß $v^{\text{net}} = v^{\rightarrow} - v^{\leftarrow}$ ist dann ein unstrittiger Kandidat für die erste Koordinate eines alternativen "Anwendungs-Flußkoordinatensystems". Gesucht ist nun eine zweite Maßzahl v^{xch} , die den bidirektionalen Stoffaustausch der entsprechenden Reaktion beschreibt. Diese wird im folgenden als Austauschfluß bezeichnet.

Austauschflüsse können im hier gegebenen Rahmen nicht über enzymkinetische oder thermodynamische Erwägungen eingeführt werden, da keine Annahmen über den Reaktionsmechanismus vorausgesetzt werden sollen. Daher empfiehlt sich eine phänomenologische Definition, die zweckmäßigerweise den folgenden Bedingungen genügt:

Definition 5.1 (Nettofluß, Austauschfluß) *Betrachtet wird ein Reaktionsschritt mit Hin- und Rückfluß $v^{\rightarrow} \geq 0$ und $v^{\leftarrow} \geq 0$. Der Nettofluß dieser Reaktion ist gegeben durch*

$$v^{\text{net}} \stackrel{\text{def}}{=} v^{\rightarrow} - v^{\leftarrow} .$$

Ein Austauschfluß v^{xch} für den betrachteten Reaktionsschritt ist eine Funktion von $v^{\rightarrow}$ und $v^{\leftarrow}$, die die nachfolgenden Bedingungen erfüllt:

Additive Skalierung: *Austauschflüsse skalieren additiv, d.h.*¹

$$\forall c > 0 \quad v^{\text{xch}}(v^{\rightarrow} + c, v^{\leftarrow} + c) = v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) + c .$$

Charakterisierung unidirektionaler Reaktionen: *Ein verschwindender Austauschfluß charakterisiert eine unidirektionale Reaktion, d.h.*

$$v^{\rightarrow} = 0 \quad \vee \quad v^{\leftarrow} = 0 \quad \Leftrightarrow \quad v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) = 0 .$$

¹Die Bedingung wird nur für $c > 0$ formuliert, damit der Definitionsbereich von $v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow})$ nicht verlassen wird.

Aufgrund dieser beiden Axiome ist v^{xch} bereits eindeutig festgelegt:

Satz 5.1 (Eindeutigkeit des Austauschflusses) v^{xch} ist durch Definition 5.1 eindeutig festgelegt als:

$$v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) = \min(v^{\rightarrow}, v^{\leftarrow})$$

Beweis:

Im Falle $v^{\rightarrow} > v^{\leftarrow}$ gilt: $v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) = v^{\text{xch}}(v^{\rightarrow} - v^{\leftarrow}, 0) + v^{\leftarrow} = v^{\leftarrow}$

Im Falle $v^{\rightarrow} < v^{\leftarrow}$ gilt: $v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) = v^{\text{xch}}(0, v^{\leftarrow} - v^{\rightarrow}) + v^{\rightarrow} = v^{\rightarrow}$

□

Abbildung 5.1a illustriert die Definition von Austauschflüssen.

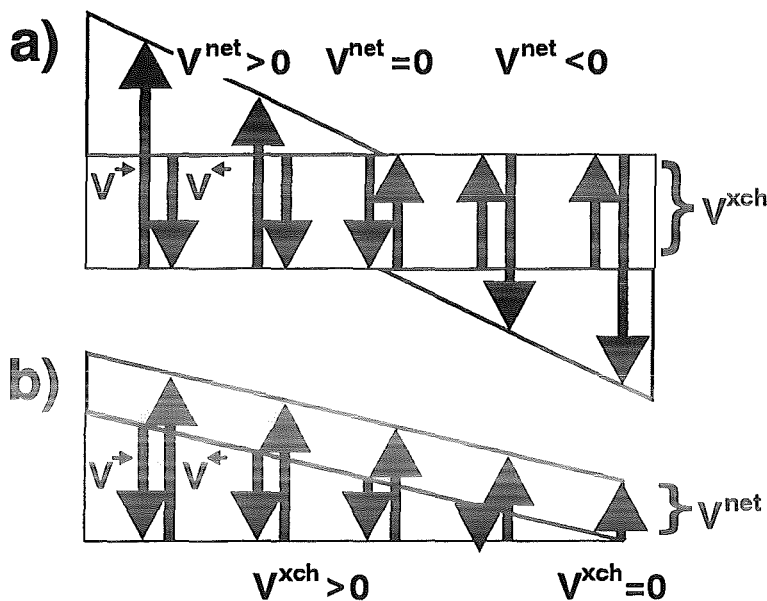


Abbildung 5.1: Veranschaulichung der Definition von Austauschflüssen für bidirektionale Reaktionsschritte: a) Variation des Nettoflusses bei festem Austauschfluß. b) Variation des Austauschflusses bei festem Nettofluß.

5.1.2 Eigenschaften von Austauschflüssen

Der folgende anhand von Abbildung 5.1 leicht zu verifizierende Satz zeigt, daß durch die Einführung von Netto- und Austauschflüssen eine alternative Koordinatendarstellung für $v^{\rightarrow}, v^{\leftarrow}$ erreicht wurde.

Satz 5.2 (Anwendungs-Flußkoordinatensystem) Die Beziehung

$$v^{\rightarrow} = v^{\text{xch}} + \max(v^{\text{net}}, 0)$$

$$v^{\leftarrow} = v^{\text{xch}} + \max(-v^{\text{net}}, 0)$$

legt eine bijektive Abbildung ϕ fest mit:

$$\phi: (v^{\text{net}}, v^{\text{xch}}) \longrightarrow (v^{\rightarrow}, v^{\leftarrow})$$

Offenbar ist diese Koordinatentransformation abschnittsweise linear, wie die Abbildung 5.2a veranschaulicht. Zum Vergleich zeigt Abbildung 5.2b das in [SSH92] vorgestellte lineare Koordinatensystem zur Beschreibung bidirektionaler Flüsse, das im wesentlichen gegeben ist durch

$$\begin{aligned} v_- &= v^{\rightarrow} - v^{\leftarrow} \\ v_+ &= \frac{1}{2} \cdot (v^{\rightarrow} + v^{\leftarrow}) \end{aligned}$$

Es wurde eingeführt, um das Verhalten eines Markierungssystems für $v_+ \rightarrow \infty$ zu untersuchen (vgl. auch Satz 8.4). Zwar ist dieses Koordinatensystem rechtechnisch gesehen angenehmer als das $(v^{\text{net}}, v^{\text{xch}})$ -System, es erfüllt jedoch die aus Anwendersicht gestellte Bedingung der additiven Skalierung nicht.

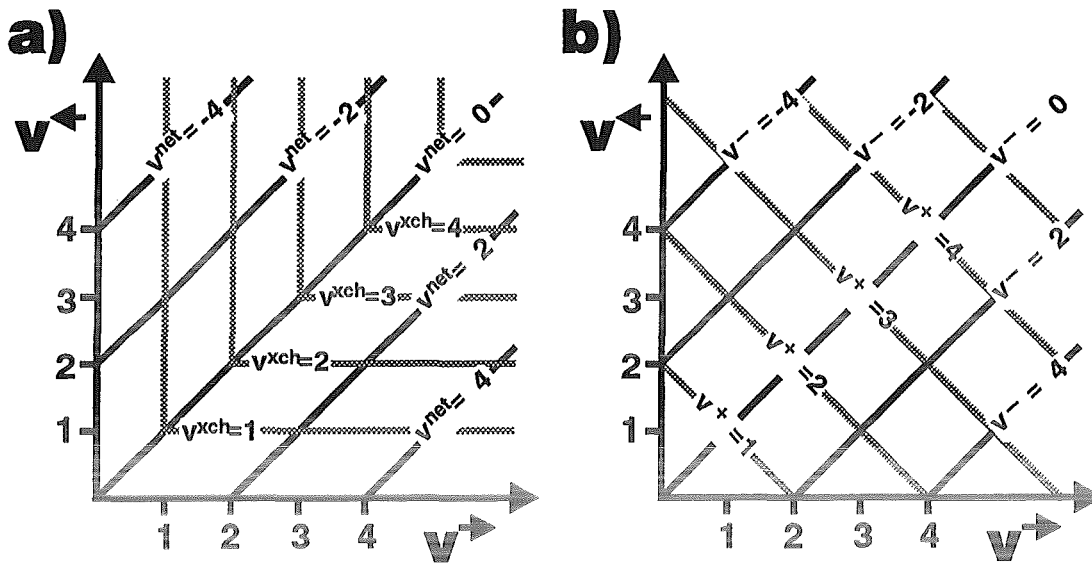


Abbildung 5.2: a) Veranschaulichung der durch die Definition 5.1 von Austauschflüssen gegebenen Koordinatentransformation Φ durch Überlagerung zweier Niveauliniendiagramme ("Gitterverzerrung"). b) Veranschaulichung des linearen Koordinatensystems aus [SSH92].

Austauschflüsse genügen noch einigen weiteren, leicht nachweisbaren Bedingungen, die belegen, daß die Definition sinnvoll gewählt wurde:

Satz 5.3 (Eigenschaften von Austauschflüssen) Für alle $v^{\rightarrow} \geq 0, v^{\leftarrow} \geq 0$ gilt:

Nicht-Negativität : Austauschflüsse sind nicht negativ, d.h.

$$v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) \geq 0 \quad .$$

Einheit : Die physikalische Einheit von v^{xch} ist die eines Flusses.

Normierung : Austauschflüsse sind bezüglich ihrer Größenordnung normiert durch die Bedingung

$$v^{\text{net}} = 0 \quad \Leftrightarrow \quad v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) = v^{\rightarrow} = v^{\leftarrow} \quad .$$

Symmetrie : Austauschflüsse verhalten sich symmetrisch bezüglich Hin- und Rückfluß, d.h.

$$v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) = v^{\text{xch}}(v^{\leftarrow}, v^{\rightarrow}) \quad .$$

Multiplikative Skalierung : *Austauschflüsse skalieren multiplikativ, d.h.*

$$\forall c > 0 \quad v^{\text{xch}}(c \cdot v^{\rightarrow}, c \cdot v^{\leftarrow}) = c \cdot v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) \quad .$$

Schnelle Gleichgewichte : *Große Austauschflüsse sind dadurch charakterisiert, daß*

$$v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow}) \rightarrow \infty \quad \Leftrightarrow \quad (v^{\rightarrow} \rightarrow \infty \quad \wedge \quad v^{\leftarrow} \rightarrow \infty) \quad .$$

Die Bedeutung von Austauschflüssen für die Beschreibung von Markierungssystemen wird in Abschnitt 5.2 noch eingehender behandelt.

5.1.3 Differenzierbare Approximation von Austauschflüssen

Für die numerische Behandlung von Markierungssystemen birgt die Definition 5.1 den Nachteil, daß die Funktion $v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow})$ für $v^{\rightarrow} = v^{\leftarrow} \Leftrightarrow v^{\text{net}} = 0$ zwar stetig, aber nicht differenzierbar ist (vgl. Abb. 5.2). Als Hilfskonstruktion wird daher eine differenzierbare Approximation eingeführt, deren Güte durch einen Parameter α variiert werden kann. Insbesondere soll im Grenzwert $\alpha \rightarrow 0$ die Approximation gegen die Originalfunktion konvergieren, während für $\alpha \rightarrow \infty$ immer "glattere", aber weniger gut approximierende Funktionen entstehen.

Eine geeignete differenzierbare Approximation resultiert aus der Feststellung, daß generell gilt:

$$\min(x, y) = \frac{1}{2} \cdot (x + y - |x - y|) \quad (5.1)$$

$$\max(x, y) = \frac{1}{2} \cdot (x + y + |x - y|) \quad (5.2)$$

$$(5.3)$$

Der Betragsterm läßt sich wiederum schreiben als

$$|x - y| = \sqrt{(x - y)^2} \quad ,$$

so daß sich unter Einführung des Glättungsparameters α die nachfolgende Approximation ergibt:

Definition 5.2 (geglättete Minimum- und Maximumfunktion, geglätteter Austauschfluß)

Sei $\alpha \geq 0$ ein Glättungsparameter. Dann ist die geglättete Maximum- und Minimumfunktion $(\min_{\alpha}, \max_{\alpha})$ festgelegt durch:²

$$\begin{aligned} \min_{\alpha}(x, y) &\stackrel{\text{def}}{=} \frac{1}{2} (x + y - \sqrt{(2\alpha)^2 + (x - y)^2}) \\ \max_{\alpha}(x, y) &\stackrel{\text{def}}{=} \frac{1}{2} (x + y + \sqrt{(2\alpha)^2 + (x - y)^2}) \end{aligned}$$

Der geglättete Austauschfluß zu $v^{\rightarrow} \geq 0, v^{\leftarrow} \geq 0$ ist hiermit gegeben durch

$$v_{\alpha}^{\text{xch}} \stackrel{\text{def}}{=} \min_{\alpha}(v^{\rightarrow}, v^{\leftarrow}) \quad .$$

Im Fall $\alpha = 0$ entspricht dies gerade der alten Definition (kein Approximationsfehler, aber Nichtdifferenzierbarkeit). Der Index α wird nachfolgend wieder weggelassen, wenn er sich aus dem Kontext ergibt.

²Der Term $(2\alpha)^2$ anstelle von α wird hier im Hinblick auf Satz 5.5 gewählt.

5.1.4 Eigenschaften des approximierten Austauschflusses

Es bleibt dem Leser überlassen, zu prüfen, welche der Bedingungen aus Definition 5.1 und Satz 5.3 für die Approximation noch exakt gelten und welche nur approximativ. Hier sei nur festgestellt, daß völlig analog zu Satz 5.2 gilt:

Satz 5.4 (Approximiertes Anwendungs-Flußkoordinatensystem) Für gegebenes $\alpha \geq 0$ definieren die Gleichungen

$$\begin{aligned} v^{\rightarrow} &= v^{\text{xch}} + \max_{\alpha} (v^{\text{net}}, 0) \\ v^{\leftarrow} &= v^{\text{xch}} + \max_{\alpha} (-v^{\text{net}}, 0) \end{aligned}$$

eine bijektive Abbildung

$$\phi_{\alpha} : (v^{\text{net}}, v^{\text{xch}}) \longrightarrow (v^{\rightarrow}, v^{\leftarrow}) .$$

Die Ableitungen von ϕ_{α} nach v^{net} und v^{xch} können berechnet werden mit Hilfe der Regel

$$\frac{\partial}{\partial x} \max_{\alpha} (x, 0) = \frac{1}{2} \left(1 + \frac{x}{2 \max_{\alpha} (x, 0) - x} \right) .$$

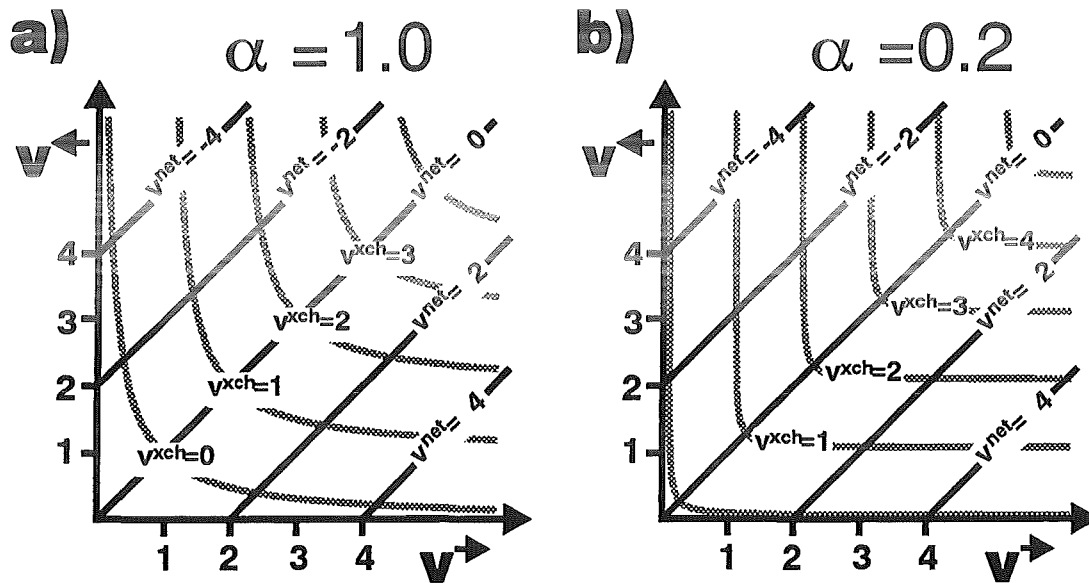


Abbildung 5.3: Veranschaulichung der durch die differenzierbare Approximation Φ_{α} des Austauschflusses (vgl. Definition 5.2) festgelegten Koordinatentransformation für verschiedene Werte des Glättungsparameters α (vgl. Abb. 5.2a).

Die approximierte Koordinatentransformation ist in Abbildung 5.3 für verschiedene Werte von α dargestellt. Für den maximalen Approximationsfehler liest man an Abbildung 5.3 ab ³:

Satz 5.5 (Maximaler Approximationsfehler des geglätteten Austauschflusses)

1. $\forall v^{\rightarrow} \geq 0, v^{\leftarrow} \geq 0 \quad |v_{\alpha}^{\text{xch}} - v^{\text{xch}}| \leq \alpha$
2. Das Maximum wird für $v^{\text{net}} = 0$ mit dem Wert α angenommen.

³Ein exakter Beweis findet sich in [Sie96].

5.2 Behandlung großer Austauschflüsse

In Abbildung 4.1 wurden anhand des Beispiels aus Abbildung 3.2 die physikalisch realisierbaren Markierungszustände eines Bilanzgleichungssystems dargestellt. Dieses Beispielsystem und insbesondere sein Verhalten in der Nähe der asymptotischen Grenzlinie wird nun mit Hilfe der neu eingeführten Flußparameter untersucht.

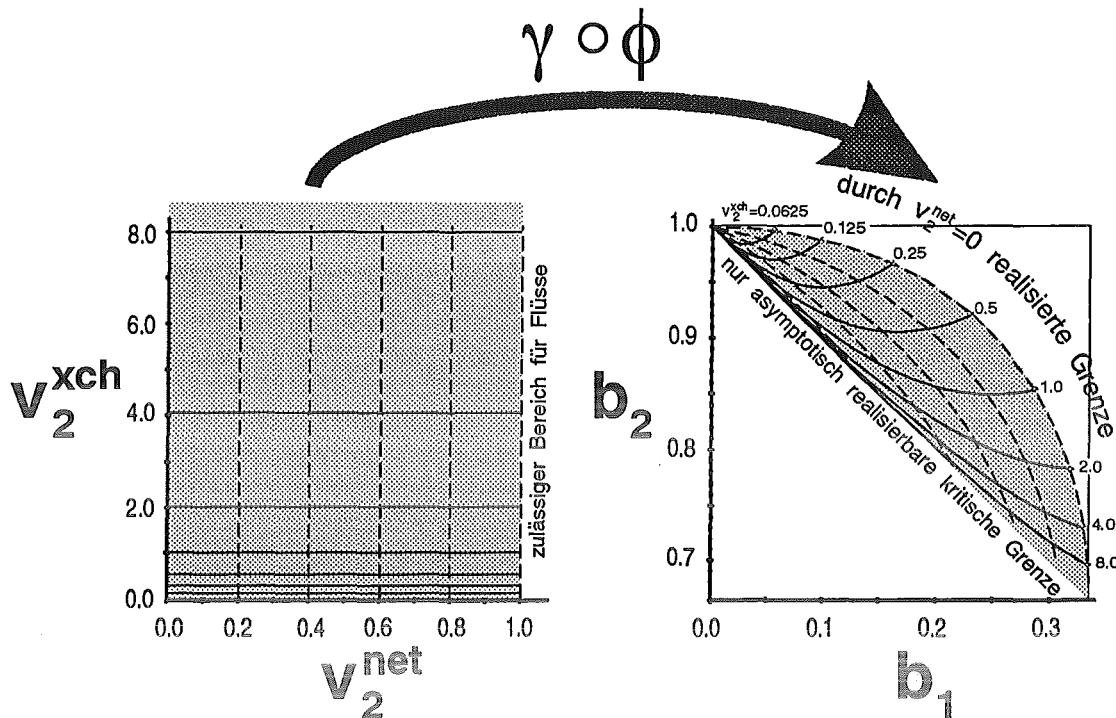


Abbildung 5.4: Realisierbare Markierungszustände für das Beispielsystem aus Abbildung 3.2 mit einigen eingezeichneten Niveaulinien bezüglich des Anwendungs-Koordinatensystems für $v_1^{\rightarrow} = 1, v_5^{\rightarrow} = 0$ (vgl. Abb. 4.1).

5.2.1 Charakterisierung der asymptotischen Grenzlinie des realisierbaren Gebietes

In Abbildung 4.1 wurde das Gebiet der erreichbaren Markierungszustände des Beispielsystems dargestellt. Dabei wurde deutlich, daß die asymptotisch realisierte Grenze des Gebiets durch die Bedingung $v_2^{\rightarrow}, v_2^{\leftarrow} \rightarrow \infty$ charakterisiert wird. Beim Übergang zu den Anwendungs-Koordinaten vermittelt

$$\gamma \circ \phi: (v_2^{\text{net}}, v_2^{\text{xch}}) \xrightarrow{1-1} (v_2^{\rightarrow}, v_2^{\leftarrow}) \xrightarrow{1-1} (b_1, b_2)$$

bleibt der Bildbereich natürlich derselbe (Abb. 5.4). Es zeigt sich jedoch, daß die physikalisch realisierbare Grenzlinie nun durch $v_2^{\text{net}} = 0$ beschrieben wird, während die asymptotische Grenzlinie einem schnellen Markierungsgleichgewicht entspricht:

$$v_2^{\text{net}} = \text{const} \wedge v_2^{\text{xch}} \rightarrow \infty \quad (5.4)$$

Für die Flußschätzung bedeutet dies, daß bei fehlerbehafteten Markierungsmeßwerten "unterhalb" der asymptotischen Grenze ein unendlicher Austauschfluß v_2^{xch} , aber ein endlicher Net-

tofluß v_2^{net} geschätzt wird. Damit kann die Behandlung solcher Grenzlagen unabhängig von den Nettoflüssen auf die Betrachtung von Austauschflüssen reduziert werden.⁴

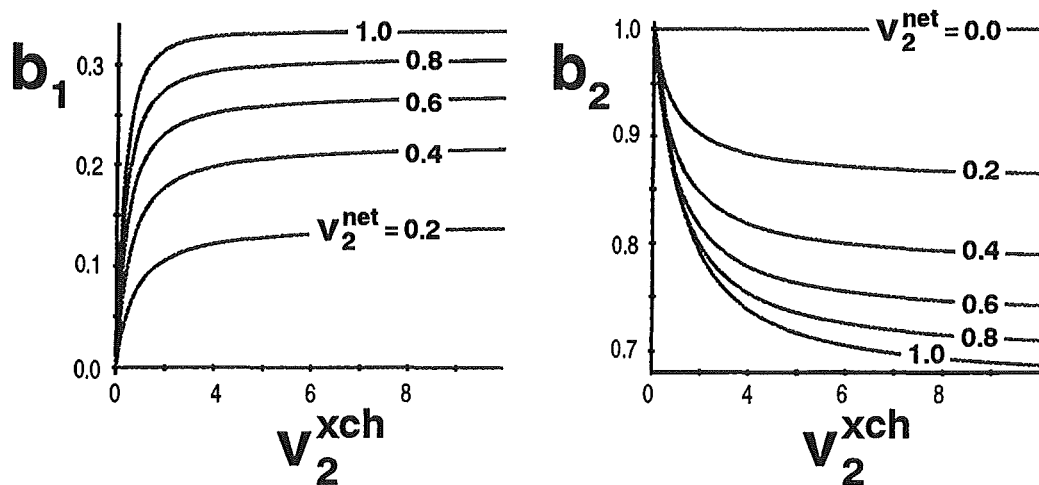


Abbildung 5.5: Auswirkung der Variation des Austauschflusses v_2^{xch} auf die C-Atom-Pools B#1 und B#2 im Beispielsystem aus Abb. 3.2 für verschiedene Werte von v_2^{net} ($v_1^{\text{net}} = 1$, $v_5^{\text{net}} = 0$).

5.2.2 Systemeigenschaften im Bereich großer Austauschflüsse

Ein unendlich großer Austauschfluß v_2^{xch} erlaubt direkte Rückschlüsse auf den Markierungszustand im C-Atom-Netzwerk. Er charakterisiert ein schnelles Gleichgewicht zwischen den durch den Fluß v_2 verbundenen C-Atom-Pools [SSH92]. Abbildung 5.5 illustriert das Grenzverhalten $v_2^{\text{xch}} \rightarrow \infty$ für verschiedene Werte von v_2^{net} . Folgende Aussagen können leicht überprüft werden (vgl. Abb. 5.5):

1. Der Markierungszustand der durch v_2 miteinander verbundenen Pools #B1 und #E1 bzw. #B2 und #E2 ist im Grenzwert identisch. Diese Aussage ist für große Austauschflüsse von allgemeiner Gültigkeit und wird in Satz 8.4 bewiesen.
2. Bei großen Austauschflüssen wirkt sich eine Variation von v_2^{xch} praktisch nicht mehr auf den Markierungszustand des Systems aus. Umgekehrt ist dann ein Austauschfluß nur noch größenordnungsmäßig aus Markierungsmeßdaten der betroffenen Pools schätzbar. Das daraus resultierende statistische Problem wird in Abschnitt 11.3 behandelt.
3. Die Grenzwerte der Markierungsanreicherung können explizit berechnet werden (vgl. Satz 8.3). Dabei zeigt sich, daß der Nettofluß v_2^{net} auch in der Grenzsituation seinen Einfluß auf das System behält, also insbesondere aus Markierungsmeßdaten schätzbar bleibt.
4. Bei Annahme extrem großer Austauschflüsse kann das C-Atom-Netz aus Abbildung 3.2 durch Zusammenziehen der austauschenden Pools #B1 und #E1 bzw. #B2 und #E2 zu je einem Pool auf ein Netzwerk mit zwei Pools reduziert werden. Daraus resultiert eine weitere, häufig angewandte Methode zur Netzwerkvereinfachung [SSH92] (vgl. auch Abb. 9.5).

⁴Offenbar strebt das System auch für einen unendlich großen Nettofluß der asymptotischen Grenzlinie zu. Jedoch ist dieser Fall für die Praxis irrelevant, da Nettoflüsse aufgrund der vorhandenen direkten Meßwerte bzw. der an sie gestellten Nebenbedingungen nur in gewissen Grenzen variabel sind (vgl. dazu Abschnitt 6.1).

5.2.3 Kompaktifizierte Definition von Austauschflüssen

Aufgrund der Analyse des Systemverhaltens bei großen Austauschflüssen ist eine erweiterte Definition von Austauschflüssen wünschenswert, die auf geeignete Weise den Wert ∞ mit in ihren Definitionsbereich einschließt. Derartige Vervollständigungen der Zahlengeraden werden in der Topologie als *Kompaktifizierungen* bezeichnet [Lan85]. Der Markierungszustand im Unendlichen soll dabei gleich dem Grenzwert sein. Für die praktische Anwendung sollte diese Festlegung darüber hinaus auch numerisch behandelbar sein, was bei Verwendung beliebig großer Zahlenwerte problematisch wäre.

Dieses Problem wird in der Analysis üblicherweise durch einen geeigneten Abschluß des Intervalls $[0, \infty[$ am rechten Rand gelöst. Zu diesem Zweck wird das Intervall $[0, 1]$ bijektiv auf $[0, \infty]$ abgebildet. In Abschnitt 8.2 wird sich zeigen, daß die nachfolgende Transformation hierfür besonders geeignet ist.⁵

Definition 5.3 ([0,1]-Kompaktifizierungs-Abbildung) Die $[0,1]$ -Kompaktifizierungs-Abbildung $[0, 1] \xrightarrow{1-x} [0, \infty]$ wird festgelegt durch die Transformation

$$\phi^{[0,1]} : \begin{array}{ll} x \in [0, 1[& \longrightarrow \frac{x}{1-x} \\ x = 1 & \longrightarrow \infty \end{array} .$$

Dies führt zur einer letzten Variante der Definition eines Austauschflusses, die im folgenden vor allem bei numerischen Rechnungen Verwendung findet.

Definition 5.4 ([0,1]-Austauschfluß) Der auf dem Intervall $[0, 1]$ definierte, bzgl. seines Wertebereichs kompaktifizierte $[0,1]$ -Austauschfluß ist gegeben durch

$$v^{\text{xch}[0,1]}(v^{\rightarrow}, v^{\leftarrow}) \stackrel{\text{def}}{=} (\phi^{[0,1]})^{-1}(v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow})) = \frac{v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow})}{1 + v^{\text{xch}}(v^{\rightarrow}, v^{\leftarrow})} .$$

Erneut ist damit eine Koordinatentransformation definiert, welche die kompaktifizierten Flußkoordinaten in die Anwendungs-Flußkoordinaten überführt:

Satz 5.6 (kompaktifiziertes Anwendungs-Flußkoordinatensystem) Die Kompaktifizierungsabbildung $\phi^{[0,1]}$ legt eine bijektive Abbildung

$$\phi^{[0,1]} : (v^{\text{net}}, v^{\text{xch}[0,1]}) \longrightarrow (v^{\text{net}}, v^{\text{xch}})$$

fest. Die Ableitung ergibt sich aus

$$\frac{\partial}{\partial x} \phi^{[0,1]}(x) = \frac{1}{(1-x)^2} .$$

⁵Allgemeiner kann statt des Terms $\frac{x}{1-x}$ auch der Term $\gamma \cdot \frac{x}{1-x}$ mit einer positiven Konstanten γ eingesetzt werden. Der Unterschied liegt in einer unterschiedlich schnellen Konvergenz gegen Unendlich mit wachsendem x . Alle im weiteren Verlauf dieser Arbeit entwickelten Konzepte können analog für die mit γ modifizierte Transformation formuliert werden. Empirische Studien haben gezeigt, daß die Parameterwahl $\gamma = 1.0$ gute Ergebnisse liefert, wenn alle Flüsse in der Größenordnung von 1.0 liegen. Da die Flüsse wie später gezeigt wird stets auf die Substrataufnahmerate 1.0 skaliert werden können, so daß alle anderen Flüsse ebenfalls in dieser Größenordnung liegen, wird in dieser Arbeit auf den formalen Aufwand der Behandlung von γ verzichtet. Eine eingehendere Untersuchung findet sich in [WG96b, WSGM96].

Die Zweckmäßigkeit der Definition 5.3 zeigt sich beim laufenden Beispiel daran, daß die Abbildung

$$\gamma \circ \phi \circ \phi^{[0,1]} : (v_2^{\text{net}}, v_2^{\text{xch}[0,1]}) \xleftrightarrow{1-1} (v_2^{\text{net}}, v_2^{\text{xch}}) \xleftrightarrow{1-1} (v_2^{\rightarrow}, v_2^{\leftarrow}) \xleftrightarrow{1-1} (b_1, b_2) \quad (5.5)$$

im Vergleich zu der Abbildung aus Gleichung (4.19) nun auf einer bzgl. $v_2^{\text{xch}[0,1]}$ kompakten Menge definiert und — wie in Satz 8.3 gezeigt — einschließlich des Randes stetig und differenzierbar ist.

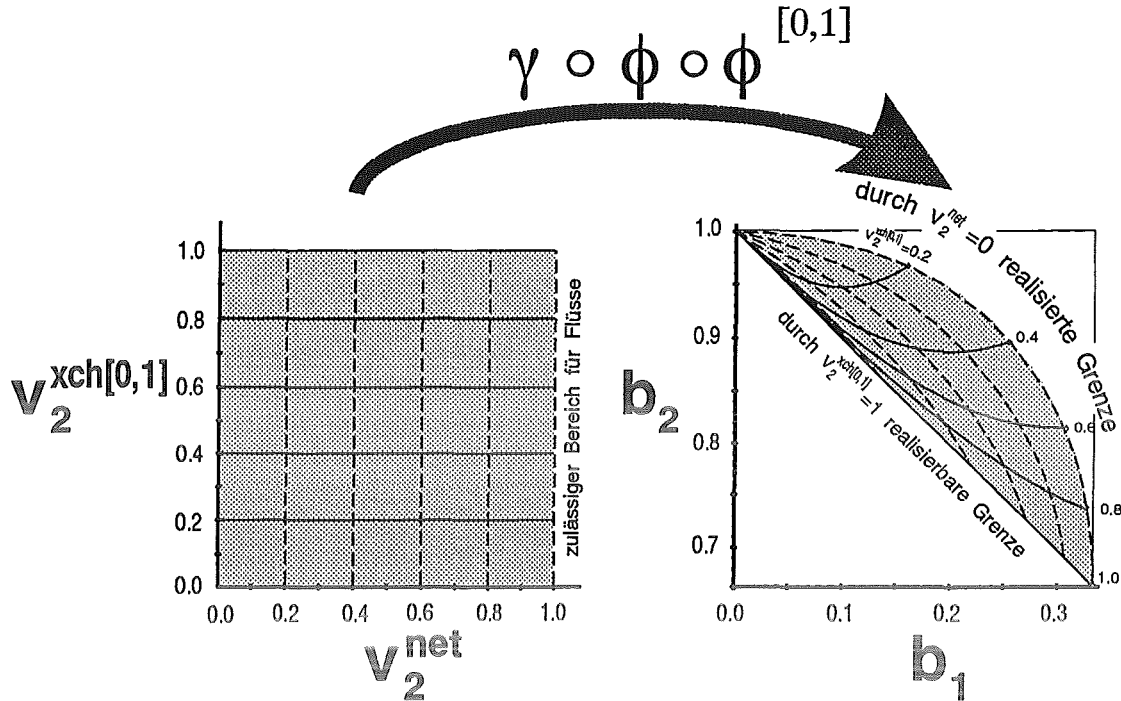


Abbildung 5.6: Realisierbare Markierungszustände für das Beispielsystem aus Abbildung 3.2 mit eingezeichneten Niveaulinien bezüglich des kompaktifizierten Anwendungskoordinatensystems. Man beachte, daß der Urbildbereich beschränkt ist und die Niveaulinien im Gegensatz zu Abb. 4.1 und 5.4 zu äquidistanten Werten von v_2^{net} und $v_2^{\text{xch}[0,1]}$ gehören.

Dies zeigen auch die Abbildungen 5.6 und 5.7. Insbesondere fällt aufgrund von Abbildung 5.7 auf, daß die Ableitung des Markierungszustands nach $v_2^{\text{xch}[0,1]}$ im Gegensatz zur Ableitung nach v_2^{xch} im Unendlichen im allgemeinen nicht verschwindet, was starke Konsequenzen für die numerische Gutartigkeit der kompaktifizierenden Koordinatentransformation hat. Es wird sich nachfolgend zeigen, daß damit sowohl die oben angedeuteten numerischen als auch die statistischen Probleme großer Austauschflüsse weitgehend entschärft sind.

5.3 Zusammenfassung der Flußkoordinatensysteme

Nachfolgend werden die Flüsse eines Markierungssystems stets in einem der oben eingeführten Flußkoordinatensysteme beschrieben. Zu diesem Zweck sollen die Flüsse und Koordinatentransformationen jetzt wieder in die vektorielle Schreibweise gebracht werden. Dabei wird die Operation der Glättung und Kompaktifizierung in einem Schritt zusammengefaßt:

Definition 5.5 (Nettoflußvektor, Austauschflußvektor, $[0,1]$ -Austauschflußvektor, Flußkoordinatentransformationen) *Ausgehend von den natürlichen Flußkoordinaten, gegeben durch*

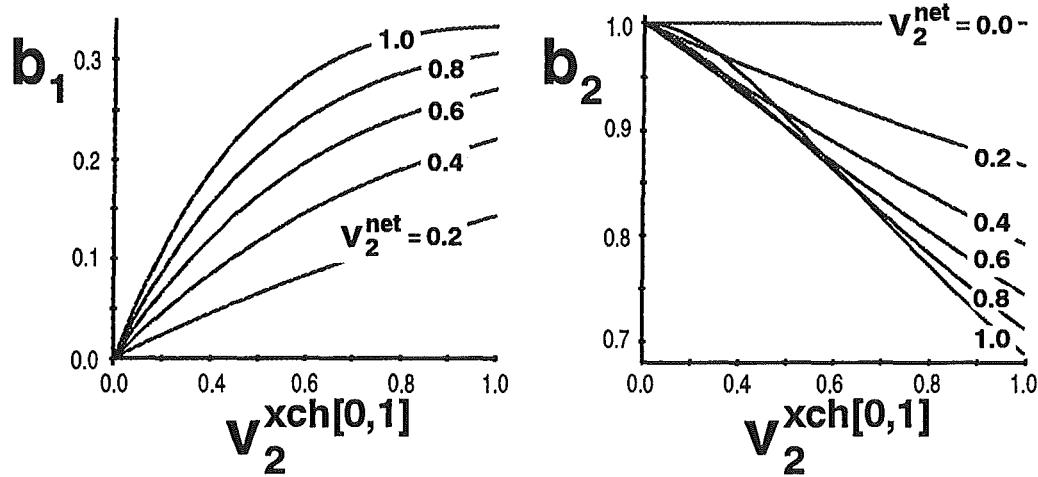


Abbildung 5.7: Auswirkung der Variation des kompaktifizierten Austauschflusses $v_2^{xch[0,1]}$ auf die C-Atom-Pools B#1 und B#2 im Beispielsystem aus Abb. 3.2 für verschiedene Werte von v_2^{net} (vgl. Abb. 5.5).

den Hinflußvektor $v^{\rightarrow}$ und den Rückflußvektor $v^{\leftarrow}$ werden die Anwendungs-Flußkoordinaten bzw. die kompaktifizierten Flußkoordinaten zur Beschreibung des Flußsystems zusammengefaßt

$$\begin{array}{lll}
 \text{zum Nettoflußvektor} & v^{net} & \stackrel{def}{=} v^{\rightarrow} - v^{\leftarrow} \\
 \text{zum (geglätteten) Austauschflußvektor} & v^{xch} & \stackrel{def}{=} \min_{\alpha} (v^{\rightarrow}, v^{\leftarrow}) \\
 \text{und zum } [0,1]\text{-Austauschflußvektor} & v^{xch[0,1]} & \stackrel{def}{=} (\phi^{[0,1]})^{-1} (v^{xch})
 \end{array}$$

Dabei sind das Minimum und die Kompaktifizierung koordinatenweise zu lesen. Entsprechende Flußkoordinatentransformationen auf den zusammengefaßten Vektoren sind festgelegt durch:⁶

$$\begin{array}{l}
 \Phi_{\alpha} : \begin{pmatrix} v^{net} \\ v^{xch} \end{pmatrix} \xleftrightarrow{1-1} \begin{pmatrix} v^{\rightarrow} \\ v^{\leftarrow} \end{pmatrix} = \begin{pmatrix} v^{xch} + \max_{\alpha} (v^{net}, 0) \\ v^{xch} + \max_{\alpha} (-v^{net}, 0) \end{pmatrix} \\
 \text{und } \Phi^{[0,1]} : \begin{pmatrix} v^{net} \\ v^{xch[0,1]} \end{pmatrix} \xleftrightarrow{1-1} \begin{pmatrix} v^{net} \\ v^{xch} \end{pmatrix} = \begin{pmatrix} v^{net} \\ v^{xch[0,1]} / (1 - v^{xch[0,1]}) \end{pmatrix}
 \end{array}$$

Diese drei Koordinatensysteme werden nachfolgend zu verschiedenen Zwecken verwendet:

1. Die natürlichen Flußkoordinaten $v^{\rightarrow}, v^{\leftarrow}$ eignen sich hervorragend zur Formulierung von Bilanzgleichungen (vgl. Kapitel 4).
2. Alle vom Experimentator getroffenen Annahmen werden im nachfolgenden Kapitel in Form linearer Gleichungen bzgl. v^{net} und v^{xch} formuliert (vgl. Abschnitt 6.1). Ebenso werden die Auswertungsergebnisse von Markierungsexperimenten im Anwendungs-Koordinatensystem interpretiert (vgl. Kapitel 12).
3. Die numerischen und statistischen Verfahren verwenden vorzugsweise die kompaktifizierten Koordinaten v^{net} und $v^{xch[0,1]}$ (vgl. Kapitel 10 und 11).

⁶Der Vektor-Bruch und die Maximumfunktion sind hier koordinatenweise zu lesen. Vgl. dazu auch die Notationskonventionen in der Einleitung.

Kapitel 6

Modellannahmen und Meßgleichungen

Wie bereits in Abschnitt 2.4 hervorgehoben, sollen Modellannahmen im Rahmen des hier vorgestellten Modellierungsansatzes präzise dokumentiert werden. Es wird sich zeigen, daß alle üblicherweise in der Literatur getroffenen Annahmen über intrazelluläre Flüsse als Nebenbedingungen an die Netto- und Austauschflüsse formuliert werden können. Zur formalen Notation dieser Annahmen werden eine Reihe von Nebenbedingungsmatrizen definiert, die vom experimentellen Anwender gemäß den gemachten Annahmen zu füllen sind. Wie dies geschieht, wird in Abschnitt 6.1 anhand des schon in den vorangehenden Kapiteln verwendeten Beispielnetzwerks aus Abbildung 3.2 vorgeführt.

Eine weitere Kategorie von Zusatzinformationen, die der experimentelle Anwender für die Stoffflußanalyse zur Verfügung stellen kann, ist natürlich durch die konkreten Meßwerte aus einem durchgeführten Experiment gegeben. Die Beziehung dieser Meßwerte zu den Zustandsvariablen der Bilanzgleichungen wird in Abschnitt 6.2 durch Meßgleichungen ausgedrückt, die außerdem den statistischen Fehler der Messung berücksichtigen. Dabei wird sich unter anderem eine quantitative Präzisierung des Begriffs der Multipllettanalyse aus Definition 2.2 ergeben.

Als Abschluß der Modellierung von metabolischen Markierungssystemen werden in Abschnitt 6.3 noch einmal alle Modellgleichungen zusammengestellt, um die Übersicht zu erleichtern.

6.1 Spezifikation von Modellannahmen

6.1.1 Formulierung von Annahmen durch lineare Gleichungen

Wie in Kapitel 5 ausgeführt, sind zur Formulierung von Modellannahmen über Markierungssysteme die Anwendungs-Flußkoordinaten v^{net} und v^{xch} am besten geeignet. Denn alle üblicherweise vom Anwender getroffenen Annahmen sind als lineare Gleichungen und Ungleichungen ausdrückbar, die entweder die Nettoflüsse oder die Austauschflüsse betreffen.

Beispielsweise werden die im Kapitel 4 für das Beispielsystem aus Abbildung 3.2 angegebenen stöchiometrischen Gleichungen (4.5) — welche die Annahme der metabolischen Stationarität ausdrücken — normalerweise nicht mit Hin- und Rückflüssen, sondern mit Nettoflüssen formuliert, also in der Form:

$$\begin{array}{lcl} \text{B : } & v_1^{\text{net}} - v_2^{\text{net}} - v_3^{\text{net}} & - v_5^{\text{net}} = 0 \\ \text{E : } & v_2^{\text{net}} + v_3^{\text{net}} - v_4^{\text{net}} & = 0 \end{array} \quad (6.1)$$

Ebenso entsprechen die dort getroffenen Direktionalitätsannahmen für die beteiligten Flüsse den

Ungleichungs-Bedingungen

$$v_1^{\text{net}} > 0 \quad v_3^{\text{net}} > 0 \quad v_4^{\text{net}} > 0 \quad v_5^{\text{net}} > 0 \quad . \quad (6.2)$$

Weiterhin lassen sich getroffene Annahmen über bidirektionale Reaktionsschritte mit Hilfe der entsprechenden Austauschflüsse schreiben. Beispielsweise können die in Gleichungen (4.2) mit Hilfe von Rückflüssen formulierten Unidirektionalitätsannahmen prägnanter geschrieben werden als

$$v_1^{\text{xch}} = 0 \quad v_3^{\text{xch}} = 0 \quad v_4^{\text{xch}} = 0 \quad v_5^{\text{xch}} = 0 \quad . \quad (6.3)$$

Schließlich kann ein Markierungsgleichgewicht aufgrund eines extrem hohen Austauschflusses formal notiert werden in der Form:

$$v_2^{\text{xch}} = \infty \quad (6.4)$$

Während die Austauschflüsse v^{xch} aus Anwendersicht die "angenehmsten" Variablen für die Spezifikation der getroffenen Annahmen darstellen, müssen alle numerischen Berechnungen im kompaktifizierten Fluß-Koordinatensystem — also mit Hilfe von $v^{\text{xch}[0,1]}$ — erfolgen (vgl. Kapitel 5). Für die praktische Umsetzung bedeutet dies, daß alle getroffenen Annahmen zunächst in das kompaktifizierte Koordinatensystem transformiert werden müssen. Aus Gleichung (6.3) und (6.4) wird also:

$$v_1^{\text{xch}[0,1]} = 0 \quad v_2^{\text{xch}[0,1]} = 1 \quad v_3^{\text{xch}[0,1]} = 0 \quad v_4^{\text{xch}[0,1]} = 0 \quad v_5^{\text{xch}[0,1]} = 0 \quad (6.5)$$

Nach Durchführung der jeweiligen numerischen Berechnungen sind dann die Ergebnisse wieder in das Anwendungs-Koordinatensystem zurückzutransformieren.

Der rein formale Aufwand der Hin- und Rücktransformation zwischen den beiden Koordinatensystemen soll in dieser Arbeit vermieden werden, um die Beschreibung der implementierten numerischen Verfahren von unnötigem formalen Ballast freizuhalten. Daher werden in diesem Abschnitt alle Annahmen über Austauschflüsse von vorn herein im kompaktifizierten Koordinatensystem formuliert. Das in Kapitel 9 beschriebene Softwaresystem ermöglicht dagegen die Formulierung aller Annahmen im Anwendungs-Koordinatensystem und führt intern die beschriebenen Transformationen durch. Der Leser mag sich davon überzeugen, daß jede der nachfolgend beschriebenen Typen von Annahmen sowohl im kompaktifizierten Koordinatensystem als auch im Anwendungs-Koordinatensystem durch lineare Gleichungen darstellbar ist, deren Koeffizienten wie beim Übergang von Gleichung (6.4) zu Gleichung (6.5) jedoch ggfs. ebenfalls transformiert werden müssen.

6.1.2 Annahmen über Nettoflüsse

Annahmen über Nettoflüsse können in verschiedenen Situationen erforderlich sein. Obligate und weitere übliche Annahmen sind:

Stationarität des Metabolit-Fluß-Systems : Diese wird grundsätzlich angenommen, wozu die stöchiometrischen Gleichungen (4.27) jetzt mit Hilfe der Nettoflüsse umformuliert werden in der Form

$$0 = N^{\text{net}} \cdot v^{\text{net}} \quad . \quad (6.6)$$

Im laufenden Beispiel entspricht Gleichung (6.1) sofort die Matrizenbeziehung:

$$\begin{pmatrix} 1 & -1 & -1 & . & -1 \\ . & 1 & . & -1 & . \end{pmatrix} \cdot (v_1^{\text{net}}, v_2^{\text{net}}, v_3^{\text{net}}, v_4^{\text{net}}, v_5^{\text{net}})^T = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Offenbar ist die Matrix N^{net} gerade die linke Hälfte der Matrix N in Gleichung (4.5).

Festsetzung von Nettoflüssen : Zu Simulationszwecken können auch einzelne Nettoflüsse v_i^{net} durch Wertvorgabe fixiert werden, etwa in der Form $v_i^{\text{net}} = n_i^{\text{net}}$ mit einem vorgegebenen Wert n_i^{net} .

Zusätzliche Bilanzgleichungen : Außerdem kann der Anwender beliebige zusätzliche lineare Gleichungen formulieren, um z.B. Annahmen über Energiebilanzen auszudrücken, wenn diese nicht bereits in der Netzwerkformulierung mit Hilfe von Metaboliten des Energiestoffwechsels (z.B. ATP, NADH) ausgedrückt sind (vgl. Abschnitt 3.2).

Die Zusammenfassung aller so entstehenden linearen Nebenbedingungen generalisiert die stöchiometrische Gleichung (6.6) zu einem allgemeinen Satz linearer Nebenbedingungen:

Definition 6.1 (Nebenbedingungsmatrix und Wertevektor für Nettoflüsse, erweiterte stöchiometrische Gleichung) Die Nebenbedingungsmatrix N^{net} für Nettoflüsse drückt zusammen mit dem Wertevektor n^{net} die stöchiometrischen Gleichungen und die vom Anwender vorgenommenen Gleichungsannahmen über Nettoflüsse in der Form der erweiterten stöchiometrischen Gleichung

$$n^{\text{net}} = N^{\text{net}} \cdot v^{\text{net}}$$

aus. Dabei entstehen die ersten $\dim X$ Zeilen von N^{net} aus der stöchiometrischen Matrix für Hin- und Rückflüsse nach Definition 4.8 durch Streichen der rechten Hälfte. Die entsprechenden ersten $\dim X$ Einträge von n^{net} sind Null. Die übrigen Zeilen von N^{net} und n^{net} sind frei belegbar, jedoch mit der Einschränkung, daß die Matrix N^{net} Maximalrang haben muß.¹

6.1.3 Annahmen über Austauschflüsse

Mögliche Annahmen über die Direktionalität eines Flusses sind:²

Unidirektionalität: $v_i^{\text{xch}[0,1]} = 0$

Schnelles Gleichgewicht: $v_i^{\text{xch}[0,1]} = 1$

Wertvorgabe: $v_i^{\text{xch}[0,1]} = n_i^{\text{xch}[0,1]}$

Weitere Typen von Annahmen über Austauschflüsse haben keine sinnvolle biologische Interpretation und werden daher nicht zugelassen. Die getroffenen Direktionalitätsannahmen lassen sich nun wieder zu einer Matrixnotation zusammenfassen:

Definition 6.2 (Nebenbedingungsmatrix und Wertevektor für Austauschflüsse, Wertsetzungen für Austauschflüsse) Die Nebenbedingungsmatrix $N^{\text{xch}[0,1]}$ für Austauschflüsse drückt zusammen mit dem Wertevektor $n^{\text{xch}[0,1]}$ die vom Anwender vorgenommenen Annahmen über Austauschflüsse in Form der Wertsetzung

$$n^{\text{xch}[0,1]} = N^{\text{xch}[0,1]} \cdot v^{\text{xch}[0,1]}$$

aus. Jede Zeile der Matrix ist ein Einheitsvektor. Zudem muß $N^{\text{xch}[0,1]}$ Maximalrang haben.

¹Dies stellt keine wesentliche Einschränkung dar, da linear abhängige Spalten der stöchiometrischen Matrix gestrichen werden können oder aber zu inkonsistenten Annahmen führen, die auf eine fehlerhafte Modellformulierung hinweisen.

²Die Wertvorgabe eines Austauschflusses ist dann sinnvoll, wenn Simulationsstudien durchgeführt werden, die den Effekt eines Austauschflusses auf das System demonstrieren sollen (vgl. z.B. Abbildung 5.5).

Im laufenden Beispiel aus Abbildung 3.2 müssen die extrazellulären Reaktionsschritte v_1 , v_3 , v_4 und v_5 gemäß Definition 3.5 als unidirektional angenommen werden, d.h. es ist

$$\begin{pmatrix} 1 & . & . & . & . \\ . & . & 1 & . & . \\ . & . & . & 1 & . \\ . & . & . & . & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1^{\text{xch}[0,1]} \\ v_2^{\text{xch}[0,1]} \\ v_3^{\text{xch}[0,1]} \\ v_4^{\text{xch}[0,1]} \\ v_5^{\text{xch}[0,1]} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} . \quad (6.7)$$

6.1.4 Freie Flüsse

Faßt man alle bisher getroffenen Annahmen zusammen, so ergibt sich die kombinierte Gleichung:

$$\begin{pmatrix} \mathbf{N}^{\text{net}} & \mathbf{0} \\ \mathbf{0} & \mathbf{N}^{\text{xch}[0,1]} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}[0,1]} \end{pmatrix} = \begin{pmatrix} \mathbf{n}^{\text{net}} \\ \mathbf{n}^{\text{xch}[0,1]} \end{pmatrix}$$

Da die hierin auftretende zusammengesetzte Matrix nach Definition Maximalrang hat, verbleiben damit immer noch $\dim \mathbf{v}^{\text{net}} - \dim \mathbf{n}^{\text{net}} + \dim \mathbf{v}^{\text{xch}[0,1]} - \dim \mathbf{n}^{\text{xch}[0,1]}$ Freiheitsgrade für die Variation von $\mathbf{v}^{\text{net}}$ und $\mathbf{v}^{\text{xch}[0,1]}$.

Diese Freiheitsgrade sollen im Hinblick auf die Durchführung von Simulationsläufen durch freie Wertvorgaben für weitere Einträge aus $\mathbf{v}^{\text{net}}$ und $\mathbf{v}^{\text{xch}[0,1]}$ ausgefüllt werden. Zu diesem Zweck müssen (mehr oder weniger willkürlich) weitere lineare Nebenbedingungen an die Flüsse formuliert werden.

Definition 6.3 (zusätzliche Nebenbedingungsmatrix und Wertevektor für Flüsse, zusätzliche Wertsetzungen für Flüsse) *Vorgegeben seien Nebenbedingungen für Netto- und Austauschflüsse gemäß Definition 6.1 und 6.2. Die zusätzliche Nebenbedingungsmatrix $\mathbf{N}^{\text{free}}$ für Netto- und Austauschflüsse drückt zusammen mit dem zusätzlichen Wertevektor $\mathbf{n}^{\text{free}}$ zusätzliche Wertsetzungen für vom Anwender ausgewählte Netto- und Austauschflüsse aus in der Form:*

$$\mathbf{n}^{\text{free}} = \mathbf{N}^{\text{free}} \cdot \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}[0,1]} \end{pmatrix}$$

Die $\dim \mathbf{v} - \dim \mathbf{n}^{\text{net}} - \dim \mathbf{n}^{\text{xch}[0,1]}$ Zeilen von $\mathbf{N}^{\text{free}}$ sind Einheitsvektoren. Ferner muß $\mathbf{N}^{\text{free}}$ Maximalrang haben.

Im laufenden Beispiel aus Kapitel 4 wurden entsprechend der Gleichung (4.11) die Flüsse $v_1^{\text{net}}, v_2^{\text{net}}, v_5^{\text{net}}, v_2^{\text{xch}[0,1]}$ durch zusätzliche Bedingungen Flüsse festgelegt in der Form

$$\begin{pmatrix} 1 & . & . & . & . & . & . & . \\ . & 1 & . & . & . & . & . & . \\ . & . & . & 1 & . & . & . & . \\ . & . & . & . & . & 1 & . & . \end{pmatrix} \cdot (v_1^{\text{net}}, \dots, v_1^{\text{xch}[0,1]}, \dots)^T = \begin{pmatrix} n_1^{\text{net}} \\ n_2^{\text{net}} \\ n_5^{\text{net}} \\ n_2^{\text{xch}[0,1]} \end{pmatrix} . \quad (6.8)$$

Durch die Festlegung zusätzlicher Nebenbedingungen ist jedoch noch nicht automatisch gesichert, daß diese zusammen mit den Nebenbedingungen für Netto- und Austauschflüsse alle Flüsse im System eindeutig bestimmen. Dazu muß das aus allen drei Typen von Nebenbedingungen kombinierte Gleichungssystem eindeutig lösbar sein. Nur in diesem Fall soll bei den durch die zusätzlichen Nebenbedingungen mit Werten belegten Flüssen von freien Flüssen gesprochen werden.

Definition 6.4 (Gesamt-Nebenbedingungsmatrix, Gesamt-Nebenbedingungsvektor, freie Flüsse) Unter den Bedingungen der Definition 6.3 ist die Gesamt-Nebenbedingungsmatrix N und der Gesamt-Wertevektor n gegeben durch

$$N = \begin{pmatrix} N^{\text{net}} & 0 \\ 0 & N^{\text{xch}[0,1]} \\ & N^{\text{free}} \end{pmatrix} \quad \text{und} \quad n = \begin{pmatrix} n^{\text{net}} \\ n^{\text{xch}[0,1]} \\ n^{\text{free}} \end{pmatrix} .$$

Die durch die zusätzliche Nebenbedingungsmatrix N^{free} ausgewählten Flüsse heißen freie Flüsse, wenn N Maximalrang hat.

Damit erhalten die linearen Nebenbedingungen für Netto- und Austauschflüsse ihre endgültige Gestalt

$$N \cdot \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} = n \quad (6.9)$$

Nachfolgend soll stets davon ausgegangen werden, daß freie Flüsse vorliegen, d.h. N invertierbar ist. Die Existenz eines solchen Satzes freier Flüsse wird in Abschnitt 8.1 diskutiert. Ferner wird dort gezeigt, wie der experimentelle Anwender beim Auffinden freier Flüsse unterstützt werden kann. In der Praxis liegen hierzu bereits sinnvolle Kandidaten in Form von charakteristischen Stoffflüssen für einzelne Stoffwechselwege vor.

Als Zusammenfassung der letzten Abschnitte ergibt sich eine durch die Freiheitsgrade n^{free} parametrisierte eindeutige Lösung für das gesamte Flußsystem. Im laufenden Beispiel kann der Maximalrang der Gesamt-Nebenbedingungsmatrix leicht geprüft werden. Hier hat Gleichung (6.9) die Gestalt:

$$\begin{pmatrix} 1 & -1 & -1 & . & -1 \\ . & 1 & . & -1 & . \\ \hline & & & 1 & . & . & . & . \\ & & & . & . & 1 & . & . \\ & & & . & . & . & 1 & . \\ \hline & & & . & . & . & . & 1 \\ 1 & . & . & . & . & . & . & . \\ . & 1 & . & . & . & . & . & . \\ . & . & . & . & 1 & . & . & . \\ . & . & . & . & . & 1 & . & . \end{pmatrix} \cdot \begin{pmatrix} v_1^{\text{net}} \\ v_2^{\text{net}} \\ v_3^{\text{net}} \\ v_4^{\text{net}} \\ v_5^{\text{net}} \\ \hline v_1^{\text{xch}[0,1]} \\ v_2^{\text{xch}[0,1]} \\ v_3^{\text{xch}[0,1]} \\ v_4^{\text{xch}[0,1]} \\ v_5^{\text{xch}[0,1]} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ \hline n_1^{\text{net}} \\ n_2^{\text{net}} \\ n_5^{\text{net}} \\ n_2^{\text{xch}[0,1]} \end{pmatrix}$$

und daher ist

$$\begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} = N^{-1} \cdot n = N^{-1} \cdot \begin{pmatrix} n^{\text{net}} \\ n^{\text{xch}[0,1]} \\ n^{\text{free}} \end{pmatrix} = f(n_1^{\text{net}}, n_2^{\text{net}}, n_5^{\text{net}}, n_2^{\text{xch}[0,1]}) \quad (6.10)$$

6.1.5 Ungleichheits-Nebenbedingungen und zulässige Flüsse

Der Definitionsbereich für die Flußvektoren v^{net} und $v^{\text{xch}[0,1]}$ muß aus verschiedenen Gründen eingeschränkt werden:

Flußrichtung extrazellulärer Flüsse : Nach den in Definition 3.5 getroffenen Annahmen werden alle extrazellulären Flüsse v_i^{net} als unidirektional mit vorgegebener Richtung angenommen: $v_i^{\text{net}} \geq 0$

Vorgabe einer Flußrichtung : Darüber hinaus kann auch die Richtung einer intrazellulären Reaktion v_j^{net} vom experimentellen Anwender vorgegeben werden: $v_j^{\text{net}} \geq 0$ bzw. $v_j^{\text{net}} \leq 0$

Wertebereich kompaktifizierter Austauschflüsse : Nach Definition der $[0,1]$ -Austauschflüssen müssen die Ungleichungen $0 \leq v_i^{\text{xch}[0,1]} \leq 1$ gelten. Die Beziehung $v_i^{\text{xch}[0,1]} \geq 0$ ist besonders wichtig, damit die Beziehung $v_i^{\rightarrow} \geq 0 \wedge v_i^{\leftarrow} \geq 0$ stets erfüllt ist (vgl. Satz 5.3).

Bereichsbeschränkte Austauschflüsse : Das zulässige Intervall für Austauschflüsse $v_i^{\text{xch}[0,1]}$ kann auch nach oben und unten enger beschränkt werden. So lassen sich etwa Unidirektionalität und schnelles Gleichgewicht ausschließen in der Form $\underline{u}_i^{\text{xch}[0,1]} \leq v_i^{\text{xch}[0,1]} \leq \overline{u}_i^{\text{xch}[0,1]}$

Bereichsbeschränkte Nettoflüsse : Numerische Optimierungsverfahren zur Flußschätzung setzen einen kompakten Suchraum voraus, d.h. die Werte der Nettoflüsse sollten in jedem Fall nach oben und unten beschränkt werden: $\underline{u}_i^{\text{net}} \leq v_i^{\text{net}} \leq \overline{u}_i^{\text{net}}$

Streng positive Nettoflüsse : In Abschnitt 7.1 wird gezeigt, daß bei der Simulation von Markierungsexperimenten gewisse pathologische Fälle ausgeschlossen werden müssen, die durch das Verschwinden eines Flusses entstehen (vgl. Definition 7.2 und Satz 7.3). Diese Situation, die nur für unidirektionale Flüsse auftritt, kann durch eine Annahme der Form $0 < \underline{v}_i^{\text{net}} \leq \overline{u}_i^{\text{net}}$ ausgeschlossen werden.

Bilanzlimitierungen : Für Nettoflüsse können darüber hinaus beliebige lineare Gleichungen formuliert werden, die gewisse Limitierungen ausdrücken können.

Ungleichungen, die eine Beziehung zwischen Netto- und Austauschflüssen fordern, haben keine sinnvolle biologische Interpretation und werden daher ausgeschlossen. Dies führt zu der nachfolgenden Darstellung der angenommenen Einschränkungen durch Matrizen:

Definition 6.5 (Ungleichungs-Nebenbedingungsmatrizen, Ungleichungs-Wertevektoren, Ungleichungs-Bedingungen, zulässige Flußvektoren) Die Ungleichungs-Nebenbedingungsmatrix $\mathbf{U}^{\text{net}}$ für Nettoflüsse faßt zusammen mit dem Ungleichungs-Wertevektor $\mathbf{u}^{\text{net}}$ für Nettoflüsse die grundsätzlich getroffenen Richtungsannahmen für extrazelluläre Flüsse und weitere vom Anwender getroffene Annahmen in der Form

$$\mathbf{U}^{\text{net}} \cdot \mathbf{v}^{\text{net}} \geq \mathbf{u}^{\text{net}}$$

zusammen. Die Ungleichungs-Nebenbedingungsmatrix $\mathbf{U}^{\text{xch}[0,1]}$ für Austauschflüsse faßt zusammen mit dem Ungleichungs-Wertevektor $\mathbf{u}^{\text{xch}[0,1]}$ für Austauschflüsse die grundsätzlich getroffene Annahme positiver Austauschflüsse und weitere Bereichsbeschränkungsannahmen für Austauschflüsse zusammen in der Form

$$\mathbf{U}^{\text{xch}[0,1]} \cdot \mathbf{v}^{\text{xch}[0,1]} \geq \mathbf{u}^{\text{xch}[0,1]} .$$

Ihre Zeilen dürfen jeweils nur einen von Null verschiedenen Eintrag ± 1 enthalten. Die Gesamt-Ungleichungs-Nebenbedingungsmatrix und der Gesamt-Ungleichungs-Wertevektor $\mathbf{u}$ sind gegeben durch

$$\mathbf{U} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{U}^{\text{net}} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}^{\text{xch}[0,1]} \end{pmatrix} \quad \text{und} \quad \mathbf{u} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{u}^{\text{net}} \\ \mathbf{u}^{\text{xch}[0,1]} \end{pmatrix} .$$

Die Beobachtungsgleichungen werden nun formuliert als:

$$\begin{aligned} \text{Fluß-Meßgleichung} \quad \mathbf{w} &= \mathbf{M}_w \cdot \mathbf{v} + \varepsilon_w, \quad \text{Cov}(\varepsilon_w) = \Sigma_w \\ \text{Markierungs-Meßgleichung} \quad \mathbf{y} &= \mathbf{M}_y \cdot \mathbf{x} + \varepsilon_y, \quad \text{Cov}(\varepsilon_y) = \Sigma_y \\ \text{Poolgrößen-Meßgleichung} \quad \mathbf{Y} &= \mathbf{M}_Y \cdot \mathbf{X} + \varepsilon_Y, \quad \text{Cov}(\varepsilon_Y) = \Sigma_Y \end{aligned} \quad (6.11)$$

Hierzu einige Anmerkungen:

1. Es sei daran erinnert daß stets nur entweder C-Atome oder Isotopomere betrachtet werden und bzgl. der Notation des Markierungsvektors nicht unterschieden wird (vgl. Definition 4.3).
2. Nur extrazelluläre Flüsse (also Nettoflüsse) sind direkt meßbar. Trotzdem wird die Fluß-Meßgleichung hier für den Gesamtflußvektor $\mathbf{v}$ formuliert, damit alle drei Gleichungen formal einheitlich sind.
3. Die Poolgrößen-Meßgleichung ist nur für isotopisch instationäre Experimente relevant. In dieser Situation muß zudem die Zeitabhängigkeit der Markierungs-Meßgleichungen berücksichtigt werden in Form einer instationären Meßgleichung $y(t_i) = \mathbf{M}_y \cdot \mathbf{x}(t_i) + \varepsilon_y(t_i)$, mit den Abtastzeitpunkten $0 \leq t_1 \leq \dots \leq t_n$.

Im laufenden Beispiel wurde angenommen (Abschnitt 4.3), daß die Flüsse v_1^{net} , v_5^{net} sowie die Markierungsanteile b_1, b_2 gemessen werden. Also sind die Meßmatrizen gegeben durch:

$$\mathbf{M}_w = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & -1 \end{pmatrix} \quad \text{und} \quad \mathbf{M}_y = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \end{pmatrix} \quad (6.12)$$

6.2.1 NMR-Multiplettanalyse

Die Aufstellung von Meßgleichungen für Isotopomeren-Systeme ist die quantitative Präzisierung des Vorgangs einer Multiplettanalyse (vgl. [KCS93, JSJ⁺93]) als Ausnutzung der aus Multiplettmessungen gewonnenen Information. Ziel ist dabei die Bestimmung möglichst vieler Isotopomeren-Anteile. Dies soll an dem in Abbildung 2.7 dargestellte Idealbeispiel eines C3-Körpers illustriert werden.

Bei der Multiplettanalyse sind die Meßmatrizen nicht nur aus Einheitsvektoren zusammengesetzt wie bei der Messung von Positionsmarkierungen, denn es besteht keine eindeutige Beziehung zwischen Multipletts und Isotopomeren (vgl. die Ausführungen zu Definition 2.2). Vielmehr wird die genaue Zuordnung durch die Meßgleichung angegeben. Betrachtet sei nun ein Spektrum eines Gemischs aus den 8 verschiedenen Isotopomeren des Moleküls I (vgl. Abb. 2.7). Bezeichnet man mit s_1, s_2, s_3 die Flächen der Singulett-Peaks im Spektrum, mit d_1, d_2, d_3 die Flächen der Doublett-Peaks und mit t_2 die Fläche des einzigen möglichen Triplett-Peaks bezogen auf die Gesamt-Isotopomerenmenge⁵ von 100%, so gilt:⁶

⁵Die Gesamt-Isotopomerenmenge kann mit Hilfe einer Kalibrationsmessung bestimmt werden. Vgl. dazu Abschnitt 2.2.

⁶Streng genommen darf die erste Gleichung nicht in die Matrix aufgenommen werden, da die Forderung, daß sich alle Anteile zu 1 addieren, keine Meßgleichung darstellt. Dies erfolgt hier lediglich der Einfachheit halber. Die Gleichung kann aber leicht aus den Meßgleichungen entfernt werden, indem man unter Ausnutzung dieser Beziehung i_{000} eliminiert.

$$M_y \cdot x + \varepsilon_y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ . & . & . & . & 1 & 1 & . & . \\ . & . & 1 & . & . & . & . & . \\ . & 1 & . & . & . & 1 & . & . \\ . & . & . & . & . & . & 1 & 1 \\ . & . & . & 1 & . & . & 1 & . \\ . & . & . & 1 & . & . & . & 1 \\ . & . & . & . & . & . & . & 1 \end{pmatrix} \cdot \begin{pmatrix} i_{000} \\ i_{001} \\ i_{010} \\ i_{011} \\ i_{100} \\ i_{101} \\ i_{110} \\ i_{111} \end{pmatrix} + \varepsilon_y = \begin{pmatrix} 1 \\ s_1 \\ s_2 \\ s_3 \\ d_1 \\ d_2 \\ d_3 \\ t_2 \end{pmatrix} = y \quad (6.13)$$

Interessanterweise hat die Meßmatrix nicht vollen Rang, d.h. mindestens eine der Gleichungen ist redundant zu den anderen. Es verbleibt jedoch lediglich ein Freiheitsgrad für die Bestimmung aller Isotopomeren-Anteile aus den Messungen, denn der Kern der Matrix wird durch den Vektor $(1, -1, 0, 0, -1, 1, 0, 0)^T$ erzeugt. Die Isotopomeren $l\#000$, $l\#001$, $l\#100$ und $l\#101$ können also nicht eindeutig voneinander getrennt werden. Diese Redundanzbeziehung drückt in einer quantitativ präzisen Form die bereits zu Abbildung 2.7 gemachten Aussagen aus.

Darüber hinaus können bei geschickter Ausnutzung der Möglichkeiten der NMR-Technik, deren Darstellung hier jedoch zu weit führen würde, noch weitere Meßwerte ermittelt werden können. Wichtig ist dabei, daß stets Messung und Isotopomeren-Anteile in einer linearen Beziehung zueinander stehen.

6.2.2 Massenspektroskopische Multipllettanalyse

Die massenspektroskopische Analyse des Isotopomeren-Gemischs kann nur Teilchen verschiedener Masse voneinander trennen. Bezeichnet m_0, m_1, m_2, m_3 den Anteil der Teilchen mit keiner, einem, zwei oder drei markierten Positionen, so resultiert die Meßgleichung

$$M_y \cdot x + \varepsilon_y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & . & . & . & . & . & . & . \\ . & 1 & 1 & . & 1 & . & . & . \\ . & . & . & 1 & . & 1 & 1 & . \\ . & . & . & . & . & . & . & 1 \end{pmatrix} \cdot \begin{pmatrix} i_{000} \\ i_{001} \\ i_{010} \\ i_{011} \\ i_{100} \\ i_{101} \\ i_{110} \\ i_{111} \end{pmatrix} + \varepsilon_y = \begin{pmatrix} 1 \\ m_0 \\ m_1 \\ m_2 \\ m_3 \end{pmatrix} = y \quad (6.14)$$

Diese Gleichung enthält offensichtlich weniger Information als mit einem NMR-Instrument erreicht werden kann. Andererseits genügt es, eine der massenspektroskopischen Messungen zu den NMR-Meßwerten hinzuzunehmen, um das gesamte Isotopomeren-Gemisch vollständig zu bestimmen.

Ein weiterer massenspektroskopischer Zugang zur Isotopomeren-Quantifizierung basiert auf einer chemischen Degradierung der untersuchten Moleküle durch Abspaltung einzelner Kohlenstoff-Atome [IL87, DRM⁺93]. Für das Beispiel sei angenommen, daß von dem C3-Körper sukzessive das dritte und dann das zweite Kohlenstoff-Atom abgespalten werden kann. Dies führt

zu fünf weiteren Meßwerten $m_0^3, m_1^3, m_2^3, m_0^{2,3}, m_1^{2,3}$ und den zusätzlichen Meßgleichungen

$$\begin{pmatrix} 1 & 1 & . & . & . & . & . & . \\ . & . & 1 & 1 & 1 & 1 & . & . \\ . & . & . & . & . & . & 1 & 1 \\ 1 & 1 & 1 & 1 & . & . & . & . \\ . & . & . & . & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} i_{000} \\ i_{001} \\ i_{010} \\ i_{011} \\ i_{100} \\ i_{101} \\ i_{110} \\ i_{111} \end{pmatrix} + \varepsilon'_y = \begin{pmatrix} m_0^3 \\ m_1^3 \\ m_2^3 \\ m_0^{2,3} \\ m_1^{2,3} \end{pmatrix} \quad (6.15)$$

Die aus den Meßmatrizen in Gleichung (6.14) und (6.15) zusammengesetzte Meßmatrix hat immer noch keinen Vollrang, jedoch hat der Kern nur die Dimension 1 und wird erzeugt von dem Vektor $(0, 0, 1, -1, -1, 1, 0, 0)$. Die Rangdefizienz läßt sich also durch einen einzigen unabhängigen Meßwert beseitigen. Dies legt nahe, daß eine Kombination von NMR-Messung und Massenspektroskopie bei Anwendung geschickter analytischer Ansätze durchaus in der Lage ist, bei Molekülen mit vier und fünf Kohlenstoff-Atomen zwischen den Isotopomeren weitgehend zu differenzieren.

6.2.3 Formulierung von Kovarianzmatrizen

Bei den meisten Meßgrößen haben die Kovarianzmatrizen der Meßfehler eine Diagonalgestalt, d.h. die einzelnen Messungen werden als voneinander unabhängig angenommen. Dies ist für Markierungsmessungen, Multiplettmessungen und Poolgrößenmessungen in der Regel der Fall.⁷ Anders ist dies bei der Messung von Flüssen. Daher soll hier das Problem der Fehlerübertragung kurz angeschnitten werden.

Aus den Gleichungen (2.1), (2.2) und (2.3) wird deutlich, daß bei einem Fermentationsexperiment die Biomasseabflüsse μ_i , die Substrataufnahmerate σ und die Produktbildungsrate π Funktionen der eigentlichen Meßgrößen X, S, P, D, c_i sind, die als unabhängig voneinander aufgefaßt werden. Damit können die daraus abgeleiteten Flußmeßgrößen jedoch untereinander korreliert sein.

Es wird angenommen, daß der Meßfehler von D im Vergleich zu den anderen Fehlern vernachlässigbar ist. Werden zudem exemplarisch nur zwei Biomasseabflüsse μ_1, μ_2 betrachtet, so liefert ein Linearisierungsansatz für die genannten Gleichungen die Jacobi-Matrix:

$$\mathbf{J} \stackrel{\text{def}}{=} \frac{\partial(\mu_1, \mu_2, \sigma, \pi)}{\partial(X, S, P, c_1, c_2)} = \frac{D}{X} \cdot \begin{pmatrix} . & . & . & X & . \\ . & . & . & . & X \\ -(S_{\text{inp}} - S)/X & -1 & . & . & . \\ -P/X & . & 1 & . & . \end{pmatrix}$$

Da eine unabhängige Messung von X, S, P, c_1, c_2 angenommen wird, ist die entsprechende Meß-Kovarianzmatrix gegeben durch

$$\Sigma = \text{Cov}(X, S, P, c_1, c_2) = \text{diag}(\sigma_X^2, \sigma_S^2, \sigma_P^2, \sigma_{c_1}^2, \sigma_{c_2}^2) \quad ,$$

und es folgt aufgrund der Linearisierung approximativ:

$$\text{Cov}(\mu_1, \mu_2, \sigma, \pi) \approx \mathbf{J} \cdot \Sigma \cdot \mathbf{J}^T \quad (6.16)$$

⁷In speziellen Situationen kann die Auswertung eines Spektrums auch zu korrelierten Meßwerten führen, nämlich dann, wenn eine Peaküberlagerung durch sukzessive Bestimmung von Peakflächen aufgelöst wird. Vgl. dazu die Beispiele in [Mö195].

Diese Matrix ist nun mit den konkret vorliegenden Daten auszuwerten. Für einen einfachen Spezialfall kann dies auch explizit durchgeführt werden. Dabei wird eine stöchiometrisch an die Substrataufnahme gekoppelte Wachstums- und Produktbildungsrate angenommen [Wie90], also

$$X = Y_{X/S} \cdot (S_{\text{inp}} - S) \quad \text{und} \quad P = Y_{P/S} \cdot (S_{\text{inp}} - S)$$

mit den (unbekannten) Ertragskoeffizienten $Y_{X/S}, Y_{P/S}$. Dann reduziert sich (6.16) auf:

$$\text{Cov} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \sigma \\ \pi \end{pmatrix} \approx D^2 \cdot \begin{pmatrix} 1/\sigma_1^2 & \cdot & \cdot & \cdot \\ \cdot & 1/\sigma_2^2 & \cdot & \cdot \\ \cdot & \cdot & \frac{(\sigma_S^2 + \sigma_X^2 Y_{X/S}^2)}{\sigma_X^2 \sigma_S^2 Y_{X/S}^4 (S_{\text{inp}} - S)^2} & \frac{Y_{P/S}}{\sigma_X^2 Y_{X/S}^4 (S_{\text{inp}} - S)^2} \\ \cdot & \cdot & \frac{Y_{P/S}}{\sigma_X^2 Y_{X/S}^4 (S_{\text{inp}} - S)^2} & \frac{(\sigma_P^2 Y_{P/S}^2 + \sigma_X^2 Y_{X/S}^2)}{\sigma_X^2 \sigma_P^2 Y_{X/S}^4 (S_{\text{inp}} - S)^2} \end{pmatrix} \quad (6.17)$$

Die einzige hierin auftretende Korrelation ungleich Null ist:

$$\begin{aligned} \text{Corr}^2(\sigma, \pi) &= \frac{\text{Cov}^2(\sigma, \pi)}{\text{Var } \sigma \cdot \text{Var } \pi} \\ &= \frac{\sigma_S^2 \sigma_P^2 Y_{P/S}^2}{(\sigma_P^2 Y_{P/S}^2 + \sigma_X^2 Y_{X/S}^2) \cdot (\sigma_S^2 + \sigma_X^2 Y_{X/S}^2)} \\ &= \frac{b^2 Y_{P/S}^2}{(b^2 Y_{P/S}^2 + a^2 Y_{X/S}^2) \cdot (1 + a^2 Y_{X/S}^2)} \end{aligned}$$

mit $a = \sigma_X/\sigma_S$ und $b = \sigma_P/\sigma_S$. Diese Korrelation tendiert für kleine a gegen 1. Dieser Fall ist jedoch unrealistisch, da σ_X in der Praxis deutlich größer als σ_S ist. Es sind also im allgemeinen nur kleine Korrelationen zu erwarten. Die Annahme unabhängiger Meßgrößen $\mu_1, \mu_2, \sigma, \pi$ ist dann näherungsweise gerechtfertigt.

6.3 Zusammenfassung der Modellgleichungen

Alle Modellgleichungen seien abschließend noch einmal in einer Übersicht zusammengestellt:

Zustandsvariablen : Diese sind nach Definition 4.1, 4.2 und 4.3 gegeben durch

$$\begin{aligned} &\text{den Markierungsvektor } \mathbf{x} \quad (\text{bzw. } \mathbf{x}^A, \mathbf{x}^I) \\ &\text{und den Gesamtflußvektor } \mathbf{v} \end{aligned}$$

Alternative Koordinatensysteme : Flüsse werden alternativ in drei Koordinatensystemen dargestellt durch:

$$\begin{aligned} &\text{den Hin- und Rückflußvektor} \\ &(\text{natürliche Flußkoordinaten nach Definition 4.1}) \quad \mathbf{v}^{\rightarrow} \quad \text{und} \quad \mathbf{v}^{\leftarrow} \\ &\text{den Netto- und Austauschflußvektor} \\ &(\text{Anwendungs-Flußkoordinaten nach Definition 5.5}) \quad \mathbf{v}^{\text{net}} \quad \text{und} \quad \mathbf{v}^{\text{xch}} \\ &\text{den Netto- und kompaktifizierten Austauschflußvektor} \\ &(\text{kompaktifizierte Flußkoordinaten nach Definition 5.5}) \quad \mathbf{v}^{\text{net}} \quad \text{und} \quad \mathbf{v}^{\text{xch}[0,1]} \end{aligned}$$

Insbesondere ist $\dim \mathbf{v}^{\rightarrow} = \dim \mathbf{v}^{\leftarrow} = \dim \mathbf{v}^{\text{net}} = \dim \mathbf{v}^{\text{xch}} = \dim \mathbf{v}^{\text{xch}[0,1]} = \dim \mathbf{v}/2$.

Koordinatentransformationen : Zwischen den einzelnen Flußkoordinatensystemen bestehen die Transformationen aus Satz 5.2, 5.4 und 5.6:

$$\Phi_{\alpha} : \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}} \end{pmatrix} \xleftrightarrow{1-1} \begin{pmatrix} \mathbf{v}^{\rightarrow} \\ \mathbf{v}^{\leftarrow} \end{pmatrix} = \begin{pmatrix} \mathbf{v}^{\text{xch}} + \max_{\alpha}(\mathbf{v}^{\text{net}}, \mathbf{0}) \\ \mathbf{v}^{\text{xch}} + \max_{\alpha}(-\mathbf{v}^{\text{net}}, \mathbf{0}) \end{pmatrix}$$

$$\Phi^{[0,1]} : \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}[0,1]} \end{pmatrix} \xleftrightarrow{1-1} \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}} \end{pmatrix} = \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}[0,1]} / (1 - \mathbf{v}^{\text{xch}[0,1]}) \end{pmatrix}$$

Systemdimensionen : Alle Systemdimensionen sind vollständig bestimmt durch die Dimension der nachfolgenden Vektoren:

Gesamtflußvektor	$\mathbf{v}$
Markierungsvektor	$\mathbf{x}$
Input-Markierungsvektor	$\mathbf{x}^{\text{inp}}$
Werte für Nettofluß-Nebenbedingungen	$\mathbf{n}^{\text{net}}$
Werte für Austauschfluß-Nebenbedingungen	$\mathbf{n}^{\text{xch}[0,1]}$
Werte für freie Flüsse	$\mathbf{n}^{\text{free}}$
Werte für Nettofluß-Ungleichungen	$\mathbf{u}^{\text{net}}$
Werte für Austauschfluß-Ungleichungen	$\mathbf{u}^{\text{xch}[0,1]}$
Flußmeßvektor	$\mathbf{w}$
Markierungsmeßvektor	$\mathbf{y}$
Poolgrößenmeßvektor	$\mathbf{Y}$

mit der Bedingung $\dim \mathbf{n}^{\text{net}} + \dim \mathbf{n}^{\text{xch}[0,1]} + \dim \mathbf{n}^{\text{free}} = \dim \mathbf{v}$.

Matrizen : Netzwerkstruktur, Annahmen und Meßgleichungen werden nach Definition 4.9, 4.10, 4.11, 6.3, 6.5 und 6.6 durch folgende Matrizen bestimmt, deren Dimensionalität jeweils in Klammern angegeben ist:

Transitionsmatrizen für C-Atome und Isotopomere:

Bimolekulare Schritte	$\mathbf{Q}$	$(\dim \mathbf{v} \times \dim \mathbf{x} \times \dim \mathbf{x} \times \dim \mathbf{x})$
Monomolekulare Schritte	$\mathbf{P}$	$(\dim \mathbf{v} \times \dim \mathbf{x} \times \dim \mathbf{x})$
Input-Schritte	$\mathbf{P}^{\text{inp}}$	$(\dim \mathbf{v} \times \dim \mathbf{x} \times \dim \mathbf{x}^{\text{inp}})$

Zugehörigkeitsmatrizen:

C-Atome zu Metaboliten	$\mathbf{Z}_M^A$	$(\dim \mathbf{X} \times \dim \mathbf{x}^A)$
Isotopomeren zu Metaboliten	$\mathbf{Z}_M^I$	$(\dim \mathbf{X} \times \dim \mathbf{x}^I)$
Isotopomeren zu C-Atomen	$\mathbf{Z}_A^I$	$(\dim \mathbf{x}^I \times \dim \mathbf{x}^A)$

Nebenbedingungsmatrizen für:

Nettoflüsse	$\mathbf{N}^{\text{net}}$	$(\dim \mathbf{n}^{\text{net}} \times \dim \mathbf{v}^{\text{net}})$
Austauschflüsse	$\mathbf{N}^{\text{xch}[0,1]}$	$(\dim \mathbf{n}^{\text{xch}[0,1]} \times \dim \mathbf{v}^{\text{xch}[0,1]})$
Freie Flüsse	$\mathbf{N}^{\text{free}}$	$(\dim \mathbf{n}^{\text{free}} \times \dim \mathbf{v})$

Ungleichungs-Nebenbedingungsmatrizen für:

Nettoflüsse	$\mathbf{U}^{\text{net}}$	$(\dim \mathbf{u}^{\text{net}} \times \dim \mathbf{v}^{\text{net}})$
Austauschflüsse	$\mathbf{U}^{\text{xch}[0,1]}$	$(\dim \mathbf{u}^{\text{xch}[0,1]} \times \dim \mathbf{v}^{\text{xch}[0,1]})$

Meßmatrizen für:

$$\begin{array}{ll} \text{Flüsse} & \mathbf{M}_w \quad (\dim w \times \dim v) \\ \text{Markierungen und Multipletts} & \mathbf{M}_y \quad (\dim y \times \dim x) \\ \text{Poolgrößen} & \mathbf{M}_Y \quad (\dim Y \times \dim X) \end{array}$$

Kovarianzmatrizen für die Messung von:

$$\begin{array}{ll} \text{Flüssen} & \Sigma_w \quad (\dim w \times \dim w) \\ \text{C-Atomen- und Multipletts} & \Sigma_y \quad (\dim y \times \dim y) \\ \text{Poolgrößen} & \Sigma_Y \quad (\dim Y \times \dim Y) \end{array} .$$

Bilanzgleichungen : Im allgemeinsten Fall gilt für Isotopomeren-Systeme die Bilanz (vgl. Gleichungen (4.29), (4.30), (4.33)):

$$\text{diag} [(\mathbf{Z}_M^I)^T \cdot \mathbf{X}] \cdot d/dt \mathbf{x} = \mathbf{P}^{\text{inp}} \odot \mathbf{v} \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v} \cdot \mathbf{x} + \mathbf{x}^T \cdot \mathbf{Q} \odot \mathbf{v} \cdot \mathbf{x} .$$

Für die Isotopomeren-Anteile muß nach Gleichung (4.31) zusätzlich gelten, daß

$$\mathbf{Z}_M^I \cdot \mathbf{x} = \mathbf{1} .$$

Für die Bilanzierung von C-Atomen gilt analog:

$$\text{diag} [(\mathbf{Z}_M^A)^T \cdot \mathbf{X}] \cdot d/dt \mathbf{x} = \mathbf{P}^{\text{inp}} \odot \mathbf{v} \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v} \cdot \mathbf{x} .$$

Annahmen über Flüsse : Alle getroffenen Annahmen werden zusammengefaßt zu den linearen Nebenbedingungen (Definition 6.1, 6.2, 6.4)

$$\mathbf{N} \cdot \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}[0,1]} \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{N}^{\text{net}} & \mathbf{0} \\ \mathbf{0} & \mathbf{N}^{\text{xch}[0,1]} \\ & \mathbf{N}^{\text{free}} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}[0,1]} \end{pmatrix} = \begin{pmatrix} \mathbf{n}^{\text{net}} \\ \mathbf{n}^{\text{xch}[0,1]} \\ \mathbf{n}^{\text{free}} \end{pmatrix} \stackrel{\text{def}}{=} \mathbf{n}$$

und den linearen Ungleichungen für zulässige Flüsse (Definition 6.5)

$$\mathbf{U} \cdot \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}[0,1]} \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{U}^{\text{net}} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}^{\text{xch}[0,1]} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{v}^{\text{net}} \\ \mathbf{v}^{\text{xch}[0,1]} \end{pmatrix} \geq \begin{pmatrix} \mathbf{u}^{\text{net}} \\ \mathbf{u}^{\text{xch}[0,1]} \end{pmatrix} \stackrel{\text{def}}{=} \mathbf{u} .$$

Meßgleichungen : Diese sind nach Gleichung (6.11) gegeben durch die:

$$\begin{array}{lll} \text{Fluß-Meßgleichung} & \mathbf{w} & = \mathbf{M}_w \cdot \mathbf{v} + \varepsilon_w, \quad \text{Cov}(\varepsilon_w) = \Sigma_w \\ \text{Markierungs-Meßgleichung} & \mathbf{y} & = \mathbf{M}_y \cdot \mathbf{x} + \varepsilon_y, \quad \text{Cov}(\varepsilon_y) = \Sigma_y \\ \text{Poolgrößen-Meßgleichung} & \mathbf{Y} & = \mathbf{M}_Y \cdot \mathbf{X} + \varepsilon_Y, \quad \text{Cov}(\varepsilon_Y) = \Sigma_Y \end{array}$$

Teil III

Simulation

Kapitel 7

Stabilitätsanalyse von Markierungssystemen

Bevor die numerischen Verfahren zur Lösung von Markierungs-Bilanzgleichungen vorgestellt werden, befaßt sich dieses Kapitel zunächst mit dem Problem ihrer eindeutigen Lösbarkeit bei Vorgabe der intrazellulären Flüsse. Für den experimentellen Anwender scheint diese Frage zunächst von geringer praktischer Bedeutung zu sein, denn er geht davon aus, daß sein Experiment unabhängig von den Startbedingungen immer denselben Ausgang nimmt. Dieses Verhalten wird in der Theorie dynamischer Systeme als globale asymptotische Stabilität bezeichnet. Bei vielen dynamischen Systemen kann diese Voraussetzung jedoch nicht als gegeben angenommen werden, d.h. nicht für jeden Flußzustand und jedes vorgelegte markierte Substrat muß sich ein eindeutig bestimmter stationärer Markierungszustand einstellen. Insbesondere ist zu klären, ob der Endzustand nicht auch von der intrazellulären Startmarkierung des Systems zum Zeitpunkt Null abhängt.

Die Existenz einer eindeutigen stabilen Lösung des jeweils untersuchten Bilanzgleichungssystems kann mit Hilfe einer mathematischen Analyse des dynamischen Verhaltens von Markierungssystemen bewiesen werden. Es wird sich zeigen, daß C-Atom-Markierungssysteme in der Regel einem durch Flüsse und Input-Markierung eindeutig bestimmten Gleichgewichtszustand zustreben. Die Ausnahme bilden einige pathologische Fälle, die für die Praxis jedoch bedeutungslos sind.

Die Stabilität der nichtlinearen Isotopomeren-Markierungssysteme wird hier erstmals untersucht. Dazu wird ein leistungsfähiges mathematisches Werkzeug bereitgestellt, das den Stabilitätsnachweis für ein gegebenes Netzwerk unter bestimmten Voraussetzungen ermöglicht. Dazu werden in Abschnitt 7.2 und 7.3 eine Reihe von mathematischen Grundlagen gelegt. Dies erfordert allerdings einige Vorkenntnisse aus der Theorie dynamischer Systeme, die hier nicht vermittelt werden können. Beispielanwendungen finden sich dann in Abschnitt 7.4.

7.1 Stabilitätsanalyse von C-Atom-Markierungssystemen

Die nachfolgend präsentierten Aussagen über lineare kompartmentelle Modelle sind seit langem bekannt, sie werden hier — dem Buch von Anderson [And83] folgend — kurz zusammengestellt. Voraussetzung ist jeweils, daß ein zulässiger Flußvektor v vorgegeben ist.

7.1.1 Kompartimentelle Matrizen

Von zentraler Bedeutung für die Stabilitätstheorie von allgemeinen Tracer-Systemen ist der Begriff der sogenannten kompartimentellen Matrix. Solche Matrizen treten ganz allgemein bei der Beschreibung von linearen Kompartimentsystemen auf.

Definition 7.1 (kompartimentelle Matrix) Eine kompartimentelle Matrix M ist eine quadratische Matrix, die folgende Eigenschaften hat:

1. Jedes Element außerhalb der Diagonalen ist nichtnegativ, d.h. $\forall_{i \neq j} M_{ij} \geq 0$.
2. Jedes Diagonalelement ist nichtpositiv, d.h. $\forall_i M_{ii} \leq 0$.
3. Jede Spaltensumme ist nichtpositiv, d.h. $\forall_j \sum_i M_{ij} \leq 0$.

Für den speziellen Fall stationärer C-Atom-Markierungssysteme gilt nun der nachfolgende Satz:

Satz 7.1 (C-Atom-Systemmatrix ist kompartimentell) Seien die stationären Bilanzgleichungen zu einem C-Atom-Markierungssystem gemäß Definition 4.9 gegeben durch

$$0 = P^{inp} \odot v \cdot x^{inp} + P \odot v \cdot x \quad .$$

Dann ist die Systemmatrix $P \odot v$ für alle $v \geq 0$ eine kompartimentelle Matrix.

Beweis:

Daß die Bedingungen 1. und 2. in Definition 7.1 erfüllt sind, ist nach Konstruktion der Atom-Transitionsmatrizen (vgl. Definition 4.9) unmittelbar klar.

Bedingung 3. ergibt sich aus der Überlegung, daß der Diagonaleintrag $(P \odot v)_{jj}$ den negativen Gesamtabfluß aus dem j -ten Pool einschließlich der extrazellulären Abflüsse angibt, während die positiven Einträge $(P \odot v)_{ij}$ mit $i \neq j$ nur diejenigen Abflüsse vom j -ten Pool zu einem anderen intrazellulären Pool i darstellen. Also ist

$$\sum_i (P \odot v)_{ij} = (P \odot v)_{jj} + \sum_{i \neq j} (P \odot v)_{ij} \leq 0 \quad .$$

□

7.1.2 Fallen in C-Atom-Netzwerken

Es gilt nun, jene Kompartimentsysteme zu charakterisieren, deren Bilanzgleichungen nicht eindeutig lösbar sind. Daß dieser Fall tatsächlich auftreten kann, belegen die Beispiele einer Insel (Abb. 7.1a) und einer Falle (Abb. 7.1b):

Definition 7.2 (Falle, Insel, fallenfreie Matrix) Ein Teilnetz eines Markierungsnetzwerkes heißt eine Falle, wenn keiner der darin befindlichen C-Atom-Pools einen Abfluß zu einem C-Atom-Pool außerhalb des Teilnetzes hat. Eine Falle heißt eine Insel, wenn es zusätzlich keinen externen Zufluß gibt.¹ Eine Matrix heißt fallenfrei, wenn sie — aufgefaßt als Adjazenzmatrix eines Netzwerks² [Chr75] — ein fallenfreies System beschreibt.

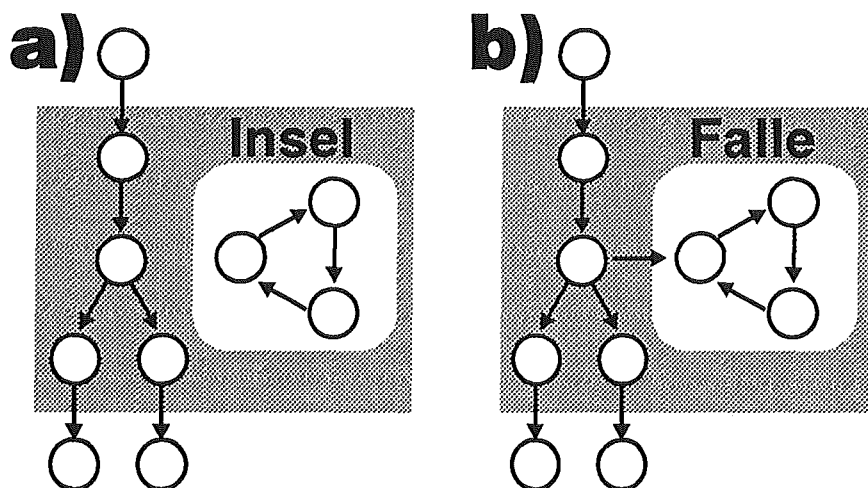


Abbildung 7.1: a) Eine Insel und b) eine Falle in einem C-Atom-Netzwerk als Ursache für einen nicht eindeutig durch den Input determinierten stationären Markierungszustand.

Eine Insel kann offenbar weder markierte C-Atome nach außen verlieren noch neue aufnehmen. Ihr Zustand ist somit abhängig von der jeweils zu Anfang des Experiments vorliegenden Startmarkierung. Diese führt zu unterschiedlichen Lösungen der Markierungsgleichungen. Offenbar ist das Auftreten einer Insel ein pathologischer Fall.

Die Existenz einer Falle im untersuchten Stoffwechselnetz ist dagegen zunächst durchaus denkbar. Ein markiertes Kohlenstoff-Atom, das in eine Falle geraten ist, kann das System nicht mehr verlassen — daher die Namensgebung. Auch hier ist die Gleichungslösung von der Startmarkierung der Falle abhängig. Allerdings kann eine Falle nur dann auftreten, wenn die dazu gehörigen C-Atom-Pools expandieren. In metabolisch stationären Stoffwechselnetzwerken ist jedoch gerade diese Situation ausgeschlossen. Es folgt der

Satz 7.2 (Äquivalenz von Falle und Insel) *In einem C-Atom-Markierungssystem ist jede Falle zugleich eine Insel, wenn die stöchiometrischen Gleichungen*

$$N \cdot v = 0$$

gemäß Definition 4.8 gelten.

Von nun ab sei das betrachtete System stets metabolisch stationär.

7.1.3 Eindeutige Lösbarkeit von C-Atom-Bilanzgleichungen

Das zentrale Ergebnis für die Lösbarkeit von C-Atom-Bilanzgleichungen besagt, daß mit den Fallen bereits alle pathologischen Situationen abgedeckt sind:

Satz 7.3 (Eindeutige Lösbarkeit von C-Atom-Bilanzen) *Sei ein metabolisch stationäres C-Atom-Markierungssystem gegeben durch:*

$$0 = P^{\text{inp}} \odot v \cdot x^{\text{inp}} + P \odot v \cdot x$$

Dann gilt:

¹In der Sprache der Graphentheorie ist dies eine Zusammenhangskomponente des Netzwerks, die weder einen Input-Pool noch einen Output-Pool enthält.

²Dabei kennzeichnet eine Spalte i mit einer negativen Summe einen Abfluß vom i -ten Pool zu einem Output-Pool.

1. Die Systemmatrix $\mathbf{P} \odot \mathbf{v}$ ist genau dann invertierbar, wenn das zugehörige C-Atom-Netzwerk fallenfrei ist. Dabei müssen Reaktionen mit $v_i = 0$ aus dem Netzwerk gestrichen werden.
2. Für fallenfreie Systeme ist der Markierungszustand eine Funktion Γ des Flußzustands, gegeben durch

$$\mathbf{x} = \Gamma(\mathbf{v}) = -(\mathbf{P} \odot \mathbf{v})^{-1} \cdot (\mathbf{P}^{\text{inp}} \odot \mathbf{v}) \cdot \mathbf{x}^{\text{inp}}.$$
3. Für fallenfreie Systeme ist Γ_i eine gebrochen rationale Funktion in den Variablen v_j , $j = 1, \dots, \dim \mathbf{v}$, deren Nenner die Determinante von $\mathbf{P} \odot \mathbf{v}$ ist.
4. $\forall \lambda > 0 \quad \Gamma(\mathbf{v}) = \Gamma(\lambda \cdot \mathbf{v})$

Beweis:

Punkt 1 wird in [And83] bewiesen, Punkt 2 folgt unmittelbar. Punkt 3 ergibt sich dann durch Anwendung der Cramerschen Regel [HJ85]. Punkt 4 ergibt sich schließlich aus $\mathbf{P} \odot (\lambda \mathbf{v}) = \lambda \cdot \mathbf{P} \odot \mathbf{v}$ und $\mathbf{P}^{\text{inp}} \odot (\lambda \mathbf{v}) = \lambda \cdot \mathbf{P}^{\text{inp}} \odot \mathbf{v}$.

□

Der letzte Punkt formuliert den bereits in Abschnitt 1.2 angesprochenen Sachverhalt, daß der Markierungszustand nur durch das gegenseitige Verhältnis der Flüsse bestimmt ist.

7.1.4 Untersuchung der Fallenfreiheit

Für große Netzwerke kann die Fallenfreiheit aufgrund von Satz 7.2 mittels eines Graphenalgorithmus zur Bestimmung von Zusammenhangskomponenten analysiert werden³ [Chr75], aber auch auf numerischem Wege (Abschnitt 8.1). Werden Simulationsstudien durchgeführt, bei denen die Flüsse $\mathbf{v} \geq 0$ variiert werden, so können zudem abhängig von den jeweils verschwindenden Flüssen $v_i = 0$ Fallen auftreten. Es muß dann durch Einschränkung der Menge der zulässigen Flüsse (vgl. Definition 6.5) sichergestellt werden, daß das System für alle zulässigen $\mathbf{v}$ fallenfrei ist.

In der Praxis ist dies wenig problematisch, da metabolische Netzwerke verhältnismäßig leicht überschaubar sind (vgl. Abb.1.2) und nach Satz 7.2 Fallen stets auch Inseln sind. Kritische Reaktionsschritte, die zum Auftreten von Fallen führen können, sind daher meist unschwer erkennbar:

1. Ein Reaktionsschritt mit positivem Austauschfluß $v_i^{\text{exh}} > 0$, kann grundsätzlich nicht zu einer Fallenbildung führen, da der Fluß in beide Richtungen stets positiv ist.
2. Ein verschwindender Output-Fluß kann zu einer Fallenbildung führen. In der Praxis kann jedoch das Verschwinden von Output-Flüssen anhand vorliegender Meßwerte sofort ausgeschlossen werden.
3. Zu untersuchen bleiben damit noch die als unidirektional angenommenen intrazellulären Reaktionsschritte. Kritisch sind solche Schritte, deren Wegfall den Zusammenhang des Netzes zerstört. Beispielsweise darf der GA3P-PEP-Schritt in Abb. 1.2 keineswegs entfallen. Dies zeigt zugleich, daß solche kritischen Schritte leicht erkennbar sind und ihr Verschwinden in der Praxis kaum in Betracht gezogen wird.

³Ein allgemeiner graphentheoretischer Ansatz zur Untersuchung der Fallenfreiheit beruht auf sogenannten *Cut-Sets*, die Punkte enthalten, deren Entfernung einen Mehrfachzusammenhang des Netzes zerstört [Chr75].

4. Schließlich können noch Kombinationen von unidirektionalen Flüssen zu einer Fallenbildung führen. Da ein Stoffwechselabschnitt nur wenige linear unabhängige Flüsse besitzt, verschwinden aufgrund der stöchiometrischen Gleichungen (4.27) meist viele Flüsse zugleich, wenn zwei und mehr der beteiligten unidirektionalen Flüsse gleich Null sind. Dies erleichtert die Untersuchung der Fallenbildung.

7.1.5 Stabilität von dynamischen Systemen

Mit der eindeutigen Lösbarkeit der Bilanzgleichungen ist noch nicht die Frage der globalen Stabilität des zugehörigen linearen Differentialgleichungssystems

$$\dot{\mathbf{x}} = \mathbf{P}^{\text{inp}} \odot \mathbf{v} \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v} \cdot \mathbf{x} \quad (7.1)$$

geklärt. Diese Eigenschaft ist für die numerische Lösung der stationären Gleichungen wünschenswert. Mit der Untersuchung dieses Differentialgleichungssystems wird zugleich die nachfolgende Analyse von Isotopomeren-Systemen vorbereitet. Es sei darauf hingewiesen, daß die Gleichung (7.1) zunächst noch nicht der Gleichung (4.33) für das isotopisch instationäre Markierungsexperiment entspricht, da die Poolgrößen fehlen. Es wird jedoch in einem nachfolgenden Abschnitt gezeigt, daß die Stabilitätsaussage für Gleichung (7.1) unmittelbar auf das instationäre System (4.33) übertragbar ist.

Definition 7.3 (globale Stabilität) Gegeben sei ein Differentialgleichungssystem

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \quad .$$

Dann heißt das System global stabil, wenn es einen Punkt $\mathbf{x}^*$ gibt, so daß für jeden Anfangszustand $\mathbf{x}(0)$ des Systems gilt:

$$\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{x}^*$$

$\mathbf{x}^*$ heißt dann der stabile Punkt des Systems.

Stabile Punkte sind insbesondere Lösungen des entsprechenden stationären Gleichungssystems, d.h. $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$, womit der Zusammenhang zur Lösung von Gleichungssystemen hergestellt ist. Aus der Theorie dynamischer Systeme [Per93] ist bekannt, daß die Stabilität linearer Systeme direkt mit den Eigenwerten der Systemmatrix zusammenhängt:

Satz 7.4 (Stabilität linearer Systeme) Ein lineares Differentialgleichungssystem

$$\dot{\mathbf{x}} = \mathbf{A} \cdot \mathbf{x} + \mathbf{b}$$

ist global stabil, wenn alle Eigenwerte λ_i der Matrix $\mathbf{A}$ einen negativen Realteil haben:

$$\forall_i \operatorname{re} \lambda_i < 0$$

In diesem Falle ist $\mathbf{x}^* = -\mathbf{A}^{-1} \mathbf{b}$ der eindeutige stabile stationäre Punkt.

7.1.6 Eigenwerte kompartmenteller Matrizen

Aufgrund der Definition 7.1 sind kompartmentelle Matrizen insbesondere diagonal-dominant, d.h.

$$\forall_i |M_{ii}| \geq \sum_{j \neq i} |M_{ij}| \quad .$$

Aus dem Satz von Gerschgorin [HJ85] folgt dann der

Satz 7.5 (Eigenwerte kompartmenteller Matrizen)

1. Jeder Eigenwert λ_i einer kompartmentellen Matrix M hat einen Realteil $\operatorname{re} \lambda_i \leq 0$.
2. Rein imaginäre Eigenwerte treten nicht auf, d.h. es gilt $\operatorname{re} \lambda_i = 0 \implies \lambda_i = 0$.

Punkt 2 wird in [And83] mit Hilfe des Perron-Frobenius-Theorems bewiesen.

Nun ist bekanntlich ist die Invertierbarkeit einer Matrix äquivalent dazu, daß der Eigenwert Null nicht auftritt, denn

$$M \text{ singular} \iff \exists_{v \neq 0} M v = 0 \iff v \text{ ist Eigenvektor zu Eigenwert } 0 \quad .$$

Mit den Sätzen 7.1, 7.3 und 7.5 folgt daher unmittelbar der Stabilitätssatz:

Satz 7.6 (globale Stabilität von C-Atom-Systemen) Die dynamischen C-Atom-Bilanzgleichungen

$$\dot{x} = P^{\text{inp}} \odot v \cdot x^{\text{inp}} + P \odot v \cdot x$$

sind genau dann global stabil, wenn das C-Atom-Netzwerk für gegebene Flüsse v fallenfrei ist.

7.1.7 Instationäre Bilanzgleichungen

Die Stabilitätsaussage für instationäre C-Atom-Markierungsexperimente kann nun leicht auf die obigen Ergebnisse für die Gleichungen (7.1) zurückgeführt werden. Dazu wird das instationäre Gleichungssystem (vgl. Gleichung (4.33))

$$\operatorname{diag} [(Z_M^A)^T \cdot X] \cdot \dot{x} = P^{\text{inp}} \odot v \cdot x^{\text{inp}} + P \odot v \cdot x$$

mit Hilfe des Koordinatenwechsels $\bar{x} = \operatorname{diag} [(Z_M^A)^T \cdot X] \cdot x$ transformiert zu

$$\dot{\bar{x}} = P^{\text{inp}} \odot v \cdot x^{\text{inp}} + (P \odot v) \cdot \operatorname{diag} [(Z_M^A)^T \cdot X]^{-1} \cdot \bar{x} \quad .$$

Dann ist die transformierte Systemmatrix $(P \odot v) \cdot \operatorname{diag} [(Z_M^A)^T \cdot X]^{-1}$ erneut eine kompartmentelle Matrix, denn die Spalten von $P \odot v$ werden lediglich mit je einer positiven Konstanten multipliziert. Da sich das zugrundeliegende Netzwerk dabei nicht verändert, hat das instationäre System dieselben Stabilitätseigenschaften wie das stationäre System. Es folgt, daß fallenfreie C-Atom-Markierungssysteme unabhängig vom Anfangszustand des Systems stets demselben stabilen stationären Zustand zustreben.

7.2 Dissipative Systeme

Im Gegensatz zu den linearen C-Atom-Bilanzgleichungen, entziehen sich die durch Gleichung (4.30) beschriebenen Isotopomeren-Bilanzgleichungssysteme

$$0 = \mathbf{P}^{\text{inp}} \odot \mathbf{v}^T \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v}^T \cdot \mathbf{x} + \mathbf{x}^T \cdot \mathbf{Q} \odot \mathbf{v} \cdot \mathbf{x}$$

— wie die meisten nichtlinearen Systeme [JS87] — einer allgemeinen Analyse. Es kann jedoch ein Instrumentarium bereitgestellt werden, daß bei konkret vorgegebenen Systemen in vielen Fällen zum Ziel führt. Dabei wird die eindeutige Lösbarkeit der Gleichungen auf das allgemeinere Problem der globalen Stabilität des entsprechenden Differentialgleichungssystems (analog zu Gleichung (7.1)) zurückgeführt.

Das wichtigste Werkzeug der hier vorgestellten Methode zur globalen Stabilitätsanalyse ist der Nachweis der sogenannten Dissipativität. Diese ist für lineare kompartmentelle Systeme stets gegeben und kann im nichtlinearen Fall, wie noch gezeigt wird, auf die Dissipativität der Linearisierung dieses Systems zurückgeführt werden. Die Dissipativität spielt bei der Analyse elektrischer Netzwerke eine entscheidende Rolle. Die Transistoren entsprechen hier den bidirektionalen Reaktionsschritten. In [MN91] werden mit großem technischen Aufwand globale Stabilitätskriterien für eine sehr weit gefaßte Klasse elektrischer Netzwerke entwickelt. Die nachfolgenden Ausführungen erweitern die dort entwickelte Theorie geeignet auf Isotopomeren-Systeme, wozu einige Hilfskonstruktionen benötigt werden.

7.2.1 Dissipative Systeme

Nichtlineare Systeme werden (im Gegensatz zu linearen Systemen) stets eingeschränkt auf einen bestimmten physikalisch sinnvollen Bereich $G \subset \mathbb{R}^n$ untersucht. Das Gebiet G kann z.B. ein Teil einer Hyperebene sein. Dieser Bereich muß so gewählt sein, daß eine Lösung des Differentialgleichungssystems mit Startpunkt in G diesen Bereich nicht verläßt (vgl. [HA83]).

Definition 7.4 (positiv invariante Menge) Sei $\dot{\mathbf{x}} = f(\mathbf{x})$ ein Differentialgleichungssystem und $G \subset \mathbb{R}^{\dim \mathbf{x}}$. Dann heißt G positiv invariant bzgl. f , wenn für jede Lösung $\mathbf{x}(t)$ gilt:

$$\mathbf{x}(0) \in G \implies \forall_{t \geq 0} \mathbf{x}(t) \in G$$

Der Begriff der Dissipativität hängt — wie sich nachfolgend zeigen wird — eng mit der Kompartimentalität von Matrizen zusammen.

Definition 7.5 (dissipative Abbildung, dissipatives System) Gegeben sei ein Differentialgleichungssystem $\dot{\mathbf{x}} = f(\mathbf{x})$ mit $\mathbf{x} \in \mathbb{R}^n$ und einer stetig differenzierbaren Funktion f sowie eine positiv invariante Menge $G \subset \mathbb{R}^n$. Dann heißt die Abbildung f bzw. das dynamische System dissipativ auf G , wenn gilt:

$$\forall_{\mathbf{x}, \mathbf{y} \in G} (\mathbf{x} - \mathbf{y})^T (f(\mathbf{x}) - f(\mathbf{y})) \leq 0$$

Der nachfolgende Satz aus [DHB94] zeigt, daß ein dissipatives System stets nichtexpansiv ist, d.h. daß sich je zwei Lösungen $\mathbf{x}(t)$ und $\mathbf{y}(t)$ des Differentialgleichungssystems nicht voneinander entfernen.

Satz 7.7 (Nichtexpansivität dissipativer Systeme) Sei $\dot{\mathbf{x}} = f(\mathbf{x})$ ein dissipatives System auf $G \subset \mathbb{R}^n$. Dann ist der zugehörige Phasenfluß nichtexpansiv auf G , d.h.

$$\forall_{\mathbf{x}, \mathbf{y} \in G} \frac{d}{dt} [\|\mathbf{x} - \mathbf{y}\|^2] \leq 0 \quad .$$

Insbesondere gilt für je zwei Lösungen $\mathbf{x}(t), \mathbf{y}(t)$ der Gleichung mit $\mathbf{x}(0), \mathbf{y}(0) \in G$, daß

$$\forall_{t \geq 0} \forall_{\mathbf{x}(0), \mathbf{y}(0) \in G} \|\mathbf{x}(t) - \mathbf{y}(t)\| \leq \|\mathbf{x}(0) - \mathbf{y}(0)\| \quad .$$

Beweis:

$$\begin{aligned} \frac{d}{dt} \|\mathbf{x} - \mathbf{y}\|^2 &= \frac{d}{dt} [(\mathbf{x} - \mathbf{y})^T \cdot (\mathbf{x} - \mathbf{y})] \\ &= 2(\mathbf{x} - \mathbf{y})^T \cdot (\dot{\mathbf{x}} - \dot{\mathbf{y}}) \\ &= 2(\mathbf{x} - \mathbf{y})^T \cdot (f(\mathbf{x}) - f(\mathbf{y})) \leq 0 \end{aligned}$$

Die zweite Beziehung folgt unmittelbar durch Integration über die Zeit.

□

Im Gegensatz zu [MN91] kann für Isotopomeren-Systeme nicht auf einfache Weise die *strenge Dissipativität*, bei der oben " $<$ " statt " $\leq$ " gilt, nachgewiesen werden. Streng dissipative Systeme sind mit demselben Beweis offenbar streng kontraktiv (d.h. in Satz 7.7 gilt ebenfalls " $<$ " statt " $\leq$ "), woraus sehr einfach die globale Stabilität folgt. Da dieser Sachverhalt jedoch bei Isotopomeren-Systemen im allgemeinen nicht gegeben ist, sind nachfolgend noch weitere Hilfskonstruktionen erforderlich.

7.2.2 Dissipativität lokal dissipativer Systeme

Für ein lineares System gilt in Anwendung der Definition 7.5 offenbar der

Satz 7.8 (Kriterium für die Dissipativität linearer Systeme) Ein lineares System $\dot{\mathbf{x}} = \mathbf{A} \cdot \mathbf{x} + \mathbf{b}$ ist dissipativ auf $G = \mathbb{R}^{\dim \mathbf{x}}$, wenn gilt:

$$\forall_{\mathbf{x}} \mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0$$

Dieser Satz motiviert die

Definition 7.6 (dissipative Matrix) Eine quadratische Matrix $\mathbf{M}$ heißt dissipativ, wenn

$$\forall_{\mathbf{x} \neq 0} \mathbf{x}^T \mathbf{M} \mathbf{x} \leq 0 \quad .$$

Der nachfolgende Satz aus [MN91] zeigt nun, daß Dissipativität eine lokale Eigenschaft ist, d.h. sie ist global erfüllt, wenn sie an jedem Punkt $\mathbf{x} \in G$ für die Linearisierung eines Systems gilt. Jedoch gilt dies nur für konvexe Gebiete G , die, wie sich noch zeigen wird, im Falle der Isotopomeren-Systeme stets angetroffen werden:

Satz 7.9 (Dissipativität als lokale Eigenschaft) Ein nichtlineares System $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ ist global dissipativ auf einem konvexen Gebiet G , wenn gilt

$$\forall \mathbf{x} \in G \quad \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}) \text{ ist eine dissipative Matrix} \quad .$$

Beweis:

Seien zwei Punkte $\mathbf{x}, \mathbf{y} \in G$ gegeben. Dann liegt aufgrund der Konvexität von G auch die Strecke $\mathbf{x} + t \cdot (\mathbf{y} - \mathbf{x})$, $t \in [0, 1]$ in G , und es gilt mit $g(t) = \mathbf{f}(\mathbf{x} + t \cdot (\mathbf{y} - \mathbf{x}))$ die Beziehung $g(1) - g(0) = \int_0^1 g'(t) dt$, also

$$\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = \int_0^1 \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x} + t \cdot (\mathbf{y} - \mathbf{x})) \right] \cdot (\mathbf{y} - \mathbf{x}) dt \quad .$$

Die Dissipativitätsbedingung ergibt sich dann mit Definition 7.5 aus:

$$(\mathbf{y} - \mathbf{x})^T \cdot (\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})) = \int_0^1 (\mathbf{y} - \mathbf{x})^T \cdot \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x} + t \cdot (\mathbf{y} - \mathbf{x})) \right] \cdot (\mathbf{y} - \mathbf{x}) dt \leq 0$$

□

7.2.3 Dissipativität lokal kompartmenteller Systeme

Ist $\mathbf{M}$ eine symmetrische Matrix, so ist die Dissipativität offenbar gleichbedeutend mit der Negativ-Semidefinitheit. Daher kann Dissipativität bei einer konkret gegebenen Matrix mit Hilfe des Hauptdeterminantenkriteriums für die Negativ-Semidefinitheit von $\frac{1}{2}(\mathbf{M} + \mathbf{M}^T)$ nachgewiesen werden [HJ85]. Bei parameterabhängigen Matrizen ist dieses Kriterium jedoch im allgemeinen nicht anwendbar, da das Vorzeichen der entstehenden parameterabhängigen Determinanten nicht effektiv angegeben werden kann.

Im Spezialfall der C-Atom-Systeme läßt sich die Dissipativität jedoch aus der Kompartimentalität der Systemmatrix folgern. Im nachfolgenden Satz aus [Cop65] wird zu Abschätzungszwecken die 1-Norm

$$\|\mathbf{x}\|_1 = \sum_i |\mathbf{x}_i| \quad (7.2)$$

eines Vektors benötigt:

Satz 7.10 (Dissipativität kompartmenteller Matrizen)

1. Für jede quadratische Matrix $\mathbf{M}$ gilt die Abschätzung⁴

$$\mathbf{x}^T \mathbf{M} \mathbf{x} \leq \|\mathbf{x}\|_1 \cdot \max_j (\mathbf{M}_{jj} + \sum_{i \neq j} |\mathbf{M}_{ij}|) \cdot \|\mathbf{x}\|_1 \quad .$$

2. Jede kompartmentelle Matrix ist dissipativ.

Beweis:

⁴Die Abschätzung ist offenbar stärker als die durch einfache Normabschätzung mit Hilfe der 1-Norm von $\mathbf{M}$ erhaltene Beziehung $\mathbf{x}^T \mathbf{M} \mathbf{x} \leq \|\mathbf{x}\|_1 \cdot \|\mathbf{M}\|_1 \cdot \|\mathbf{x}\|_1$.

Punkt 1: Es gilt:⁵

$$\begin{aligned}
 \mathbf{x}^T \mathbf{M} \mathbf{x} &= \sum_j \mathbf{x}_j \cdot \sum_i \mathbf{M}_{ij} \mathbf{x}_i \\
 &\leq \sum_j |\mathbf{x}_j| \cdot \left[\mathbf{M}_{jj} |\mathbf{x}_j| + \sum_{i \neq j} |\mathbf{M}_{ij}| |\mathbf{x}_i| \right] \\
 &\leq \sum_j |\mathbf{x}_j| \cdot \left[\mathbf{M}_{jj} + \sum_{i \neq j} |\mathbf{M}_{ij}| \right] \cdot \|\mathbf{x}\|_1 \\
 &\leq \sum_j |\mathbf{x}_j| \cdot \max_j \left[\mathbf{M}_{jj} + \sum_{i \neq j} |\mathbf{M}_{ij}| \right] \cdot \|\mathbf{x}\|_1 \\
 &= \|\mathbf{x}\|_1 \cdot \max_j \left[\mathbf{M}_{jj} + \sum_{i \neq j} |\mathbf{M}_{ij}| \right] \cdot \|\mathbf{x}\|_1
 \end{aligned}$$

Punkt 2 ist eine unmittelbare Anwendung der Definition 7.1 kompartmenteller Matrizen.

□

In Analogie zum Prinzip der lokalen Dissipativität in Satz 7.9 kann nun die lokale Kompartimentalität eines nichtlinearen Systems ausgenutzt werden:

Satz 7.11 (Dissipativität lokal kompartmenteller Systeme) *Ein nichtlineares System $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ ist global dissipativ auf einem konvexen Gebiet G , wenn gilt*

$$\forall \mathbf{x} \in G \quad \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}) \text{ ist eine kompartmentelle Matrix} \quad .$$

Beweis:

Da lokal kompartmentelle Systeme nach Satz 7.10 auch lokal dissipativ sind, folgt nach Satz 7.9 die globale Dissipativität. □

Leider läßt sich aus der Kompartimentalität nicht die strenge Dissipativität folgern und damit die globale Stabilität (vgl. die Anmerkung zu Satz 7.7). Dazu müßten *alle* Spaltensummen der Matrix negativ sein, was in [MN91] vorausgesetzt wird, aber für Markierungssysteme unrealistisch ist. Der Grund liegt darin daß nicht jeder Pool einen externen Abfluß hat, so daß die entsprechende Spaltensumme häufig exakt gleich Null ist (vgl. den Beweis von Satz 7.1).

7.2.4 Ein globales Stabilitätskriterium für dissipative Systeme

Das nachfolgende Stabilitätskriterium wird hier im Gegensatz zu [MN91] zusätzlich benötigt und liefert den Schlüssel für einen globalen Stabilitätsbeweis bei Isotopomeren-Systemen:

⁵Entscheidend ist die Gültigkeit der trivialen Beziehung $\mathbf{x}_j \cdot [\mathbf{M}_{jj} \cdot \mathbf{x}_j] = |\mathbf{x}_j| \cdot [\mathbf{M}_{jj} \cdot |\mathbf{x}_j|]$. Daher können auf der rechten Seite der ersten Abschätzung die Betragstriche bei $\mathbf{M}_{jj}$ weggelassen werden.

Dann gilt unter Anwendung der Dreiecksungleichung und Ausnutzung der Dissipativität:

$$\begin{aligned} \|y(T)\| - \|y(0)\| &\leq \|y(T) - x(T)\| + \|x(T)\| - \|y(0)\| \\ &\leq \|y(0) - x(0)\| + \|x(0)\| - \|y(0)\| - \varepsilon/2 = -\varepsilon/2 \end{aligned}$$

□

7.3 Stabilitätsanalyse von Isotopomeren-Markierungssystemen

In diesem Abschnitt sei

$$\dot{x} = f(x, x^{\text{inp}}) \stackrel{\text{def}}{=} P^{\text{inp}} \odot v^T \cdot x^{\text{inp}} + P \odot v^T \cdot x + x^T \cdot Q \odot v \cdot x$$

ein Isotopomeren-System gemäß Definition 4.10. Die vorangehenden Sätze legen eine Methodik für den Beweis der globalen Stabilität eines solchen Systems nahe:

1. Es ist eine positiv invarianter konvexe Menge $G \subset \mathbb{R}^n$ zu finden, auf dem eine Stabilitätsanalyse sinnvoll ist. Dieser wird nachfolgend als die zulässige Markierungsmenge bezeichnet.
2. Auf diesem Bereich ist überall die lokale Kompartimentalität des Systems nachzuweisen.
3. Es muß die Existenz eines stationären Punktes gezeigt werden. Dies ist mit Hilfe eines Standardarguments der Theorie dynamischer Systeme leicht möglich.
4. Schließlich muß die Linearisierung des Systems in jedem Punkt des Phasenraums (und damit auch im stationären Punkt) fallenfrei und damit auch invertierbar sein.

7.3.1 Die zulässige Markierungsmenge eines Isotopomeren-Systems

In Ausführung des oben vorgegebenen Programms soll zunächst ein sinnvoller Bereich G definiert werden, auf dem ein Isotopomeren-System $\dot{x} = f(x, x^{\text{inp}})$ untersucht wird. Das Interesse gilt vor allem der eindeutigen und physikalisch sinnvollen Lösbarkeit des entsprechenden stationären Gleichungssystems.

Bei der Suche nach einer geeigneten Menge G hilft Satz 4.1, der die Beziehung zwischen der Lösung eines Isotopomeren-Bilanzgleichungssystems und dem entsprechenden C-Atom-Bilanzgleichungssystem charakterisiert. Naheliegenderweise gilt eine völlig analoge Beziehung zwischen den Lösungen der entsprechenden Differentialgleichungssysteme, d.h. aus einer Lösung $x^I(t)$ der Isotopomeren-Differentialgleichungen (4.34) läßt sich unter einfacher Anwendung der Zugehörigkeitsmatrix Z_A^I eine Lösung $x^A(t)$ der C-Atom-Differentialgleichungen (4.33) gewinnen. Folgende Einschränkungen an das Isotopomeren-System sind daher naheliegend:

1. Isotopomeren-Markierungsanteile haben Werte zwischen Null und Eins.
2. Die Summe aller Markierungsanteile eines Metaboliten im Markierungs-Input $x^{\text{inp}, I}$ muß Eins betragen. Die analoge Aussage muß für den Startwert $x^I(0)$ gelten und sodann für den zeitabhängigen Systemzustand $x^I(t)$ bewiesen werden.

3. Da die Lösung $\mathbf{x}^A(t)$ des C-Atom-Systems stets auf einen eindeutigen stabilen stationären Punkt $\mathbf{x}^{A,*}$ zuläuft, kann die Suche nach stationären Lösungen $\mathbf{x}^{I,*}$ für das Isotopomerensystem nach Satz 4.1 auf das Urbild von $\mathbf{x}^{A,*}$ unter $\mathbf{Z}_A^I$ eingeschränkt werden.

Diese Beziehungen werden nun unter Anwendung der Isotopomeren-Metabolit-Zugehörigkeitsmatrizen $\mathbf{Z}_M^I, \mathbf{Z}_M^{I,\text{inp}}$ und der Isotopomeren-Atom-Zugehörigkeitsmatrizen $\mathbf{Z}_A^I, \mathbf{Z}_A^{I,\text{inp}}$ aus Definition 4.11 wie folgt formalisiert. Alle Punkte des nachfolgenden Satzes sind aufgrund der obigen Ausführungen intuitiv klar. Auf den rein formalen Aufwand eines Beweises wird jedoch hier verzichtet.

Satz 7.13 (Beziehung zwischen C-Atom- und Isotopomeren-Bilanzgleichungen) Sei $\mathbf{x}^I(t)$ eine Lösung des Isotopomeren-Systems $\dot{\mathbf{x}}^I = f(\mathbf{x}^I, \mathbf{x}^{I,\text{inp}})$, und der Markierungs-Input $\mathbf{x}^{I,\text{inp}}$ erfülle die Bedingung (vgl. Gleichung (4.31))

$$\mathbf{Z}_M^{I,\text{inp}} \mathbf{x}^{I,\text{inp}} = \mathbf{1} \quad .$$

Ferner sei das zugehörige C-Atom-System fallenfrei und $\mathbf{x}^{A,*}$ seine nach Satz 7.3 eindeutige stabile stationäre Lösung. Dann gilt:

1. $\mathbf{Z}_A^I \mathbf{x}^I(t)$ ist eine Lösung des C-Atom-Bilanzgleichungssystems.
2. Für jede stationäre Lösung $\mathbf{x}^{I,*}$ von $f(\mathbf{x}, \mathbf{x}^{\text{inp}})$ gilt $\mathbf{Z}_A^I \mathbf{x}^{I,*} = \mathbf{x}^{A,*}$.
3. $G_1 \stackrel{\text{def}}{=} [0, 1]$ ist positiv invariant bzgl. f .
4. $G_2 \stackrel{\text{def}}{=} \{ \mathbf{x}^I \mid \mathbf{Z}_M^I \mathbf{x}^I = \mathbf{1} \}$ ist positiv invariant bzgl. f .
5. $G_3 \stackrel{\text{def}}{=} \{ \mathbf{x}^I \mid \mathbf{Z}_A^I \mathbf{x}^I = \mathbf{x}^{A,*} \}$ ist positiv invariant bzgl. f .

Satz 7.13 schränkt die Menge der möglichen Lösungen für die Isotopomeren-Bilanzgleichungen erheblich ein. Er motiviert die Definition:

Definition 7.7 (zulässige Lösung und zulässiger Input eines Isotopomeren-Systems) Sei $\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{x}^{\text{inp}})$, ein Isotopomeren-System und G_1, G_2, G_3 festgelegt wie in Satz 7.13. Dann heißt $\mathbf{x}^{\text{inp}}$ ein zulässiger Input, wenn

$$\mathbf{Z}_A^{I,\text{inp}} \cdot \mathbf{x}^{\text{inp}} = \mathbf{1}$$

und $\mathbf{x}(t)$ heißt eine zulässige Lösung, wenn

$$\mathbf{x}(0) \in G \stackrel{\text{def}}{=} G_1 \cap G_2 \cap G_3 \quad .$$

Aufgrund von Satz 7.13 ist G positiv invariant bzgl. des Isotopomeren-Systems. Außerdem nähert sich jede Lösung von $\mathbf{x}^I(t) \in G_1 \cap G_2$ stets asymptotisch an G_3 an. Stabile stationäre Punkte sind also sinnvollerweise nur in G zu suchen.

7.3.2 Ein Stabilitätskriterium für Isotopomeren-Systeme

Der Nachweis der Existenz eines stationären Punktes gelingt stets mit Hilfe eines Standardlemmas aus der Theorie dynamischer Systeme, das in [HA83] mit Hilfe des Brouwerschen Fixpunktsatzes bewiesen wird. Dazu ist lediglich festzustellen, daß die Menge G aus Definition 7.7 ein *Simplex* ist, d.h. eine durch lineare Gleichungen und Ungleichungen festgelegte kompakte und konvexe Menge.

Satz 7.14 (Existenz eines stationären Punktes) Sei $\dot{x} = f(x)$ ein nichtlineares Differentialgleichungssystem auf einer durch lineare Ungleichungen definierten Menge G . Ist G positiv invariant unter f , so existiert eine stationäre Lösung $x^* \in G$.

Nun kann ein Kriterium für die globale Stabilität bewiesen werden:

Satz 7.15 (Stabilitätskriterium für Isotopomeren-Systeme) Sei $\dot{x} = f(x, x^{\text{inp}})$ ein Isotopomeren-System mit der Menge G zulässiger Lösungen und einem zulässigen Input x^{inp} . Ferner gelte für alle $x \in G$:

$$\frac{\partial f}{\partial x}(x) \text{ ist eine fallenfreie kompartmentelle Matrix}$$

Dann ist das System global stabil und besitzt insbesondere eine eindeutige stationäre Lösung.

Beweis:

Nach Satz 7.14 existiert ein stationärer Punkt. Die Jacobi-Matrix in diesem Punkt ist nach Voraussetzung und Satz 7.6 eine stabile Matrix. Aus der Linearisierungstheorie für nichtlineare dynamische Systeme [Per93] folgt dann, daß der stationäre Punkt lokal stabil ist. Da aufgrund der lokalen Kompartimentalität auf G auch die lokale Dissipativität auf G folgt (Satz 7.11) und daraus nach Satz 7.9 die globale Dissipativität, ist nun Satz 7.12 anwendbar.

□

7.4 Beispiele für Stabilitätsanalysen

In der Praxis liefert Satz 7.15 zwar ein Rezept für den Stabilitätsbeweis, jedoch ist im einzelnen erst einige Vorarbeit zu leisten, bis er anwendbar ist. Auf die formale Darstellung der weiteren technischen Details soll hier verzichtet werden. Stattdessen soll die Stabilitätsanalyse an einigen Beispielen vorgeführt werden.

7.4.1 Eine Rechenhilfe

Das bereits in den Kapiteln 3-6 diskutierte Beispielsystem aus Abbildung 3.2 — dessen eindeutige Lösbarkeit aufgrund der Gleichungen 4.13 bereits nachgewiesen ist — soll als erstes mit den

oben beschriebenen Methoden behandelt werden. Linearisierung der Isotopomeren-Bilanzgleichungen (4.7) liefert die Matrix:

$$\begin{pmatrix} -A & 0 & 0 & 0 & v_2^{\leftarrow} & 0 & 0 & 0 \\ 0 & -A & 0 & 0 & 0 & v_2^{\leftarrow} & 0 & 0 \\ 0 & 0 & -A & 0 & 0 & 0 & v_2^{\leftarrow} & 0 \\ 0 & 0 & 0 & -A & 0 & 0 & 0 & v_2^{\leftarrow} \\ \alpha + v_2^{\rightarrow} & 0 & \alpha & 0 & \gamma - B & \gamma & 0 & 0 \\ \beta & v_2^{\rightarrow} & \beta & 0 & 0 & -B & \gamma & \gamma \\ 0 & \alpha & v_2^{\rightarrow} & \alpha & \delta & \delta & -B & 0 \\ 0 & \beta & 0 & \beta + v_2^{\rightarrow} & 0 & 0 & \delta & \delta - B \end{pmatrix}$$

$$\begin{aligned} A &= v_1^{\rightarrow} + v_2^{\leftarrow} \\ B &= v_2^{\leftarrow} + v_1^{\rightarrow} - v_5^{\rightarrow} \\ \alpha &= (e_{00} + e_{01}) \cdot (v_1^{\rightarrow} + v_2^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \\ \beta &= (e_{10} + e_{11}) \cdot (v_1^{\rightarrow} + v_2^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \\ \gamma &= (b_{00} + b_{10}) \cdot (v_1^{\rightarrow} + v_2^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \\ \delta &= (b_{01} + b_{11}) \cdot (v_1^{\rightarrow} + v_2^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \end{aligned}$$

mit

Da offenbar die Diagonaleinträge stets negativ und die übrigen Einträge positiv sind (was allgemein gezeigt werden kann), kann die Untersuchung auf die Überprüfung der Spaltensummen und der Zeilenfreiheit beschränkt werden. Letztere prüft man leicht nach. Aufgrund von (4.3) ist ferner die Spaltensumme der ersten vier Spalten jeweils $-v_5^{\rightarrow} \leq 0$, die der letzten vier $v_2^{\leftarrow} - v_2^{\rightarrow} \leq 0$ (vgl. Gleichung (4.20)). Folglich ist dieses System global stabil auf der zulässigen Menge G .

An diesem Beispiel kann eine hilfreiche Rechenregel für die Untersuchung der Linearisierung von Isotopomeren-Systemen abgelesen werden, die in vielen Fällen die explizite Konstruktion der Jacobi-Matrix erspart. Offenbar ergänzen sich in jeder Spalte die Isotopomeren-Anteile von B bzw. E zu eins. Dies ist eine leicht zu beweisende Gesetzmäßigkeit, denn bei partieller Ableitung nach b_i verbleiben von allen Produkten $b_i \cdot e_j$ gerade die zweiten Faktoren e_j , die alle in einer Spalte versammelt sind und sich zu Eins addieren. Diese Beziehung gilt allgemein, da ein Isotopomer in einer bimolekularen Reaktion stets mit jeder Markierungsform seines Partners zusammenkommt. Mit einer analogen Überlegung sieht man, daß die Spaltensumme der B betreffenden ersten vier Spalten und der E betreffenden letzten vier Spalten jeweils gleich ist. Auf den formalen Beweis dieser allgemein gültigen Beziehungen wird hier verzichtet.

Wenn also nur zwei der Spaltensummen interessieren und diese zudem keine Markierungsvariablen mehr enthalten, so kann nach einem einfacheren alternativen Weg gesucht werden, um die Summen zu berechnen. Man überlegt sich, daß die nachfolgende formale Konstruktion zum Ziel führt. Sie beruht darauf, daß die sich zu Eins addierenden Terme bereits vor der Ableitungsbildung zusammengefaßt werden. Zunächst werden dazu die Variablen

$$\begin{aligned} a &= a_{00} + a_{01} + a_{10} + a_{11} \\ b &= b_{00} + b_{01} + b_{10} + b_{11} \\ e &= e_{00} + e_{01} + e_{10} + e_{11} \end{aligned}$$

für die Isotopomeren-Summen eingeführt und dann alle Bilanzen zu je einem Metaboliten aus Gleichung (4.7) addiert in der Form (vgl. hierzu den vollkommen analogen Satz 4.1):

$$\begin{aligned} \text{B: } v_1^{\rightarrow} \cdot a + v_2^{\leftarrow} \cdot e - (v_2^{\rightarrow} + v_3^{\rightarrow} + v_5^{\rightarrow}) \cdot b &= 0 \\ \text{E: } v_2^{\rightarrow} \cdot b + v_3^{\leftarrow} \cdot b \cdot e - (v_2^{\leftarrow} + v_3^{\rightarrow} + v_4^{\rightarrow}) \cdot e &= 0 \end{aligned}$$

Offenbar sind dies für $a = b = e = 1$ gerade die stöchiometrischen Gleichungen. Die Jacobi-Matrix bzgl. b und e an der Stelle $b = 1, e = 1$ ist nun aufgrund von Gleichung (4.11) gegeben durch:

$$\begin{pmatrix} -v_1^{\rightarrow} - v_2^{\leftarrow} & v_2^{\leftarrow} \\ v_1^{\rightarrow} + v_2^{\leftarrow} - v_5^{\rightarrow} & -v_2^{\rightarrow} \end{pmatrix}$$

Die Spaltensummen dieser Matrix sind gerade die Spaltensummen der großen Jacobi-Matrix.

7.4.2 Isotopomeren-Stabilität im Citratzyklus

Die beschriebene Konstruktion zum Nachweis der Stabilität von Isotopomeren-Systemen zeigt, daß diese im wesentlichen durch die Struktur des Metabolit-Netzwerks bestimmt wird. Lediglich die Fallenfreiheit der Jacobi-Matrix beruht auf den Isotopomeren-Transitionen. Sie ist bei nicht-pathologischen Systemen stets gegeben. Dies macht die Anwendung der Theorie verhältnismäßig einfach.

Als ein komplexeres Beispiel wurde in Kapitel 3 wurde Netzwerk aus Abbildung 3.2 durch Vereinfachung des Netzwerks aus Abbildung 3.1 gewonnen. Man überlegt sich leicht, daß auch das ursprüngliche Isotopomeren-System stabil ist, da beide Systeme durch eine lineare Umformung ineinander überführbar sind, die der durchgeführten Netzwerkvereinfachung entspricht.

Betrachtet man nur das Metabolit-Netzwerk des Systems aus Abbildung 3.1, so hat es die gleiche Struktur wie der Citratzyklus mit einem bidirektionalen anaplerotischen Reaktionsschritt (vgl. Abb. 1.2) (dabei ist der Glyoxylatzyklus zunächst ausgenommen). Dann liefert aber die oben beschriebene Rechenhilfe dasselbe Ergebnis unabhängig davon, wieviele C-Atome die einzelnen Stoffe haben. Daraus folgt die Stabilität des Isotopomeren-Systems des Citratzyklus sofern die Linearisierung dieses Systems fallenfrei ist. Dies läßt sich jedoch leicht nachweisen.

7.4.3 Eine Transformationsmethode

Das Stabilitätskriterium 7.15 in Kombination mit der oben beschriebenen Rechenhilfe kann schon bei einfachen Netzwerken scheitern. Dies zeigt das elementare Beispiel eines Systems mit zwei Input-Metaboliten A und B, die (via C und D) zu dem Intermediat E zusammengefügt werden, welches sodann wieder zu F und G zerfällt.

$$\begin{array}{llll} v1: & A & > & C \\ & \#1 & > & \#1 \\ v2: & B & > & D \\ & \#1 & > & \#1 \\ v3: & C + D & > & E \\ & \#1 + \#2 & > & \#12 \\ v4: & E & > & F + G \\ & \#12 & > & \#1 + \#2 \\ v5: & F & > & H \\ & \#1 & > & \#1 \\ v6: & G & > & K \\ & \#1 & > & \#1 \end{array}$$

das bei unidirektional angenommenen Schritten im wesentlichen eine lineare Sequenz darstellt und daher trivialerweise stabil ist. Es gelte der Einfachheit halber $v_1 = v_2 = v_3 = v_4 = 1$, so daß für die Anwendung der Rechenhilfe die nachfolgenden Gleichungen zu linearisieren sind:

$$\begin{array}{lll} C: & 0 & = a - c \\ D: & 0 & = b - d \\ E: & 0 & = cd - e \\ F: & 0 & = e - f \\ G: & 0 & = e - g \end{array}$$

Die Jacobi-Matrix in $c = d = e = 1$ ist

$$\begin{pmatrix} -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot & \cdot \\ 1 & 1 & -1 & \cdot & \cdot \\ \cdot & \cdot & 1 & -1 & \cdot \\ \cdot & \cdot & 1 & \cdot & -1 \end{pmatrix}.$$

Offenbar ist diese Matrix nicht kompartmentell, denn die dritte Spalte hat eine positive Summe. Hier hilft jedoch eine lineare Variablentransformation $\bar{x} = D \cdot x$ mit der diagonalen Matrix

$$D = \text{diag} \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 1, 1 \right).$$

Das Resultat ist die kompartmentelle Matrix:

$$\begin{pmatrix} -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot & \cdot \\ 1 & 1 & -1 & \cdot & \cdot \\ \cdot & \cdot & \frac{1}{3} & -1 & \cdot \\ \cdot & \cdot & \frac{1}{3} & \cdot & -1 \end{pmatrix}$$

Eine analoge Transformation (nämlich $(Z_M^I)^T \cdot D$) hat die entsprechende Wirkung auf die Jacobi-Matrix des Isotopomeren-Systems. Allerdings verändert eine solche Transformation auch das zulässige Gebiet G . Jedoch müssen alle Sätze völlig analog für das transformierte System mit dem transformierten Gebiet gelten, was hier nicht mehr ausgeführt werden soll. Die globale Stabilität des Beispiels ist also erwiesen.

Kapitel 8

Numerische Lösung von Markierungs-Bilanzgleichungen

Dieses Kapitel beschäftigt sich mit den numerischen Verfahren zur praktischen Lösung von Markierungs-Bilanzgleichungssystemen. Die im vorangehenden Kapitel erzielten Eindeutigkeitsaussagen über die prinzipielle Lösbarkeit solcher Systeme haben für deren numerische Behandlung die unmittelbare Konsequenz, daß eine Lösung der stationären Bilanzgleichungssysteme auch durch Lösung des entsprechenden instationären Differentialgleichungssystems bestimmt werden kann. Anschaulich gesprochen, wird dabei das Experiment im Rechner nachvollzogen und der stationäre Zustand abgewartet. Allerdings ist diese Methode wenig effizient, wenn nur der stationäre Zustand des Systems interessiert und nicht das Übergangsverhalten. Es wurde gezeigt, daß die C-Atom-Bilanzen viel einfacher durch Lösung eines linearen Gleichungssystems berechnet werden können. Dagegen ist dieses Verfahren bei den nichtlinearen Isotopomeren-Systemen nicht mehr anwendbar, d.h. hier muß ein iteratives Verfahren eingesetzt werden, das im Prinzip einem Differentialgleichungslöser verwandt ist.

Nachdem im Abschnitt 8.1 zunächst die erforderlichen Rechenschritte zur Berechnung der stationären bzw. instationären Lösungen beschrieben werden, befaßt sich Abschnitt 8.2 mit dem Problem der numerischen Stabilität von Lösungsmethoden für stationäre C-Atom-Bilanzgleichungen. Dabei wird sich herausstellen, daß die üblichen numerischen Verfahren zur Lösung von linearen Gleichungssystemen in Schwierigkeiten geraten, wenn große Austauschflüsse im System vorliegen. In diesem Falle müssen geeignete Maßnahmen getroffen werden, um die numerische Stabilität zu gewährleisten. Zum Verständnis der in diesem Kapitel vorgestellten Methoden werden Grundkenntnisse der numerischen Analysis vorausgesetzt [SB73].

8.1 Numerische Lösung der Bilanzgleichungssysteme

Die numerische Lösung der stationären bzw. instationären Bilanzgleichungssysteme ist der rechnerische Nachvollzug des praktischen Experimentes, d.h. eine Simulation. Das Ziel der Simulation ist die Berechnung des Markierungszustands x in Abhängigkeit von den Flüssen v . Letztere müssen zudem die in Kapitel 6 vorgestellten Nebenbedingungen, also die stöchiometrischen Gleichungen und die vom Anwender getroffenen Annahmen erfüllen. Als Simulationseingabe genügt die Vorgabe von Werten für die freien Flüsse, also des Wertevektors n^{free} , alle anderen Strukturen sind dann durch die Formulierung der Bilanzgleichungen und Nebenbedingungen eindeutig festgelegt.

8.1.1 Rechenschritte einer Simulation

Ausgegangen wird von den im Abschnitt 6.3 zusammengestellten Strukturen. Der rechnerische Ablauf einer Simulation ergibt sich dann wie folgt:

1. Aufgrund der vorausgesetzten Invertierbarkeit der Nebenbedingungsmatrix N kann man schreiben

$$N^{-1} = \begin{pmatrix} K^{\text{net}} & K^{\text{xch}[0,1]} & K^{\text{free}} \end{pmatrix} \quad (8.1)$$

mit $\text{coldim } K^{\text{net}} = \dim n^{\text{net}}$, $\text{coldim } K^{\text{xch}[0,1]} = \dim n^{\text{xch}[0,1]}$ und $\text{coldim } K^{\text{free}} = \dim n^{\text{free}}$. Aus Gleichung (6.9) folgt dann

$$\begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} = N^{-1} \begin{pmatrix} n^{\text{net}} \\ n^{\text{xch}[0,1]} \\ n^{\text{free}} \end{pmatrix} = K^{\text{net}} \cdot n^{\text{net}} + K^{\text{xch}[0,1]} \cdot n^{\text{xch}[0,1]} + K^{\text{free}} \cdot n^{\text{free}}.$$

Mit der Setzung $k^{\text{free}} \stackrel{\text{def}}{=} K^{\text{net}} \cdot n^{\text{net}} + K^{\text{xch}[0,1]} \cdot n^{\text{xch}[0,1]}$ folgt, daß die kompaktifizierten Flußkoordinaten eine affin lineare Funktion Ψ von n^{free} sind:

$$\begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} = \Psi(n^{\text{free}}) \stackrel{\text{def}}{=} K^{\text{free}} \cdot n^{\text{free}} + k^{\text{free}} \quad (8.2)$$

2. Es ist durch Einsetzen zu prüfen, ob die berechneten Flüsse zulässig sind, d.h. ob gilt

$$U \cdot \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} \geq u.$$

Ist U entsprechend gewählt, so folgt nach den Ausführungen in Abschnitt 7.1 auch die Fallenfreiheit des Systems.

3. Die Koordinatentransformationen $\Phi^{[0,1]}$ aus Satz 5.6 überführt nun die kompaktifizierten Koordinaten in die Anwendungs-Koordinaten v^{net} und v^{xch} :

$$\begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}} \end{pmatrix} = \Phi^{[0,1]} \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} = \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} / (1 - v^{\text{xch}[0,1]}) \end{pmatrix}$$

4. Daraus werden durch Anwendung der Transformation Φ_α aus Satz 5.2 die natürlichen Koordinaten $v^\rightarrow$ und $v^\leftarrow$ und damit der Gesamt-Flußvektor v :

$$v = \begin{pmatrix} v^\rightarrow \\ v^\leftarrow \end{pmatrix} = \Phi_\alpha \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}} \end{pmatrix} = \begin{pmatrix} v^{\text{xch}} + \max_\alpha(v^{\text{net}}, 0) \\ v^{\text{xch}} + \max_\alpha(-v^{\text{net}}, 0) \end{pmatrix}$$

5. Nach Satz 7.3 ist der Markierungszustand x für C-Atom-Markierungssysteme stets eine Funktion Γ von v und diese Beziehung ist explizit gegeben durch

$$x = \Gamma(v) = -(P \odot v)^{-1} \cdot (P^{\text{inp}} \odot v) \cdot x^{\text{inp}} \quad (8.3)$$

Bei Isotopomeren-Systemen muß dagegen – unter Voraussetzung der Stabilität – eine iterative Lösungsstrategie zur Berechnung von $\Gamma(v)$ verwendet werden. Für die instationäre Simulation nach Gleichung (4.33) muß schließlich ein Differentialgleichungslöser angewandt werden.

Kurz gefaßt ist eine Simulation ein Hintereinanderschalten von Abbildungen

$$n^{\text{free}} \xrightarrow{\Psi} \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} \xrightarrow{\Phi^{[0,1]}} \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}} \end{pmatrix} \xrightarrow{\Phi_\alpha} \begin{pmatrix} v^\rightarrow \\ v^\leftarrow \end{pmatrix} \xrightarrow{\Gamma} x$$

zur Berechnung von x als Funktion von n^{free} :

$$x = \Gamma \circ \Phi_\alpha \circ \Phi^{[0,1]} \circ \Psi(n^{\text{free}}) \quad (8.4)$$

8.1.2 Beurteilung und Lösung des Flußsystems

Der erste Rechenschritt ist die Berechnung von $\Psi(n^{\text{free}})$. In der Praxis wird dazu die Gleichung (8.2) nicht durch eine explizite Invertierung von

$$N = \begin{pmatrix} N^{\text{net}} & 0 \\ 0 & N^{\text{xch}[0,1]} \\ & & N^{\text{free}} \end{pmatrix}$$

gelöst, sondern durch eine Zerlegungsmethode. Dabei hängt es von der Kondition der Matrix N ab, ob eine einfache LR-Zerlegung nach dem Gauß-Algorithmus genügt oder die numerisch stabile QR-Zerlegung verwendet werden muß [DHB94].

Während N^{net} und $N^{\text{xch}[0,1]}$ durch die Systemstruktur festgelegt sind, kann N^{free} noch durch Wahl der freien Flüsse gestaltet werden. Konkret sollte diese Wahl im Hinblick auf eine möglichst günstige numerische Kondition der Matrix N getroffen werden. Nach erfolgter Festlegung von N^{free} kann als Erfolgskontrolle die Konditionszahl von N mit Hilfe einer Singulärwertzerlegung (vgl. Definition 11.3) bestimmt werden. Die nachfolgenden Überlegungen helfen bei der Auswahl der freien Flüsse:

1. In Bezug auf das Erzielen einer guten Kondition von N bereiten die freien Austauschflüsse keine Schwierigkeiten, denn die Matrix $N^{\text{xch}[0,1]}$ ist gemäß Definition 6.2 aus Zeilen-Einheitsvektoren zusammengesetzt. Als optimale Wahl für die freien Austauschflüsse werden standardmäßig die in $N^{\text{xch}[0,1]}$ noch fehlenden Zeilen-Einheitsvektoren in N^{free} ergänzt. Konkret bedeutet dies, daß alle nicht festgesetzten Austauschflüsse automatisch als freie Flüsse gewählt werden.
2. Es verbleibt das Problem, ein optimales Vektorraum-Komplement zu den Zeilen von N^{net} zu finden, d.h. eine möglichst günstige Wahl freier Nettoflüsse zu treffen. In der Praxis wird man dabei so verfahren, daß zunächst biologisch aussagekräftige Nettoflüsse als freie Flüsse gewählt werden (vgl. die Wahl der Flüsse v_1^{net} und v_5^{net} in Abbildung 9.1). In der Regel führt diese intuitive Wahl auch zu einer guten numerischen Kondition von N .
3. Wurde damit noch keine akzeptable Kondition erreicht, so kann mit etwas mehr Aufwand systematisch vorgegangen werden, indem die durch die Systemannahmen unveränderlich festgelegten Matrix N^{net} mit Hilfe einer Singulärwertzerlegung (vgl. Definition 11.3) analysiert wird [Val91]. Sollte die auf diese Weise berechnete Konditionszahl von N^{net} bereits sehr groß sein, so läßt sich an der schlechten Kondition der größeren Matrix N durch Wahl von N^{free} grundsätzlich nichts mehr ändern. Für die Praxis muß dann eine starke Fehlerübertragung bei der Flußschätzung in Kauf genommen werden (vgl. Kapitel 11).
4. Trotzdem gibt es stets eine optimale Wahl für N^{free} , welche die Kondition der Matrix N gegenüber N^{net} zumindest nicht verschlechtert. Diese kann mit Hilfe der Singulärwertzerlegung aus einem orthogonalen Vektorraumkomplement des durch die Zeilen von N^{net} aufgespannten Raums berechnet werden.¹ Leider ist die auf diese Weise erhaltene Wahl der Matrix N^{free} in der Regel nicht zulässig, da sie nach Definition 6.3 aus Einheitsvektoren zusammengesetzt sein soll. Es kann jedoch versucht werden, der optimalen Wahl möglichst nahe zu kommen. Oft bietet die optimale Matrix hierfür bereits gute Anhaltspunkte, weil in ihr nur an wenigen Stellen große Werte und an allen übrigen Stellen Werte nahe Null auftreten.

¹Auf die Details soll hier verzichtet werden. Vgl. aber auch die Ausführungen nach Definition 11.3.

5. Ganz allgemein stellt die Lösung der Frage nach der Wahl optimaler freier Flüsse ein kombinatorisches Problem dar. Dieses kann grundsätzlich nur durch ein diskret arbeitendes Optimierungsverfahren gelöst werden, daß gezielt nach einer optimalen aus Einheitsvektoren zusammengesetzten Matrix N^{free} sucht. Dazu müssen insgesamt

$$\begin{pmatrix} \dim v^{\text{net}} - \dim n^{\text{net}} \\ \dim v^{\text{net}} \end{pmatrix}$$

Möglichkeiten durchsucht werden. Eine solch exakte Lösung der Fragestellung ist aber für die praktische Arbeit jedoch kaum von Interesse, da nur die Größenordnung der Matrixkondition von Bedeutung ist.

8.1.3 Lösung der stationären C-Atom-Bilanzgleichungen

Offenbar bereitet die Berechnung der Transformationen $\Phi^{[0,1]}$ und Φ_α keine Probleme. Nach dieser Berechnung von v aus v^{net} und $v^{\text{sch}[0,1]}$ ist nun $\Psi(v)$ zu berechnen. Dazu müssen die jeweiligen Markierungs-Bilanzgleichungen (4.29) bzw. (4.30) gelöst werden. Für das Kohlenstoff-Markierungssystem gelingt dies nach Satz 7.3 durch Lösung der bzgl. x linearen Bilanzgleichungen (4.29).

Natürlich wird man die Systemmatrix $P \odot v$ dazu nicht explizit invertieren, sondern erneut eine Zerlegungstechnik anwenden. In Abschnitt 8.2 wird gezeigt, daß diese Matrix schlecht konditioniert ist, wenn große Austauschflüsse vorliegen. Daher ist grundsätzlich eine QR-Zerlegung empfehlenswert.

Einen alternativen Zugang zur Lösung der Bilanzgleichungen stellt ein iteratives Verfahren dar, wie z.B. das in [ZS94] verwendete Gauß-Seidel-Verfahren. Die Konvergenz eines solchen Verfahrens wird durch die Stabilität des entsprechenden Differentialgleichungssystems (4.33) garantiert. Jedoch sei darauf hingewiesen, daß nach allgemeiner Erfahrung [Hac93] erst ab etwa der Dimension 100 die Überlegenheit eines iterativen Lösungsansatzes gegenüber der QR-Zerlegung festgestellt werden kann. Dies konnte bei der Lösung der in der Praxis aufgetretenen Bilanzgleichungssysteme verifiziert werden.

8.1.4 Lösung der stationären Isotopomeren-Bilanzgleichungen

Die bzgl. x nichtlinearen Isotopomeren-Bilanzgleichungen erfordern einen iterativen Lösungsansatz. Unter Voraussetzung der Stabilität des Systems kann grundsätzlich ein Differentialgleichungslöser (mit zulässigen Startwerten) zur Lösung der Gleichungen verwendet werden.

Da die Lösung von Isotopomeren-Bilanzgleichungen im aktuellen Stadium des Jülicher Stoffflußanalyse-Projektes noch nicht in großem Umfang erforderlich war, genügte zunächst die Implementierung des Euler-Verfahrens als dem einfachsten denkbaren Differentialgleichungslöser [PFTV88]. Wichtig ist dabei die Projektion auf den Raum G der zulässigen Lösungen (Definition 7.7) nach jedem Schritt, wodurch die Stabilität des Verfahrens erhöht wird. Zusätzlich kann die Schrittweite erfolgsabhängig gesteuert werden.

Das Verfahren lief bei dem in Kapitel 9 vorgestellten Beispielsystem mit mehreren tausend Isotopomeren-Transitionen zuverlässig und konvergierte bei unidirektional angenommenen Flüssen startwertabhängig nach durchschnittlich etwa 200 Schritten. Ein günstiger Startwert kann dabei die Zahl der Iterationsschritte erheblich reduzieren. Ein solcher Startwert läßt sich aus der Lösung der entsprechenden C-Atom-Bilanzgleichungen gewinnen. Dazu setze man die Anteile der einfach markierten Isotopomeren auf die berechneten Positions-Markierungsanteile und normiere auf die Gesamtsumme Eins.

Bei großen Austauschflüssen sind jedoch erheblich längere Rechenzeiten zu erwarten (vgl. dazu Abschnitt 8.2), so daß in Zukunft effizientere Verfahren implementiert werden müssen [BL94, Deu95].

8.1.5 Lösung der instationären Bilanzgleichungen

Zum Nachvollzug isotopisch instationärer Experimente muß schließlich die numerische Lösung des Differentialgleichungssystems (4.33) mit hoher Genauigkeit angenähert werden. Bei nicht allzu schlecht konditionierten Systemen hat sich hierfür ein spezielles Runge-Kutta-Fehlberg-Verfahren für steife Systeme [KR79] als leistungsfähig erwiesen.

Treten jedoch große Austauschflüsse auf, so führt die resultierende Steifheit des Systems zu einer extrem kleinen Schrittweitenwahl durch das Verfahren und damit zu immer längeren Rechenzeiten. In dieser Situation kann wie in [SSH92] zu einer Näherungslösung mit Hilfe einer Quasi-Stationaritätsannahme für die in schnellem Austausch stehenden Pools übergegangen werden. Dabei werden die austauschenden Pools zu einem neuen Pool zusammengefaßt, dessen Größe der Summe der zusammengefaßten Poolgrößen entspricht (Pool-Lumping). Die Gültigkeit einer solchen Quasi-Stationaritätsannahme werden in Abschnitt 9.2 noch ausführlicher diskutiert. Eine andere Vereinfachungsoperation für ein bestimmtes Netzwerk werden in [CB95] diskutiert.

8.2 Ableitungen und große Austauschflüsse

Für die in Kapitel 10 vorgestellte numerische Parameteranpassung und die in Kapitel 11 behandelte statistische Auswertung wird nicht nur der Markierungszustand $\mathbf{x} = \Gamma \circ \Phi_\alpha \circ \Phi^{[0,1]} \circ \Psi(\mathbf{n}^{\text{free}})$ aus Gleichung (8.4) sondern auch seine Ableitung nach dem Vektor $\mathbf{n}^{\text{free}}$ der freien Flüsse benötigt. Hierbei sind die Ableitungen der Abbildungen Φ_α , $\Phi^{[0,1]}$ und Ψ offenbar wenig problematisch. Bei der Auswertung von Γ und der entsprechenden Ableitung stellt sich jedoch heraus, daß große Austauschflüsse zu numerischen Konditionsproblemen führen. Nachfolgend wird gezeigt, wie die C-Atom-Bilanzgleichungen geeignet vorkonditioniert, numerisch stabil gelöst und auch abgeleitet werden können. Ein wichtiges Ergebnis ist, daß eine explizit berechenbare Ableitung von Γ für "unendlich große" Austauschflüsse (im Sinne von Definition 5.3) existiert. Auf die Diskussion der technischen Details der Rechnerimplementierung wird hier verzichtet. Diese können in [Sie96] nachgeschlagen werden.

8.2.1 Ableitung des C-Atom-Markierungszustands nach den Flüssen

Vernachlässigt man zunächst das Problem der numerischen Kondition, so ergibt sich folgende exakte Lösung für die Ableitung von Γ :

Satz 8.1 (Ableitung des C-Atom-Markierungszustands) *Sei mit*

$$\mathbf{0} = \mathbf{P}^{\text{inp}} \odot \mathbf{v} \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v} \cdot \mathbf{x}$$

ein stationäres C-Atom-Bilanzgleichungssystem gegeben. Dann ist $\Gamma(\mathbf{v})$ für alle zulässigen $\mathbf{v}$ differenzierbar und die Ableitung ist gegeben durch

$$\frac{\partial \Gamma}{\partial \mathbf{v}}(\mathbf{v}) = -(\mathbf{P} \odot \mathbf{v})^{-1} \cdot (\mathbf{P} \cdot \mathbf{x} + \mathbf{P}^{\text{inp}} \cdot \mathbf{x}^{\text{inp}}) \quad .$$

Beweis:

Die Differenzierbarkeit folgt aus dem Satz über implizite Funktionen. Die Ableitungsformel ergibt sich dann sofort durch implizite Differentiation der Bilanzgleichung. Dabei beachte man, daß aufgrund der 3D-Matrixnotation $\mathbf{P} = (\mathbf{P}_1, \dots, \mathbf{P}_{\dim \mathbf{v}})$ (siehe Abschnitt 4.4) gilt:

$$\frac{\partial(\mathbf{P} \odot \mathbf{v})}{\partial \mathbf{v}} = \left(\frac{\partial(\mathbf{P} \odot \mathbf{v})}{\partial v_1}, \dots, \frac{\partial(\mathbf{P} \odot \mathbf{v})}{\partial v_{\dim \mathbf{v}}} \right) = (\mathbf{P}_1, \dots, \mathbf{P}_{\dim \mathbf{v}}) = \mathbf{P}$$

□

Dieser Satz hat weitreichende Konsequenzen für die numerische Berechnung der Ableitungen im gut konditionierten Fall. Es zeigt sich, daß dieselbe Matrixfaktorisierung für $\mathbf{P} \odot \mathbf{v}$ eingesetzt werden kann, um sowohl $\mathbf{x}$ als auch $\frac{\partial \mathbf{x}}{\partial \mathbf{v}}$ zu berechnen. Daher ist die Ableitungsberechnung für Markierungssysteme äußerst effizient durchführbar und kostet nur einen Bruchteil der Rechenzeit für die Berechnung von $\Gamma(\mathbf{v})$.

Für das Isotopomeren-System gilt leider kein entsprechendes Resultat. Hier ist zur Ableitungsberechnung die Lösung eines sehr großen linearen Gleichungssystems erforderlich. Der Beweis ist dem obigen vollkommen analog.

Satz 8.2 (Ableitung des Isotopomeren-Markierungszustands) Sei mit

$$\mathbf{0} = \mathbf{P}^{\text{inp}} \odot \mathbf{v} \cdot \mathbf{x}^{\text{inp}} + \mathbf{P} \odot \mathbf{v} \cdot \mathbf{x} + \mathbf{x}^T \cdot \mathbf{Q} \odot \mathbf{v} \cdot \mathbf{x}$$

ein für alle zulässigen $\mathbf{v}$ eindeutig lösbares Isotopomeren-Bilanzgleichungssystem gegeben. Dann ist $\Gamma(\mathbf{v})$ für alle zulässigen $\mathbf{v}$ differenzierbar und die Ableitung ist gegeben durch

$$\frac{\partial \mathbf{x}}{\partial \mathbf{v}}(\mathbf{v}) = - [\mathbf{x}^T \mathbf{Q} \odot \mathbf{v} + \mathbf{P} \odot \mathbf{v}]^{-1} (\mathbf{x}^T \cdot \mathbf{Q} \cdot \mathbf{x} + \mathbf{P} \cdot \mathbf{x} + \mathbf{P}^{\text{inp}} \cdot \mathbf{x}^{\text{inp}}) \quad .$$

8.2.2 Lösung gewichteter Gleichungssysteme

Die Untersuchung gewichteter Gleichungssysteme bereitet die Behandlung eines Markierungssystems im Falle großer Austauschflüsse vor. Ein solches gewichtetes Gleichungssystem hat die Gestalt:

$$[(1 - \lambda) \mathbf{A} + \lambda \mathbf{B}] \cdot \mathbf{x} = (1 - \lambda) \mathbf{y} \quad (8.5)$$

Wenn $\mathbf{B}$ nicht invertierbar ist, so stellt sich die Frage, wie sich die Lösung des Systems für $\lambda \rightarrow 1$ verhält. Dann tendiert die Kondition der Matrix $(1 - \lambda) \mathbf{A} + \lambda \mathbf{B}$ offenbar gegen unendlich, so daß eine numerische Lösung von Gleichung (8.5) zunehmend schwieriger wird. Wie nachfolgend gezeigt wird, tritt diese Situation gerade im Fall großer Austauschflüsse (bzw. von $[0, 1]$ -Austauschflüssen nahe 1) auf. Der folgende Satz zeigt, daß die entstehende numerische Instabilität mit Hilfe einer geeignet gewählten Transformation beseitigt werden kann.

Satz 8.3 (Lösung eines gewichteten Gleichungssystems) Gegeben seien Matrizen $\mathbf{A}(\mu)$, $\mathbf{B}$, $\mathbf{T}$ und ein Vektor $\mathbf{y}(\mu)$ mit den folgenden Eigenschaften:

1. $\mathbf{A} = \mathbf{A}(\mu)$ und $\mathbf{y} = \mathbf{y}(\mu)$ hängen stetig differenzierbar von μ ab.

2. $\mathbf{A}(\mu) + \frac{\lambda}{1-\lambda} \mathbf{B}$ ist invertierbar für alle λ in einer Umgebung von $\lambda = 1$ (was für $\mathbf{A}$ und $\mathbf{B}$ allein nicht gelten muß).

3. $\mathbf{T} = \begin{pmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \end{pmatrix}$, ist eine Transformationsmatrix mit ²

$$\mathbf{T}_2 \mathbf{B} = \mathbf{0} \quad \text{und invertierbarer Matrix} \quad \begin{pmatrix} \mathbf{T}_1 \mathbf{B} \\ \mathbf{T}_2 \mathbf{A}(\mu) \end{pmatrix}.$$

Dann gilt:

1. Die Lösung des gewichteten Gleichungssystems (8.5), nämlich

$$\mathbf{x}(\mu, \lambda) \stackrel{\text{def}}{=} \left[\mathbf{A}(\mu) + \frac{\lambda}{1-\lambda} \mathbf{B} \right]^{-1} \cdot \mathbf{y}(\mu),$$

ist stetig auf den durch $\lambda = 1$ gegebenen Rand des Definitionsbereichs fortsetzbar.

2. Die Fortsetzung ist explizit gegeben durch:

$$\mathbf{x}(\mu, 1) = \begin{pmatrix} \mathbf{T}_1 \mathbf{B} \\ \mathbf{T}_2 \mathbf{A}(\mu) \end{pmatrix}^{-1} \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{T}_2 \mathbf{y}(\mu) \end{pmatrix}$$

3. $\mathbf{x}(\mu, 1)$ ist stetig differenzierbar bzgl. μ mit der Ableitung:

$$\frac{\partial \mathbf{x}}{\partial \mu}(\mu, 1) = \begin{pmatrix} \mathbf{T}_1 \mathbf{B} \\ \mathbf{T}_2 \mathbf{A}(\mu) \end{pmatrix}^{-1} \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{T}_2 \frac{\partial \mathbf{y}}{\partial \mu}(\mu) - \mathbf{T}_2 \frac{\partial \mathbf{A}}{\partial \mu}(\mu) \cdot \mathbf{x}(\mu, 1) \end{pmatrix}$$

4. $\mathbf{x}(\mu, \lambda)$ ist stetig differenzierbar bzgl. λ in $\lambda = 1$ mit der Ableitung:

$$\frac{\partial \mathbf{x}}{\partial \lambda}(\mu, 1) = \begin{pmatrix} \mathbf{T}_1 \mathbf{B} \\ \mathbf{T}_2 \mathbf{A}(\mu) \end{pmatrix}^{-1} \cdot \begin{pmatrix} \mathbf{T}_1 \cdot [(\mathbf{A}(\mu) - \mathbf{B}) \cdot \mathbf{x}(\mu, 1) - \mathbf{y}(\mu)] \\ \mathbf{T}_2 \cdot [-\mathbf{y}(\mu)] \end{pmatrix}$$

5. In der Nähe von $\lambda = 1$ ist $\mathbf{S} = \begin{pmatrix} (1-\lambda) \mathbf{T}_1 \\ \mathbf{T}_2 \end{pmatrix}$ eine Vorkonditionierungsmatrix für $\mathbf{A}(\mu) + \frac{\lambda}{1-\lambda} \mathbf{B}$, d.h. für eine Konditionszahl Cond gilt:

$$\text{Cond } \mathbf{S} \cdot \left[\mathbf{A}(\mu) + \frac{\lambda}{1-\lambda} \mathbf{B} \right] \ll \text{Cond} \left[\mathbf{A}(\mu) + \frac{\lambda}{1-\lambda} \mathbf{B} \right]$$

Beweis:

Es gilt die Umformung:

$$\begin{aligned} & \left[(1-\lambda) \mathbf{A}(\mu) + \lambda \mathbf{B} \right] \cdot \mathbf{x}(\lambda) = (1-\lambda) \mathbf{y}(\mu) \\ \Leftrightarrow & \begin{bmatrix} (1-\lambda) \mathbf{T}_1 \mathbf{A}(\mu) + \lambda \mathbf{T}_1 \mathbf{B} \\ (1-\lambda) \mathbf{T}_2 \mathbf{A}(\mu) + \lambda \mathbf{T}_2 \mathbf{B} \end{bmatrix} \cdot \mathbf{x}(\lambda) = \begin{bmatrix} (1-\lambda) \mathbf{T}_1 \mathbf{y}(\mu) \\ (1-\lambda) \mathbf{T}_2 \mathbf{y}(\mu) \end{bmatrix} \\ \Leftrightarrow & \begin{bmatrix} (1-\lambda) \mathbf{T}_1 \mathbf{A}(\mu) + \lambda \mathbf{T}_1 \mathbf{B} \\ \mathbf{T}_2 \mathbf{A}(\mu) \end{bmatrix} \cdot \mathbf{x}(\lambda) = \begin{bmatrix} (1-\lambda) \mathbf{T}_1 \mathbf{y}(\mu) \\ \mathbf{T}_2 \mathbf{y}(\mu) \end{bmatrix} \end{aligned}$$

²Eine solche Matrix $\mathbf{T}$ kann stets aus einer LR- oder QR-Zerlegung gewonnen werden, jedoch wird für Markierungssysteme eine spezielle Transformation $\mathbf{T}$ intendiert. Vgl. dazu die folgenden Ausführungen.

Der Grenzübergang $\lambda \rightarrow 1$ ist nun für die obere und untere Gleichung getrennt vollziehbar und liefert Punkt 1 und 2. Nun ergeben sich die Formeln aus Punkt 3 und 4 durch implizite Differentiation.

Schließlich folgt Punkt 5, weil Konditionszahlen stetige Funktionen auf dem Raum der Matrizen sind und daher

$$\text{Cond} \left[\mathbf{A}(\mu) + \frac{\lambda}{1-\lambda} \mathbf{B} \right] \rightarrow \infty \quad \text{für } \lambda \rightarrow 1$$

gilt, während die Kondition der vorkonditionierten Matrix endlich bleibt, da der Grenzwert eine invertierbare Matrix ist.

□

Nach Satz 7.3 setzt sich die Lösung der Gleichung (8.5) stets aus gebrochen rationalen Funktionen zusammen, die von λ bzw. μ abhängen. Der Satz 8.3 zeigt im wesentlichen, daß diese Funktionen in $\lambda = 1$ einen hebbaren Pol haben. Die Stetigkeit und Differenzierbarkeit der Lösung in $\lambda = 1$ ist daher grundsätzlich klar. Der Satz zeigt darüber hinaus, wie die jeweiligen Ableitungen unter Verwendung von $\mathbf{T}$ effektiv berechnet werden können. Dazu ist offenbar nur eine einzige Matrix-Faktorisierung erforderlich. Seine Anwendung findet der Satz bei der numerisch stabilen Lösung gewichteter Gleichungssysteme für $\lambda \approx 1$, indem er eine Vorkonditionierung des Gleichungssystems ermöglicht.

8.2.3 Anwendung auf Markierungssysteme

Die Anwendung des Satzes auf C-Atom-Markierungssysteme erweitert die Überlegungen in [SSH92], wo eine analoge Vorgehensweise für instationäre Systeme (allerdings nur im Grenzfall $\lambda = 1$) beschrieben ist. Sie erfolgt anhand des schon in den vorangehenden Kapiteln ausführlich behandelten Beispiels aus Abbildung 3.2. Hier wird nun untersucht, wie sich das System verhält, wenn v_2 eine schnelle Gleichgewichtsreaktion ist. Hierzu werden zunächst die in Gleichung (4.28) definierten Matrizen

$$\mathbf{P}_2^{\rightarrow} = \begin{pmatrix} -1 & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \end{pmatrix} \quad \text{und} \quad \mathbf{P}_2^{\leftarrow} = \begin{pmatrix} \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix}$$

betrachtet, die die den Flüssen $v_2^{\rightarrow}$ und $v_2^{\leftarrow}$ zugeordneten Einträge der C-Atom-Transitionsmatrix $\mathbf{P}$ aus Definition 4.9 darstellen. Eine Transformation in den $(v_2^{\text{net}}, v_2^{\text{xch}[0,1]})$ -Raum (vgl. Definition 5.4 und Satz 5.4) liefert dann:

$$\begin{aligned} \mathbf{P} \odot \mathbf{v} &= \dots + \mathbf{P}_2^{\rightarrow} \cdot v_2^{\rightarrow} + \mathbf{P}_2^{\leftarrow} \cdot v_2^{\leftarrow} + \dots \\ &= \dots + \mathbf{P}_2^{\rightarrow} \cdot (v_2^{\text{xch}} + \max_{\alpha}(v_2^{\text{net}}, 0)) + \mathbf{P}_2^{\leftarrow} \cdot (v_2^{\text{xch}} + \max_{\alpha}(-v_2^{\text{net}}, 0)) + \dots \\ &= \dots + (\mathbf{P}_2^{\rightarrow} + \mathbf{P}_2^{\leftarrow}) \cdot v_2^{\text{xch}} + \mathbf{P}_2^{\rightarrow} \cdot \max_{\alpha}(v_2^{\text{net}}, 0) + \mathbf{P}_2^{\leftarrow} \cdot \max_{\alpha}(-v_2^{\text{net}}, 0) + \dots \\ &= \dots + \frac{v_2^{\text{xch}[0,1]}}{1 - v_2^{\text{xch}[0,1]}} (\mathbf{P}_2^{\rightarrow} + \mathbf{P}_2^{\leftarrow}) + \mathbf{P}_2^{\rightarrow} \cdot \max_{\alpha}(v_2^{\text{net}}, 0) + \mathbf{P}_2^{\leftarrow} \cdot \max_{\alpha}(-v_2^{\text{net}}, 0) + \dots \end{aligned}$$

Damit ist die Ausgangssituation für die Anwendung des Satzes 8.3 gegeben. Nun ist eine geeignete Transformation $\mathbf{T}$ zu der Matrix

$$\mathbf{P}_2^{\rightarrow} + \mathbf{P}_2^{\leftarrow} = \begin{pmatrix} -1 & . & 1 & . \\ . & -1 & . & 1 \\ 1 & . & -1 & . \\ . & 1 & . & -1 \end{pmatrix}$$

gesucht. Offenbar sind in dieser Matrix die erste und dritte sowie die zweite und vierte Spalte jeweils linear abhängig. Die Transformationsmatrix $\mathbf{T}$ wird nun so gewählt, daß die beiden letzten Zeilen (also $\mathbf{T}_2$) diese linear abhängigen Zeilen addieren während die ersten beiden Zeilen (also $\mathbf{T}_1$) zu je einer Zeile von $\mathbf{T}_2$ orthogonal sind:

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \end{pmatrix} = \begin{pmatrix} 1 & . & -1 & . \\ . & 1 & . & -1 \\ 1 & . & 1 & . \\ . & 1 & . & 1 \end{pmatrix}$$

Dann gilt

$$\mathbf{T} \cdot (\mathbf{P}_2^{\rightarrow} + \mathbf{P}_2^{\leftarrow}) = 2 \cdot \begin{pmatrix} -1 & . & 1 & . \\ . & -1 & . & 1 \\ . & . & . & . \\ . & . & . & . \end{pmatrix} . \quad (8.6)$$

Damit können für große Austauschflüsse v_2^{sch} der zugehörige Markierungsvektor $\mathbf{x}$ und dessen Ableitung nach allen im System auftretenden Flüssen numerisch stabil berechnet werden.

In der Praxis bedeutet die allgemeine Durchführung des oben skizzierten Programms einigen technische Aufwand, denn bei einem gegebenen Flußzustand können auch mehrere Austauschflüsse des Systems zugleich gegen Unendlich streben. In diesem Falle sind zunächst die Zusammenhangskomponenten aller über schnelle Austauschflüsse miteinander in Verbindung stehenden Pools zu berechnen. Die technischen Details dieser Vorgehensweise werden in [Sie96] beschrieben. An praxisrelevanten Beispielen hat sich gezeigt, daß die Vorkonditionierung selbst großer Markierungssysteme stets zu kleinen Konditionszahlen führt.³

8.2.4 Die Grenzgleichung für schnelle Gleichgewichte

Interessant ist nun das Grenzverhalten des Systems, d.h. seine Lösung für $v_2^{\text{sch}[0,1]} = 1$. In diesem Fall gilt nach Satz 8.3 die *Grenzgleichung* (vgl. Gleichung (8.6))

$$\mathbf{0} = \mathbf{T}_1 (\mathbf{P}_2^{\rightarrow} + \mathbf{P}_2^{\leftarrow}) \cdot \mathbf{x} = \begin{pmatrix} -1 & . & 1 & . \\ . & -1 & . & 1 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} b_1 - e_1 \\ b_2 - e_2 \end{pmatrix} , \quad (8.7)$$

d.h. der Markierungszustand der in schnellem Austausch stehenden C-Atom-Pools ist im Grenzwert gleich. Das System kann daher durch Pool-Lumping von B#1 und E#1 bzw. B#2 und E#2 auf ein System mit zwei Pools reduziert werden (vgl. dazu auch Abb. 9.5). Da das so reduzierte System keine großen Austauschflüsse mehr aufweist, ist auch anschaulich klar, warum auf diese Weise eine numerisch stabile Lösung erzielt wurde.

³Dazu wurden die Konditionszahlen testweise mit Hilfe einer Singulärwertzerlegung berechnet (vgl. auch [FR95]).

Die Bilanzgleichungen für das reduzierte System werden im wesentlichen durch die Produkte $T_2 \cdot P_i^{\rightarrow}$ und $T_2 \cdot P_i^{\leftarrow}$ mit $i \neq 2$ festgelegt. Diese auf beliebige Systeme verallgemeinerbaren Überlegungen bilden die theoretische Grundlage für das Pool-Lumping bei schnell austauschenden Flüssen.

Durch die oben getroffene spezielle Wahl von T ergeben sich die Beziehungen $b_1 = e_1$ und $b_2 = e_2$ in Gleichung (8.7) unmittelbar. Es kann gezeigt werden, daß sich ein solches T stets finden läßt, so daß gilt:

Satz 8.4 (Grenzgleichung bei schnellen Markierungsgleichgewichten) Sei x der Markierungszustand und v der Flußvektor eines fallenfreien C-Atom-Bilanzgleichungssystems. Der i -te Fluß verbinde den j -ten und den k -ten C-Atom-Pool miteinander (d.h. $(P_i)_{k,j} > 0$ in Definition 4.9). Dann gilt die Grenzgleichung:

$$\lim_{v_i^{\text{xch}} \rightarrow \infty} [x_j(v_i^{\text{xch}}) - x_k(v_i^{\text{xch}})] = 0$$

Im allgemeinen sind die Nettoflüsse in einem metabolischen System bereits durch die in Abschnitt 6.1 gestellten Nebenbedingungen auf ein kompaktes Gebiet beschränkt. Aufgrund von Satz 8.4 folgt dann, daß die Menge

$$\{ \Gamma(v) \mid v \text{ zulässig} \wedge v^{\text{xch}} \leq \infty \} \quad ,$$

bei der unendliche Werte für Austauschflüsse zugelassen sind, stets wieder eine kompakte Menge ist. Diese Beziehung erweitert den in Abschnitt 5.2 am Beispiel eingeführten Begriff des realisierbaren bzw. asymptotisch realisierbaren Markierungszustands auf allgemeine Markierungssysteme. Die asymptotischen Grenzen werden dabei durch die Beziehungen $v_i^{\text{xch}} = \infty$ definiert.

Kapitel 9

Simulation von Markierungssystemen

Die numerischen Aspekte der Durchführung einer Simulation wurden bereits in den Abschnitten 8.1 und 8.2 ausführlich diskutiert. Diese rechen- und verwaltungstechnischen Details sollen natürlich vor dem experimentell arbeitenden Anwender eines Simulators verborgen bleiben. Nachfolgend werden nun die in Jülich entwickelten Simulationswerkzeuge vorgestellt und exemplarisch angewendet.

Der Simulator erlaubt die Festlegung einer Netzwerkstruktur und aller für eine Simulation relevanten Annahmen in einem übersichtlichen Eingabeformat. Diese Eingabe spiegelt unmittelbar die in den Kapiteln 3-6 eingeführten Strukturelemente wieder. Auf der Grundlage dieser Eingabe werden alle für die Simulation erforderlichen Matrizen und Vektoren automatisch generiert. In diesem Kapitel wird anhand eines konkreten Beispiels der Ablauf einer Simulationsstudie dargestellt.

Die Verfügbarkeit einer derartigen Simulationsumgebung stellt einen großen Fortschritt dar, da sie eine einfache Modifikation der Netzwerkstrukturen erlaubt, ohne den Anwender mit der expliziten Formulierung von Bilanzgleichungen oder Transitionsmatrizen zu belasten. Die Architektur und Implementierung des in Jülich entwickelten Simulationssystems wird in Abschnitt 9.3 kurz beschrieben.

9.1 Durchführung von Simulationen

9.1.1 Simulationsstrategie

Die generelle Simulationsstrategie sollte mit Rücksicht auf die Bedürfnisse des Anwenders wie folgt aussehen:

1. Formulierung eines biochemischen Netzwerks.
2. Formulierung der Annahmen über die im System vorhandenen Netto- und Austauschflüsse in Form von linearen Beziehungen für Netto- und Austauschflüsse (Anwendungs-Flußkoordinaten).
3. Auswahl von freien Netto- und Austauschflüssen.
4. Vorgabe der Eingangsmarkierung, also x^{inp} .
5. Wertvorgaben für die freien Flüsse, also n^{free} .
6. Prüfung der Zulässigkeit des dadurch bestimmten Flußzustands.

7. Berechnung von v und x gemäß der Berechnungsreihenfolge aus Abschnitt 8.1.

Zur Durchführung einer Variationsstudie, werden schließlich die Schritte 4 bis 7 mit einer Reihe verschiedener Werte für die Eingangsmarkierungen x^{inp} und die freien Flüsse n^{free} wiederholt.

9.1.2 Anforderungen an ein Simulationswerkzeug

Ein Simulationswerkzeug sollte jeden der genannten Schritte (soweit erforderlich) unterstützen. Insbesondere heißt das:

1. Alle Eingaben sollten in einem für den experimentell arbeitenden Anwender leicht verständlichen Texteingabeformat erfolgen.
2. Jede Eingabe muß auf logische Konsistenz mit den übrigen Daten geprüft werden.
3. Nur die unbedingt erforderlichen Informationen sind einzugeben. Alle Systemmatrizen und -vektoren sowie grundsätzlich gemachten Annahmen (z.B. die der Unidirektionalität extrazellulärer Flüsse) sollten daraus automatisch generiert werden.
4. Bei willkürlichen Festlegungen, wie z.B. der Auswahl der freien Flüsse, sollte eine Hilfestellung erfolgen.
5. Die in Abschnitt 8.1 dargestellten Rechenschritte sollten numerisch stabil implementiert werden.
6. Ausnahmesituationen (pathologische Fälle, systembedingte numerische Instabilitäten) sollten automatisch diagnostiziert werden.
7. Die Ausgabe der Simulationsergebnisse sollte für die Weiterverarbeitung geeignet vorformatiert sein.

Die Durchführung der einzelnen Schritte der Simulationsstrategie wird nachfolgend anhand eines Beispiels vorgestellt. Die oben formulierten Anforderungen werden dabei noch an verschiedenen Stellen diskutiert. Bei dem vorgestellten Beispiel steht weniger die umfassende Untersuchung des Systems im Vordergrund als vielmehr eine überblickartige Einführung in die Möglichkeiten der Simulation. Eine tiefergehende Diskussion der systemtheoretischen Eigenschaften dieses Beispiels und deren praktische Konsequenzen findet sich in [WG96b].

9.2 Ein Beispiel

Als Beispiel für eine Simulationsstudie wird die Markierungsverteilung im zyklischen Pentosephosphatweg studiert. Der zyklischen Weg tritt beim Wachstum auf Pentosen auf, beispielsweise bei der xyloseverwertenden Hefe *Pichia stipitis*¹ [LPdPB88] oder einer mit gentechnischen Methoden hergestellte Rekombinante von *Zymomonas mobilis* [ZED⁺95]. Letztere hat in jüngster Zeit eine große Bedeutung für die Untersuchung des Zuckerstoffwechsels erlangt.

¹Die Auswahl dieses etwas "exotischen" Organismus hat ihre Ursache darin, daß *Pichia stipitis* im Jülicher mikrobiologischen Praktikum als Beispiel für die NMR-Anwendung diente.

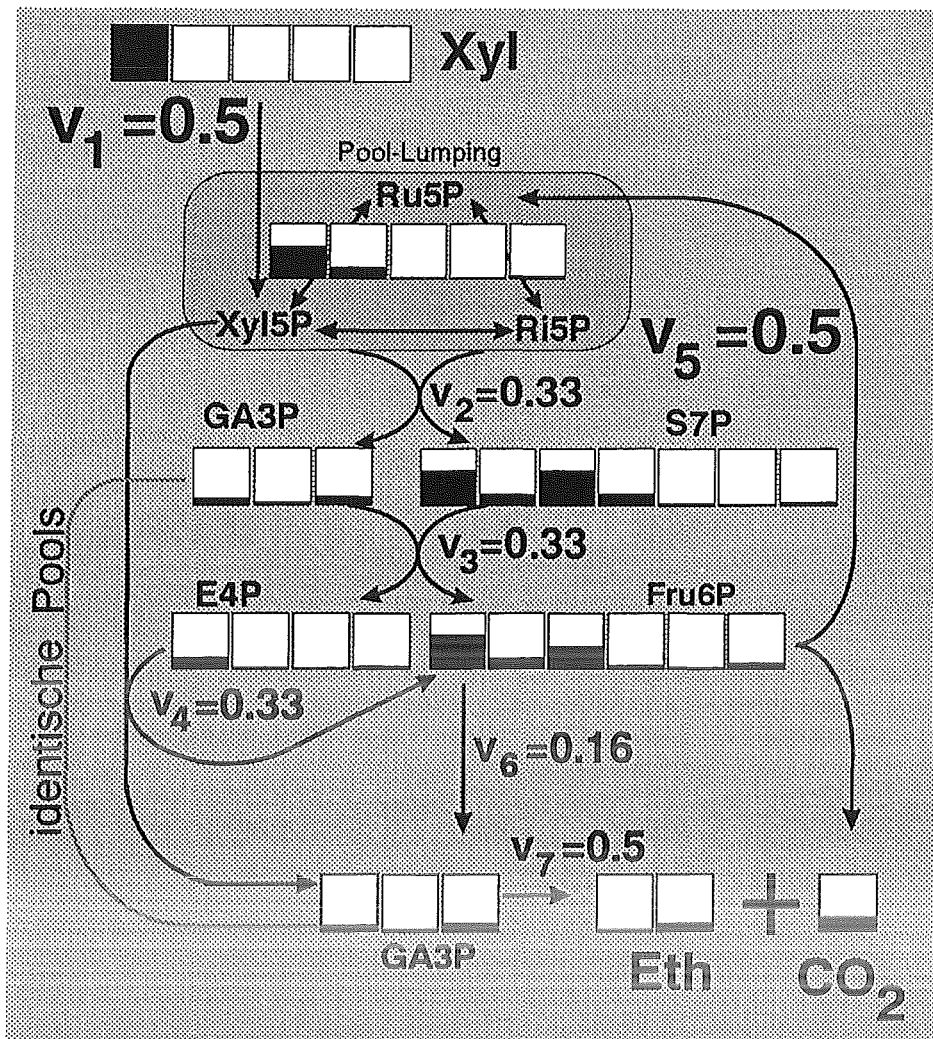


Abbildung 9.1: Zyklisch ablaufender Pentosephosphatweg in der xyloseverwertenden Hefe *Pichia stipitis* und Simulationsergebnis für ein stationäres Markierungsexperiment mit freien Flüssen $v_1 = 0.5$ und $v_5 = 0.5$. Alle Flüsse wurden als unidirektional angenommen. Die extrazellulären Flüsse v_8 und v_9 wurden aus Gründen der Übersichtlichkeit weggelassen.

Das Stoffwechselnetzwerk des zyklischen Pentosephosphatwegs zeigt die Abbildung 9.1. Beim Wachstum auf Xylose nimmt der Organismus den C5-Körper direkt über den Pentosephosphatweg auf. Somit ist mit einer hohen Aktivität dieses Stoffwechselabschnitts zu rechnen. Im Gegensatz zum Wachstum auf Glucose müssen nun aber die für den Zellwandaufbau und andere Zwecke benötigten C6-Körper via Fructose-6-phosphat neu synthetisiert werden. Hierzu läuft ein Teil der Glykolyse in umgekehrter Richtung ab. Diese sogenannten gluconeogenetischen Schritte [Str88] sind die umgekehrt zur Tabelle 3.2 ablaufenden Schritte Gly2 und Gly1. Die neu synthetisierten C6-Körper können dann über den oxidativen Pentosephosphatweg PPW1 unter Abspaltung von CO_2 wieder in den Pentose-5-Phosphat-Pool gelangen, wodurch ein Zyklus geschlossen wird.

Um die Notation kurz zu halten, sind die Flüsse durchnummeriert worden. Zudem wurden einige Flußsequenzen aus Tabelle 3.2 und 3.3 zusammengefaßt. Folgende Überlegungen liegen dieser Auswahl und Zusammenfassung von Flüssen zugrunde (vgl. Abb. 9.1):

1. Die Pentose-5-Phosphat-Pools (Ri5P, Ru5P und Xy5P) wurden als im schnellen Gleich-

gewicht befindlich angenommen, was durch die in der Regel anzutreffende sehr hohe Aktivität der entsprechenden Isomerasen und Epimerasen [Woo85] gerechtfertigt ist. Es entsteht ein zusammengefaßter Pentose-5-Phosphat-Pool.

2. Die drei Zyklus-Reaktionsschritte der Gluconeogenese und des oxidativen Pentosephosphatwegs bilden eine lineare Sequenz und können daher zu V5 zusammengefaßt werden.
3. Die Biomasseabflüsse aus der Glykolyse werden in dieser Studie zur Vereinfachung mit dem Abfluß V7 des Glycerinaldehyd-3-phosphats in das Produkt Ethanol zusammengefaßt. Der geringe Abfluß des Pentosephosphatweg-Intermediats Erythrose-4-phosphat in die Biomasse wird dagegen vernachlässigt.

Es resultiert die nachfolgende Zuordnung der Bezeichnungen (vgl. Abb. 9.1)

- V1 $\hat{=}$ PPW0, PPW2, PPW3
- V2 $\hat{=}$ PPW4
- V3 $\hat{=}$ PPW5
- V4 $\hat{=}$ PPW6
- V5 $\hat{=}$ Gly2, Gly1, PPW1
- V6 $\hat{=}$ Gly3, Gly4
- V7 $\hat{=}$ Gly5, Gly6, PB1

Intrazelluläre Stoffkonzentrationen können je nach betrachtetem Metaboliten zwischen 100 mmol/l (Glutamat-Pool in *C. glutamicum*) [Sch95] und einer nicht mehr nachweisbaren Konzentration von 10 μ mol/l für einen Übergangsmetaboliten in einer schnell ablaufenden Reaktionssequenz variieren. Dabei wird bei Intermediärmetaboliten im Zuckerstoffwechsel im allgemeinen von (im Rahmen der metabolischen Konzentrationsskala) kleinen Konzentrationen ausgegangen. Hier sollen daher zunächst Poolgrößen von 1 mmol/l angenommen werden (vgl. auch [FHBS90, WGKF⁺92, CFGC95]).

Da Hefen einen im Vergleich zu tierischen Zellen sehr großen Stoffdurchsatz haben, wird eine Substrataufnahmerate von 0.5 mol/(l · h) bezogen auf das Zellvolumen eingesetzt. Geht man von einem Zellvolumen von 0.002 l/g (Trockenmasse) aus [Sch95], so entspricht dies bei einem Reaktor mit einem Volumen von 1 l und einer Zelldichte von 10 g/l etwa einer Xylose-Aufnahmerate von 0.01 mol/(l · h) = 1.5 g/(l · h) bezogen auf das Reaktorvolumen. Bedenkt man, daß schnell wachsende Mikroorganismen innerhalb von zwei Stunden ihre Zellmasse verdoppeln können, so ist dies für die biotechnologische Stoffproduktion ein eher kleiner Wert.

9.2.1 Simulatoreingabe

Die hierarchisch strukturierte Texteingabe für den Simulationslauf zeigt Abb. 9.2. Die Einträge dieser Datei sind wie folgt zu interpretieren (vgl. dazu Abschnitt 6.3):

1. Das biochemisches Netzwerk ist in der Sektion EQUATIONS mit Hilfe der in Definition 3.1 eingeführten Notation definiert. Hieraus werden die Transitionsmatrizen P , P^{inp} , Q generiert sowie die grundsätzlich angenommenen stöchiometrischen Gleichungen als erster Bestandteil von N^{net} .
2. In den Untersektionen CONSTRAINTS von NET_FLUXES und XCH_FLUXES können beliebige lineare Gleichungen für Nettoflüsse und Wertsetzungen für Austauschflüsse vorgenommen werden. Die Unidirektionalität extrazellulärer Flüsse (also von V1, V8 und V9) ist automatisch vorausgesetzt. Im Beispiel werden außerdem die Flüsse V5, V6, V7

```

{
  EQUATIONS :=
  {
    V1 := Xyl    > Xyl5P,
          #ABCDE > #ABCDE;

    V2 := Xyl5P + Xyl5P > S7P      + GA3P,
          #ABCDE + #abcde > #ABabcde + #CDE;

    V3 := GA3P + S7P    > E4P    + Fru6P,
          #ABC + #abcdefg > #defg + #abcABC;

    V4 := Xyl5P + E4P    > GA3P + Fru6P,
          #ABCDE + #abcd  > #CDE + #ABabcd;

    V5 := Fru6P > CO2 + Xyl5P,
          #ABCDE > #A  + #BCDEF;

    V6 := Fru6P > GA3P + GA3P,
          #ABCDE > #CBA + #DEF;

    V7 := GA3P > CO2 + Eth,
          #abc > #a  + #bc;

    V8 := Eth > EthOut,
          #ab > #ab ;

    V9 := CO2 > CO2OUT,
          #a  > #a;
  };

  NET_FLUXES :=
  {
    CONSTRAINTS :=
    {
    };
    INEQUALITIES :=
    {
      NET_INEQ5A := V5 >= 0.0;
      NET_INEQ5B := V5 <= 1.0;
    };
    FREE :=
    {
      V1 := 0.5;
      V5 := 0.5;
    };
  };

  XCH_FLUXES :=
  {
    CONSTRAINTS :=
    {
      XCH_CONSTR5 := V5 = 0.0 ;
      XCH_CONSTR6 := V6 = 0.0 ;
      XCH_CONSTR7 := V7 = 0.0 ;
    };
    INEQUALITIES :=
    {
    };
    FREE :=
    {
      V2 := 0.0;
      V3 := 0.0;
      V4 := 0.0;
    };
  };
};

POOLS :=
{
  Xyl :=
  {
    SIMULATION :=
    {
      SIZE := 0.001;
      LABEL :=
      {
        #1 := 1.0 ;
      }
    };
  };
  Xyl5P :=
  {
    SIMULATION :=
    {
      SIZE := 0.001;
    } ;
  };
  S7P :=
  {
    SIMULATION :=
    {
      SIZE := 0.001;
    } ;
  };
  GA3P :=
  {
    SIMULATION :=
    {
      SIZE := 0.001;
    } ;
  };
  E4P :=
  {
    SIMULATION :=
    {
      SIZE := 0.001;
    } ;
  };
  Fru6P :=
  {
    SIMULATION :=
    {
      SIZE := 0.001;
    } ;
  };
  CO2 :=
  {
    SIMULATION :=
    {
      SIZE := 0.001;
    } ;
  };
};
};

```

Abbildung 9.2: Beispiel einer Simulatoreingabe zur Simulation von Markierungsverteilungen im Pentosephosphatweg (vgl. Abb. 9.1). Erklärung der Einträge im Text. Die Unterteilung SIMULATION unterhalb von POOLS scheint an dieser Stelle noch überflüssig zu sein. Dies liegt daran, daß das Format in Abbildung 12.2 noch um einen weiteren Abschnitt für Meßdaten (MEASUREMENT) ergänzt wird.

als unidirektional angenommen. Aus diesen Angaben können dann N^{net} , N^{xch} und n^{net} , n^{xch} generiert werden.

3. In den Untersektionen INEQUALITIES von NET_FLUXES und XCH_FLUXES sind beliebige Ungleichungen für Nettoflüsse und Bereichseinschränkungen für Austauschflüsse vorgebar. Die Richtung extrazellulärer Flüsse wird automatisch angenommen. Im Beispiel ist außerdem der Nettofluß v_5^{net} auf den Wertebereich $[0, 1]$ eingeschränkt. Daraus werden dann U^{net} , U^{xch} und u^{net} , u^{xch} generiert.
4. Die gewählten freien Flüsse befinden sich in den Unterabschnitten FREE von NET_FLUXES und XCH_FLUXES. Sie müssen den Anforderungen der Definition 6.4 genügen, was automatisch überprüft wird. Im Beispiel sind sieben stöchiometrische Gleichungen formulierbar. Weitere sechs Nebenbedingungen wurden durch die Richtungsannahmen eingeführt. Es verbleiben noch $\dim v - \dim n^{\text{net}} - \dim n^{\text{xch}} = 18 - 7 - 6 = 5$ Freiheitsgrade zur eindeutigen Bestimmung des Flußsystems. Gemäß der in Abschnitt 8.1 beschriebenen Strategie werden daher zunächst die noch nicht festgesetzten Austauschflüsse $v_2^{\text{xch}[0,1]}$, $v_3^{\text{xch}[0,1]}$ und $v_4^{\text{xch}[0,1]}$ zu freien Flüssen gemacht.

Es verbleiben zwei Freiheitsgrade für die Nettoflüsse. Als aussagekräftige Flüsse werden die Werte der Substrataufnahme v_1^{net} und des Zyklusflusses v_5^{net} gewählt. Aus den stöchiometrischen Gleichungen berechnet der Simulator dann:

$$\begin{array}{rclcl} v_2^{\text{net}} & = & v_3^{\text{net}} & = & v_4^{\text{net}} & = & \frac{1}{3} (v_1^{\text{net}} + v_5^{\text{net}}) \\ v_6^{\text{net}} & & & & & = & \frac{1}{3} (2 v_1^{\text{net}} - v_5^{\text{net}}) \\ v_7^{\text{net}} & = & v_8^{\text{net}} & = & & = & v_1^{\text{net}} \\ v_9^{\text{net}} & & & & & = & v_1^{\text{net}} + v_5^{\text{net}} \end{array}$$

Da somit alle abhängigen Flüsse durch die freien Flüsse ausgedrückt werden können, ist die Invertierbarkeit der automatisch generierten Nebenbedingungsmatrix N gegeben.

Die nachfolgenden Simulationsläufe gehen zunächst von unidirektionalen Flüssen aus, was in der Texteingabe durch die Vorgabewerte $v_2^{\text{xch}} = v_3^{\text{xch}} = v_4^{\text{xch}} = 0$ ausgedrückt wird. Beim letzten Simulationsbeispiel werden diese Flüsse dann ebenfalls variiert.

5. Der Input-Metabolit Xyl ist auf dem ersten Kohlenstoff-Atom markiert, wie man aus der POOLS-Sektion entnimmt. Nicht angegebene Markierungsanteile werden dabei automatisch als 0.0 angenommen.
6. Die Poolgrößen aller intrazellulären Metabolite sind ebenfalls in der POOLS-Sektion unter SIZE angegeben.

9.2.2 Stationäre C-Atom-Markierungsanteile

Mit der hier vorgenommenen Wahl der freien Flüsse wird zunächst die stationäre Markierungsverteilung berechnet, die in Abbildung 9.1 visualisiert ist. Es wird anschaulich klar, daß die unmarkierten C-Atome im System gegenüber den markierten C-Atomen bei weitem überwiegen. Dies ist auch zu erwarten, da nur 20% der zugeführten Atome markiert sind. Markierungsmessungen liefern also in der Regel kleine Markierungsmeßwerte. Ferner verbleibt der Hauptanteil der Markierung an den jeweils ersten Positionen. Die Wahrscheinlichkeit, daß ein markiertes

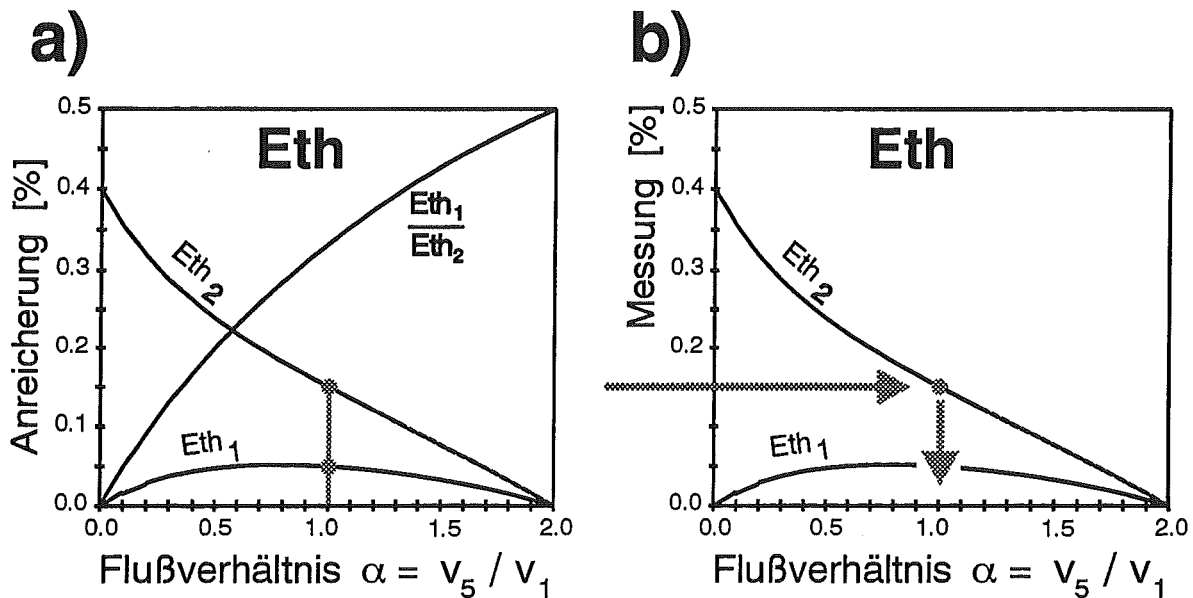


Abbildung 9.3: a) Stationäre Markierungen in Eth#1 und Eth#2 sowie deren Quotient in Abhängigkeit von $v_5^{\text{net}}/v_1^{\text{net}}$ (übrige Parameter wie in Abb. 9.1). Der in Abbildung 9.1 dargestellte spezielle Zustand ist gekennzeichnet. b) Flußbestimmung anhand eines Meßwertes für Eth#2

Teilchen in die letzten C-Atom-Positionen von S7P oder Fru6P gelangt, ist sehr gering, da dies nur über den Zyklusschritt V5 möglich ist.

Um ein etwas vollständigeres Bild zu erhalten, werden nun die freien Nettoflüsse v_1^{net} und v_5^{net} variiert. Nach Satz 7.3 wird der Markierungszustand des Systems bereits durch den Quotienten $v_5^{\text{net}}/v_1^{\text{net}}$ bestimmt. Es genügt also die Variation von v_5^{net} . Das Resultat ist exemplarisch für die beiden Kohlenstoffatome des Ethanol in Abbildung 9.3 dargestellt.

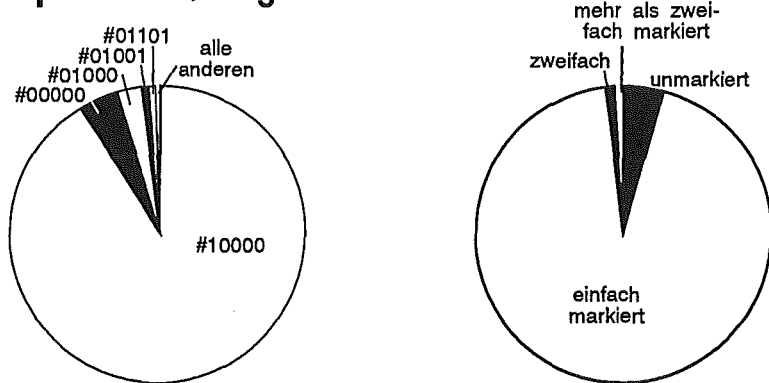
Die Variationsstudie verdeutlicht einige Aspekte des Problems der Flußbestimmung aus gemessenen Daten:

1. Ohne die Kenntnis mindestens eines absoluten Flußwerts können nur Flußverhältnisse relativ zu einem Referenzfluß bestimmt werden.
2. Für eine Flußbestimmung aus gemessenen Größen muß die Zahl der Meßwerte mindestens so groß sein wie die Zahl der freien Flüsse. Wird $v_2^{\text{xch}} = v_3^{\text{xch}} = v_4^{\text{xch}} = 0$ angenommen,² so kann offenbar $v_5^{\text{net}}/v_1^{\text{net}}$ anhand einer Messung von eth_2 , wie in Abbildung 9.3b gezeigt, bestimmt werden.
3. Selbst wenn genügend viele Werte gemessen wurden, kann es Redundanzen zwischen den Meßwerten geben, die ihren Informationsgehalt einschränken. Beispielsweise ist es unter der Annahme $v_2^{\text{xch}} = v_3^{\text{xch}} = v_4^{\text{xch}} = 0$ überflüssig, beide Ethanol-Markierungen zu vermessen.
4. Schließlich garantieren auch redundanzfreie Meßwerte noch keine eindeutige Flußbestimmung. Dies zeigt der Versuch einer Bestimmung von $v_5^{\text{net}}/v_1^{\text{net}}$ aus eth_1 (vgl. auch das Beispiel in [ZS94]).

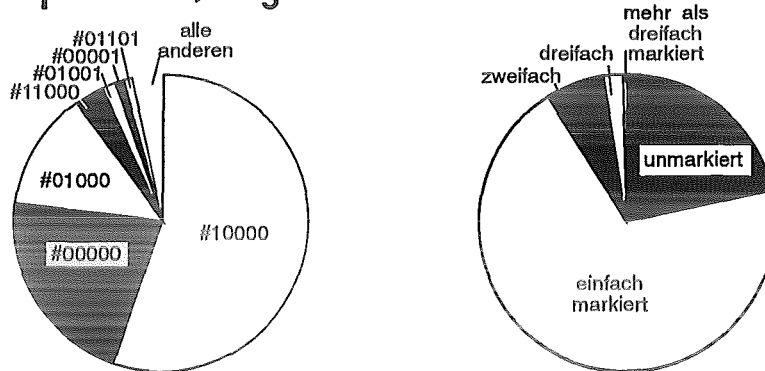
²Die Problematik der sehr unrealistischen Unidirektionalitätsannahme für die Pentosephosphatweg-Flüsse wird in Abschnitt 13 noch kritisch diskutiert (vgl. die Ausführungen zu Gleichung (13.1)). Mit dem Beispiel sollen vor allem einige weitere Eigenschaften von Markierungssystemen illustriert werden.

Die hier angeschnittenen Fragen berühren die allgemeinen Probleme der Identifizierbarkeit und Redundanz in Zustandsraummodellen, die in der vorliegenden Arbeit nicht behandelt werden sollen (siehe dazu [Wie95a]).

a) $v_1 = 0.5, v_5 = 0.05$



b) $v_1 = 0.5, v_5 = 0.50$



b) $v_1 = 0.5, v_5 = 0.95$

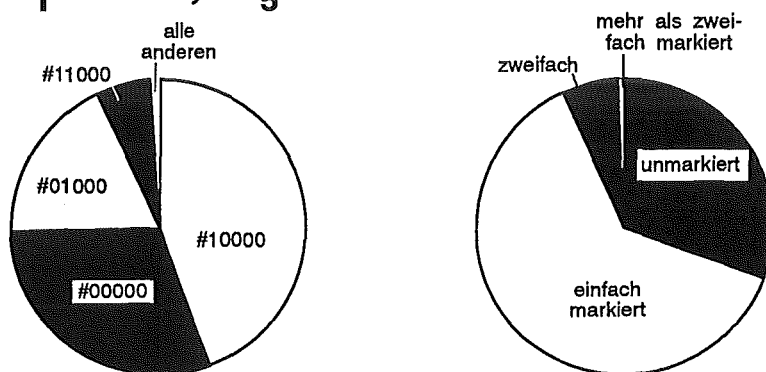


Abbildung 9.4: Isotopomeren-Verteilung im Xyl5P-Pool des Pentosephosphatweg-Beispiels (Parameter und Annahmen wie in Abb. 9.3).

9.2.3 Isotopomeren-Verteilungen

Für einen weiteren stationären Simulationslauf werden mit drei ausgewählten Werten für das Flußverhältnis $v_5^{\text{net}}/v_1^{\text{net}}$ die entsprechenden Isotopomeren-Verteilungen im Xyl5P-Pool berech-

net (Abbildung 9.4). Erneut sind alle Flüsse als unidirektional angenommen. Es zeigt sich, daß bei unidirektionalen Flüssen V2, V3 und V4 praktisch nur wenige Isotopomeren in quantitativ relevanten Mengen entstehen. Dazu gehören auch einige mehrfach markierte Isotopomere, die direkten Aufschluß über die zyklischen Abläufe geben können. Treten in den Zwischenschritten V2, V3 und V4 hohe Austauschflüsse auf (wovon auszugehen ist), so sind weitere mengenmäßig relevante Isotopomeren-Pools zu erwarten.

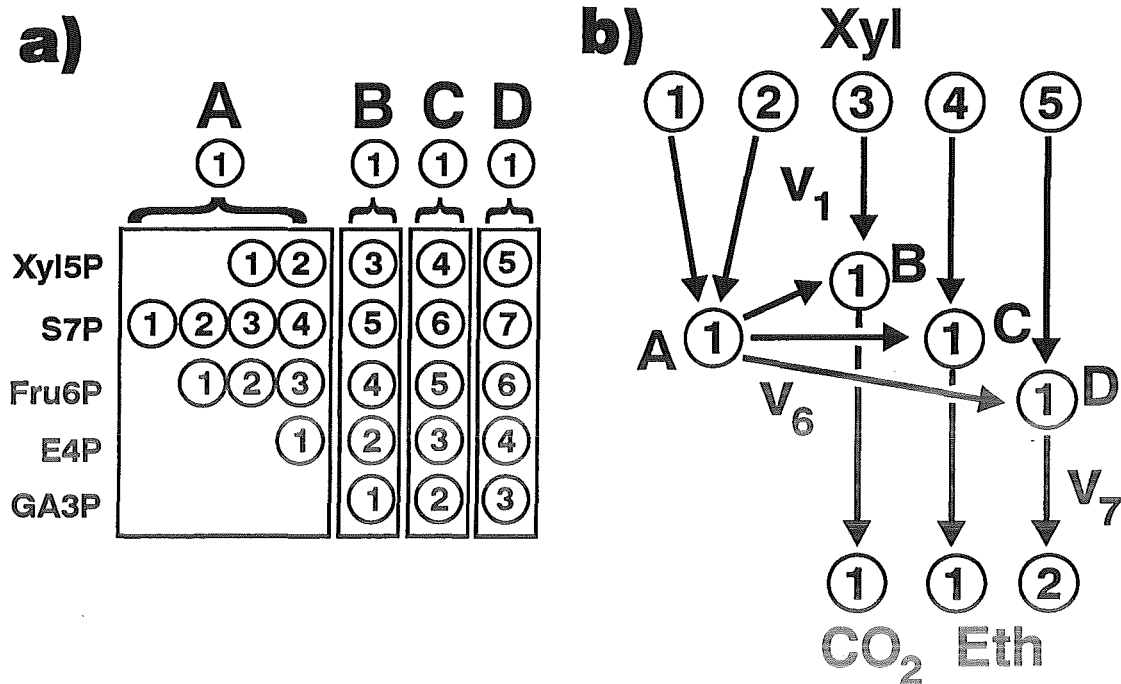


Abbildung 9.5: a) Bei hohen Austauschflüssen in V2, V3 und V4 in schnellem Markierungsgleichgewicht stehende C-Atom-Pools des Pentosephosphatwegs. b) Durch Pool-Lumping auf die Zusammenhangskomponenten A, B, C, D bzgl. dieser Reaktionsschritte reduziertes C-Atom-Netzwerk des Pentosephosphatwegs (nach [SSH92]).

9.2.4 Instationäre Simulationsergebnisse

Das Ergebnis eines instationären Simulationslaufs mit denselben Parametern wie in Abbildung 9.1 ist in Abbildung 1.4 dargestellt. Die benötigte Rechenzeit für diesen Simulationslauf ist in der Praxis vernachlässigbar. Das Runge-Kutta-Fehlberg-Verfahren für steife Systeme [KR79] benötigt nur etwa 10-20 Schritte bis zum Erreichen des stationären Zustands.

Diese Simulation gibt eine Vorstellung davon, wie schnell sich der isotopisch stationäre Zustand im Zuckerstoffwechsel eines schnell wachsenden Mikroorganismus einstellt, nämlich nach bereits etwa 5 Minuten. Dieser rasche Übergang ist mit der heute verfügbaren NMR-Meßtechnik kaum noch auflösbar. Die in der Literatur dokumentierten instationären Experimente gehen daher stets von Organismen mit geringeren Wachstumsraten, höheren Zellkonzentrationen oder größeren Pools aus. Beispielsweise vervielfachen sich alle Zeitkonstanten um den Faktor 10, wenn alle intrazellulären Poolgrößen mit 0.01 mol/l angenommen werden.

Die Abbildung 1.4 zeigt auch, daß das dynamische Verhalten von Markierungssystemen im Vergleich zu anderen dynamischen Systemen eher "uninteressant" ist. Die Zeitverlaufskurven zeigen überwiegend ein monoton ansteigendes Verhalten. Ein leichtes Überspringen wird nur

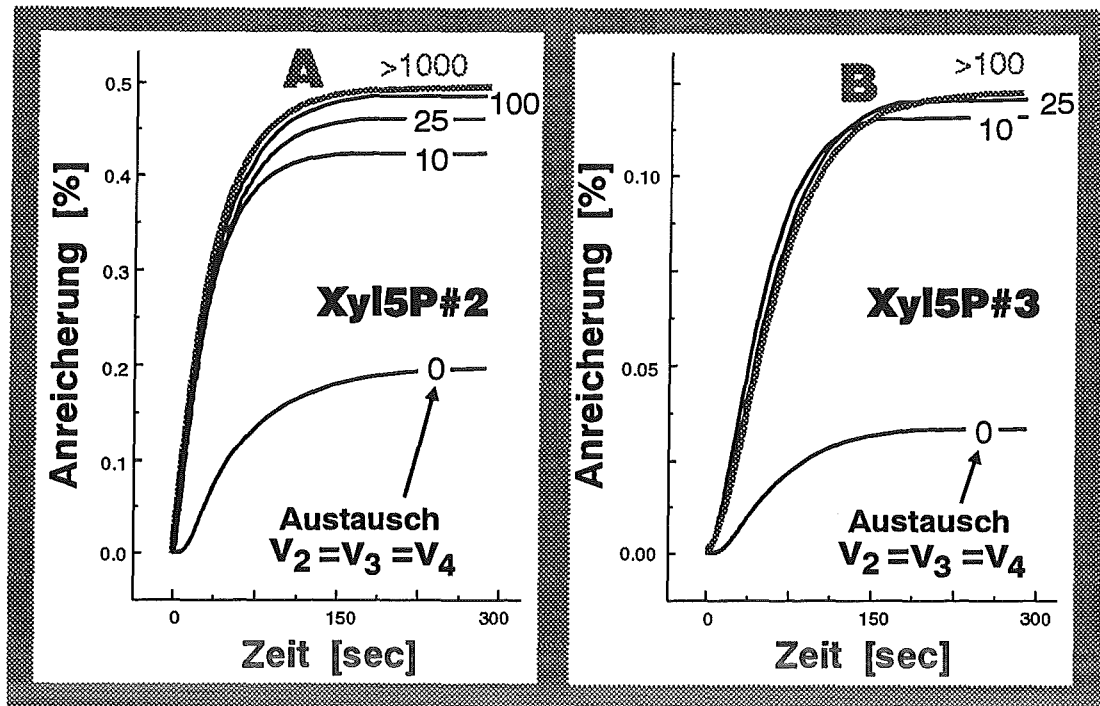


Abbildung 9.6: Vergleich der isotopisch instationären Zeitverläufe bei gleichzeitiger Variation der Austauschflüsse in V2, V3 und V4. Dargestellt sind die Markierungsverläufe von Xyl5P#2 und Xyl5P#3. Im Vergleich dazu die Zeitverläufe der Markierungsanteile in den Pools A und B des reduzierten Netzwerks aus Abb. 9.5b.

selten beobachtet. Dies war jedoch aufgrund der in Kapitel 7 gezeigten Dissipativität von C-Atom-Markierungssystemen bereits zu erwarten.

Variationsstudien bzgl. der Poolgrößen (hier nicht dargestellt) zeigen, daß bei verschiedenen Poolgrößen der größte im System auftretende Pool dominant ist, d.h. er bestimmt weitgehend die Zeitkonstanten der Markierungsanreicherung. Die Größen kleinerer Pools können dann in relativ weiten Grenzen variiert werden, ohne daß deutliche Veränderungen im instationären Verlauf auftreten (vgl. auch [WGKF⁺92, CFGC95]). Diese Beobachtung stützt die in der Literatur oft getroffene pauschale Annahme einer festen Poolgröße für kleine Pools, während große Pools immer quantitativ festgelegt werden müssen [WGKF⁺92].

Zum Abschluß wird nun die Auswirkung einer Variation der (üblicherweise angenommenen) Austauschflüsse in den Reaktionsschritten V2, V3, V4 des instationären Systems studiert. Dazu werden $v_2^{\text{sch}} = v_3^{\text{sch}} = v_4^{\text{sch}} = v_{2,3,4}^{\text{sch}}$ gesetzt und der Wert $v_{2,3,4}^{\text{sch}}$ ausgehend von Null schrittweise vergrößert. Aufgrund von Satz 8.4 ist dann eine zunehmende Angleichung der Markierungsanteile in den durch die Austauschflüsse miteinander verbundenen Pools zu erwarten.

Abbildung 9.5 zeigt das Ergebnis des dann durchführbaren Pool-Lumpings bei sehr großen Austauschflüssen [SSH92]. Es wird deutlich, daß das System im wesentlichen auf vier intrazelluläre Pools A, B, C und D reduzierbar ist, für die exemplarisch Xyl5P#2, Xyl5P#3, Xyl5P#4 und Xyl5P#5 stehen können. Aus Abbildung 9.5b wird ferner deutlich, daß sich bei einer Input-Markierung auf der ersten Position von Xyl die zusammengefaßten Pools B, C und D gleich verhalten. Daher sind in Abbildung 9.6 nur die Zeitverläufe für A und B gezeigt.

Abbildung 9.6 zeigt die instationären Verläufe der Markierungsanteile in den beiden exemplarischen Pools für zunehmende Austauschflüsse und zum Vergleich das Verhalten des durch Pool-Lumping reduzierten Systems aus Abbildung 9.5. Es wird deutlich, daß ab einem Aus-

tauschfluß von etwa 50 mol/h zum reduzierten System übergegangen werden kann. Dies ist zunächst ein erstaunlich hoher Wert, denn der Austauschfluß von 50 mol/h ist um den Faktor 100 größer als der Systemfluß von 0.5 mol/h! Dies wird jedoch dadurch relativiert, daß zugleich die Poolgrößen — als Zusammenfassung mehrerer C-Atom-Pools (vgl. Abb. 9.5) — um etwa den Faktor 10 anwachsen.

Trotzdem sind bei einem Austauschfluß von 50 mol/h bereits sehr lange Rechenzeiten für den Simulationslauf erforderlich, da das Differentialgleichungssystem dann offenbar steif wird.³ Dagegen kann das durch Pool-Lumping reduzierte System ohne merklichen Zeitaufwand berechnet werden. Zur Behandlung von instationären Systemen mit hohen Austauschflüssen sind also effiziente Integratoren für steife Systeme erforderlich [Gea71, HNW87, HW91, DHB94].

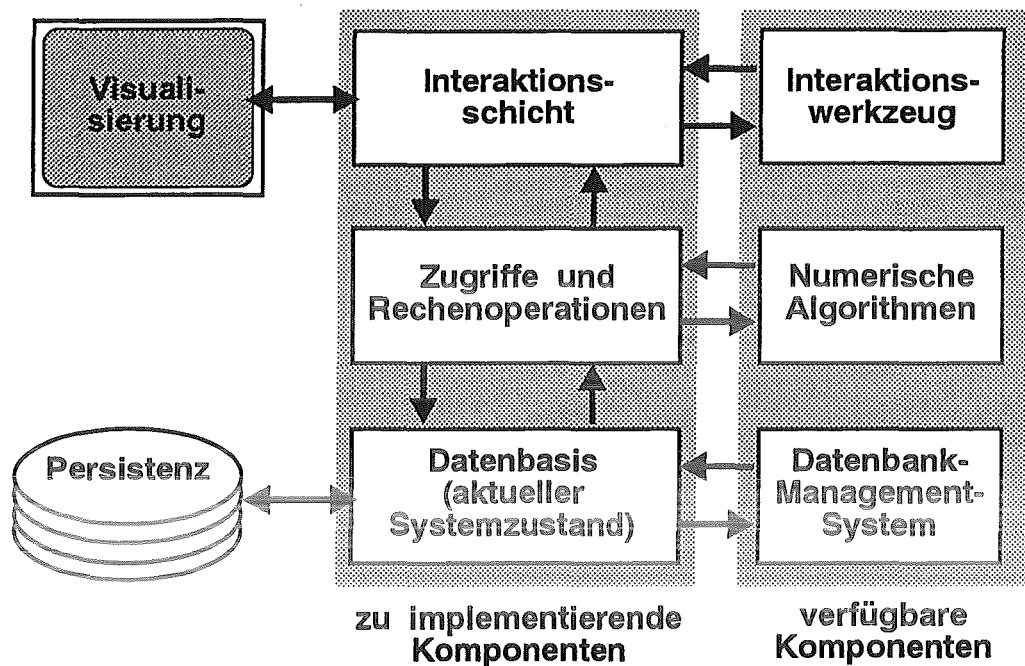


Abbildung 9.7: Allgemein anwendbares Architekturschema für Simulations- und Datenanalyse-Systeme zur Unterstützung experimenteller Arbeiten. Zugrunde liegt eine Datenbasis zur persistenten Speicherung des Systemzustandes. Darauf setzen problemspezifische Berechnungsoperationen auf. Die oberste Ebene bewerkstelligt die Benutzerinteraktion.

9.3 Implementierung und Architektur des Simulationssystems

9.3.1 Generische Architektur eines Datenanalysesystems

Für die Erstellung von Datenanalysesystemen kann eine generelle Architektur angegeben werden, die nicht nur auf das hier beschriebene Programmsystem zur Stoffflußanalyse, sondern auch auf viele andere, insbesondere kommerzielle Datenanalysesysteme zutrifft (Abb. 9.7, vgl. auch [WMWG95, WJW⁺95]).

1. Das Fundament des Systems wird durch eine Datenbasis gelegt, die alle für die aktuelle Datenanalysesitzung relevanten Informationen aufnimmt und verwaltet. Im konkreten

³Bei einem Austauschfluß größer 100 mol/h kann von mehr als einer Stunde Rechenzeit auf einem 486er PC ausgegangen werden.

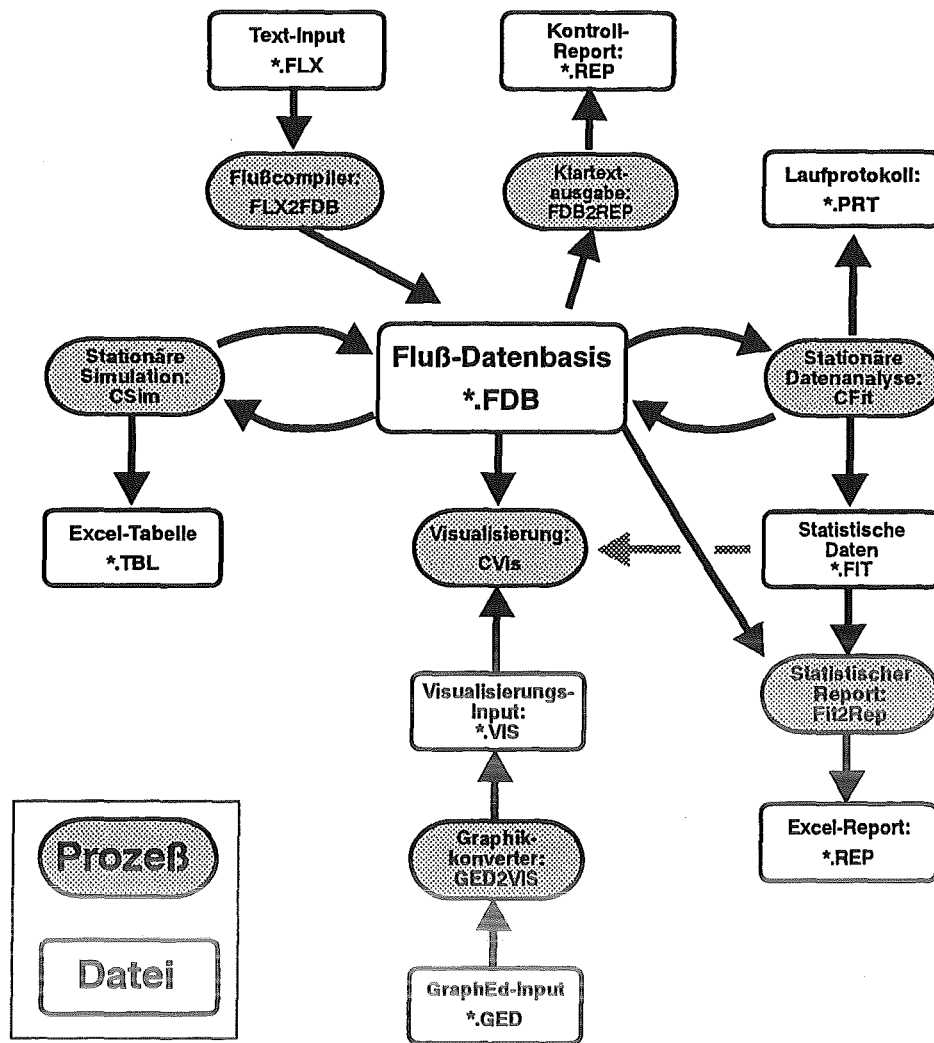


Abbildung 9.8: Datenflußdiagramm für die im Rahmen des Stoffflußanalyse-Projekts erstellten Softwarekomponenten zur stationären Simulation, Datenanalyse und Visualisierung (die Erklärung der einzelnen Komponenten erfolgt im Text).

Fall der Stoffflußanalyse enthält die Datenbasis alle Strukturmatrizen, Zustandsvektoren und Meßdaten des aktuell untersuchten Markierungssystems sowie weitere Parameter für spezielle Zwecke. Diese Daten werden für andere im Softwaresystem vorhandene Komponenten verfügbar gemacht. Umgekehrt können die Datenbestände unter Berücksichtigung bestimmten Zugriffsrechte und Konsistenzbedingungen von diesen verändert werden. Die Datenbasis sorgt dabei für die Konsistenz des Datenbestandes und verfügt zudem über die Möglichkeit, den gesamten aktuellen Systemzustand persistent abzulegen. Bei größeren Datenbeständen kann dafür ein Datenbanksystem verwendet werden [WJW⁺95]. Dies war jedoch bei der Flußanalyse nicht sinnvoll, weil keine Interdependenzen zwischen verschiedenen Experimenten bestehen und die Portierung des Datenbanksystems auf mehrere räumlich verteilte Rechnerplattformen einen zu hohen Arbeitsaufwand darstellt.

2. Auf der Datenbasis setzt eine Schicht von problemspezifischen Algorithmen auf, die in diesem Fall die Methoden zur Simulation, Parameteranpassung und Netzwerkanalyse enthält. Diese Verfahren können sich wiederum vorgefertigter, allgemein verfügbarer nume-

rischer oder algorithmischer Programmkomponenten bedienen. Die Werkzeuge dieser Schicht operieren auf der Datenbasis, indem sie die jeweils benötigten Informationen abfragen und die Ergebnisse ihrer Berechnungen wieder dort ablegen. Die Anwendung eines Werkzeugs verändert also den Zustand der Datenbasis. Beispielsweise bewirkt die Simulation eines Markierungsexperiments, daß anschließend die entsprechenden stationären Markierungsanteile in der Datenbasis abgelegt sind.

3. Eine dritte Architecturebene ermöglicht die Benutzerinteraktion und Visualisierung der Ergebnisse. Wichtig ist, daß solche Ein-Ausgabe-Operationen strikt von den beiden anderen Ebenen getrennt werden.⁴ Auf diese Weise kann das System leicht portiert und die Ein-Ausgabe gegebenenfalls später durch graphische Komponenten ergänzt werden. Die Jülicher Softwarekomponenten arbeiten zumeist mit Textdialogen. Inzwischen steht jedoch auch der Prototyp einer graphischen Visualisierungskomponente für metabolische Netzwerke bereit (vgl. Abb. 9.10).

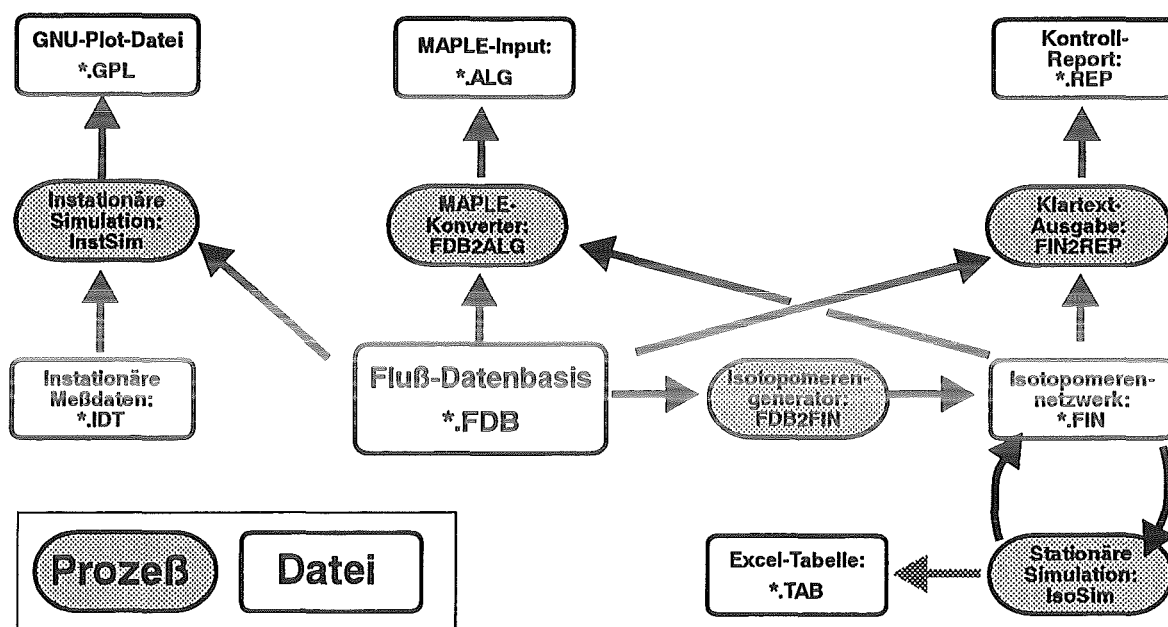


Abbildung 9.9: Datenflußdiagramm für die im Rahmen des Stoffflußanalyse-Projekts erstellten Softwarekomponenten zur Generierung von Isotopomeren-Netzwerken, instationären Simulation und algebraischen Analyse. (die Erklärung der einzelnen Komponenten erfolgt im Text).

9.3.2 Programmkomponenten für das Jülicher Stoffflußanalyse-Projekt

Das Jülicher Stoffflußanalysesystem ist nicht als ein einziger monolithischer Block implementiert worden. Vielmehr erlaubt der Datenbasisansatz ein räumlich verteiltes Arbeiten und die separate Erstellung einzelner Module durch verschiedene Mitarbeiter des Projekts. Gerade bei einem solchen Vorgehen ist die Möglichkeit zur persistenten Sicherung des Systemzustandes von zentraler Bedeutung. Denn nur auf diese Weise kann das Ergebnis eines Rechenschrittes in einer Datei abgespeichert und für den nächsten Schritt verfügbar gemacht werden.

⁴Hier sei auf die bekannte Model-Controller-View-Architektur der objektorientierten Programmierung verwiesen [GR89].

Sollen — wie auf Dauer intendiert — alle Komponenten zu einem integrierten System verbunden werden, so verlangsamt dieser immer wiederkehrende Zwischenspeicherschritt natürlich den interaktiven Prozeß. Daher kann er nach Fertigstellung aller Komponenten leicht aus jedem Modul entfernt werden, so daß sie sodann zu einem einheitlichen Ganzen verbunden werden.⁵ Aufgrund der objektorientierten Implementierung aller Systemkomponenten erfordert diese Verbindung zu einem Gesamtsystem nur einen geringen Zeitaufwand.

Derzeit liegen die Werkzeuge als eine Reihe von einzelnen Programmen vor, deren Datenaustausch in dem Datenflußdiagramm von Abbildung 9.8 und 9.9 dargestellt ist (vgl. hierzu [Wie94a]). Solche Diagramme (vgl. auch [You89]) haben sich bei der Durchführung des Projekts als nützliche Strukturierungshilfe erwiesen, da sie den einzelnen Mitarbeitern eine Positionsbestimmung innerhalb des Gesamtzusammenhangs ermöglichen. Ferner kann der Anwender mit ihrer Hilfe den vorhandenen Programmprototyp bedienen, zumal das Aufrufformat jedes Einzelprogramms einheitlich gestaltet wurde und die jeweils verwendeten Dateinamen eine bestimmte Endung erhalten. Einzelne Komponenten des Systems sollen nachfolgend kurz beschrieben werden.

9.3.3 Netzwerksynthese

Die Metabolit-, C-Atom- und Isotopomeren-Netzwerke können in der Praxis sehr komplex sein. Das Zentralstoffwechselmodell aus Abbildung 1.2 führt zu etwa 20 Metaboliten, 80 C-Atomen und 600 Isotopomeren. Außerdem sind 25 Reaktionsschritte, 120 C-Atom-Transitionen und 3200 Reaktionen zwischen Isotopomeren (vgl. auch [CFG95]) zu berücksichtigen. Während bei C-Atom-Matrizen eine manuelle Eingabe von Bilanzgleichungen und Matrizen noch denkbar wäre, ist dies für Isotopomeren-Netzwerke kaum noch durchführbar (vgl. dazu auch [WGKF⁺92] [CSKW83, HSG⁺93, Lee93, ZS94]). Ihre Komplexität kann daher nur mit Hilfe von Werkzeugen zur automatisierten Netzwerksynthese gehandhabt werden (siehe dazu [Wie94a, Sch96]).

Grundsätzlich birgt die explizite manuelle Formulierung von Gleichungen bzw. Matrizen die Gefahr einer erhöhten Fehlerquote. Letzterer kann jedoch durch Reduktion der Eingabe auf möglichst wenige Beschreibungselemente in Verbindung mit einer automatischen Konsistenzprüfung begegnet werden. Das in Definition 3.1 und Abbildung 9.2 präsentierte Eingabeformat stellt eine solche Form der Netzwerkrepräsentation dar. Ein speziell hierfür erstellter "Fluocompiler" (vgl. [ASU86]) generiert daraus automatisch die in den Kapiteln 4-6 beschriebenen Matrizen und legt sie in der Datenbasis ab. Die Netzwerksynthese basiert auf einer graphentheoretischen Beschreibung der Metabolit-, C-Atom- und Isotopomeren-Netzwerke, die in [Wie94a] vorgestellt ist (vgl. hierzu auch [MSS90, Val91, Hof93, Sch96]).

Zur leichteren Erstellung und Bearbeitung des in Abbildung 9.2 dargestellten, hierarchisch strukturierten Text-Eingabeformats wäre ein graphisches Werkzeug geeignet, wie es in vielen Betriebssystemen zur Konfigurationsverwaltung und Layout-Spezifikation von Benutzungsschnittstellen eingesetzt wird [You91]. Ein solches Werkzeug befindet sich aktuell noch in der Konzeptionsphase. Die Systemeingabedateien werden daher derzeit mit einem normalen Texteditor bearbeitet.

Nachdem die Netzwerkstruktur in Form des Metabolit- und C-Atom-Netzwerks in der Datenbasis abgespeichert wurde, kann das Isotopomeren-Netzwerk automatisch generiert werden. Aufgrund seines Umfangs in Verbindung mit der Überlegung, daß viele Algorithmen nur das C-Atom-Netzwerk benötigen, geschieht dieser Verarbeitungsschritt in einem separaten Modul. Das

⁵Im Prinzip entspricht dies der aus dem Software-Engineering bekannten Vorgehensweise des Rapid Prototyping [WP89, CS94, Mu190].

Isotopomeren-Netzwerk wird daher in einer eigenen Datei abgespeichert, die jederzeit zur Datenbasis hinzugeladen werden kann. Optional kann die Erzeugung des Isotopomeren-Netzwerkes auf die tatsächlich erreichbaren Isotopomeren beschränkt werden.

Ein weiteres Werkzeug, das der Fehlersuche und der Kontrolle der Netzwerkgeneratoren dient, ist ein Reportgenerator, der aus einer bestehenden Ausgabe des Flußcompilers und des Netzwerkgenerators, also der internen Datenrepräsentation, wieder eine Ausgabe in lesbarem Text erzeugt.

9.3.4 Numerische und algebraische Komponenten

Die in dieser Arbeit beschriebenen numerischen Verfahren zur Simulation und Auswertung (vgl. Kapitel 8, 10, 11) von Experimenten sind über mehrere (nicht immer als eigenständige Programme implementierte) Module verteilt, die jeweils auf der Flußdatenbasis operieren. Allen Werkzeugen gemeinsam ist eine Komponente zur Berechnung des Flußsystems aus gegebenen freien Flüssen nach Gleichung (8.2). Da alle numerischen Komponenten in Form von C++-Klassen implementiert sind, ist die Mehrfachnutzung der entsprechenden Quelltexte einfach möglich.

Die statistischen Daten, die von der Datenauswertung erzeugt werden, lassen sich stets auf die sogenannte Output-Sensitivitätsmatrix reduzieren, wie in Kapitel 11 (speziell Satz 11.2) erläutert wird. Diese Sensitivitätsmatrix wird von einem eigenen Werkzeug zur Erzeugung statistischer Reports verarbeitet, das entsprechend universell einsetzbar ist.

Ein großer Vorteil der gewählten Netzwerkrepräsentation durch eine Datenbasis besteht darin, daß auf dieser Grundlage leicht Formatkonverter programmiert werden können, die die vorhandene Information für andere — insbesondere kommerziell verfügbare — Werkzeuge zugänglich machen. Um die in [Wie95a, Wie96a] beschriebenen algebraischen Algorithmen anwenden zu können, wurde z.B. ein Formatkonverter für das Computeralgebra-System Maple [Kof94] implementiert.

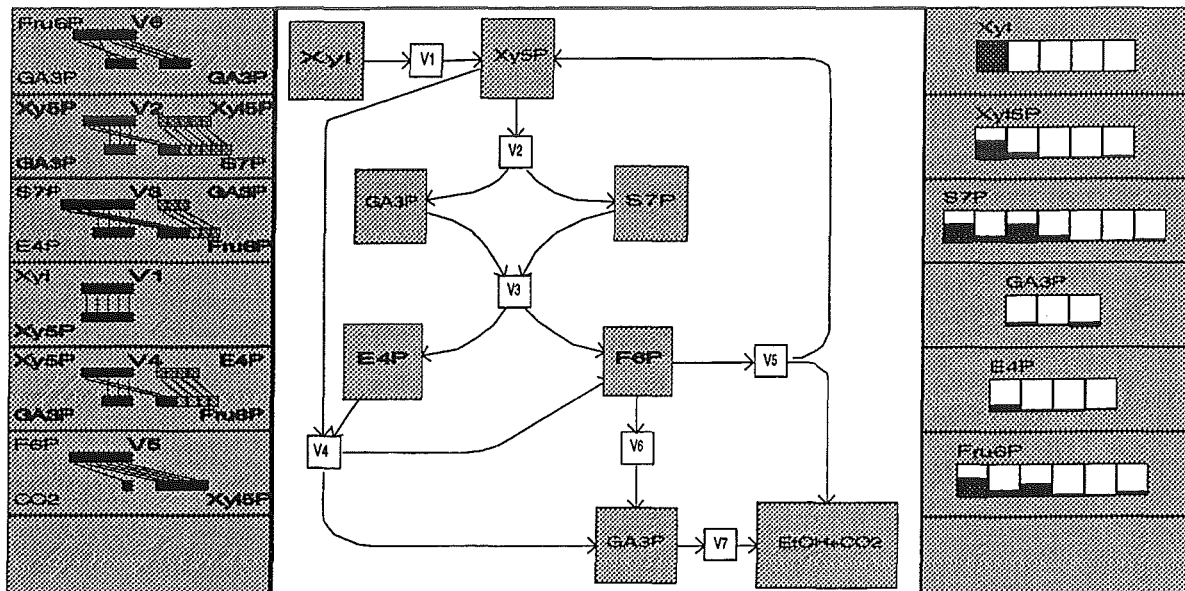


Abbildung 9.10: Prototyp eines Visualisierungswerkzeugs für stationäre Markierungsexperimente (Bildschirmabzug). Links: Auswählbare Darstellungen der C-Atom-Transitionen. Mitte: Metabolisches Netzwerk. Rechts: Auswählbare Darstellungen der prozentualen Anreicherungen. (vgl. auch Abb. 9.1)

9.3.5 Visualisierung

Für die graphische Netzwerkrepräsentation und die Visualisierung der Simulations- und Datenanalyseergebnisse wurde der Prototyp einer graphischen Oberfläche erstellt. Besonders arbeitsaufwendig ist hierbei die Programmierung einer Programmkomponente zum Erstellen und Editieren von graphischen Netzwerkrepräsentationen auf dem Bildschirm. Ein solches Werkzeug fällt in die allgemeine Klasse der Grapheneditoren.

In jüngster Zeit sind gleich mehrere Arbeiten zur Problematik solcher Grapheneditoren erschienen, und es liegen einige Implementationen für Unix-Systeme vor [Him89, Pau93]. Aus den genannten Arbeiten wird deutlich, daß die Mensch-Maschine-Schnittstelle bei Grapheneditoren ein bis heute nicht zufriedenstellend gelöstes Problem ist. Die interaktive Erstellung von Graphenvisualisierungen steht dabei vor zwei zentralen Problemen:

1. Umfangreiche Netzwerke erfordern effiziente Methoden zur ihrer Darstellung auf dem beschränkten Raum eines Bildschirms bei gleichzeitiger Gewährleistung eines schnellen interaktiven Zugriffs auf alle Komponenten.
2. Die Layout-Vorstellungen des Anwenders sind hochgradig spezifisch für das jeweilige Anwendungsgebiet. Beispielsweise bestehen für die Darstellung von Stoffwechselnetzwerken völlig andere Konventionen und ästhetische Vorstellungen als für das Zeichnen von elektrischen Schaltplänen. Aus diesem Grunde ist die Layout-Gestaltung bis heute nicht zufriedenstellend automatisierbar.

Bereits aus der knappen Skizzierung dieser Schwierigkeiten wird deutlich, daß die Entwicklung eines speziellen Grapheneditors für Stoffwechselnetzwerke einen im Rahmen eines interdisziplinären Projektes nicht vertretbarer Zeitaufwand erfordert. Daher wurde im Sinne des Prototyping auf den verfügbaren Netzwerkditor GraphEd zurückgegriffen [Him89]. Dies birgt gegenüber einer problemspezifischen Lösung den Nachteil, daß das Fremdwerkzeug nur in einem unidirektionalen Austausch mit dem Datenanalysesystem stehen kann. Insbesondere kann es nicht auf die in der Flußdatenbasis gespeicherten Netzwerkinformationen zugreifen. Daher muß der Anwender das graphische Netzwerk separat von dem durch die formale Repräsentation dargestellten Netzwerk ein zweites Mal entwerfen. Änderungen in der Netzwerkstruktur müssen ebenso in der Visualisierung nachgetragen werden.

Aufgrund dieses Kompromisses ließ sich eine Visualisierungskomponente für das Stoffflußanalysesystem relativ schnell entwickeln. Die Abbildung 9.10 zeigt einen Bildschirmabzug des derzeitigen Prototyps der Visualisierungskomponente.

9.3.6 Objektorientierte Methoden

Auf die implementierungstechnischen Details der einzelnen für das Stoffflußanalyseprojekt erstellten Programme soll hier nicht eingegangen werden. Hervorgehoben sei aber, daß der Einsatz objektorientierter Methoden [RBP⁺91, WJW⁺95] entscheidend zur Erstellung einer modularen Architektur beigetragen hat, die eine weitgehend voneinander entkoppelte Entwicklung der einzelnen Komponenten erlaubt.

Zur Implementierung wurde durchgehend die Programmiersprache C++ eingesetzt [ES90, Wie95b]. Von besonderem Nutzen bei der Erstellung der numerischen Komponenten war die Verwendung sogenannter Entwurfsmuster [Wie96b, GHJV95], die in besonderem Maße die Programmierung wiederverwendbarer Softwarekomponenten ermöglichen.

Teil IV

Auswertung von Experimenten

Kapitel 10

Stationäre Flußschätzung

Die Schätzung intrazellulärer Stoffflüsse auf der Basis gemessener Fluß- und Markierungsdaten kehrt die Fragestellung der Simulation um. Während bei einer Simulation der Markierungszustand des Systems aus vorgegebenen Flüssen berechnet wird, sind bei der Flußschätzung solche Flußparameter zu ermitteln, die einen durch Meßwerte gegebenen experimentellen Befund reproduzieren.¹ In der Terminologie des wissenschaftlichen Rechnens handelt es sich also um ein sogenanntes *inverses Problem* [Gro93]. Nachfolgend werden die in Jülich entwickelten numerischen Verfahren zur Lösung des stationären Schätzproblems vorgestellt. Dabei soll vor allem die Natur der auftretenden Probleme deutlich gemacht werden. Grundkenntnisse der Optimierung werden in diesem Kapitel allerdings vorausgesetzt. Auf die technischen Details der Implementierung kann hier nicht eingegangen werden (vgl. dazu [Sie96, Möl95]).

10.1 Das allgemeine Schätzproblem

10.1.1 Verschiedene Ansätze

In der Literatur können drei grundsätzliche Ansätze zur Lösung des Flußschätzproblems unterschieden werden:

1. Die weitaus meisten Autoren leiten explizite Formeln für die Berechnung der intrazellulären Flüsse aus den Markierungsdaten her. Dies entspricht der in Abschnitt 4.3 für das Beispiel aus Abbildung 3.2 entwickelten Gleichung (4.18). Dieser Ansatz verlangt für reale Stoffwechselsysteme die Lösung eines sehr großen algebraischen Gleichungssystems. Es ist bekannt, daß derartige algebraische Probleme eine extrem hohe algorithmische Komplexität haben (vgl. dazu [Wie95a]). Daher wurden die in der Literatur dokumentierten algebraischen Umformungen in der Regel nur für einen Abschnitt des Stoffwechsel (meist den Citratzyklus) und unter Rückgriff auf stark vereinfachende Annahmen durchgeführt [Coh83, MSJ88, Lee93, MCDB94, STM⁺94, Szy95] (vgl. dazu auch Abschnitt 13).

Es läßt sich zeigen [Wie96a, PK92, Mes93], daß die Berücksichtigung bidirektionaler Reaktionsschritte die algorithmische Komplexität der Behandlung eines Netzwerkes dramatisch erhöht, da die Zahl der möglichen geschlossenen Zyklen stark anwächst. Die meisten in der Literatur belegten expliziten Ergebnisse beruhen daher auf der Annahme ausschließlich unidirektionaler Flüsse und gegebenenfalls schneller Gleichgewichte. Ein weiterer

¹In älteren Veröffentlichungen wird dieses Problem noch über Versuch und Irrtum gelöst [KKW79, SB79].

Nachteil algebraischer Methoden besteht darin, daß redundante Meßwerte nicht zur Verbesserung der statistischen Qualität des Ergebnisses herangezogen werden können (vgl. auch Abschnitt 12.2).

2. In wenigen Fällen wurde ein visueller Ansatz gewählt, bei dem eine geeignete graphische Auftragung der zu erwartenden Meßgrößen in Abhängigkeit von den freien intrazellulären Flüssen durchgeführt wurde [Lee93, ZS94] (vgl. z.B. Abb. 4.1 oder 5.4). Derartige Graphiken können leicht durch Simulationsläufe mit Parametervariationen gewonnen werden (vgl. Abb. 9.3). Der Ansatz wird häufig mit ergänzenden expliziten Rechnungen kombiniert. Das Ausgangsproblem wird dabei zunächst auf einige wenige Gleichungen mit geringer Parameterzahl reduziert, die explizit nur schwer oder gar nicht lösbar sind. Erst auf dieses modifizierte Problem wird eine graphische Methode angewandt.

Die Möglichkeiten graphischer Auftragung sind limitiert durch die verfügbaren Visualisierungstechniken. Daher können kaum mehr als zwei unbekannte Flüsse zur gleichen Zeit geschätzt werden. Darüber hinaus ist die graphische Methode genausowenig wie der explizite Ansatz in der Lage, redundante Meßwerte zu berücksichtigen. Andererseits erlaubt sie jedoch bei gleicher Anzahl von Unbekannten und Meßwerten eine elegante graphische Analyse der Sensitivität geschätzter Größen gegenüber Meßfehlern und die einfache Konstruktion von Vertrauensbereichen. In Abbildung 11.5 wird daher eine graphische Auftragung verwendet, um die Problematik der Gewinnung statistischer Vertrauensbereiche zu veranschaulichen.

3. Der einzige universell anwendbare Ansatz besteht in der Flußschätzung auf der Basis einer numerischen Parameteranpassung [WGKF⁺92, KCS93, MGWE⁺96]. Diese Methode beruht auf wiederholten Simulationsläufen mit dem Ziel, einen Flußzustand zu finden, der die gemessenen Daten möglichst gut reproduziert. Die Übereinstimmung zwischen der Vorhersage einer Simulation und den gemessenen Daten wird dabei durch ein gewichtetes Abstandsmaß quantifiziert, d.h. es wird eine Parameteranpassung vorgenommen. Zur Suche nach einem Flußzustand, der dieses Abstandsmaß minimiert, sind entsprechende Optimierungsalgorithmen zur gezielten Variation der freien Flüsse erforderlich.

Der Parameteranpassungsansatz ist in der Lage, beliebig viele (auch redundante) Meßwerte zu berücksichtigen. Andererseits tritt hierbei stets das Problem der Auffindung eines globalen Minimums des Abstandsmaßes auf. Dieses Problem ist mit vertretbarem Rechenaufwand bis heute nur empirisch lösbar (vgl. dazu auch [RR95, RR93]). Es bleibt letztlich der Intuition des Anwenders überlassen, ob er den gefundenen Flußzustand als globale Lösung akzeptiert. Die Zuverlässigkeit der Methode ist natürlich um so größer, je kleiner die Zahl der zu schätzenden Parameter (also die Dimension des Suchraums) und je größer die Zahl der Meßwerte (also die bereitgestellte Information) ist. Daher finden sich in der Literatur auch Kombinationen aus expliziten Ansätzen zur Reduktion der Dimension des Parameterraums und Parameteranpassungsansätzen [MSJ88].

In der vorliegenden Arbeit wird der Weg über eine Parameteranpassung beschritten, während graphische Methoden eher zu Veranschaulichungszwecken verwendet werden. Darüber hinaus konnten in [Wie95a, Wie96a] einige weitreichende explizite Aussagen für die algebraische Behandlung von Markierungssystemen gewonnen werden, deren detaillierte Ausführung den Rahmen dieser Arbeit sprengen würde. Zwar ist die explizite Methode auf komplexe Netzwerke nicht mehr anwendbar, jedoch verschafft sie bei Detailanalysen wertvolle Einsichten in die Natur von globalen Problemen, die bei der numerischen Flußschätzung auftreten (vgl. z.B. [Wie95a]).

Durch die Kombination von Simulationsstudien, Parameteranpassung und algebraischer Analyse kann dann eine gewisse Sicherheit bei der Bestimmung global optimaler Flußschätzer gewonnen werden.

10.1.2 KQ-Flußschätzung

Nachfolgend soll der in der Regressionsanalyse [SW89, BW88] üblicherweise verwendete Kleinste-Quadrate-Schätzer oder kurz KQ-Schätzer zur Lösung des Flußschätzproblems eingesetzt werden. Die Ausführungen beschränken sich dabei auf die Behandlung isotopisch stationärer C-Atom-Markierungsexperimente, die für die praktische Anwendung in der Jülicher Arbeitsgruppe die größte Relevanz haben. Zur Problematik instationärer Systeme sei auf [FHBS90, TDHW91, SSH92, CFGC95] verwiesen.

Das dem KQ-Schätzer zugrundeliegende Abstandsmaß ist die Fehlerquadratsumme. Diese bewertet die Residuen ε_w und ε_y , also die Differenzen zwischen den gemessenen Werten für Flüsse und Markierungsanteile und dem durch eine Simulation und die Meßgleichungen (6.11) vorhergesagten Ausgang des Experiments:

$$\varepsilon_w = w - M_w \cdot v \quad \text{und} \quad \varepsilon_y = y - M_y \cdot x$$

Dabei müssen die unterschiedlich großen statistischen Fehler der einzelnen Messungen berücksichtigt werden. Der Meßfehler eines bekanntermaßen stark verrauschten Meßwerts darf nicht so stark gewichtet werden wie dieselbe Abweichung bei einem hochgenauen Meßinstrument. Im Spezialfall unabhängiger Messungen mit Standardabweichungen $\sigma_{w,i}$ und $\sigma_{y,i}$ liegt eine Bewertung in Form einer gewichteten Quadratsumme nahe:

$$\sum_{i=1}^{\dim w} \left(\frac{\varepsilon_{w,i}}{\sigma_{w,i}} \right)^2 + \sum_{i=1}^{\dim y} \left(\frac{\varepsilon_{y,i}}{\sigma_{y,i}} \right)^2 \quad (10.1)$$

Wie in Abschnitt 6.2 diskutiert können im allgemeinen Fall darüber hinaus auch Korrelationen zwischen einzelnen Meßwerten vorliegen (vgl. Gleichung (6.16)), so daß die entsprechende Fehlerkovarianzmatrix Σ_w bzw. Σ_y in die Bewertung eingehen muß. Der Spezialfall unabhängiger Meßwerte entspricht gerade dem einfachsten Fall diagonalen Fehlerkovarianzmatrizen. Für beliebige Kovarianzmatrizen der Meßgrößen kann der Term (10.1) verallgemeinert werden zu:

$$\varepsilon_w^T \cdot \Sigma_w^{-1} \cdot \varepsilon_w + \varepsilon_y^T \cdot \Sigma_y^{-1} \cdot \varepsilon_y \quad (10.2)$$

Dies motiviert die Einführung einer abkürzenden Notation:

Definition 10.1 (Wurzel einer Matrix, gewichtete Norm) Sei Σ eine symmetrische positiv definite Matrix. Dann wird mit $\sqrt{\Sigma}$ die durch die Cholesky-Zerlegung [HJ85] definierte Matrix mit

$$\sqrt{\Sigma}^T \cdot \sqrt{\Sigma} = \Sigma$$

bezeichnet. Die durch Σ festgelegte gewichtete Norm ist dann gegeben durch:

$$\|\xi\|_{\Sigma}^2 = \xi^T \cdot \Sigma^{-1} \cdot \xi = \left(\sqrt{\Sigma}^{-1} \cdot \xi \right)^T \cdot \left(\sqrt{\Sigma}^{-1} \cdot \xi \right)$$

Damit kann die gewichtete Fehlerquadratsumme (10.2) kurz geschrieben werden als

$$\|\varepsilon_w\|_{\Sigma_w}^2 + \|\varepsilon_y\|_{\Sigma_y}^2 = \|w - M_w \cdot v\|_{\Sigma_w}^2 + \|y - M_y \cdot x\|_{\Sigma_y}^2 \quad .$$

Sie ist nach dem Gauß-Markov-Theorem [Arn90] das in einem gewissen Sinne optimale Fehlermaß zur Berechnung eines Parameterschätzers.

Die in Abschnitt 6.1 entwickelten, durch Gleichungen und Ungleichungen formulierten Nebenbedingungen schränken die Variationsmöglichkeiten für v und x zusätzlich ein. Die Definition des Flußschätzers muß diese Nebenbedingungen berücksichtigen. Im Fall eines stationären Markierungsexperiments ergibt sich unter Berücksichtigung der in Abschnitt 6.3 zusammengestellten Strukturen die nachfolgende²

Definition 10.2 (KQ-Flußschätzer, Fehlerquadratsumme) *Der Kleinste-Quadrate-Schätzer (KQ-Schätzer) für die Stoffflußanalyse unter stationären Bedingungen ist das (globale) Minimum des nachfolgenden Optimierungsproblems:*

$$\begin{aligned} &\text{Minimiere die Funktion } \kappa(v, x) \stackrel{\text{def}}{=} \|w - M_w \cdot v\|_{\Sigma_w}^2 + \|y - M_y \cdot x\|_{\Sigma_y}^2 \\ &\text{unter den Bedingungen} \quad 1. \quad 0 = P^{\text{inp}} \odot v \cdot x^{\text{inp}} + P \odot v \cdot x^{\text{inp}} \\ &\quad 2. \quad v = \Phi_\alpha \circ \Phi^{[0,1]} \left(\begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} \right) \\ &\quad 3. \quad N \cdot \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} = n \\ &\quad 4. \quad U \cdot \begin{pmatrix} v^{\text{net}} \\ v^{\text{xch}[0,1]} \end{pmatrix} \geq u \\ &\quad 5. \quad 0 \leq x \leq 1 \quad . \end{aligned}$$

Die zu optimierende Funktion $\kappa(v, x)$ heißt Fehlerquadratsumme. Ihr globales Minimum unter den oben aufgeführten Nebenbedingungen — also der KQ-Schätzer für die wahren Werte von (v, x) — wird mit $(\hat{v}, \hat{x})$ bezeichnet.

10.1.3 Vereinfachung des Schätzproblems

Bevor das entwickelte numerischen Verfahren zur Berechnung des Schätzers vorgestellt wird, soll das Problem zunächst vereinfacht werden. Dazu können die Nebenbedingungen 1., 2. und 3. mittels Parametrisierung des Lösungsraums durch die freien Flüsse n^{free} nach Gleichung (8.2) und (8.3) eliminiert werden in der Form:

$$x = x(n^{\text{free}}) = x = \Gamma \circ \Phi_\alpha \circ \Phi^{[0,1]} \circ \Psi(n^{\text{free}})$$

Schließlich gilt nach Satz 7.13, daß die Ungleichheitsbeziehung $0 \leq x \leq 1$ für $x = \Gamma(v)$ stets erfüllt ist, sie kann daher entfallen. Damit reduziert sich das Optimierungsproblem aus Definition 10.2 auf das gleichwertige Problem:

²Bei Hinzufügen der bimolekularen Isotopomeren-Transitionsmatrix Q ist diese Definition auch für Isotopomeren-Anteile zutreffend.

Satz 10.1 (äquivalentes KQ-Schätzproblem) Das KQ-Flußschätzproblem aus Definition 10.2 ist äquivalent zur der Lösung $\hat{n}^{\text{free}}$ des Problems

$$\begin{aligned} \text{Minimiere} \quad \kappa(n^{\text{free}}) &= \|w - M_w \cdot \Phi_\alpha \circ \Phi^{[0,1]} \circ \Psi(n^{\text{free}})\|_{\Sigma_w}^2 \\ &\quad + \|y - M_y \cdot \Gamma \circ \Phi_\alpha \circ \Phi^{[0,1]} \circ \Psi(n^{\text{free}})\|_{\Sigma_y}^2 \\ \text{unter der Bedingung} \quad U \cdot \Psi(n^{\text{free}}) &\geq u \end{aligned}$$

und anschließender Setzung von $\hat{v} = \Phi_\alpha \circ \Phi^{[0,1]} \circ \Psi(n^{\text{free}})$ und $\hat{x} = \Gamma(\hat{v})$.

10.1.4 Das allgemeine Regressionsmodell

Wie gezeigt, ist das KQ-Schätzproblem im wesentlichen durch n^{free} bestimmt. In der Statistik hat sich für den zu schätzenden Parameter der Buchstabe Θ und entsprechend $\hat{\Theta}$ für den KQ-Schätzer eingebürgert. Daher soll im folgenden Θ anstelle von n^{free} verwendet werden. Um die Notation weiter zu entlasten werden folgende Zusammenfassungen vorgenommen:

1. Der Fluß- und Markierungsmeßvektor, die Kovarianzmatrizen der Messung und die Meßfehlervektoren werden jeweils zusammengefaßt zu

$$\eta = \begin{pmatrix} w \\ y \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma_w & 0 \\ 0 & \Sigma_y \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_w \\ \varepsilon_y \end{pmatrix}.$$

2. Der nichtlineare Zusammenhang zwischen dem System-Input Θ und dem (unverrauschten) System-Output wird beschrieben durch die Funktion

$$F(\Theta) = \begin{pmatrix} M_w \cdot \Phi_\alpha \circ \Phi^{[0,1]} \circ \Psi(n^{\text{free}}) \\ M_y \cdot \Gamma \circ \Phi_\alpha \circ \Phi^{[0,1]} \circ \Psi(n^{\text{free}}) \end{pmatrix}.$$

3. Die Nebenbedingungen an Θ werden mit einer Funktion H formuliert als:

$$H(\Theta) \stackrel{\text{def}}{=} U \cdot \Psi(\Theta) - u \geq 0$$

Damit reduzieren sich die Modellgleichungen auf den folgenden allgemeinen Typ [Páz93]:

Definition 10.3 (nichtlineares Regressionsmodell, KQ-Schätzung unter Nebenbedingungen)
Ein nichtlineares Regressionsmodell ist gegeben durch

$$\eta = F(\Theta) + \varepsilon \quad \text{mit} \quad \text{Cov}(\varepsilon) = \Sigma$$

mit einer stetig differenzierbaren Funktion F . Der KQ-Schätzer für Θ unter den stetig differenzierbaren Nebenbedingungen $H(\Theta) \geq 0$ ist definiert durch³

$$\hat{\Theta} = \arg \min_{H(\Theta) \geq 0} \|\eta - F(\Theta)\|_{\Sigma}^2.$$

Außer dieser allgemeiner Problemstruktur wird nachfolgend nur verwendet, daß $H(\Theta)$ hier speziell eine lineare Nebenbedingung ist. Jedoch lassen sich die nachfolgend beschriebenen Methoden auch auf den nichtlinearen Fall erweitern.

³Das Argument-Minimum $\arg \min$ meint die Stelle, an der das globale Minimum angenommen wird.

10.1.5 Numerische Problematik des Schätzproblems

Die Schwierigkeiten, die bei der numerischen Behandlung des Schätzproblems entstehen, resultieren aus den in den vorangehenden Kapiteln durchgeführten Analysen und sind nachfolgend noch einmal kurz zusammengestellt:

1. Die Verwendung der nicht-kompaktifizierten Austauschflüsse v^{xch} ist zur Berechnung des Flußschätzers nicht empfehlenswert, da diese bei ungünstig verteilten Meßfehlern beliebig groß werden können (vgl. Abschnitt 5.2, 8.2) und dabei zugleich ihren Einfluß auf das System verlieren. Stattdessen ist — wie hier durchgeführt — die Verwendung kompaktifizierter $[0,1]$ -Austauschflüsse angebracht. Insbesondere verschwindet im allgemeinen die Ableitung nach $v_i^{xch[0,1]}$ am Rand des Intervalls $[0, 1]$ nicht (vgl. Satz 8.3 und Abbildung 5.7), d.h. das Sensitivitätsproblem bei großen Austauschflüssen tritt nicht auf.
2. Kompaktifizierte Austauschflüsse $v^{xch[0,1]}$ nahe oder gleich Eins führen allerdings trotzdem zu Konditionsproblemen bei der Lösung des Flußgleichungssystems, die jedoch mit den Methoden aus Satz 8.3 effizient gelöst werden können.
3. In der nicht geglätteten Formulierung (d.h. mit dem Glättungsparameter $\alpha = 0$) in Definition 5.2 ist die Transformation Φ_α nicht überall differenzierbar. Dieses Problem kann in der Praxis leicht durch glatte Approximation von $F(\Theta)$ mit Hilfe des in Definition 5.2 eingeführten Glättungsparameters $\alpha > 0$ gelöst werden. Sodann wird eine Folge von Problemen mit monoton fallendem α gelöst. Dieser Approximationsansatz ist eine in der Optimierungstheorie übliche Vorgehensweise [Spe93].

In der Praxis kann wie folgt vorgegangen werden: Nach Satz 5.1 und 5.3 ist die Nichtdifferenzierbarkeitsstelle für einen Reaktionsschritt durch einen verschwindenden Nettofluß gekennzeichnet. Es ist aber kaum zu erwarten, daß ein Nettofluß innerhalb der numerischen Genauigkeit exakt gleich Null ist (es sei denn, es liegt ein systematischer Fehler bei der Modellformulierung vor). In der Praxis genügt es daher, mit Werten $\alpha > 0$ zunächst die Richtung der im System auftretenden Flüsse grob zu bestimmen und diese Richtung dann durch Hinzunahme einer neuen Ungleichheits-Nebenbedingung für die entsprechenden Nettoflüsse zu fixieren. Damit wird die Nichtdifferenzierbarkeitsstelle aus dem zulässigen Bereich ausgeschlossen. Danach kann $\alpha = 0$ gesetzt werden.

4. Optimierungsalgorithmen mit hoher Konvergenzordnung nahe des Optimums benötigen die Ableitung $\partial F / \partial \Theta$ [Fle87]. Diese kann auf der Basis von Gleichung (8.4) mit Hilfe der Kettenregel sowie den Sätzen 5.4, 8.1 und 8.3 auf numerisch stabile Weise und zeiteffizient berechnet werden.
5. Auf die Berücksichtigung der Nebenbedingungen kann allerdings nicht verzichtet werden (vgl. Abschnitt 6.1). Dies ist das einzige verbleibende Problem, das mit den bisher entwickelten Methoden noch nicht gelöst werden kann.

Die in den letzten Jahren gesammelte Erfahrung zeigt, daß die Bewältigung der genannten Schwierigkeiten für die Praxis von großer Relevanz ist. Aus ihrer unvollständigen Berücksichtigung resultieren numerische Konditions- und Konvergenzprobleme. Beispielsweise zeigten erste, mit einfachen Methoden durchgeführte Ad-hoc-Flußschätzungen mit vorliegenden Datensätzen zwar stabile Werte für die Nettoflußschätzung, jedoch eine extrem schlechte Kondition

bzgl. der Austauschflüsse. Darüber hinaus kam es bei Anwendung des einfachen Nelder-Mead-Verfahrens [PFTV88] zu frühzeitiger Konvergenz aufgrund inexakter Markierungsberechnungen, Überschreitung des zulässigen Bereichs und extrem langen Rechenzeiten. Das nachfolgend beschriebene neu entwickelte Verfahren benötigt dagegen nur wenige (d.h. etwa 10-20), jedoch rechenaufwendigere Schritte für ein realistisches Stoffwechselnetz unter voller Berücksichtigung der Nebenbedingungen [Sie96].

10.2 Optimierung unter Nebenbedingungen

10.2.1 Mathematische Programmierung

Das Flußschätzproblem stellt ein Problem der Mathematischen Programmierung dar, d.h. ein Optimierungsproblem unter Gleichungs- und Ungleichungs-Nebenbedingungen [Zou91]. Dieses Problem wird in Abbildung 10.1 veranschaulicht.

Definition 10.4 (Nichtlineares Programm) Ein Nichtlineares Programm ist ein Minimierungsproblem auf dem Raum $\mathbb{R}^n$, gegeben durch:

$$\begin{array}{ll} \text{Minimiere die Funktion} & f(\Theta) \\ \text{unter den Gleichheitsbedingungen} & g(\Theta) = 0 \\ \text{und den Ungleichheitsbedingungen} & h(\Theta) \geq 0 \end{array} .$$

Dabei ist f eine skalare und g, h sind vektorwertige Funktionen eines vektoriellen Arguments Θ , die hinreichend oft⁴ differenzierbar sind.

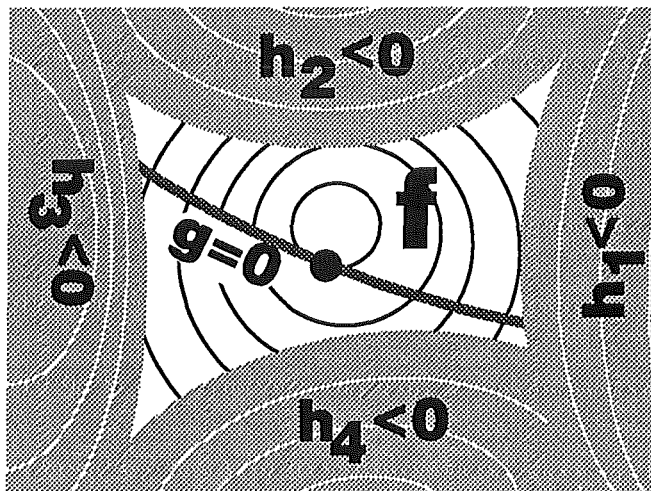


Abbildung 10.1: Problemstellung der Nichtlinearen Programmierung, dargestellt mit Hilfe von Niveauliniendiagrammen. Es ist das Minimum einer skalaren Funktion $f(x)$ entlang der durch die Beziehung $g(x) = 0$ definierten Menge zu finden. Das Suchgebiet wird durch die Bedingungen $h_i(x) \geq 0$ eingeschränkt. Der Punkt kennzeichnet das gesuchte Optimum.

⁴Auf die Einzelheiten der Problemformulierung und die präzisen Voraussetzungen der entsprechenden Verfahren wird hier nicht eingegangen. Siehe dazu [Fle87].

Sind die Gleichheits-Nebenbedingungen linear so können sie leicht durch Parametrisierung des entsprechenden linearen Lösungsraumes eliminiert werden. Dazu ist die Singulärwertzerlegung als numerisch stabiles Verfahren einsetzbar [DHB94]. Dies ist wie in Abschnitt 8.1 beschrieben im Falle der stöchiometrischen Gleichungen bereits erfolgt. Da im konkreten Fall somit die Gleichheitsbedingungen bereits eliminiert wurden, wird $g(\Theta)$ nachfolgend meist nicht eigens aufgeführt.

10.2.2 Aktive und inaktive Nebenbedingungen

Durch die Nebenbedingungen $h(\Theta) \geq 0$ wird im Raum $\mathbb{R}^n$ ein Gebiet eingegrenzt, in dem die Lösung des Optimierungsproblems zu suchen ist. Ein beliebiger Punkt $\Theta \in \mathbb{R}^n$ kann entweder im Inneren dieses Gebietes, auf seinem Rand oder außerhalb liegen. Dafür haben sich in der Mathematischen Programmierung entsprechende Begriffe eingebürgert (vgl. Abb. 10.2):

Definition 10.5 (zulässiger Punkt, aktive Nebenbedingungen) Sei durch f, g, h ein Nichtlineares Programm gemäß Definition 10.4 spezifiziert. Dann heißen die Punkte in

$$\{\Theta \in \mathbb{R}^n \mid g(\Theta) = 0 \wedge h(\Theta) \geq 0\}$$

zulässig und für einen zulässigen Punkt Θ heißt

$$\text{Act}(\Theta) \stackrel{\text{def}}{=} \{i \mid h_i(\Theta) = 0\} \subset \{1, \dots, \dim h\}$$

die Indexmenge der aktiven Nebenbedingungen.

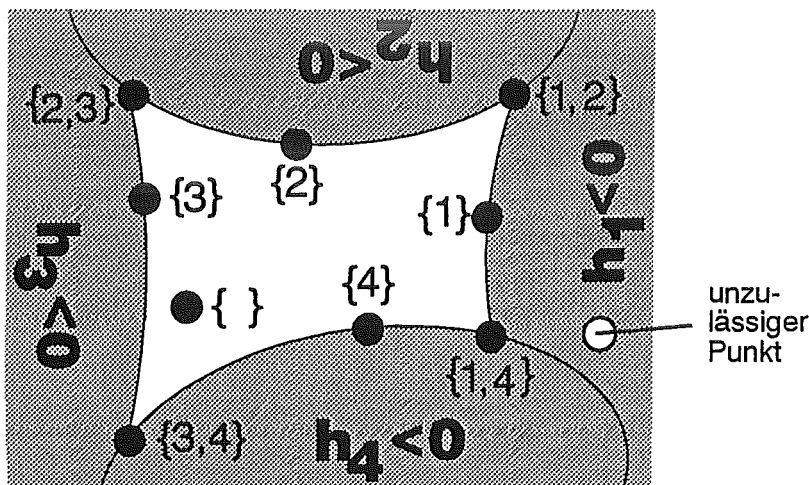


Abbildung 10.2: Indexmengen der aktiven Nebenbedingungen für verschiedene Punkte des eingeschränkten Suchraums.

In der Praxis der Nichtlinearen Programmierung wird meist vorausgesetzt, daß zumindest ein zulässiger Punkt existiert und explizit angebbar ist [Spe93]. Ferner sollte das zulässige Gebiet zusammenhängend und kompakt sein. Im Falle linearer Nebenbedingungen können Zusammenhang und Kompaktheit effektiv entschieden werden [PFTV88]. Insbesondere ist dann die Menge der zulässigen Lösungen konvex.

10.2.3 Die Kuhn-Tucker-Bedingung

Wie schon von der Optimierung ohne Nebenbedingungen bekannt, suchen auch Lösungsverfahren für allgemeine mathematische Programme nach einem lokalen Optimum der Funktion f auf dem durch die Nebenbedingungen eingeschränkten Bereich. Für eine solche Lösung kann die globale Optimalität nur in speziellen Fällen (z.B. bei konvexen Problemen) garantiert werden.⁵ In der Praxis "vergewissert" man sich durch mehrfaches Aufsetzen des jeweiligen Optimierungsverfahrens mit verschiedenen Startpunkten über die Eindeutigkeit des Optimums.

Ein lokales Optimum eines Nichtlinearen Programms ist durch die sogenannte Kuhn-Tucker-Bedingung charakterisiert [Spe93]:

Satz 10.2 (Kuhn-Tucker-Bedingung) Für jedes lokale Minimum $\Theta^* \in \mathbb{R}^n$ eines Nichtlinearen Programms gemäß Definition 10.4 gibt es Vektoren $\lambda \in \mathbb{R}^{\dim g}$ und $\mu \geq 0 \in \mathbb{R}^{\dim h}$, für die gilt:

$$\begin{aligned} 1. \quad 0 &= \frac{\partial f}{\partial \Theta}(\Theta^*)^T - \frac{\partial g}{\partial \Theta}(\Theta^*)^T \cdot \lambda - \frac{\partial h}{\partial \Theta}(\Theta^*)^T \cdot \mu \\ 2. \quad 0 &= h(\Theta^*)^T \cdot \mu \end{aligned}$$

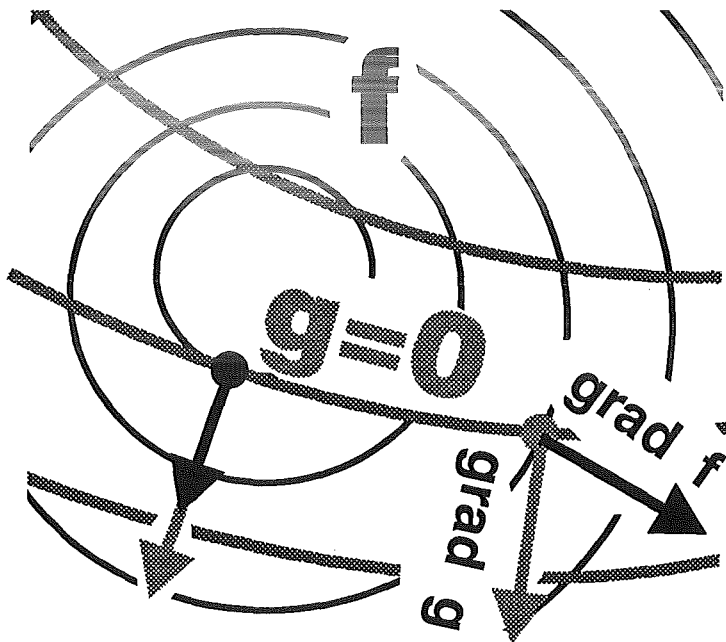


Abbildung 10.3: Lagrange-Kriterium für ein Optimum einer Funktion $f(x)$ entlang der durch $g(x) = 0$ definierten Menge. Im Beispiel liegt das Optimum an der Stelle, an der die Gradienten von f und g kollinear sind.

Die beiden Bedingungen sollen kurz erläutert werden.

⁵Das globale Optimierungsproblem für stetige Funktionen weist im allgemeinen eine extrem hohe Komplexität auf und ist heute nur unter schwachen Zusatzvoraussetzungen und für kleine Dimensionen algorithmisch lösbar [RR95].

1. Falls keine Nebenbedingungen gegeben sind, so ist dies genau die bekannte Gradientenbedingung $\frac{\partial f}{\partial \Theta}(\Theta^*) = 0$ für ein lokales Optimum.
2. Falls nur Gleichheits-Nebenbedingungen gegeben sind, so ist dies genau die Lagrange-Multiplikator-Regel für ein lokales Optimum unter Nebenbedingungen (Abbildung 10.3), also:

$$\frac{\partial f}{\partial \Theta}(\Theta^*)^T - \frac{\partial g}{\partial \Theta}(\Theta^*)^T \cdot \lambda = 0 \quad (10.3)$$

3. Kommen noch Ungleichheitsbedingungen hinzu, so können für jede Komponente h_i von h zwei Fälle auftreten:
 - (a) h_i ist inaktiv. Dann ist der entsprechende Eintrag des Multiplikators gerade $\mu_i = 0$.
 - (b) h_i ist aktiv. Dann kann h_i genau wie eine Gleichheitsbedingung gehandhabt werden, d.h. es gibt einen zugehörigen Lagrange-Multiplikator $\mu_i \neq 0$. Nun zeigt der Gradient $\frac{\partial h_i}{\partial \Theta}$ stets ins Innere des zulässigen Gebietes, woraus gefolgert werden kann [Spe93], daß für ein Minimum am Rand stets sogar $\mu_i > 0$ gelten muß.

Die Einträge $\mu_i \neq 0$ kennzeichnen somit gerade die aktiven Nebenbedingungen, womit auch die zweite Bedingung in des Satzes plausibel ist.

Die Kuhn-Tucker-Bedingungen stellen lediglich eine notwendige Bedingung für das Vorliegen eines lokalen Minimums dar, denn sie gelten unter Umständen auch für Maxima und Sattelpunkte. Trotzdem versuchen numerische Verfahren aus dem Bereich der Mathematischen Programmierung zunächst, Kuhn-Tucker-Punkte für Nichtlineare Programme zu finden. Da dabei zur Schrittweitenkontrolle stets ein Abstiegskriterium verwendet wird [Fle87, Spe93, BT95], ist jedoch die Konvergenz gegen einen nicht minimalen Kuhn-Tucker-Punkt fast sicher ausgeschlossen. Spätestens bei mehrfachem Starten des Verfahrens mit verschiedenen Startpunkten werden Sattelpunkte ausgeschlossen.

10.2.4 Optimierungsverfahren mit hoher Konvergenzordnung

Die Praxis der Optimierung lehrt, daß der entscheidende Zeitfaktor bei der Lösung stetiger Optimierungsprobleme die Konvergenzgeschwindigkeit in der Nähe des Optimums ist. Während selbst einfache Verfahren (bei gut konditionierten Problemen) schon nach wenigen Schritten die Nähe des Optimums erreichen, kann der weitere Verlauf noch hunderte von Iterationsschritten erfordern bis schließlich Konvergenz durch das Verfahren diagnostiziert wird. Bei Konditionsproblemen in der Nähe des Optimums — die bei KQ-Problemen häufig auftreten — verschärft sich diese Problematik erheblich [Spe93].

Alle Optimierungsverfahren mit superlinearer Konvergenzordnung in der Nähe des Optimums benötigen in irgendeiner Form die Ableitung $\partial f / \partial \Theta$ der Funktion $f(\Theta)$. Schon bei den klassischen Verfahren der Optimierung ohne Nebenbedingungen (z.B. dem Verfahren mit konjugierten Gradienten, dem BFGS-Verfahren oder dem Levenberg-Marquardt-Algorithmus [Fle87]) hat sich jedoch gezeigt, daß eine hohe Konvergenzgeschwindigkeit auf der Basis einer zusätzlichen Approximation der zweiten Ableitung $\partial^2 f / \partial \Theta^2$ von $f(\Theta)$ (also der Hesse-Matrix) erzielt werden kann [Spe93]. Eine explizite Berechnung der Hesse-Matrix ist aber schon bei Problemen mittlerer Größe nicht mehr praktikabel. Dagegen kann man ausgehend von ersten Ableitungen eine geeignete Approximation auf sehr verschiedenen Wegen gewinnen. Die bekanntesten Vorgehensweisen sind durch die cg- und BFGS-Update-Formeln sowie die Gauß-Newton-Linearisierung im Levenberg-Marquardt-Verfahren (vgl. Gleichung (10.8)) gegeben [Fle87].

Liegt eine Approximation der zweiten Ableitung vor, so kann die Funktion im aktuellen Arbeitspunkt Θ des Suchalgorithmus durch die Taylor-Reihe 2. Ordnung angenähert werden:

$$f(\Theta + \Delta\Theta) \approx f(\mathbf{x}) + \frac{\partial f}{\partial \Theta}(\Theta) \cdot \Delta\Theta + \frac{1}{2} \cdot \Delta\Theta^T \cdot \frac{\partial^2 f}{\partial \Theta^2}(\Theta) \cdot \Delta\Theta \quad (10.4)$$

Ersetzt man nun die Funktion $f(\Theta)$ durch diese quadratische Approximation, so ist ein guter Kandidat für den nächsten Arbeitspunkt durch den Newton-Schritt, also den Schritt zum Optimum der rechten Seite, gegeben:

$$\Delta\Theta = -\frac{1}{2} \left(\frac{\partial^2 f}{\partial \Theta^2}(\Theta) \right)^{-1} \cdot \frac{\partial f}{\partial \Theta}(\Theta) \quad (10.5)$$

Allerdings ist hierbei darauf zu achten, daß die Approximation (10.4) nicht über den zulässigen Bereich hinaus zur Extrapolation verwendet wird, d.h. der Newton-Schritt $\Delta\Theta$ im Vergleich zum Vorhersagehorizont nicht zu groß ist. Dieses Problem wird von den klassischen Algorithmen im Prinzip durch den Verzicht auf die zweite Ableitung in großer Entfernung vom Optimum in Kombination mit einer geeigneten Schrittweitensteuerung gelöst. Liegt jedoch das gesuchte Optimum innerhalb des Vorhersagehorizonts, so erzielt das reine Newton-Verfahren eine hohe Konvergenzordnung. Entscheidend ist nun, daß alle Verfahren in der Nähe des Optimums näherungsweise den unmodifizierten Newton-Schritt verwenden können, was letztlich die Ursache ihrer schnellen Konvergenz ist.

10.2.5 Sequentielle Quadratische Programmierung

Der derzeit in der Mathematischen Programmierung dominierende Lösungsansatz für Nichtlineare Programme ist die sogenannte Sequentielle Quadratische Optimierung (SQP) [Spe86, BT95]. Dabei wird das Prinzip der Taylor-Approximation auf die Nebenbedingungen erweitert. Allerdings werden diese nur mittels einer Taylor-Reihe erster Ordnung approximiert:

$$\begin{aligned} g(\Theta + \Delta\Theta) &\approx g(\Theta) + \frac{\partial g}{\partial \Theta}(\Theta) \cdot \Delta\Theta \\ h(\Theta + \Delta\Theta) &\approx h(\Theta) + \frac{\partial h}{\partial \Theta}(\Theta) \cdot \Delta\Theta \end{aligned}$$

Sodann wird das Nichtlineare Programm aus Definition 10.4 durch ein quadratisches Optimierungsproblem bzgl. $\Delta\Theta$ angenähert, das hier nur für den Fall linearer Nebenbedingungen angegeben wird:⁶

$$\begin{aligned} &\text{Minimiere die Funktion} && f(\mathbf{x}) + \frac{\partial f}{\partial \Theta}(\Theta) \cdot \Delta\Theta + \frac{1}{2} \cdot \Delta\Theta^T \cdot \frac{\partial^2 f}{\partial \Theta^2}(\Theta) \cdot \Delta\Theta \\ &\text{unter den Gleichheitsbedingungen} && g(\mathbf{x}) + \frac{\partial g}{\partial \Theta}(\Theta) \cdot \Delta\Theta = 0 \\ &\text{und den Ungleichheitsbedingungen} && h(\mathbf{x}) + \frac{\partial h}{\partial \Theta}(\Theta) \cdot \Delta\Theta \geq 0 \quad . \end{aligned} \quad (10.6)$$

Die Lösung dieses sogenannten Quadratischen Programms wird in den den folgenden Abschnitten behandelt. Ist eine solche Lösung $\Delta\Theta^*$ gefunden, so ist daraus ein geeigneter Schritt zum nächsten Arbeitspunkt zu ermitteln. Ausgehend vom aktuellen Arbeitspunkt Θ wird dazu eine Liniensuche entlang der Halbgeraden

$$\Theta + t \cdot \Delta\Theta^* \quad (t > 0)$$

⁶Bei nichtlinearen Nebenbedingungen muß auch die Hesse-Matrix von g mit einbezogen werden [Fle87].

durchgeführt. Dabei wird häufig eine sogenannte *Vorteilsfunktion* zur Beurteilung der Verbesserung verwendet. Diese berücksichtigt nicht nur den Funktionswert, sondern auch die Erfüllung der Nebenbedingungen. Details über Vorteilsfunktionen und das Problem der Schrittweitensteuerung können in [BT95] nachgelesen werden.

10.2.6 Quadratische Programmierung

Es soll nachfolgend davon ausgegangen werden, daß die Approximation der Hesse-Matrix in jedem Schritt so berechnet wird, daß sie stets positiv-definit ist. Die üblicherweise verwendeten Approximationsansätze erfüllen diese Voraussetzung. Das Nichtlineare Programm aus Gleichung (10.6) hat dann die folgende spezielle Gestalt:

Definition 10.6 (Quadratisches Programm) Ein Quadratisches Programm auf $\mathbb{R}^n$ ist ein Optimierungsproblem der Form:

$$\begin{array}{ll} \text{Minimiere die Funktion} & \mathbf{a}^T \cdot \Delta\Theta + \Delta\Theta^T \cdot \mathbf{A} \cdot \Delta\Theta \\ \text{unter den Gleichheitsbedingungen} & \mathbf{B} \cdot \Delta\Theta = \mathbf{b} \\ \text{und den Ungleichheitsbedingungen} & \mathbf{C} \cdot \Delta\Theta \geq \mathbf{c} \end{array}$$

Dabei ist $\mathbf{A}$ eine symmetrische positiv-definite Matrix. $\mathbf{B}, \mathbf{C}$ sind beliebige Matrizen, und $\mathbf{a}$ ist ein Vektor.

Nachfolgend werden zur Vereinfachung wieder die Gleichheitsbedingungen weggelassen, d.h. $\mathbf{B} = \mathbf{0}$. Das durch die Ungleichheits-Nebenbedingungen definierte Gebiet ist — als Schnittmenge von linearen Halbräumen — stets konvex. Es ist also das Minimum einer konvexen quadratischen Funktion in einem konvexen Gebiet zu suchen. Dieses Problem hat bekanntlich eine eindeutige Lösung [Fle87].

Gesucht sind nun Algorithmen, die das Minimum zuverlässig und mit wenigen Schritten finden. Bekanntlich hat bereits das Problem der Linearen Programmierung (also für $\mathbf{A} = \mathbf{0}$) eine hohe algorithmische Komplexität [PFTV88]. Dies gilt um so mehr für das quadratische Problem, wie nun kurz illustriert werden soll.

Sei dazu $\Delta\Theta^*$ die Lösung des Quadratischen Programms. Angenommen wird, daß die Menge $\text{Act}(\Delta\Theta^*)$ bekannt ist und die Matrix $\mathbf{P}_{\text{Act}(\Delta\Theta^*)}$ den Raum $\mathbb{R}^{\dim c}$ auf die aktiven Koordinaten projiziert.⁷ Dann reduziert sich das Quadratische Programm auf das gleichheitsrestringierte Problem:

$$\begin{array}{ll} \text{Minimiere die Funktion} & \mathbf{a}^T \cdot \Delta\Theta + \Delta\Theta^T \cdot \mathbf{A} \cdot \Delta\Theta \\ \text{unter den Gleichheitsbedingungen} & \mathbf{P}_{\text{Act}(\Delta\Theta^*)} \cdot \mathbf{C} \cdot \Delta\Theta = \mathbf{P}_{\text{Act}(\Delta\Theta^*)} \cdot \mathbf{c} \end{array}$$

Dieses ist mit Hilfe des Lagrange-Kriteriums aus Gleichung (10.3) mit einem einzigen Matrizen-Rechenschritt lösbar. Leider ist jedoch $\text{Act}(\Theta^*)$ nicht bekannt, so daß eine Suchstrategie konstruiert werden muß, die das erschöpfende Durchsuchen aller $2^{\dim c}$ Teilmengen der Indexmenge $\{1, \dots, \dim c\}$ erspart.

⁷ Anders gesagt operiert $\mathbf{P}_{\text{Act}(\Delta\Theta^*)}$ auf einem Vektor $\mathbf{c}$ durch Streichen der nicht aktiven Einträge und auf einer Matrix durch Streichen der entsprechenden Zeilen.

10.2.7 Das Verfahren von Goldfarb und Idnani

Bis in die jüngste Zeit hinein wird an der Verbesserung von Verfahren zur Lösung Quadratischer Programme gearbeitet [LC80, GI83, AHRT93]. Dabei haben sich viele Ansätze als in der Praxis unbrauchbar erwiesen, weil sie zu anfällig gegen numerische Fehler sind (vgl. [Zou91]). Ein heute vielfach verwendetes, robustes und numerisch stabiles Verfahren ist die Methode von Goldfarb und Idnani [GI83], die zur Behandlung der während des Optimierungsprozesses entstehenden linearen Gleichungen in starkem Maße die Singulärwertzerlegung nutzt. Daraus resultiert die Einschränkung, daß dieses Verfahren nur für "mittelgroße" Probleme mit einer Dimension von etwa $\dim \Theta + \dim c < 100$ geeignet ist. Diese Dimension wird jedoch von Flußschätzproblemen nicht erreicht. Neuere Entwicklungen auf dem Gebiet der Quadratischen Programmierung versuchen vor allem, den Rechenaufwand zu verringern und parallele Ansätze zu nutzen, um auch große Probleme des wissenschaftlichen Rechnens behandeln zu können [DF95].

Das Goldfarb-Idnani-Verfahren versucht, sukzessive die Menge der für das Minimum aktiven Nebenbedingungen $\text{Act}(\Theta^*)$ zu bestimmen, indem eine Ausgangshypothese $\mathcal{A}$ für diese Menge schrittweise variiert wird. Dabei können sowohl neue Indizes hinzukommen als auch alte verworfen werden. Zu Details des Verfahrens sei verwiesen auf [GI83]. Hier sollen nur einige für die Einbettung des Verfahrens in die Sequentielle Quadratische Programmierung wichtigen Eigenschaften erläutert werden.

1. Eine dieses Verfahren auszeichnende Eigenschaft besteht darin, daß es mit einem beliebigen Startpunkt arbeiten kann, der die Nebenbedingungen nicht unbedingt erfüllen muß.
2. Das Verfahren terminiert nach endlich vielen Schritten mit dem Auffinden eines Kuhn-Tucker-Punktes. Empirische Studien zeigen, daß im Durchschnitt sogar nur wenige Schritte erforderlich sind.
3. Besteht bereits eine Hypothese, welche Nebenbedingungen für das gesuchte Optimum aktiv sind, so kann dies als Ausgangsinformation vorgegeben werden. War die Vermutung richtig, so konvergiert das Verfahren in einem Schritt. War sie zumindest annähernd richtig, so ist eine Konvergenz in wenigen Schritten zu erwarten, da nur wenige Elemente der Menge $\mathcal{A}$ verändert werden müssen. Hinsichtlich der Einbettung des Verfahrens in die Sequentielle Quadratische Programmierung liegt es nahe, bei jedem Optimierungsschritt von den für den vorangehenden Schritt ermittelten aktiven Nebenbedingungen auszugehen.
4. Das Verfahren verwendet in starkem Maße die QR-Zerlegung von B und A . Es kann daher zusätzlich die Kondition der auftretenden linearen Teilprobleme beurteilen. Damit ist es in der Lage, frühzeitig abubrechen, wenn eine weitere Verbesserung im Rahmen der numerischen Genauigkeit nicht mehr erzielt werden kann. Ferner stehen die entsprechenden Konditionszahlen zur weiteren Verwendung bei der Sequentielle Quadratischen Programmierung zur Verfügung.

10.3 Numerische Lösung des Flußschätzproblems

Während lineare KQ-Probleme unter linearen Nebenbedingungen bereits sehr früh untersucht wurden [Sto71, Mif79, SS79], sind allgemein anwendbare Verfahren zur Lösung nichtlinearer Kleinste-Quadrat-Probleme unter Nebenbedingungen in der Literatur äußerst selten anzutref-

fen.⁸ Auf der einen Seite verwundert dies, weil bei vielen Modellen in den Naturwissenschaften physikalisch unzulässige Bereiche für die Parameter existieren (z.B. müssen Konzentrationen stets positiv sein). Auf der anderen Seite muß jedoch bedacht werden, daß die statistische Theorie für die Parameterschätzung unter Nebenbedingungen kaum entwickelt ist, d.h. die Analyse der Optimierungsergebnisse Schwierigkeiten bereitet. Auf dieses Problem wird in Abschnitt 11.4 noch genauer eingegangen.

10.3.1 Der Levenberg-Marquardt-Algorithmus

Spezielle Ansätze der Sequentiellen Quadratischen Programmierung für Kleinste-Quadrate-Probleme sind nach Kenntnis des Autors bisher nicht explizit publiziert (vgl. [Spe86, Spe93, Ni93, BT95]). Dies verwundert, denn mit dem Levenberg-Marquardt-Algorithmus ist ein speziell auf die Lösung von Kleinste-Quadrate-Problemen zugeschnittenes Verfahren verfügbar. Zur Lösung des Flußschätzproblems wurde daher das Levenberg-Marquardt-Verfahren geeignet modifiziert. Dabei stand zunächst weniger die Effizienz als vielmehr die Robustheit der Methode im Vordergrund.

Bei einem KQ-Problem hat die zu minimierende Funktion die spezielle Gestalt

$$f(\Theta) = \|\eta - F(\Theta)\|_{\Sigma}^2, \quad (10.7)$$

wobei F die vektorwertige Modellfunktion, η der Meßwertvektor und Σ die Gewichtungsmatrix ist (vgl. Definition 10.3). Der zur Quadratsummenminimierung eingesetzte Levenberg-Marquardt-Algorithmus [Mar63] in seiner modernen Implementierung von Moré [Mor78] ist durch nachfolgende Besonderheiten charakterisiert.

1. Es wird in der Taylor-Reihe (10.4) eine spezielle Approximation für die zweite Ableitung verwendet, nämlich die Gauß-Newton-Linearisierung der zu optimierenden Funktion. Diese entsteht durch Linearisierung der Modellfunktion innerhalb der Norm, also

$$\begin{aligned} f(\Theta + \Delta\Theta) &\approx \|\eta - F(\Theta) - \frac{\partial F}{\partial \Theta}(\Theta) \cdot \Delta\Theta\|_{\Sigma}^2 \\ \Rightarrow \frac{\partial f}{\partial \Theta}(\Theta)^T &\approx 2 \frac{\partial F}{\partial \Theta}(\Theta)^T \cdot \Sigma^{-1} \cdot (\eta - F(\Theta)) \\ \Rightarrow \frac{\partial^2 f}{\partial \Theta^2}(\Theta)^T &\approx 2 \frac{\partial F}{\partial \Theta}(\Theta)^T \cdot \Sigma^{-1} \cdot \frac{\partial F}{\partial \Theta}(\Theta) \end{aligned} \quad (10.8)$$

Die erste Ableitung entspricht genau dem Gradienten der ursprünglichen Quadratsumme (10.7), während die approximierte zweite Ableitung sich um den Term⁹

$$2 \frac{\partial^2 F}{\partial \Theta^2}(\Theta)^T \cdot \Sigma^{-1} \cdot (\eta - F(\Theta)) \quad (10.9)$$

von der wahren zweiten Ableitung unterscheidet. Ist das Modell in der Lage, die beobachteten Daten gut zu beschreiben, so liegt der Ausdruck $\eta - F(\Theta^*)$ in einem Minimum Θ^* im Vergleich zu den übrigen Termen der approximierten zweiten Ableitung nahe bei Null. Im Falle fehlerfreier Messung verschwindet er sogar. Dies erklärt, warum die Gauß-Newton-Approximation bei sogenannten kleinen Modellabweichungen gute Resultate liefert.¹⁰ Für

⁸Erst im derzeit neuesten Buch zu diesem Thema [Bjo96] — das für diese Arbeit jedoch nicht mehr herangezogen werden konnte — wird dieses Thema eigens behandelt.

⁹ $\partial^2 F / \partial \Theta^2(\Theta)$ ist hierin die auf naheliegende Weise aus den zweiten partiellen Ableitungen gebildete 3D-Matrix.

¹⁰Im Falle großer Abweichungen kann sich die Konvergenzgeschwindigkeit des Levenberg-Marquardt-Verfahrens deutlich verringern [Spe93]. In dieser Situation sollte aber ohnehin nach einem besser geeigneten Modell gesucht werden.

die Anwendung bei der Sequentiellen Quadratischen Programmierung ist weiterhin wichtig, daß die Gauß-Newton-Approximation der zweiten Ableitung stets symmetrisch und positiv definit ist.

2. Das Verfahren arbeitet mit einem sogenannten Vertrauensbereich, d.h. es läßt nur Schrittweiten innerhalb eines vorgegebenen Extrapolationshorizonts um den derzeitigen Arbeitspunkt zu. Auf diese Weise wird dafür gesorgt, daß die Taylor-Approximation nicht durch unzulässig große Schritte überfordert wird. Vertrauensbereichsverfahren erweisen sich in der Praxis als äußerst robust und zuverlässig, da sie keine großen "Sprünge" im Suchraum zulassen (vgl. auch [MS83, CGST93, Yua95]).

Beim Levenberg-Marquardt-Verfahren wird der Vertrauensbereich durch eine Ellipse um den aktuellen Arbeitspunkt festgelegt (vgl. Abb. 10.4a):

$$E = \{ \Delta\theta \mid \Delta\theta^T \cdot D^{-1} \cdot \Delta\theta \leq r \} \quad , \quad (10.10)$$

wobei D eine symmetrische positiv definite Skalierungsmatrix ist, die zu Anfang vorzugeben ist. Über sie kann Vorwissen bezüglich der erwarteten Größenordnung der Suchschritte eingebracht werden. In der Regel wird mit einer Diagonalmatrix gearbeitet.

Das Verfahren sucht nun ein Minimum der Gauß-Newton-Approximation auf dem elliptischen Gebiet E , d.h. es löst (mit der Gauß-Newton-Linearisierung der zweiten Ableitung) das spezielle Nichtlineare Programm:

$$\begin{aligned} &\text{Minimiere} && f(\theta) + \frac{\partial f}{\partial \theta}(\theta) \cdot \Delta\theta + \frac{1}{2} \Delta\theta^T \cdot \frac{\partial^2 f}{\partial \theta^2}(\theta) \cdot \Delta\theta \\ &\text{unter der Nebenbedingung} && \Delta\theta^T \cdot D^{-1} \cdot \Delta\theta \leq r \quad . \end{aligned}$$

Hierfür gibt es ein effizientes Näherungsverfahren [Spe93].

3. Die Vertrauensbereichsparameter D und r werden im Laufe des Verfahrens dynamisch angepaßt. Dabei wird die aus dem zuletzt durchgeführten Suchschritt verfügbare Information über die aktuelle Güte der Gauß-Newton-Approximation ausgenutzt. Ist $t \cdot \Delta\theta^*$ die gewählte Schrittweite und wurde $f(\theta + t \cdot \Delta\theta^*)$ bestimmt, so wird hierzu das Kriterium

$$\frac{f(\theta + t \cdot \Delta\theta^*) - f(\theta)}{f(\theta) + \frac{\partial f}{\partial \theta}(\theta) \cdot t \cdot \Delta\theta^* + \frac{1}{2} t \cdot \Delta\theta^{*T} \cdot \frac{\partial^2 f}{\partial \theta^2}(\theta) \cdot t \cdot \Delta\theta^* - f(\theta)}$$

ausgewertet, d.h. das Verhältnis des tatsächlichen Fortschritts zu dem durch die Approximation vorhergesagten Fortschritt. Diese Maßzahl sollte nahe bei Eins liegen.

Die genauen Vorschriften für die Anpassung der einzelnen Verfahrensparameter nach jedem Schritt sind aus der Literatur zu entnehmen [Mor78, Spe93, Möl95].

10.3.2 Adaptation des Verfahrens für restringierte KQ-Probleme

Die beschriebene Vorgehensweise des Levenberg-Marquardt-Algorithmus soll nun auf das KQ-Problem mit Ungleichheits-Nebenbedingungen aus Definition 10.3 übertragen werden. Dabei bereitet die Form des Vertrauensbereiches die Hauptschwierigkeit. Bei einer direkten Übertragung wäre als quadratisches Subproblem eine quadratische Funktion unter einer quadratischen und weiteren linearen Ungleichheits-Nebenbedingungen zu lösen. Hierzu gibt es derzeit keine Lösungsansätze.

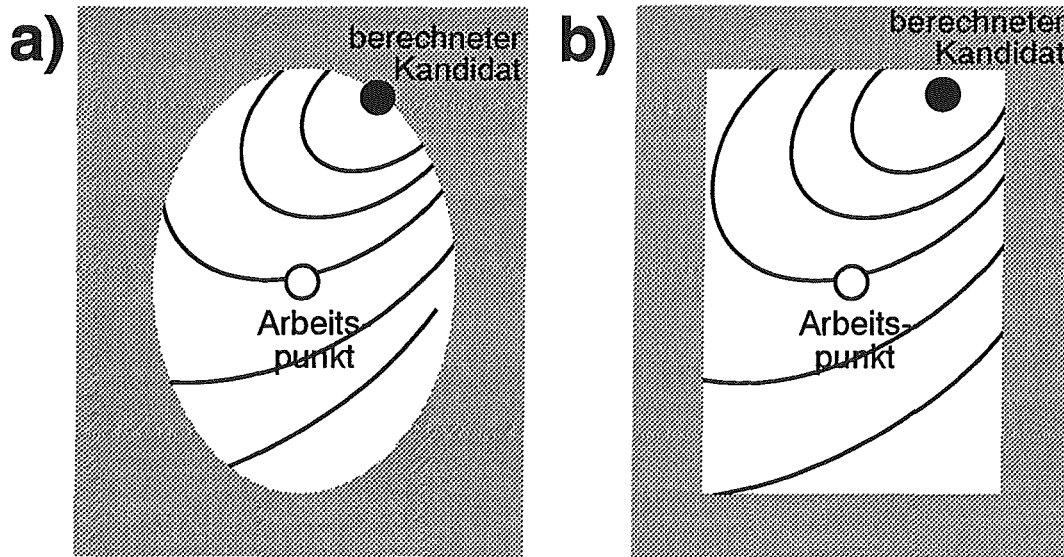


Abbildung 10.4: a) Elliptischer und b) intervallförmiger Vertrauensbereich des Levenberg-Marquardt-Verfahrens und des für die Optimierung unter Nebenbedingungen modifizierten Levenberg-Marquardt-Verfahrens. In beiden Fällen wird die Form des Gebietes durch dieselbe Diagonalmatrix $\mathbf{D}$ und die Größe durch denselben Parameter r festgelegt.

Allerdings muß der Vertrauensbereich nicht notwendig eine elliptische Gestalt haben. Es genügt auch ein mehrdimensionales Intervall mit dem Mittelpunkt $\mathbf{0}$, das mit Hilfe einer Diagonalmatrix $\mathbf{D} = \text{diag}(d_1, \dots, d_{\dim \Theta})$ durch lineare Ungleichungen beschrieben wird:¹¹

$$Q = [-r d_1, r d_1] \times \dots \times [-r d_{\dim \Theta}, r d_{\dim \Theta}] = \{ \Delta \Theta \mid -r \cdot \mathbf{1} \leq \mathbf{D}^{-1} \cdot \Delta \Theta \leq r \cdot \mathbf{1} \} \quad (10.11)$$

Da man jeder achsenparallelen Ellipse einen solchen quaderförmigen Bereich sowohl ein- als auch umschreiben kann, wird klar, daß diese Maßnahme die Konvergenzeigenschaften des Verfahrens nicht berührt. Der genaue Ablauf des modifizierten Levenberg-Marquardt-Algorithmus kann in [Sie96] nachgelesen werden.

10.3.3 Ausschluß von Nebenbedingungen

Zwar treten mit der Wahl eines intervallförmigen Vertrauensbereichs gleich $2 \cdot \dim \Theta$ neue Nebenbedingungen auf, doch überwiegt — zumindest bei kleiner und mittlerer Dimension des Parameterraums — der Vorteil einer größeren Robustheit des Verfahrens durch die Verwendung eines Vertrauensbereichs. Außerdem ist zu bedenken, daß — einmal in der Nähe des Optimums angelangt — die neu hinzugenommenen Nebenbedingungen nicht mehr aktiviert werden. Dabei hilft das nachfolgende Kriterium (vgl. Abb. 10.5).

Satz 10.3 (Ausschluß linearer Nebenbedingungen) Seien ein mehrdimensionales Intervall Q und ein linearer Halbraum L im Raum $\mathbb{R}^n$ gegeben durch:

$$Q = \{ \Delta \Theta \mid -r \cdot \mathbf{1} \leq \mathbf{D}^{-1} \cdot \Delta \Theta \leq r \cdot \mathbf{1} \} \quad \text{und} \quad L = \{ \Delta \Theta \mid \mathbf{c}^T \cdot \Delta \Theta \leq b \}$$

mit $b > 0$ und es gelte die Ungleichung

$$\|\mathbf{D} \cdot \mathbf{c}\| \leq \frac{b}{r \cdot \sqrt{\dim \Delta \Theta}} \quad .$$

¹¹ Anders ausgedrückt, wird die L_2 -Norm in Gleichung (10.10) gegen die L_∞ -Norm ausgetauscht.

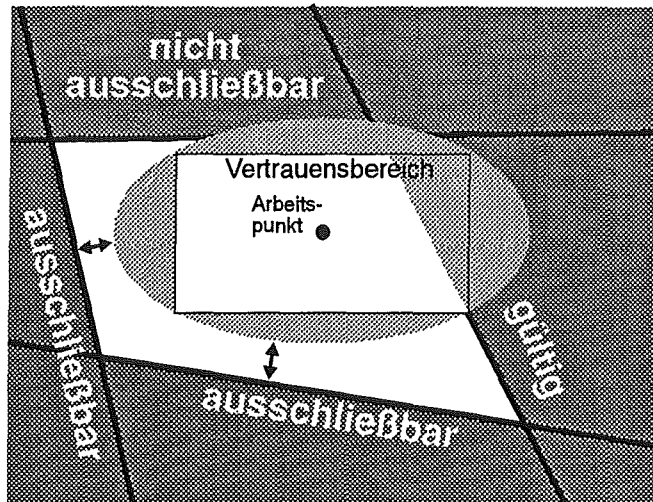


Abbildung 10.5: Ausschluß von Nebenbedingungen mit Hilfe einer elliptischen Umschließung des Vertrauensbereichs. Durch die Ellipse bereits implizierte lineare Nebenbedingungen können ausgeschlossen werden.

Dann folgt $Q \subset L$.

Beweis:

Da D eine Diagonalmatrix ist, wird Q durch eine Transformation mit D^{-1} in einen Kubus überführt:

$$D^{-1} \cdot Q = r \cdot [-1, 1] = [-r, r] \times \dots \times [-r, r]$$

Dieser ist in der Kugel mit Radius $r \cdot \sqrt{\dim \Theta}$ enthalten, d.h. es gilt

$$\Delta\Theta \in Q \implies \|D^{-1} \cdot \Delta\Theta\| \leq r \cdot \sqrt{\dim \Theta}$$

Mit dieser Beziehung, der vorausgesetzten Ungleichung und der Cauchy-Schwarz-Ungleichung ergibt sich nun für $\Delta\Theta \in Q$:

$$\begin{aligned} |c^T \Delta\Theta| &= |(c^T D) \cdot (D^{-1} \Delta\Theta)| \\ &\leq \|D c\| \cdot \|D^{-1} \Delta\Theta\| \\ &\leq \frac{b}{r \cdot \sqrt{\dim \Delta\Theta}} \cdot r \cdot \sqrt{\dim \Delta\Theta} = b \end{aligned}$$

□

10.3.4 Anmerkungen zur Konvergenztheorie

Zum Abschluß sollen nun noch einige Anmerkungen zur Konvergenztheorie der Sequentiellen Quadratischen Programmierung gemacht werden. Die klassische Konvergenztheorie für die unrestringierte Optimierung konnte in den letzten Jahren weitgehend auf Probleme mit Nebenbedingungen übertragen werden [Spe86]. Es konnte gezeigt werden, daß eine superlineare Konvergenzordnung bei exakt berechenbaren Gradienten erzielt wird [BT95]. Dies ist im vorliegenden Fall durch Satz 8.3 gewährleistet. Darüber hinaus gibt es eine große Zahl von empirischen

Studien, die spezielle Probleme der restringierten Optimierung aufzeigen, wie zum Beispiel das "Zickzacklaufen" von Suchverfahren im Falle spitz zulaufender Ecken des zulässigen Gebietes [Spe93]. Zur Vermeidung derartiger Probleme wurde in den letzten Jahren durch Einführung entsprechender Vorteilsfunktionen und Vertrauensbereiche ein großer Fortschritt erzielt [BT95].

Die Entwicklung einer Konvergenztheorie für das oben entwickelte modifizierte Levenberg-Marquardt-Verfahren würde den Rahmen dieser Arbeit sprengen. In der praktischen Anwendung ist jedoch die Konvergenz gegen einen Kuhn-Tucker-Punkt überprüfbar. Zudem ist die Form des durch die Nebenbedingungen bei Markierungssystemen definierten zulässigen Gebietes erfahrungsgemäß nicht stark entartet, so daß keine Zickzacklauf-Probleme entstehen könnten. In der praktischen Anwendung des Verfahrens bei realistisch großen Stoffwechselnetzwerken trat Konvergenz schon nach etwa 10 – 20 Iterationsschritten ein [Sie96]. Dabei sind auch Situationen mit Konvergenz auf dem Rand des zulässigen Gebietes aufgetreten.

Kapitel 11

Statistische Auswertung

Aus statistischer Sicht wird die Kleinste-Quadrate-Schätzung als Regressionsanalyse bezeichnet [BW88, SW89]. Sie hat nicht nur das Ziel, eine Flußschätzung durchzuführen, sondern vor allem auch Aussagen über die Güte dieser Schätzung zu machen. Dabei interessieren besonders nachfolgende Fragestellungen, die anschließend in diesem Kapitel behandelt werden.¹ Die zitierte Literatur nennt jeweils Beispiele aus dem Bereich der Stoffflußanalyse:

- Untersuchung des Einflusses der Flußparameter auf die gemessenen Größen, d.h. Berechnung der Output-Sensitivitäten des Modells [CB83, Lee93, CFGC95, CB95, Wie96a].
- Quantifizierung des Einflusses einzelner Meßwerte auf den berechneten Flußschätzer, d.h. Berechnung von Parameter-Sensitivitäten.
- Berechnung von Gütemaßen für die erhaltene Flußschätzung, insbesondere des Erwartungswertes und der Kovarianzmatrix des Flußschätzers [RB85, KCS93, Wie96a].
- Quantifizierung der Bestimmtheit der Schätzung, d.h. Berechnung von Konfidenzbereichen für die erhaltene Flußschätzung.
- Aufdeckung von Modellfehlern, die dazu führen, daß sich die gemessenen Daten nicht durch die gegebene Modellstruktur beschreiben lassen [RB85, WGKF⁺92, RDC⁺95] [FR95].
- Ermittlung redundanter Meßwerte, d.h. solcher Meßwerte, deren Information ganz oder zum großen Teil bereits in anderen Meßwerten enthalten ist [Val91, HRHH⁺94, Wie95a].
- Identifizierung von Ausreißern unter den Meßwerten, d.h. solcher Daten, die systematische Verzerrungen aufweisen oder die Normalverteilungsannahme verletzen.
- Versuchsplanung im Sinne der Bestimmung weiterer Meßgrößen, die – falls verfügbar – die Güte des Schätzers verbessern würde.²

¹Ausführliche Darstellungen der vorgestellten Methoden finden sich in [BKW80, Páz86, CH88, Arn90, Páz93].

²Es sei darauf hingewiesen, daß die in Abschnitt 2.5 vorgestellte Entkopplungstechnik (Definition 2.6) die nachträgliche Vermessung ausgewählter Pools erlaubt, ohne den Fermentationslauf wiederholen zu müssen.

11.1 Linearisierung nichtlinearer Modelle

11.1.1 Zur Problematik biometrischer Methoden

Angesichts dieser Vielfalt an Fragestellungen ist allerdings einschränkend zu beachten, daß nie alle diese Fragen *zugleich* mit exakten statistischen Methoden beantwortet werden können. Dies liegt darin begründet, daß derselbe Datensatz mit statistischen Methoden in der Regel nicht mehrfach "befragt" werden darf. Eine Aussage, die aufgrund einer statistischen Analyse gemacht wird, verwendet stets einen Teil der im Datensatz enthaltenen Information. Dieser steht — in gewissem Sinne — für weitere Analysen nicht mehr zur Verfügung.

Das Hauptproblem einer Regressionsanalyse in der Praxis besteht stets in der Gültigkeit der zugrundeliegenden Voraussetzung, daß das betrachtete nichtlineare Regressionsmodell (vgl. Definition 10.3)

$$\eta = F(\Theta) + \varepsilon \quad \text{mit} \quad \varepsilon = N(0, \Sigma)$$

korrekt ist. Im Kontext biologischer Experimente (vgl. auch Kapitel 2) muß aus verschiedenen Gründen jedoch ganz im Gegenteil davon ausgegangen werden, daß dies nicht der Fall ist, denn

1. die Modellfunktion beschreibt die komplexen realen Zusammenhänge nur näherungsweise (Modellfehler).
2. die Meßwerte weisen häufig systematische Verzerrungen auf, d.h. für den Erwartungswert der Residuen gilt $E \varepsilon \neq 0$.
3. die Kovarianzmatrix Σ der Meßfehler ist nur approximativ bestimmbar bzw. ist ein Erfahrungswert.
4. der Meßfehler ist meist nicht normalverteilt.

Vom strengen statistischen Standpunkt aus gesehen, stellen diese Einflüsse zusammengenommen (wovon meist auszugehen ist) die erzielten statistischen Ergebnisse grundsätzlich in Frage. In der Praxis führen die genannten Effekte in der Regel zu einer Verschlechterung der berechneten Gütemaße. Statistische Ansätze, die eine der vier oben genannten Ausnahmen zulassen, sind für hochdimensionale nichtlineare Modelle nicht praktikabel.³

Trotzdem sollten diese Argumente nicht zu einer völligen Verwerfung des statistischen Ansatzes führen. Vielmehr stellen statistische Methoden heute die einzigen verfügbaren und ausgereiften Werkzeuge dar, die eine quantifizierte Bewertung von Regressionsergebnissen ermöglichen. Die erzielten Resultate müssen jedoch mit Sachverstand und Systemkenntnis interpretiert werden. Die Anwendung statistischer Verfahren in der Biologie stellt somit niemals eine wertfreie Interpretation der Meßergebnisse dar. Vielmehr liegt die sorgfältige und nachvollziehbare Auswertung des Datenmaterials stets in der Verantwortung des Modellierers. Darüber hinaus muß der Anwender in die Lage versetzt werden, die erzielten Ergebnisse nachzuvollziehen. In der Regel interessiert dabei vor allem die Größenordnung der aufgrund der Analyse ermittelten Gütemaße und weniger die genauen Zahlenwerte. Dies wird in Kapitel 12 noch im einzelnen zu diskutieren sein.

Schließlich sei auf einen wichtiger Unterschied zwischen einer Regressionsanalyse im experimentellen Umfeld und einer statistischen Analyse anhand nicht reproduzierbarer Datensätze (z.B. in der Medizin) hingewiesen. Im letzteren Fall müssen höchste Ansprüche an die Signifikanz der gemachten Aussagen gestellt werden, während im experimentellen Kontext ein Folgeexperiment unter verbesserten Bedingungen die erzielten Aussagen erhärten kann. Aus diesem

³Verwiesen sei auf nichtparametrische oder robuste statistische Verfahren [Hub81, BD93].

Grunde wird in dieser Arbeit stets das Konfidenzniveau von 90% als zufriedenstellend erachtet und nicht wie vielfach verwendet das 95%- oder gar 99%-Niveau.

11.1.2 Der Linearisierungsansatz

Für lineare statistische Modelle mit normalverteilten Zufallsgrößen können alle eingangs genannten Probleme unter Rückgriff auf Methoden der Linearen Algebra explizit beantwortet werden (vgl. auch [VS93, HHRH⁺94, HRHH⁺94]), während für nichtlineare Modellstrukturen oder nicht-normalverteilte Größen explizite (d.h. durch Formeln beschreibbare) Aussagen kaum noch zu gewinnen sind. Aus diesem Grunde wird in der Praxis meist versucht, lineare Approximationen für die betrachteten Systeme zu finden, um darauf die Ergebnisse der linearen statistischen Regressionstheorie anzuwenden.

Definition 11.1 (lineares statistisches Modell) Ein lineares statistisches Modell ist gegeben durch

$$\eta = \mathbf{A} \cdot \Theta + \varepsilon \quad \text{mit} \quad \varepsilon = N(\mathbf{0}, \Sigma) \quad .$$

Dabei sind die Systemmatrix $\mathbf{A}$ und die Kovarianzmatrix Σ bekannt.⁴ Es wird zusätzlich vorausgesetzt, daß $\dim \eta \geq \dim \Theta$ gilt und $\mathbf{A}$ maximalen Rang hat.⁵

Die Linearisierung eines allgemeinen nichtlinearen Modells

$$\eta = F(\Theta) + \varepsilon \quad \text{mit} \quad \varepsilon = N(\mathbf{0}, \Sigma)$$

gemäß Definition 10.3 geht nun wie folgt vor. Zunächst wird der entsprechende Kleinste-Quadrate-Schätzer $\hat{\Theta}$ für den Parametervektor Θ berechnet, also (vgl. Abb. 11.1a)

$$\hat{\Theta} = \arg \min_{\Theta} \|\eta - F(\Theta)\|_{\Sigma}^2 \quad .$$

Unter der Annahme, daß der Schätzer "nahe genug" am wahren Wert Θ^* des Parametervektors liegt, kann dann das nichtlineare Modell approximativ durch seine Taylor-Entwicklung erster Ordnung im Punkte $\hat{\Theta}$ ersetzt werden [Páz93] (vgl. Abb. 11.1b):

$$\eta \approx F(\hat{\Theta}) + \frac{\partial F}{\partial \Theta}(\hat{\Theta}) \cdot (\Theta - \hat{\Theta}) + \varepsilon$$

Dies ist bereits ein lineares statistisches Modell

$$\begin{aligned} \varepsilon &= N(\mathbf{0}, \Sigma) \\ \Delta \eta &= \mathbf{A} \cdot \Delta \Theta + \varepsilon \quad \text{mit} \quad \Delta \eta = \eta - F(\hat{\Theta}) + \frac{\partial F}{\partial \Theta}(\hat{\Theta}) \cdot \hat{\Theta} \quad . \quad (11.1) \\ \mathbf{A} &= \frac{\partial F}{\partial \Theta}(\hat{\Theta}) \end{aligned}$$

⁴Vielfach findet man anstelle von Σ im Regressionsmodell den Term $\sigma^2 \cdot \Sigma$, wobei σ ein zu schätzender Skalierungsfaktor ist. Wenn alle Messungen mit demselben Meßinstrument gewonnen wurden und a priori nichts über den Meßfehler bekannt ist, so kann hiermit zugleich der Parametervektor Θ und der Standardmeßfehler σ geschätzt werden [Arn90]. Bezogen auf die experimentelle Praxis ist es jedoch eher unrealistisch anzunehmen, daß das Größenverhältnis der Meßfehler (gegeben durch Σ) bekannt ist, nicht aber deren absolute Werte (gegeben durch $\sigma^2 \cdot \Sigma$). Bezogen auf die Stoffflußanalyse kann die Fehlercharakteristik der Instrumente aufgrund von Messungen mit standardisierten Proben im voraus ermittelt werden.

⁵Die Theorie der linearen Regression umfaßt auch den entarteten Fall, daß $\mathbf{A}$ nicht Maximalrang hat (vgl. [Páz93]). Eine solche Rangdefizienz weist auf ein unvollständig bestimmtes Modell hin (Nichtidentifizierbarkeit). In der Praxis sollte diese Fragestellung besser vor einer Regressionsanalyse untersucht werden. Wird die Regressionsanalyse — wie nachfolgend beschrieben — mit Hilfe einer Singulärwertzerlegung durchgeführt, so sind Rangdefizienzen ferner numerisch leicht feststellbar [LH74].

Wie sofort ersichtlich ist (vgl. auch Abb. 11.1b), liefert die KQ-Schätzung, angewandt auf das bzgl. $\Delta\Theta$ lineare Modell (11.1), das Ergebnis $\widehat{\Delta\Theta} = 0$ (vgl. auch Abb. 11.1b). Der Übergang zur Linearisierung führt also nicht zu methodisch bedingten Inkonsistenzen.

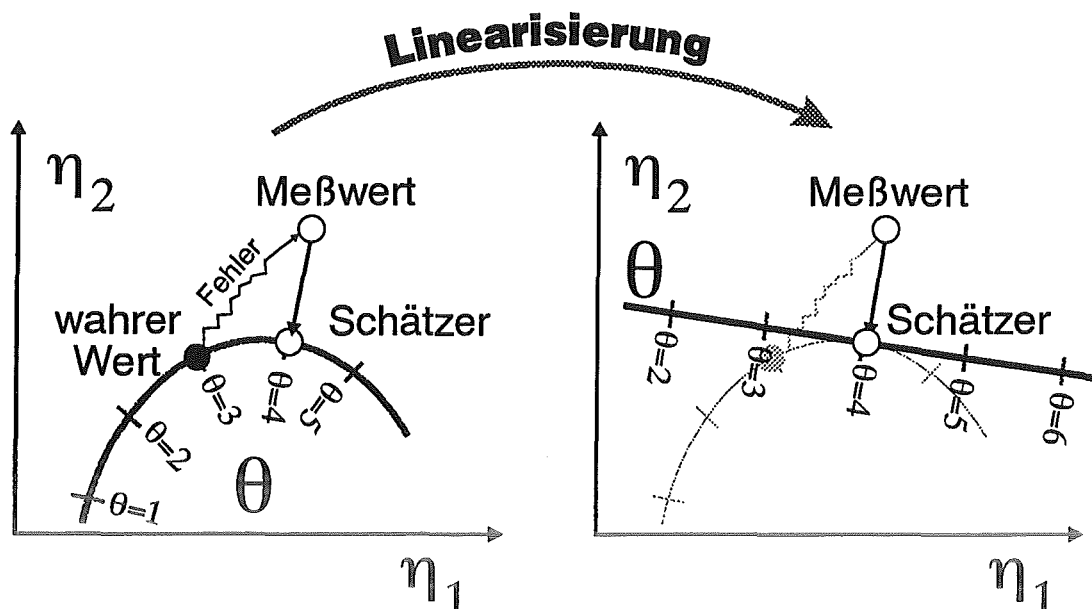


Abbildung 11.1: a) Nichtlineare Regression als Abstandsminimierung, d.h. als Projektion auf die Erwartungsfläche des Modells. b) Linearisierung als Ersetzung der durch das Modell beschriebenen Fläche durch deren Tangentialebene im durch den Schätzer definierten Punkt.

Der Linearisierungsansatz ist nicht unproblematisch und sollte daher mit Vorsicht angewendet werden (vgl. dazu Abb. 11.1). Die Schwierigkeit liegt darin, die Gültigkeit der gewählten linearen Approximation innerhalb der statistischen Schwankungsbreite der Meßwerte zu verifizieren. Entscheidend ist dabei das Größenverhältnis zwischen dem Bereich um Θ^* , in dem die Linearisierung hinreichend gut ist und dem Bereich, in dem die Messung statistisch schwankt. Kritisch sind somit stark nichtlineare Modelle oder große Fehlereinflüsse. In jedem Fall trägt der entstehende Linearisierungsfehler weiter zur Verzerrung der erzielten statistischen Resultate bei. Solche Effekte werden in der nichtlinearen Statistik studiert [Páz93]. In Bezug auf Markierungsexperimente wird hierauf in Abschnitt 11.3 noch detaillierter eingegangen.

11.2 Lineare Regressionsanalyse

11.2.1 Eigenschaften linearer KQ-Schätzer

Die grundlegenden Resultate der linearen Regressionsanalyse sollen nun kurz zusammengefaßt werden (vgl. [Arn90]). Zunächst wird die Verteilung des Schätzers durch den nachfolgenden Satz charakterisiert.⁶ Daraus ergibt sich bereits eine Antwort auf einige der eingangs gestellten Fragen.

⁶Man beachte, daß Σ als bekannt angenommen wird. Damit ändern sich einige Verteilungsaussagen leicht gegenüber der Standardliteratur (vgl. [Páz93]).

Satz 11.1 (Verteilung des linearen KQ-Schätzers) Gegeben sei ein lineares Regressionsmodell $\eta = \mathbf{A} \cdot \Theta + \varepsilon$ mit $\varepsilon = N(0, \Sigma)$ gemäß Definition 11.1, und Θ^* sei der (unbekannte) wahre Parametervektor. Dann gilt:

1. Der KQ-Schätzer für das lineare Modell ist gegeben durch

$$\hat{\Theta} = (\mathbf{A}^T \Sigma^{-1} \mathbf{A})^{-1} \cdot \mathbf{A}^T \Sigma^{-1} \eta \quad .$$

2. $\hat{\Theta}$ ist unverzerrt und gemäß $N(\hat{\Theta}^*, \text{Cov}(\hat{\Theta}))$ verteilt mit der Kovarianzmatrix

$$\text{Cov} \hat{\Theta} = (\mathbf{A}^T \Sigma^{-1} \mathbf{A})^{-1} \quad .$$

3. $\hat{\Theta}$ ist (in gewissem Sinne) der beste lineare unverzerrte Schätzer für Θ (Gauß-Markov-Theorem).

4. Ein Konfidenzbereich für $\hat{\Theta}$ zum Konfidenzniveau $(1 - \beta)$ ist gegeben durch ⁷

$$\begin{aligned} \text{Conf}_{\hat{\Theta}}(1 - \beta) &= \left\{ \Theta \mid \|\mathbf{A}(\hat{\Theta} - \Theta)\|_{\Sigma}^2 < \chi_{\dim \Theta}^2(1 - \beta) \right\} \\ &= \left\{ \Theta \mid (\hat{\Theta} - \Theta)^T [\text{Cov} \hat{\Theta}]^{-1} (\hat{\Theta} - \Theta) < \chi_{\dim \Theta}^2(1 - \beta) \right\} \quad . \end{aligned}$$

5. Ein Konfidenzbereich für eine einzelne Komponente $\hat{\Theta}_i$ zum Konfidenzniveau $(1 - \beta)$ ist gegeben durch das Intervall

$$\text{Conf}_{\hat{\Theta}}(1 - \beta) = \left[\hat{\Theta}_i - \sqrt{\chi_1^2(1 - \beta) \cdot \text{Var} \hat{\Theta}_i}, \hat{\Theta}_i + \sqrt{\chi_1^2(1 - \beta) \cdot \text{Var} \hat{\Theta}_i} \right] \quad .$$

6. Ein kritischer Bereich für den Test der Hypothese, daß die Daten durch das Modell beschreibbar sind, ist gegeben durch

$$\left\{ \Theta \mid \|y - \mathbf{A} \hat{\Theta}\|_{\Sigma}^2 > \chi_{\dim \eta - \dim \Theta}^2(1 - \beta) \right\} \quad .$$

7. Die Korrelationsmatrix

$$\text{Corr}(\hat{\Theta}) = \left[\text{diag Var}(\hat{\Theta}) \right]^{-1} \cdot \text{Cov}(\hat{\Theta}) \cdot \left[\text{diag Var}(\hat{\Theta}) \right]$$

ist eine entlang den Koordinatenachsen reskalierte Kovarianzmatrix.

Offenbar hat das durch den Satz gegebene Konfidenzgebiet $\text{Conf}_{\hat{\Theta}}(1 - \beta)$ die Gestalt eines mehrdimensionalen Ellipsoids (vgl. Abb. 11.2a). Der entsprechende Ellipsoid zur Korrelationsmatrix wird von einem Kubus eingeschlossen, so daß die geometrische Form des Ellipsoids aufgrund der skalierten Achsen einfacher bewertet werden kann (vgl. Abb. 11.2b).

Im Hinblick auf den Test der Beschreibungsgüte des Modells (Punkt 6 des obigen Satzes) sei jedoch auf die Aussagen aus dem einleitenden Abschnitt verwiesen. Dieser Test scheitert aufgrund der dort vorgestellten Einwände häufig bei biologischen Datensätzen (vgl. auch Abschnitt 12.2), was die dort skizzierten Ursachen haben kann.

⁷ χ_m^2 bezeichnet die Chi-Quadrat-Verteilung mit m Freiheitsgraden.

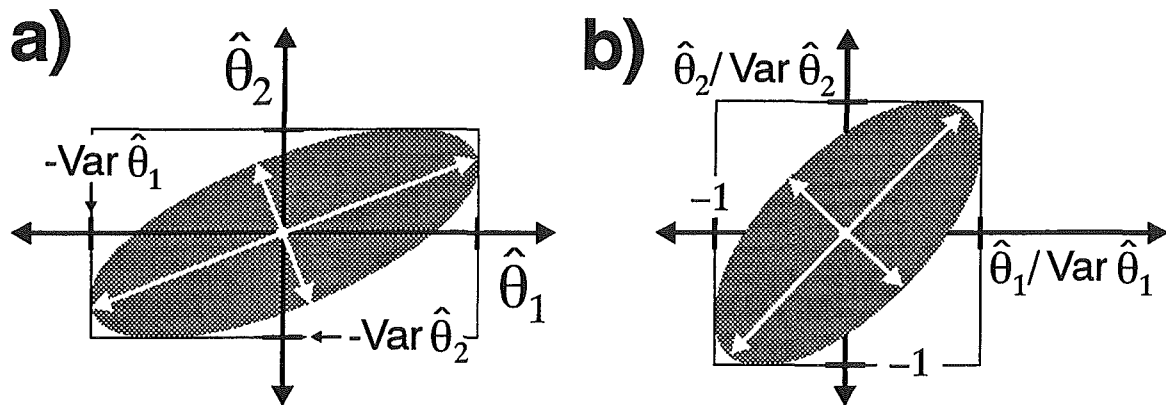


Abbildung 11.2: a) Elliptisches 40%-Konfidenzgebiet (für $\beta = 0.4$ gilt gerade $\chi_2^2(1 - \beta) = 1$) für einen zweidimensionalen Schätzer $\hat{\Theta}$. Ein Gebiet mit 90%-Konfidenz ist um den Faktor 2.14 größer, ein Gebiet mit 95% um den Faktor 2.44. b) Durch die Korrelationsmatrix definiertes reskaliertes Konfidenzgebiet. Die Stärke der Korrelation ist an der Abweichung von der Kreisform ablesbar.

11.2.2 Sensitivitätsmaße

Zu den wichtigsten Maßzahlen für die praktische Analyse von Regressionsmodellen [CH88] gehören die Sensitivität der Modellvorhersage gegenüber Variationen der Parameter (Output-Sensitivität) und die Sensitivität des berechneten Parameterschätzers gegenüber Variationen der Meßdaten (Parameter-Sensitivität). Beide Maße können zur Ermittlung von Haupteinflussfaktoren auf den KQ-Schätzer herangezogen werden. Sensitivitätsmaße stehen aufgrund ihrer differentiellen Definition in engem Zusammenhang mit der Linearisierung des Modells:

Definition 11.2 (Output-Sensitivität, Parameter-Sensitivität) Gegeben sei ein nichtlineares Regressionsmodell $\eta = F(\Theta) + \varepsilon$ mit $\varepsilon = N(0, \Sigma)$. Die Output-Sensitivität des Modells an der Stelle Θ ist gegeben durch

$$\text{Sens}_{\Theta}^{\eta} = \frac{\partial F}{\partial \Theta}(\Theta) \quad .$$

Die Parameter-Sensitivität des KQ-Schätzers $\hat{\Theta}(\eta)$ bei vorliegenden Meßdaten η ist gegeben durch

$$\text{Sens}_{\eta}^{\hat{\Theta}} = \frac{\partial \hat{\Theta}}{\partial \eta}(\eta) \quad .$$

Offenbar ist die Output-Sensitivitätsmatrix gerade die Systemmatrix des linearisierten Modells (vgl. Gleichung (11.1)). Für ein lineares Modell folgt dann unmittelbar nach Satz 11.1 der

Satz 11.2 (Output-Sensitivität und Parameter-Sensitivität eines linearen Modells) Sei ein lineares Regressionsmodell gegeben durch $\eta = A \cdot \Theta + \varepsilon$ mit $\varepsilon = N(0, \Sigma)$. Dann gilt unabhängig vom Parameterschätzer $\hat{\Theta}$ bzw. dem Meßvektor η :

$$\begin{aligned} \text{Sens}_{\Theta}^{\eta} &= A \\ \text{Sens}_{\eta}^{\hat{\Theta}} &= (A^T \Sigma^{-1} A)^{-1} \cdot A^T \Sigma^{-1} \end{aligned}$$

Die Output-Sensitivität legt in dieser Situation also bereits alle statistischen Eigenschaften des Schätzers fest. Im nichtlinearen Fall ist diese Aussage jedoch nur noch approximativ richtig:

Satz 11.3 (Parameter-Sensitivität nichtlinearer Regressionsmodelle) *Gegeben sei ein nicht-lineares Regressionsmodell $\eta = F(\Theta) + \varepsilon$ mit $\varepsilon = N(0, \Sigma)$. Dann ist die Parameter-Sensitivität gegeben durch*

$$\text{Sens}_{\eta}^{\hat{\Theta}} = \left[\frac{\partial F}{\partial \Theta}(\hat{\Theta})^T \Sigma^{-1} \frac{\partial F}{\partial \Theta}(\hat{\Theta}) - \frac{\partial^2 F}{\partial \Theta^2}(\hat{\Theta})^T \Sigma^{-1} (\eta - F(\hat{\Theta})) \right]^{-1} \cdot \frac{\partial F}{\partial \Theta}(\hat{\Theta})^T \Sigma^{-1}.$$

Beweis:

Ergibt sich aus einer impliziten Differentiation der bestimmenden Gleichung für den KQ-Schätzer:

$$\left[\eta - F(\hat{\Theta}(\eta)) \right]^T \Sigma^{-1} \frac{\partial F}{\partial \Theta}(\hat{\Theta}(\eta)) = 0$$

□

Der Term $\frac{\partial^2 F}{\partial \Theta^2}(\hat{\Theta})(\eta - F(\hat{\Theta}))$, der die Differenz zum linearisierten Modell repräsentiert, ist schon aus Kapitel 10 bekannt (vgl. Gleichung (10.9)). Er stellt gerade den Approximationsfehler der Gauß-Newton-Approximation für die zweite Ableitung der Fehlerquadratsumme dar. Wie dort beschrieben, wird dieser Term bei kleinen Modellfehlern als vernachlässigbar angesehen. In der Praxis wird die Parameter-Sensitivität daher aufgrund der Linearisierung (11.1) mit Hilfe von Satz 11.2 approximiert.⁸ Zusammenfassend kann gesagt werden, daß die gesamte Information für die linearisierte Analyse eines nichtlinearen Regressionsmodells in der Output-Sensitivitätsmatrix und der Kovarianzmatrix des Meßfehlers enthalten ist.

11.2.3 Skalierung von Sensitivitäts- und Kovarianzmatrizen

Sensitivitätsmatrizen sind in der Praxis nicht leicht zu interpretieren, weil die Größenordnungen ihrer Einträge nicht miteinander vergleichbar sind. In der Literatur finden sich daher verschiedene Vorschläge für gewichtete Sensitivitäten der allgemeinen Form

$$\frac{\partial F_i / \bar{F}_i}{\partial \Theta_j / \bar{\Theta}_j}$$

mit gewissen Konstanten $\bar{F}_i$ und $\bar{\Theta}_j$. Damit können sowohl die Funktionswerte als auch die Parameterwerte größenordnungsmäßig reskaliert werden. Der bekannteste Ansatz ist die Verwendung relativer Sensitivitäten $\frac{\partial f_i / f_i}{\partial \theta_j / \theta_j}$ (z.B. in der Metabolic Control Theory [Red88]). Relative Skalen sind jedoch bekanntermaßen dann problematisch, wenn Werte nahe Null zu erwarten sind. In der Praxis muß eine adäquate Wahl der Skalierungsfaktoren also vom Anwender getroffen werden.

Aus statistischer Sicht ist ein unmittelbar durch die Modellformulierung vorgegebener Skalierungsansatz tragfähiger. Dabei sollte die Skalierung den durch die Fehlerquadratsumme gegebenen tatsächlichen Einfluß der Meßwerte auf den KQ-Schätzer widerspiegeln. Die entsprechende Gewichtung für den Meßvektor η ist gegeben durch $\sqrt{\Sigma}^{-1} \cdot \eta$, denn diese Skalierung

⁸Ein differentialgeometrischer Ansatz, der auch Terme zweiter Ordnung berücksichtigt, wird in [Páz93] beschrieben.

entkoppelt und normiert die Meßwerte. Das Resultat der Skalierung ist dann $N(0, 1)$ -verteilt. Anders ausgedrückt wird das lineare Modell

$$\eta = \mathbf{A} \cdot \Theta + \varepsilon \quad \text{mit} \quad \varepsilon = N(0, \Sigma)$$

mittels der Zuordnungen

$$\bar{\eta} = \sqrt{\Sigma}^{-1} \cdot \eta \quad \bar{\mathbf{A}} = \sqrt{\Sigma}^{-1} \cdot \mathbf{A} \quad \bar{\varepsilon} = \sqrt{\Sigma}^{-1} \cdot \varepsilon$$

transformiert in das reskalierte lineare Modell

$$\bar{\eta} = \bar{\mathbf{A}} \cdot \Theta + \bar{\varepsilon} \quad \text{mit} \quad \bar{\varepsilon} = N(0, 1)$$

Die entsprechenden reskalierten statistischen Kenngrößen sind

$$\text{Sens } \bar{\eta}_{\Theta} = \bar{\mathbf{A}} \quad \text{Sens } \hat{\Theta}_{\bar{\eta}} = (\bar{\mathbf{A}}^T \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \quad \text{Cov } \hat{\Theta} = (\bar{\mathbf{A}}^T \bar{\mathbf{A}})^{-1} \quad (11.2)$$

In diesem reskalierten Modell sind also alle Meßfehler unabhängig und größenordnungsmäßig direkt vergleichbar. Allerdings ist $\bar{\eta}$ bei allgemeinem Σ ein "künstlicher" Meßfehler, der nicht mehr physikalisch interpretierbar ist. Im Falle der Stoffflußanalyse wird Σ jedoch als Diagonalmatrix gewählt (vgl. Abschnitt 6.2), so daß bzgl. der Interpretation der Sensitivitäten keine Schwierigkeiten auftreten.

Eine entsprechende systemimmanente Skalierung für den Parametervektor Θ gibt es nicht, weil Θ selbst keine Zufallsvariable ist. Ein möglicher Ansatz wäre, nach erfolgter Schätzung durch $\hat{\Theta}$ die Matrix $\sqrt{\text{Cov}(\hat{\Theta})}^{-1}$ zur Skalierung zu verwenden. Damit ergeben sich die zusätzlichen Transformationen

$$\bar{\Theta} = \sqrt{\text{Cov}(\hat{\Theta})}^{-1} \cdot \Theta \quad \bar{\bar{\mathbf{A}}} = \sqrt{\Sigma}^{-1} \cdot \mathbf{A} \cdot \sqrt{\text{Cov}(\hat{\Theta})}$$

Zusammen mit den Gleichungen (11.2) entsteht das bzgl. der Meßwerte und Parameter reskalierte lineare Modell

$$\bar{\eta} = \bar{\bar{\mathbf{A}}} \cdot \bar{\Theta} + \bar{\varepsilon} \quad \text{mit} \quad \bar{\varepsilon} = N(0, 1)$$

mit den Kenngrößen

$$\text{Sens } \bar{\eta}_{\bar{\Theta}} = \bar{\bar{\mathbf{A}}} \quad \text{Sens } \hat{\bar{\Theta}}_{\bar{\eta}} = \bar{\bar{\mathbf{A}}}^T \quad \text{Cov } \hat{\bar{\Theta}} = \mathbf{1}$$

Dieses Ergebnis hat für die praktische Anwendung keine Aussage mehr, da $\bar{\Theta}$ ein "künstlicher" Parametervektor mit nicht mehr physikalisch interpretierbaren Größen ist. Die für den Anwender relevante Information ist dagegen in der Systemmatrix $\bar{\bar{\mathbf{A}}}$ enthalten. Im nichtlinearen Fall sind darüber hinaus zwei verschiedene Datensätze nicht mehr im Vergleich auswertbar, da sie zu unterschiedlichen Transformationen der Systemmatrix führen.

Die Transformation zeigt jedoch einen anderen Effekt, der für die nichtlineare Anwendung von großer Bedeutung ist: Eine Parameter-Reskalierung kann grundsätzlich die Gütemaße der Schätzung verändern. Insbesondere kann bei nichtlinearen Modellen nach "günstigen" Reskalierungen des Parameterraums gesucht werden [Páz93]. Eine solche Reskalierung wird in Abschnitt 11.3 diskutiert.

Zusammenfassend gilt, daß eine geeignete Größenskalierung im Parameterraum wenn überhaupt stets vom Anwender vorgenommen werden muß. Auf der Output-Seite wird das Modell dagegen nachfolgend stets mit $\sqrt{\Sigma}^{-1}$ reskaliert.

11.2.4 Numerische Analyse von Sensitivitäts- und Kovarianzmatrizen

Die vorangehenden Abschnitte (insbesondere Satz 11.2) betonen die Bedeutung der Output-Sensitivitätsmatrix als Grundlage für alle weiteren statistischen Analysen. Für die numerische Implementierung der entsprechenden Rechenoperationen ist die Singulärwertzerlegung von zentraler Bedeutung [LH74, PFTV88].

Definition 11.3 (Singulärwertzerlegung, Singulärwerte) Eine Singulärwertzerlegung einer Matrix A mit $\text{rowdim } A \geq \text{coldim } A$ ist eine Faktorisierung der Gestalt

$$A = U \cdot \begin{pmatrix} D \\ 0 \end{pmatrix} \cdot V^T$$

mit orthonormalen Matrizen U, V und einer Diagonalmatrix

$$D = \text{diag}(d_1, d_2, \dots, d_{\text{coldim } A})$$

mit $d_1 \geq d_2 \geq \dots \geq d_{\text{coldim } A}$. Die Einträge von D heißen die Singulärwerte von A . Im Fall $\text{rowdim } A \leq \text{coldim } A$ ist die Definition entsprechend auf A^T anzuwenden.

Es kann gezeigt werden, daß eine solche Zerlegung stets existiert und eindeutig ist. Ferner kann sie in kleinen und mittleren Dimensionen (bis etwa 100) numerisch stabil berechnet werden [LH74, DHB94]. Ein solches Verfahren ist in allen Standard-Numerikpaketen für die Lineare Algebra (z.B. LINPAK, LAPAK [Don79, Hag88]) implementiert. Die Anwendung auf das lineare Regressionsmodell ist unmittelbar klar [PFTV88]:

Satz 11.4 (Numerisch stabile Berechnung von Regressionsschätzern) Seien ein lineares Regressionsmodell $\eta = A \cdot \Theta + \varepsilon$ mit $\varepsilon = N(0, \Sigma)$ und eine Singulärwertzerlegung $\sqrt{\Sigma}^{-1} \cdot A = U \cdot \begin{pmatrix} D & 0 \end{pmatrix}^T \cdot V^T$ der reskalierten Systemmatrix gegeben. Dann gilt:

1. Die Quadrate der Singulärwerte von $\sqrt{\Sigma}^{-1} A$ sind die Eigenwerte von $A^T \Sigma^{-1} A$, und die Spaltenvektoren von V sind die zugehörigen (normierten) Eigenvektoren.
2. Der Quotient aus dem größten und dem kleinsten Singulärwert ist ein Maß für die numerische Kondition von $\sqrt{\Sigma}^{-1} A$. Diese Matrix hat genau dann Maximalrang, wenn die Diagonale von D voll besetzt ist.
3. Der KQ-Schätzer ist gegeben durch $\hat{\Theta} = V(D^{-1} 0) U^T \sqrt{\Sigma}^{-1} \eta$ und damit auch die Parameter-Sensitivität.
4. $\text{Cov } \hat{\Theta} = V D^{-2} V^T$.
5. Die Spalten von V geben die Richtung der Hauptachsen des in Satz 11.1 beschriebenen ellipsoiden Konfidenzgebiets $\text{Conf}_{\hat{\Theta}}(1 - \beta)$ an, die reziproken Singulärwerte ihre Längen.

Hierzu einige Anmerkungen:

1. In der Praxis muß bei schlechter Konditionierung von einer Entartung von A ausgegangen werden, d.h. der Rang ist im Bereich der numerischen Genauigkeit nicht mehr maximal. Die Rangdefizienz entspricht der Zahl der in diesem Sinne sehr kleinen Singulärwerte. Mit Hilfe der Singulärwertzerlegung kann auch in dieser Situation eine Regressionsschätzung durchgeführt werden [LH74, Páz93].

2. Aus den berechneten Hauptachsen kann unmittelbar die genaue geometrische Beschreibung des Parameter-Konfidenzbereichs abgelesen werden (Abb. 11.2a). Dabei sind nur die längsten Achsen von Interesse, da diese für die Unbestimmtheit der Parameterschätzung hauptsächlich verantwortlich sind. Die schwer zu schätzenden Parameter sind dann durch große Koordinateneinträge in den längsten Achsen gekennzeichnet.
3. Oft sind viele Koordinateneinträge einer langen Hauptachse sehr klein im Verhältnis zu den großen Einträgen, d.h. die Achse hat das Besetzungsmuster

$$(0, \dots, 0, *, 0, \dots, 0, *, 0, \dots)$$

mit $*$ $\gg 0$. Eine solche Achse trägt in starkem Maße zu den Konfidenzintervallen der durch $*$ markierten Parameter bei. Zudem kann in diesem Fall auf eine Korrelation zwischen diesen Parametern geschlossen werden (vgl. Abb. 11.2). Kommen mehrere Achsen mit dem gleichen Besetzungsmuster vor, so spannen diese einen Unterraum auf, an dem sich der Konfidenzellipsoid ausrichtet.

Findet man beispielsweise in einem dreidimensionalen Parameterraum eine einzige lange Achse mit dem Muster $(*, 0, 0)$ vor, so hat der Ellipsoid eine Spindelform und richtet sich an der ersten Achse aus. Findet man zwei lange Achsen mit dem Muster $(*, *, 0)$, so liegt eine Linsenform ausgerichtet an den beiden ersten Koordinatenachsen vor.

4. Die Korrelationsmatrix $\text{Corr}(\hat{\Theta})$ ist für die Analyse des Konfidenzbereichs nur eingeschränkt nutzbringend. Bekanntermaßen charakterisieren Einträge nahe ± 1 eine Entartung des Kovarianzellipsoids (Abb. 11.2b). Jedoch gilt dies nur für die Spindelform, d.h. dann wenn eine einzige dominierende Achse auftritt, deren Besetzungsmuster genau zwei Einträge hat. Liegen mehrere lange Achsen mit gleichem Muster vor, so treten zwar deutliche Korrelationen auf, jedoch liegen diese nicht mehr bei ± 1 , da sich die gegenseitige Abhängigkeit auf mehr als zwei Parameter verteilt.
5. Bei Vorliegen großer Korrelationen oder allgemeiner beim Vorliegen mehrerer langer Achsen mit gleichem Besetzungsmuster kann die Parameter-Sensitivitätsmatrix herangezogen werden, um gemeinsame Einflußfaktoren der korrelierten Parameter zu ermitteln. Meist tritt der Fall auf, daß n der Parameter gemeinsam im wesentlichen durch $m < n$ der Meßwerte beeinflußt werden und nur schwach durch weitere Parameter. Diese Haupteinflussfaktoren können dann bestimmt werden und darüber hinaus kann nach zusätzlichen Meßwerten gesucht werden, welche den gemeinsamen Einfluß entkoppeln.

11.2.5 Redundanzanalyse

In der Praxis liegen meist mehr Meßwerte als Parameter vor, d.h. $\dim \eta > \dim \Theta$. Bei einem linearen Modell mit einer Maximalrangmatrix A müssen folglich $\dim \eta - \dim \Theta$ Meßwerte redundant sein. Allerdings ist es falsch, von *den* redundanten Meßgrößen zu sprechen, da die Redundanz über alle einzelnen Meßwerte η_i "verteilt" ist. Dies wird durch eine sogenannte Redundanzmatrix [HHHR⁺94] präzisiert:

Definition 11.4 (Redundanzmatrix) Eine Redundanzmatrix R für ein lineares Modell $\eta = A \cdot \Theta + \varepsilon$ mit $\varepsilon = N(0, \Sigma)$ ist eine Matrix mit $\text{rang } R = \dim \eta - \dim \Theta$ Zeilen, für die gilt

$$R \cdot A = 0.$$

Im Falle einer fehlerfreien Messung η (d.h. $\varepsilon = 0$) ist also $\mathbf{R} \cdot \eta = \mathbf{R} \cdot \mathbf{A} \cdot \Theta = 0$. Unter Berücksichtigung des Meßrauschens kann ein redundanter Wert zwar keine zusätzliche Information zur Bestimmung der zu schätzenden Parameter liefern, jedoch die Genauigkeit der Schätzung erhöhen. Wie man leicht nachrechnet, erlaubt die Singulärwertzerlegung eine effiziente Berechnung von Redundanzbeziehungen:

Satz 11.5 (Konstruktion von Redundanzbeziehungen) Seien ein lineares Regressionsmodell $\eta = \mathbf{A} \cdot \Theta + \varepsilon$ mit $\varepsilon = N(0, \Sigma)$ und eine Singulärwertzerlegung $\sqrt{\Sigma}^{-1} \cdot \mathbf{A} = \mathbf{U} \cdot (\mathbf{D} \ 0)^T \cdot \mathbf{V}^T$ der reskalierten Systemmatrix gegeben. Dann ist eine Redundanzmatrix zu $\mathbf{A}$ gegeben durch die $\dim \eta - \dim \Theta$ -zeilige Matrix

$$\mathbf{R} = \sqrt{\Sigma} \cdot \left(\underbrace{0}_{\dim \Theta} \quad \underbrace{1}_{\dim \eta - \dim \Theta} \right) \cdot \mathbf{U}^T.$$

Für die Praxis der nichtlinearen Regression ist diese Form der Redundanzanalyse jedoch kaum einsetzbar, weil die Systemmatrix $\mathbf{A}$ als Linearisierung einer nichtlinearen Funktion F von $\hat{\Theta}$ abhängt und damit a priori nicht gegeben ist. Das Konzept der Redundanzanalyse kann jedoch in speziellen Situationen auf den nichtlinearen Fall erweitert werden (vgl. dazu [Wie95a, Wie96a] und die Ausführungen in Abschnitt 4.3).

11.2.6 Ausreißererkennung

Bekanntlich gewichtet die Fehlerquadratsumme Ausreißer in den Meßwerten besonders stark. Einzelne Meßwerte haben dann eine besonders große Auswirkung auf die Fehlerquadratsumme, so daß in der Folge angenommen werden muß, daß das Modell die Daten nicht adäquat beschreibt (vgl. Satz 11.1). Eine andere Sichtweise stellt der Versuch dar, Ausreißerwerte zu identifizieren und aus dem Datensatz zu entfernen [Him78].

Der in [HRHH⁺94] durchgeführte Ansatz nutzt hierzu die nach Satz 11.5 konstruierten Redundanzrelationen. Dabei wird versucht, sowohl das Problem der Erkennung von Ausreißermessungen als auch der Erkennung von Modellfehlern zu behandeln. Dazu wird konkret die bekannte Normalverteilung von $\mathbf{R} \cdot \eta = \mathbf{R} \cdot \varepsilon$ zur Formulierung eines χ^2 -Tests herangezogen. Zu diesem Zweck wird ein Modell im Vergleich zu einem um einen Ausreißerkandidaten reduzierten Modell getestet.

11.2.7 Versuchsplanung

Im Falle der Stoffflußanalyse ist ein wesentliches Element der Versuchsplanung die Auffindung solcher Meßgrößen, die möglichst viel Information über den Parametervektor enthalten. Dabei ist zu berücksichtigen, daß bei der in Abschnitt 2.5 beschriebenen Entkopplungsmethode durchaus nachträgliche Markierungsmessungen zu einem bereits durchgeführten Experiment gemacht werden können. Für das Modell bedeutet dies die Hinzunahme weiterer Meßgleichungen.

Wird angenommen, daß der aufgrund des vorhandenen Datensatzes η geschätzte Parameter $\hat{\Theta}(\eta)$ dem wahren Parameter schon recht nahe kommt, so kann hypothetisch die Auswirkung einer Erweiterung des Meßvektors zu η' untersucht werden. Wenn sich der Flußschätzer hierdurch nicht wesentlich ändert, also $\hat{\Theta}(\eta) \approx \hat{\Theta}(\eta')$ gilt, kann die entsprechend zu erwartende Kovarianzmatrix $\text{Cov } \hat{\Theta}(\eta')$ näherungsweise berechnet werden. Daraus läßt sich ablesen, inwieweit eine entsprechende Messung die statistische Güte des Schätzers verbessern würde. Der Aufwand

einer zusätzlichen Messung lohnt sich nur, wenn eine wesentliche Verbesserung zu erwarten ist. Dazu wird ein Vergleichsmaß für Kovarianzmatrizen benötigt.

Als Informationsmaß wird in der Versuchsplanung meist die sogenannte *Fisher-Information* verwendet [Páz86, Mun89]. Diese ist proportional zur Fläche des Konfidenzellipsoids

$$\left\{ \Theta \mid (\hat{\Theta} - \Theta)^T [\text{Cov } \hat{\Theta}]^{-1} (\hat{\Theta} - \Theta)^T < \chi_{\dim \Theta}^2 (1 - \beta) \right\}$$

aus Satz 11.1. Da diese Fläche wiederum proportional zur Determinante von $\text{Cov } \hat{\Theta}$ ist, eignet sich letztere als Maßzahl zum Vergleich von Kovarianzmatrizen [TWWB96]. Es sollte gelten

$$\det \text{Cov } \hat{\Theta}(\eta) \ll \det \text{Cov } \hat{\Theta}(\eta') \quad ,$$

damit sich eine zusätzliche Messung lohnt. In der Praxis ergibt sich die Fisher-Information bis auf einen konstanten Faktor sofort aus dem Produkt der Singulärwerte der Systemmatrix. Die Versuchsplanung läuft also auf eine Simulationsstudie auf der Basis bereits gemessener Daten hinaus.

11.3 Ein Beispielsystem

In diesem Abschnitt wird anhand eines einfachen Beispiels untersucht, inwieweit der vorgestellte Linearisierungsansatz für die Flußschätzung zu zufriedenstellenden Ergebnissen führt. Es wird sich dabei zeigen, daß die Anwendung der Linearisierung bei der Schätzung von Austauschflüssen stark verzerrte Ergebnisse liefern kann, während das Modell bzgl. der Nettoflüsse im allgemeinen gut linearisierbar ist. Dieses Beispiel wird in [WG96b, WSGM96] noch eingehender analysiert.

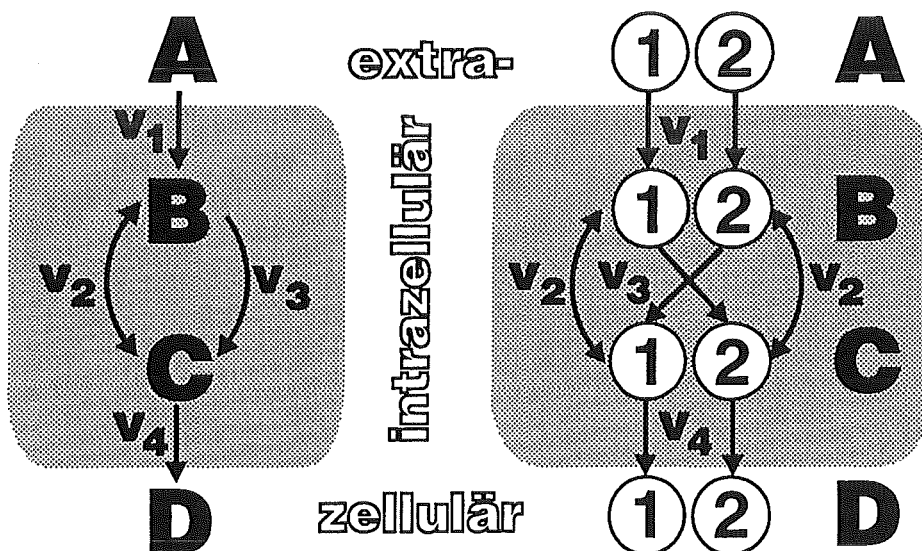


Abbildung 11.3: Beispielnetzwerk zur Veranschaulichung der statistischen Eigenschaften der Flußschätzung. Links: Metabolit-Netzwerk. Rechts: C-Atom-Netzwerk. Die Flüsse v_1 , v_3 , v_4 werden als unidirektional angenommen. Der Fluß v_1 und die Markierungsanteile in B#1, C#1 sind meßbar. Das Substrat ist auf dem ersten C-Atom markiert.

11.3.1 Ein Modellsystem

Als Beispiel wird das in Abb. 11.3 dargestellte einfache Netzwerk betrachtet. Es wurde so gewählt, daß ein Austauschfluß auftritt und zugleich eine zweidimensionale graphische Darstellung der Lösungsgleichungen möglich ist. Die Netzwerkstruktur ist gegeben durch:

$$\begin{array}{ll} v1: & A > B \\ & \#ab > \#ab ; \\ v2: & B > C \\ & \#ab > \#ab ; \\ v3: & B > C \\ & \#ab > \#ba ; \\ v4: & C > D \\ & \#ab > \#ab ; \end{array}$$

Es wird angenommen, daß die Reaktion v3 unidirektional ist, während v2 in beiden Richtungen abläuft. Die Substrataufnahme v1 und die Produktexkretion v4 sind nach den grundsätzlichen Annahmen aus Definition 3.5 unidirektional:

$$v_1^{\leftarrow} = 0 \quad v_3^{\leftarrow} = 0 \quad v_4^{\leftarrow} = 0$$

Aufgrund dieser Annahmen und der stöchiometrischen Gleichungen

$$\begin{array}{l} B: v_1^{\rightarrow} = v_2^{\rightarrow} - v_2^{\leftarrow} + v_3^{\rightarrow} \\ C: v_4^{\rightarrow} = v_2^{\rightarrow} - v_2^{\leftarrow} + v_3^{\rightarrow} \end{array} \quad (11.3)$$

können nun die Flußvariablen $v_3^{\leftarrow}$ und $v_4^{\rightarrow}$ eliminiert werden:

$$\begin{array}{l} v_4^{\rightarrow} = v_1^{\rightarrow} \\ v_3^{\rightarrow} = v_1^{\rightarrow} - v_2^{\rightarrow} + v_2^{\leftarrow} \end{array} \quad (11.4)$$

Das Flußsystem ist damit auf die freien Flüsse $v_1^{\rightarrow}$, $v_2^{\rightarrow}$, $v_2^{\leftarrow}$ reduziert. Die Markierungs-Bilanzgleichungen sind gegeben durch:

$$\begin{array}{ll} B\#1: 0 = & v_1^{\rightarrow} A_1 - v_2^{\rightarrow} B_1 + v_2^{\leftarrow} C_1 - v_3^{\rightarrow} B_1 \\ B\#2: 0 = & v_1^{\rightarrow} A_2 - v_2^{\rightarrow} B_2 + v_2^{\leftarrow} C_2 - v_3^{\rightarrow} B_2 \\ C\#1: 0 = & v_2^{\rightarrow} B_1 - v_2^{\leftarrow} C_1 + v_3^{\rightarrow} B_2 - v_4^{\rightarrow} C_1 \\ C\#2: 0 = & v_2^{\rightarrow} B_2 - v_2^{\leftarrow} C_2 + v_3^{\rightarrow} B_1 - v_4^{\rightarrow} C_2 \end{array} \quad (11.5)$$

Die Eingangsmarkierung des Systems sei der Einfachheit halber vorgegeben mit

$$a_1 = 1 \quad a_2 = 0 \quad .$$

11.3.2 Explizite Lösungen

Da im Beispiel drei freie Flüsse zu schätzen sind, müssen mindestens drei Meßwerte gegeben sein. Durch Addition von Gleichungen im Gleichungssystem (11.5) erhält man sofort die *Markierungs-Erhaltungsgleichungen*

$$b_1 + b_2 = 1 \quad c_1 + c_2 = 1 \quad ,$$

aus denen hervorgeht, daß b_1 mit b_2 und c_1 mit c_2 redundant ist. Da ferner nach Gleichung (11.4) die beiden extrazellulären Flüsse $v_1^{\rightarrow}$ und $v_4^{\rightarrow}$ übereinstimmen, ist eine Flußschätzung nur auf der Basis einer Messung von $v_1^{\rightarrow}$, b_1 , c_1 zu erreichen. Die Meßgleichungen seien also:

$$w = \begin{pmatrix} 1 & . & . & . & -1 & . & . & . \end{pmatrix} \cdot v + \varepsilon_w \quad \text{und} \quad y = \begin{pmatrix} 1 & . & . & . \\ . & . & 1 & . \end{pmatrix} \cdot x + \varepsilon_y$$

Da in dem Beispiel die Zahl der Parameter mit der der Meßwerte übereinstimmt, entspricht der KQ-Schätzer genau der expliziten Lösung von Gleichung (11.5) bzgl. der Flüsse $v_2^{\rightarrow}$ und $v_2^{\leftarrow}$. Die Lösung des Flußschätzproblems ist also explizit gegeben durch die Gleichungen [Wie95a]

$$\begin{aligned} v_2^{\leftarrow} &= -\frac{1-B_1}{1-B_1-C_2} \cdot v_1^{\rightarrow} \\ v_2^{\rightarrow} &= \frac{C_2-B_1}{1-2B_1} \cdot (v_2^{\leftarrow} + v_1^{\rightarrow}) \end{aligned} \quad (11.6)$$

wobei für die Variablen $v_1^{\rightarrow}$, b_1 , c_1 die entsprechenden verrauschten Meßwerte einzusetzen sind.

Andererseits ist die Lösung der Bilanzgleichungen bzgl. des Markierungszustands gegeben durch:

$$\begin{aligned} B_1 &= \frac{-v_2^{\leftarrow 2} - 2v_2^{\leftarrow} + v_2^{\rightarrow} v_2^{\leftarrow 2} - v_1^{\rightarrow}}{-2v_2^{\leftarrow 2} - 3v_2^{\leftarrow} + 2v_2^{\rightarrow} v_2^{\leftarrow 2} - v_1^{\rightarrow}} \quad , \quad B_2 = 1 - B_1 \\ C_1 &= \frac{-v_2^{\leftarrow 2} - v_2^{\leftarrow} + v_2^{\rightarrow} v_2^{\leftarrow 2} - v_2^{\rightarrow}}{-2v_2^{\leftarrow 2} - 3v_2^{\leftarrow} + 2v_2^{\rightarrow} v_2^{\leftarrow 2} - v_1^{\rightarrow}} \quad , \quad C_2 = 1 - C_1 \end{aligned} \quad (11.7)$$

Von nun an wird zur Vereinfachung

$$v_1^{\rightarrow} = 1$$

angenommen, da $v_1^{\rightarrow}$ nach Satz 7.3 lediglich die Skala für die Größenordnung der Flüsse vorgibt.

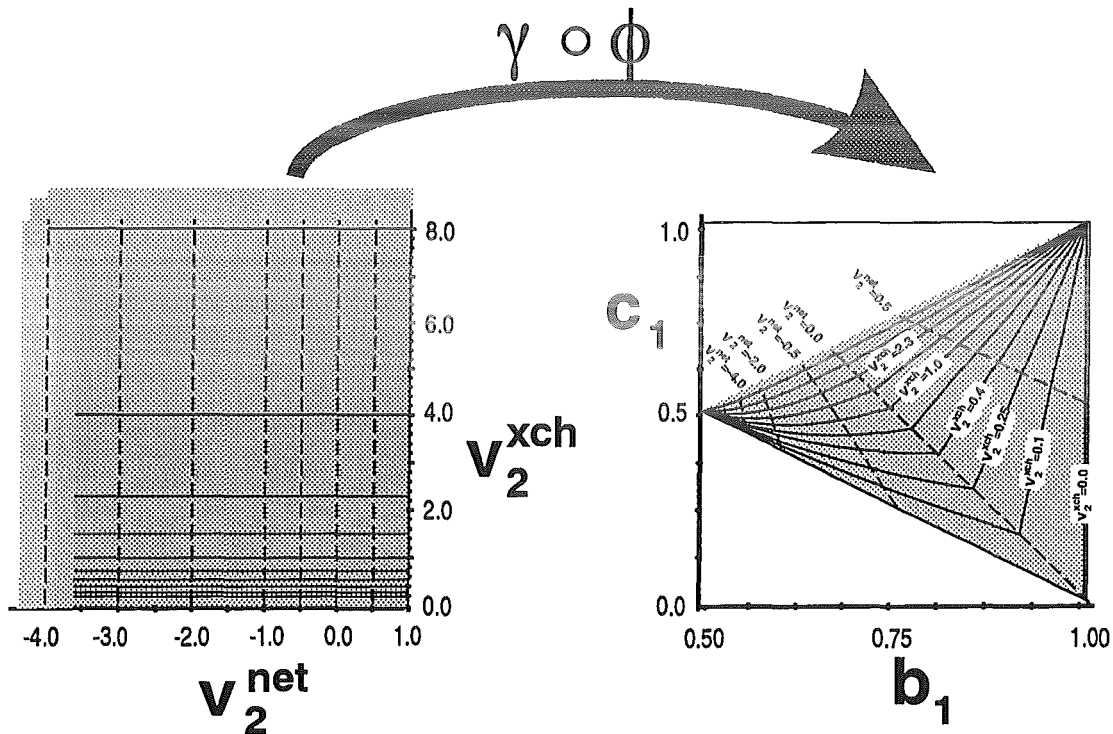


Abbildung 11.4: Definitions- und Bildbereich der Abbildung $\gamma \circ \phi : (v_2^{\text{net}}, v_2^{\text{xch}}) \rightarrow (b_1, c_1)$ für das Beispielnetzwerk aus Abb. 11.3. Es ist $v_1^{\rightarrow} = 1$ angenommen.

11.3.3 Realisierbare Markierungszustände

Analog zur Behandlung des Beispiels in Abschnitt 5.2 wird für die graphische Darstellung der Zusammenhänge zunächst das $(v_2^{\text{net}}, v_2^{\text{xch}})$ -Koordinatensystem betrachtet. Wenn γ die durch die Gleichungen (11.7) und (11.6) festgelegte Bijektion zwischen Flüssen und Markierungsanteilen und ϕ die Flußkoordinaten-Transformation entsprechend Satz 5.4 ist, so ergibt sich die Bijektion

$$\gamma \circ \phi : (v_2^{\text{net}}, v_2^{\text{xch}}) \xrightarrow{1-1} (v_2^{\rightarrow}, v_2^{\leftarrow}) \xrightarrow{1-1} (b_1, c_1) \quad .$$

Die zulässigen Flüsse (also der Definitionsbereich von $\gamma \circ \phi$) sind dann aufgrund von Gleichung (11.4) charakterisiert durch:

$$0 \leq v_2^{\text{sch}} \quad \text{und} \quad v_2^{\text{net}} \leq 1$$

Die realisierbaren und asymptotisch realisierbaren (b_1, c_1) -Markierungen (also der Bildbereich von $\gamma \circ \phi$) ergeben sich als das durch

$$b_1 \geq \frac{1}{2}, \quad b_1 \geq c_1, \quad b_1 + c_1 \geq 1$$

festgelegte Dreieck (vgl. Abb. 11.4). Die asymptotische Grenze für $v_2^{\text{sch}} \rightarrow \infty$ ist die durch $b_1 = c_1$ bestimmte Dreiecksseite.

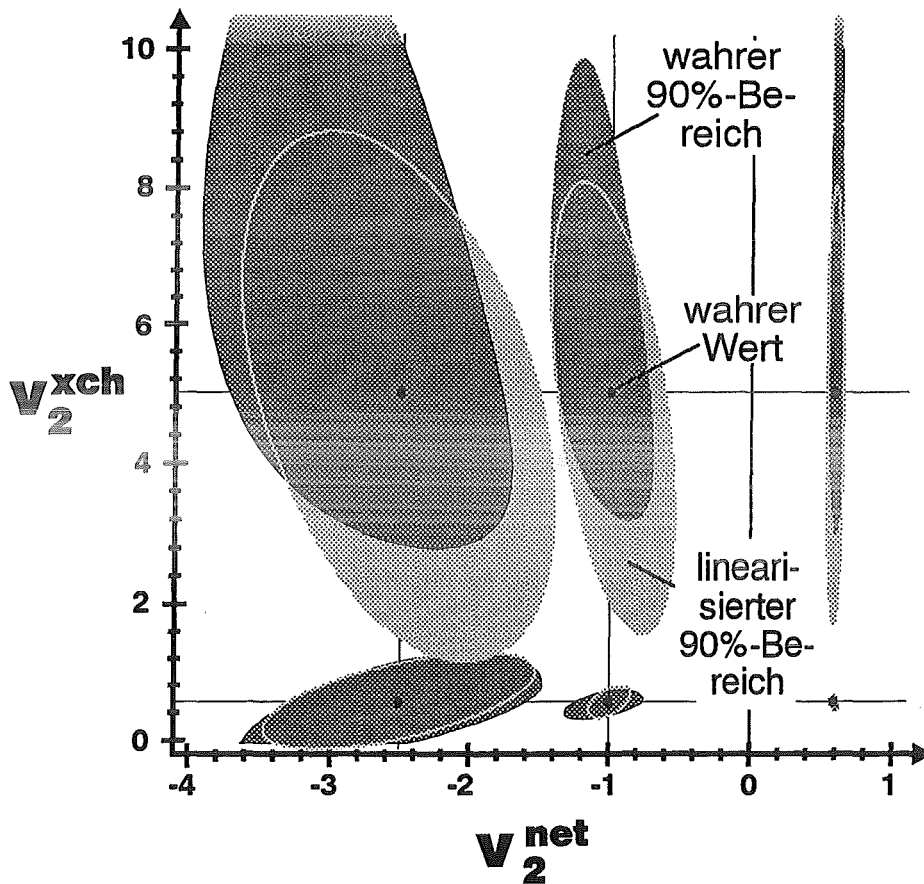


Abbildung 11.5: Verteilung des Flußschätzers im Beispiel aus Abb. 11.3 ($v_1^* = 1$). Gezeigt ist jeweils der 90%-Aufenthaltsbereich des Schätzers $\hat{\Theta}$ bei bekanntem wahren Wert von v_2^{net} und v_2^{sch} . Zugrundegelegt ist dabei ein Meßfehler mit Standardabweichung 0.008 für b_1 und c_1 . Im Vergleich dargestellt sind die elliptischen 90%-Konfidenzbereiche der Linearisierung des Modells um die wahren Werte von v_2^{net} und v_2^{sch} .

11.3.4 Exakte Verteilung des Schätzers

Bezeichnen ε_1 bzw. ε_2 die (unabhängigen) Meßfehler der Messung von b_1 bzw. c_1 , so ist offenbar

$$(\gamma \circ \phi)^{-1}(b_1 + \varepsilon_b, c_1 + \varepsilon_c) \quad (11.8)$$

der zugehörige KQ-Schätzer der intrazellulären Flüsse.

Es wird nun eine Meßgenauigkeit von 0.008 (Standardabweichung) für die Messung der beiden Markierungsanteile zugrundegelegt. Damit liegt die Markierungsmessung mit $(1 - \beta)$ -prozentiger Wahrscheinlichkeit in dem Kreis [Arn90] (vgl. Abb. 11.2)

$$K_\beta \stackrel{\text{def}}{=} \{ (b_1 + \varepsilon_1, c_1 + \varepsilon_2) \mid \varepsilon_1^2 + \varepsilon_2^2 \leq 0.008^2 \cdot \chi_2^2(1 - \beta) \}$$

mit dem Mittelpunkt (b_1, c_1) . Folglich liegt der Flußschätzer mit derselben Wahrscheinlichkeit in dem Gebiet

$$(\gamma \circ \phi)^{-1}(K_\beta) \quad .$$

Dies liefert eine exakte quantitative Vorstellung von der wahren Verteilung des Schätzers. Abbildung 11.5 stellt die 90%-Verteilungsbereiche für einige repräsentative Punkte im $(v_2^{\text{net}}, v_2^{\text{xch}})$ -Raum dar. Es wird deutlich, daß mit Ausnahme der extremen Situation zweier im Verhältnis zu $v_1^{\rightarrow}$ großen Flüsse $v_2^{\text{net}}, v_2^{\text{xch}}$ die Nettoflußrate v_2^{net} mit vertretbar großen Konfidenzintervallen schätzbar ist. Dagegen ist die Schätzung großer Austauschflüsse nur innerhalb weiter Bereiche möglich. Diese Bereiche haben eine unsymmetrische Lage in Bezug auf den wahren Wert.

Die Flußkoordinatentransformation ϕ entkoppelt also den (relativ) gut bestimmbaren Nettofluß von dem schlecht bestimmbaren Austauschfluß. Dies ist ein weiterer Beleg für die Vorteile der alternativen Koordinatendarstellung. Der Austauschfluß ist dagegen in der Regel nur innerhalb einer Größenordnung schätzbar. Immerhin sind damit unidirektionale Flüsse unterscheidbar von schwach und stark austauschenden Reaktionen. Diese Unterscheidung ist für die praktische Anwendung wichtiger als die Kenntnis exakter Zahlenwerte. Die aus dem Beispiel hergeleiteten Ergebnisse können zwar nicht ohne weiteres auf die allgemeine Situation übertragen werden, die Erfahrung zeigt jedoch, daß das Beispiel wichtige Aspekte veranschaulicht.

11.3.5 Linearisierungsergebnisse

Die erzielten exakten Ergebnisse bzgl. der Verteilung des Schätzers im $(v_2^{\text{net}}, v_2^{\text{xch}})$ -Raum sollen nun mit den Ergebnissen eines Linearisierungsansatzes verglichen werden. Dabei wird das Regressionsmodell direkt im $(v_2^{\text{net}}, v_2^{\text{xch}})$ -Raum um den wahren Wert der Flüsse linearisiert.

Die Linearisierung der Funktion $\gamma \circ \phi$ ist anhand der expliziten Formeln (11.7) und (11.6) zu berechnen. Die daraus nach Satz 11.1 bestimmten Konfidenzbereiche bei Linearisierung um den wahren Wert zeigt Abbildung 11.5. Im Vergleich zu den in derselben Abbildung dargestellten exakten Bereichen wird sofort deutlich, daß

1. die Linearisierung bzgl. der Nettoflußraten im allgemeinen zulässig ist, während sie
2. bzgl. großer Austauschflüsse zu stark verzerrten Resultaten führt. Es kommt dann zu einer Überschätzung des Konfidenzgebietes nach unten und einer Unterschätzung nach oben.⁹

Offenbar ist die durch die Linearisierung hervorgerufene statistische Verzerrung sehr ungünstig für die Bestimmung der Größenordnung von Austauschflüssen.

⁹Derartige Diskrepanzen sind auch von anderen nichtlinearen statistischen Modellen bekannt (vgl. z.B. [BW88]).

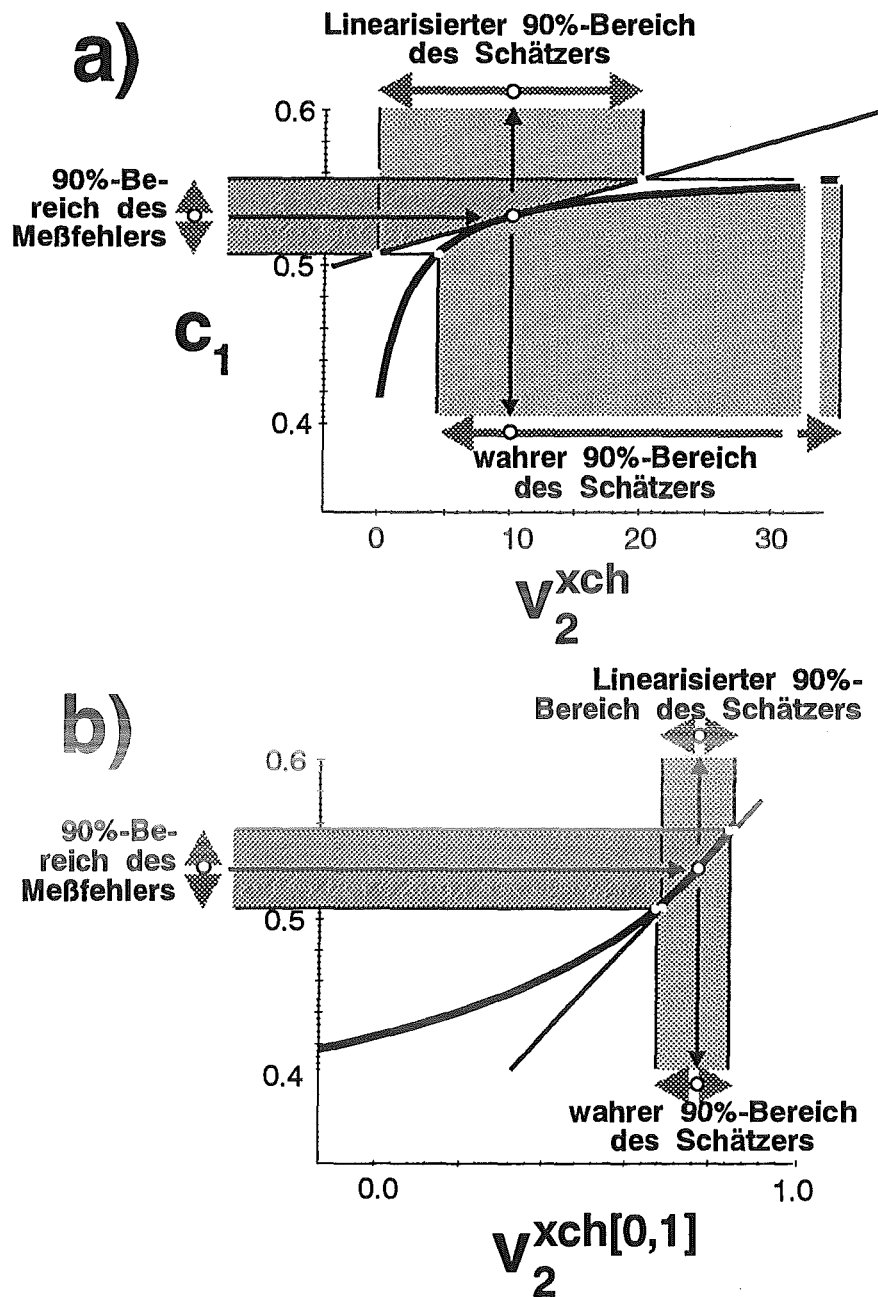


Abbildung 11.6: a) Abhängigkeit des Markierungsanteils c_1 im Beispielnetzwerk aus Abb. 11.3 vom Austauschfluß v_2^{xch} für $v_1^{net} = 1, v_2^{net} = -2.5$. Dazu der Vergleich des wahren 90%-Aufenthaltsbereichs des Austauschflußschätzers mit seiner Linearisierung. Für den Meßfehler von c_1 ist ein Standardabweichung von 0.012 angenommen. Dies entspricht dem Radius der zweidimensionalen Aufenthaltsbereiche in Abbildung 11.5. b) Analoge Darstellung für das $v_2^{xch[0,1]}$ -Koordinatensystem.

11.3.6 Nichtlineare Parameterreskalierung

Die Ursache des offenbar großen Approximationsfehlers soll hier noch einmal in einer eindimensionalen Perspektive dargestellt werden. Dazu wird der durch Messung direkt zugängliche Nettofluß mit $v_2^{\text{net}} = -2.5$ ebenfalls als bekannt angenommen und der Meßwert b_1 entfällt. Die Standardabweichung des Meßwertes für c_1 wird zur Kompensation etwas größer angenommen (vgl. Abb. 11.5 und Abb. 11.6). In dieser Situation zeigt Abbildung 11.6a die Abhängigkeit des Markierungsanteils c_1 vom Austauschfluß v_2^{xch} . Es ergibt sich das allgemein in Satz 8.4 bewiesene asymptotische Verhalten.

Bei hohen Austauschflüssen hat v_2^{xch} praktisch keine Wirkung mehr auf die betroffenen Markierungsanteile, d.h. die Output-Sensitivität $\partial c_1 / \partial v_2^{\text{xch}}$ tendiert gegen Null. Die Linearisierung ersetzt daher den Verlauf der Kurve durch eine annähernd waagerechte Tangente, so daß die Parameter-Sensitivität und zugleich die geschätzte Varianz mit $v_2^{\text{xch}} \rightarrow \infty$ gegen Unendlich laufen.

Die Auswirkung der Linearisierung auf die Bestimmung der Größe und Lage der Konfidenzintervalle kann nun leicht an der graphischen Darstellung aus Abbildung 11.6a abgelesen werden. Wird der Wert auf der Ordinate innerhalb der durch den Meßfehler gegebenen Grenzen variiert, so können die zugehörigen Austauschflüsse auf der Abszisse abgelesen werden. Ausgehend von der nichtlinearen Originalkurve ergibt sich dann auf der Abszisse ein unsymmetrischer Bereich, der auf der linken Seite des geschätzten Flusses wesentlich kürzer ist als auf der rechten. Im Extremfall erstreckt sich dieser Bereich nach rechts bis ins Unendliche.

Im Vergleich dazu liefert der Linearisierungsansatz stets symmetrische Konfidenzintervalle, die links bis in den negativen Bereich hineinreichen können, aber rechts immer endlich sind. Wie Abbildung 11.6a zeigt, führt die Linearisierung für große v_2^{xch} also zu einer unzulässig weiten Extrapolation des lokalen Kurvenverlaufs. Nur wenn der wahre Austauschfluß klein ist, stimmen Linearisierung und Originalkurve annähernd überein.¹⁰

Im Vergleich zur Linearisierung im v_2^{xch} -Raum illustriert die Abbildung 11.6b dieselbe Vorgehensweise im $v_2^{\text{xch}[0,1]}$ -Raum, d.h. nach einer nichtlinearen Parameter-Reskalierung (vgl. Definition 5.3 und 5.4). Die graphische Auftragung zeigt, daß die Ableitung im Punkt $v_2^{\text{xch}[0,1]} = 1 \Leftrightarrow v_2^{\text{xch}} = \infty$ jetzt von Null verschieden ist, der oben beschriebene Effekt verschwindender Sensitivitäten tritt hier also nicht mehr auf. Entsprechend sind nichtlineare und linearisierte Konfidenzintervalle in wesentlich besserer Übereinstimmung.

Für die Anwendung muß im Anschluß allerdings noch das bzgl. $v_2^{\text{xch}[0,1]}$ ermittelte Konfidenzgebiet mit Hilfe von $\phi^{[0,1]}$ in den v_2^{xch} -Raum zurücktransformiert werden:

$$v_2^{\text{xch}[0,1]} \in [\hat{v}_2^{\text{xch}[0,1]} - \sigma, \hat{v}_2^{\text{xch}[0,1]} + \sigma] \xrightarrow{\phi^{[0,1]}} v_2^{\text{xch}} \in \left[\frac{(\hat{v}_2^{\text{xch}[0,1]} - \sigma)}{1 - (\hat{v}_2^{\text{xch}[0,1]} - \sigma)}, \frac{(\hat{v}_2^{\text{xch}[0,1]} + \sigma)}{1 - (\hat{v}_2^{\text{xch}[0,1]} + \sigma)} \right]$$

Dabei ist σ die jeweils ermittelte Größe des symmetrischen Vertrauensintervalls für $v_2^{\text{xch}[0,1]}$ (vgl. Satz 11.1). Es entsteht offenbar ein unsymmetrisches Konfidenzintervall für v_2^{xch} , das der wahren Verteilung wesentlich näher kommt als das auf einer Linearisierung im v_2^{xch} -Raum beruhende symmetrische Intervall.

11.3.7 Zur Wahl einer nichtlinearen Reskalierung

Allgemein kann die Frage gestellt werden, ob die Konfidenzbereichsschätzung durch nichtlineare Reskalierung im Parameterraum verbessert werden kann [Páz93]. Das Beispiel verdeutlicht

¹⁰In der Terminologie der nichtlinearen Statistik gibt die sogenannte *Parameter-Effekt-Krümmung* ein gutes Maß für die Güte eines Linearisierungsansatzes. Für große v_2^{xch} tendiert diese Krümmung gegen unendlich [Páz93].

jedoch, daß die Frage in dieser Weise schlecht gestellt ist, denn es gibt im Beispiel eine triviale nichtlineare Reskalierung des Parameterraums, die das Modell vollständig linear macht, nämlich $(\gamma \circ \phi)^{-1}$ (vgl. Gleichung (11.8)). Die neuen Parameter des Modells sind dann ganz einfach b_1 und c_1 , und die Modellfunktion ist die Identität!

Dies zeigt, daß eine nichtlineare Parameter-Reskalierung für die Praxis eine wichtige Bedingung erfüllen muß: Das Ergebnis der Reskalierung muß bei Rücktransformation in den Originalraum praktisch interpretierbar bleiben, d.h. durch wenige Maßzahlen, deren Bedeutung anschaulich verstanden werden kann, repräsentierbar sein. Aus diesem Grunde können im wesentlichen nur solche Transformationen verwendet werden, die auf jeder Koordinate getrennt arbeiten. Deshalb ist im Beispiel eine Transformation, die lediglich die Austauschfluß-Koordinate linearisiert, nicht möglich. Eine solche Transformation müßte nämlich auch von v_2^{net} abhängen (vgl. Abb. 5.5). Die gefundene $[0, 1]$ -Reskalierung ist also eine gute Wahl für die praktische Anwendung.

11.3.8 Praktische Vorgehensweise

Zusammenfassend wird bei der praktischen statistischen Auswertung von Markierungsexperimenten der folgende Weg beschritten:

1. Die Parameteranpassung und anschließende linearisierte Analyse erfolgt im kompaktifizierten $(v^{\text{net}}, v^{\text{xch}[0,1]})$ -Raum.
2. Der resultierende linearisierte Konfidenzbereich wird in den $(v^{\text{net}}, v^{\text{xch}})$ -Raum zurücktransformiert. Dies liefert symmetrische Konfidenzbereiche für v^{net} und unsymmetrische Konfidenzbereiche für v^{xch} .
3. Treten zu große Konfidenzbereiche auf, so muß geprüft werden, ob dies grundsätzlich an einer nicht ausreichenden Meßinformation liegt. Dabei helfen die Analysemethoden für lineare Modelle aus Abschnitt 11.2 und die in [Wie95a, Wie96a] vorgestellten algebraischen Methoden.
4. Darüber hinaus können nichtlineare Methoden zur Vertrauensbereichsschätzung für einzelne Parameter eingesetzt werden. Dazu gehören Monte-Carlo-Simulationsmethoden (vgl. hierzu [Möl95]) oder nichtlinear definierte Konfidenzbereiche [Har64, Sun78, Páz93]. Die praktische Umsetzung derartiger Ansätze ist allerdings nur für niedrige Parameterdimensionen sinnvoll und erfordert hohe Rechenaufwände sowie geeignete Visualisierungstechniken [Sco92, Möl95].

11.4 Regression unter Nebenbedingungen

11.4.1 Restringsierte KQ-Schätzung

Bisher wurde das Problem der Ungleichheits-Nebenbedingungen bei der Flußschätzung nicht berücksichtigt. Vielmehr wurde davon ausgegangen, daß sich der Parameter Θ im Raum frei bewegen kann. Eine statistische Theorie für Probleme mit Ungleichheits-Nebenbedingungen

$$h(\Theta) > 0$$

ist allerdings kaum entwickelt. Hier soll das Problem kurz für ein (bzgl. der Meßwerte skaliertes) einparametrisches lineares Modell mit einer einzigen linearen Nebenbedingung angeschnitten werden ($\mathbf{a}$ ist hier ein Vektor, Θ ein Skalar):

$$\begin{aligned} \eta &= \mathbf{a}^T \cdot \Theta + \varepsilon \\ \Theta &\geq b \end{aligned} \quad \text{mit} \quad \varepsilon = N(0, 1)$$

Der Einfachheit halber sei zudem $\|\mathbf{a}\| = 1$. In der unrestringierten Situation ist dann der KQ-Schätzer nach Satz 11.1 gegeben durch

$$\hat{\Theta} = \|\mathbf{a}\|^{-2} \mathbf{a}^T \cdot \eta = \mathbf{a}^T \cdot \eta \quad .$$

Dies Abbildung projiziert den Punkt η orthogonal auf die durch $\mathbf{a}$ bestimmte Ursprungsgerade und ermittelt dann die zugehörige eindimensionale Koordinate $\hat{\Theta}$ (vgl. auch Abb. 11.7).

Die Erweiterung auf den unrestringierten Fall ist unmittelbar klar: Falls ein Punkt η einen Wert $\hat{\Theta} < b$ liefert, so ist $\hat{\Theta} = b$ der im Sinne kleinster Quadrate nächstgelegene Parameter (Abb. 11.7a). Der restringierte KQ-Schätzer lautet also:

$$\hat{\Theta} = \begin{cases} \|\mathbf{a}\|^{-2} \mathbf{a}^T \cdot \eta & \text{falls } b \leq \|\mathbf{a}\|^{-2} \mathbf{a}^T \cdot \eta \\ b & \text{sonst} \end{cases} \quad (11.9)$$

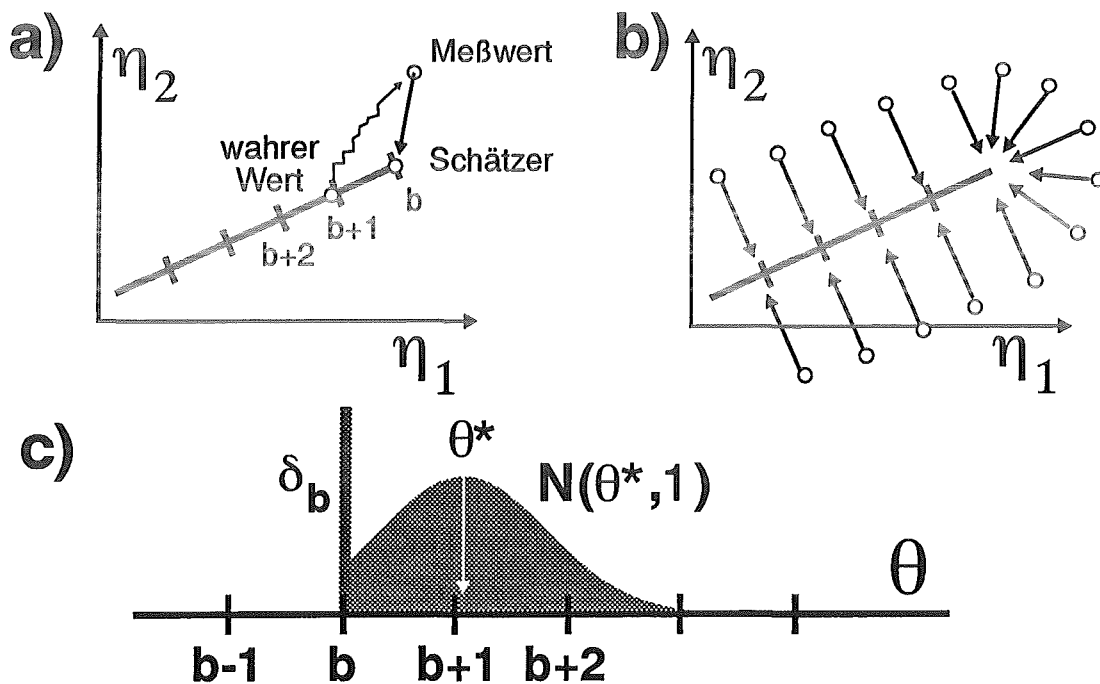


Abbildung 11.7: a) Restringierte KQ-Schätzung für ein einfaches lineares Modell mit $\dim \Theta = 1$ und $\dim \eta = 2$ sowie einer einzelnen linearen Nebenbedingung. b) KQ-Schätzer für verschiedene Meßwerte. c) Verteilung des Schätzers.

11.4.2 Verteilung des restringierten KQ-Schätzers

Aus der expliziten Formel (11.9) für den restringierten KQ-Schätzer ergibt sich sofort dessen Verteilung. Es handelt sich um eine "abgeschnittene" Normalverteilung (Abb. 11.7b). Nimmt man an, daß Θ^* der wahre Wert des Parameters ist, so folgt mit

$$p = P(N(0, 1) > b) \quad ,$$

daß die Dichte der Verteilung von $\hat{\Theta}$ gegeben ist durch:

$$f_{\Theta}(\xi) = (1 - p) \cdot f_{N(\Theta^*, 1)}(\xi) + p \cdot \delta_b(\xi)$$

Hierbei ist $f_{N(\Theta^*, 1)}$ die Normalverteilungsdichte mit Erwartungswert Θ^* und Varianz Eins und δ_b die Punktmasse Eins auf b .

Ein auf der Basis dieser Verteilung bestimmter Konfidenzbereich ist stets kleiner als das durch unrestringierte Schätzung erhaltene Gebiet. Im Extremfall eines Meßpunktes η für den gilt $\|\mathbf{a}\|^{-2} \mathbf{a}^T \cdot \eta \ll b$ schrumpft dieser Bereich sogar auf die einpunktige Menge $\{b\}$ zusammen. In dieser Situation muß allerdings untersucht werden, ob die Modellannahme für η korrekt war, d.h. es ist zu testen, ob η dem auf den Rand eingeschränkten Modell

$$\eta = \mathbf{a}^T \cdot \mathbf{b} + \varepsilon$$

genügt. Hierzu kann wieder ein χ^2 -Test herangezogen werden.

11.4.3 Anwendung auf die Flußanalyse

Aufgrund des Beispiels zeigt sich, daß der restringierte KQ-Schätzer stets "dichter" am wahren Wert liegt als der entsprechende unrestringierte Schätzer. Folglich führt die Übertragung statistischer Methoden aus der unrestringierten Theorie im Vergleich zur restringierten Theorie stets zu einer Überschätzung des exakten Konfidenzgebietes. Für die praktische Anwendung ist dieses Ergebnis vorteilhaft, denn die aus der unrestringierten Theorie erhaltenen Resultate sind auf jeden Fall richtig, wenn sie auch bei vollständiger Ausnutzung der Modellinformation noch verbesserbar sind. Dies ist eine Problemstellung der mathematischen Statistik, die den Rahmen dieser Arbeit übersteigt. Daher wird die unrestringierte Theorie ohne Bedenken angewandt.

Es kann auch ohne Berechnung der exakten Verteilungen leicht gezeigt werden, daß die oben für das Beispiel erzielten Resultate auch für höhere Dimensionen sowie im nichtlinearen Fall zutreffen. Stets erhält man durch "Beschneidung" eines unrestringierten Gebietes ein restringiertes Konfidenzgebiet.

Für den Fall, daß der Parameterschätzer auf dem Rand des physikalisch realisierbaren Gebietes liegt, muß untersucht werden, ob dies durch die Hypothese eines Meßfehlers noch erklärt werden kann. Dazu genügt die Festsetzung dieses Parameters (d.h. die Einschränkung des Modells auf den Rand) und eine erneute statistische Analyse.

Kapitel 12

Beispielauswertung eines Markierungsexperiments

Das abschließende Kapitel befaßt sich mit der Auswertung eines experimentellen Datensatzes, der bei *Corynebacterium glutamicum* gemessen wurde. Dieser Organismus ist ein industriell eingesetzter Aminosäureproduzent. Verschiedene Mutanten produzieren unter anderem Glutamat, Lysin, Isoleucin, Valin, Histidin und Threonin [Eik93, EG96]. Da für die Wirtschaftlichkeit der Aminosäureproduktion bereits geringe Verbesserungen der Ausbeute große Auswirkungen zeigen, wurde der Stofffluß in *C. glutamicum* in den letzten Jahren von vielen verschiedenen Arbeitsgruppen untersucht [ISNY⁺91, VS93, Dom93, Hol94, CB94, SSGM⁺95, RMGGK95]. Der Jülicher Stamm *C. glutamicum* MH20-22B ist ein Lysin-Produzent, der bei dem ersten der hier diskutierten Experimente etwa 20 % der aufgenommenen Glucose-Moleküle als Lysin ausschied. Der Hauptteil dieses Kapitels befaßt sich mit der statistischen Auswertung eines Datensatzes aus einem isotopisch stationären Experiment.

Nach der Auswertung dieses Experiments wird es einem instationären Experiment gegenübergestellt, das in einem innerhalb des NMR-Gerätes betriebenen Bioreaktor bei hoher Zelldichte durchgeführt wurde. Bereits eine grobe Inspektion der Daten und ihr Vergleich mit den in Kapitel 9 vorgestellten Simulationsergebnissen wird zeigen, daß die instationäre Kohlenstoff-Markierungsanreicherung im Zentralstoffwechsel schnell wachsender Mikroorganismen selbst mit einer derart aufwendigen Anordnung kaum zugänglich ist.

12.1 Stoffflußanalyse bei *Corynebacterium glutamicum*

Der hier vorgestellte Datensatz wurde mit Hilfe der in Kapitel 2 beschriebenen experimentellen Entkopplungstechnik (Definition 2.6) produziert, d.h. die extrazellulären Flußdaten sind mit Hilfe von Standard-Meßinstrumenten in einer stabilen kontinuierlichen Kultur gemessen, wogegen die Markierungsmeßdaten nach Erreichen eines annähernd isotopisch stationären Zustands aus dem Hydrolysat verschiedener Zellkomponenten gewonnen wurden. Seit seiner ersten Publikation in [MGWE⁺96] konnte der Datensatz noch durch die aus der DNA/RNA-Fraktion gewonnenen Meßwerte des Ribose-5-phosphat-Pools vervollständigt werden. Es zeigt sich, daß dadurch die geschätzten Nettoflüsse im Vergleich zu [MGWE⁺96] fast unbeeinflusst geblieben sind, jedoch verbessert sich die statistische Qualität der Austauschfluß-Schätzung.

Dieser Abschnitt konzentriert sich zunächst auf die Vorstellung des Datenmaterials und des Ergebnisses der Parameteranpassung. Bzgl. der experimentellen Details sei auf die ausführliche Darstellung in [Mar94, MGWE⁺96] verwiesen. Im folgenden Abschnitt wird dann erstmals die

statistische Qualität der Flußschätzung diskutiert. Weitere Ausführungen hierzu finden sich in [WSGM96].

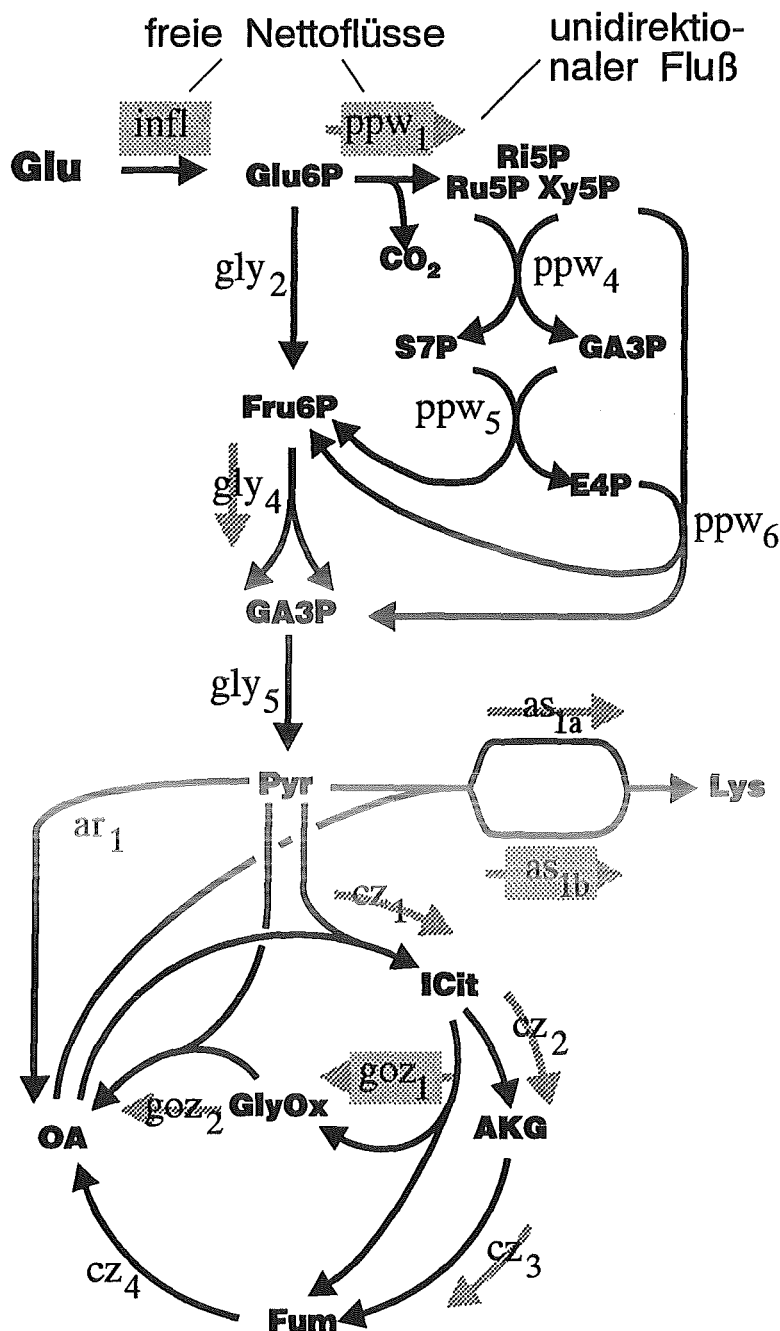


Abbildung 12.1: Metabolisches Netzwerk und Flußnamen bei *C. glutamicum*. Freie Nettoflüsse und als unidirektional angenommenen Reaktionsschritte sind markiert.

12.1.1 Das metabolische Netzwerk

Das der Modellierung zugrundeliegende metabolische Netzwerk ist das Resultat eines lang andauernden und intensiven Diskussionsprozesses zwischen den beteiligten Experimentatoren und Modellierern, der hier im Detail nicht wiedergegeben werden kann. Es ist in Abbildung 12.1 dar-

gestellt und entspricht im wesentlichen der Darstellung in Abbildung 1.2 (vgl. auch Tabelle 3.2 und 3.3). Jedoch wurden folgende Veränderungen vorgenommen:

1. Die drei Pentose-5-phosphat-Pools im Pentosephosphatweg wurden — wie in der Literatur üblich (vgl. [Woo85]) — als in schnellem Austausch befindlich angenommen, d.h. die Reaktionsschritte PPW2 und PPW3 in Tabelle 3.2 entfallen.
2. Für die Scrambling-Reaktionen im Citratzyklus wird grundsätzlich ein 50%-Scrambling angenommen, denn bei Prokaryonten sind bisher keine gegenteiligen Befunde bekannt (vgl. dazu die Diskussion in Kapitel 3.2 und Abb. 3.3).
3. Die Netzwerkformulierung aus Abbildung 1.2 enthält drei anaplerotische Reaktionsschritte, von denen zwei möglicherweise bidirektional ablaufen [Val91]. Diese werden durch die PEP-Carboxylase, die PEP-Carboxykinase und das Malat-Enzym katalysiert. Sogar ein viertes Enzym wird vermutet [TMT79]. Eine erste numerische Analyse zeigte jedoch, daß die Output-Sensitivitätsmatrix bei Einbeziehung aller drei Reaktionsschritte singulär wird, d.h. eine linear unabhängige Schätzung der anaplerotischen Flüsse ist nicht möglich. Eine algebraische Analyse dieses Systemabschnitts in [Wie96a] zeigte dann, daß eine Zusammenfassung des PEP- und Pyr-Pools einerseits und des Mal- und OA-Pools andererseits die übrigen Flußschätzungen nicht beeinflußt. Daher wird hier eine einzige bidirektionale anaplerotische Reaktion zwischen den jeweiligen zusammengefaßten Pools angenommen.
4. Die Lysinproduktion erfolgt in *C. glutamicum* auf zwei verschiedenen Wegen, wovon der eine einen Scrambling-Schritt involviert [SSK⁺91]. Dies führt zu einer Unterscheidbarkeit beider Wege anhand der Lysin-Markierung [SEGS93] (vgl. auch Abschnitt 1.2). Die beiden Wege sind in Tabelle 3.3 mit AS1a und AS1b bezeichnet.
5. Im Citratzyklus wurden die linearen Sequenzen CZ4-CZ5-CZ6 (CO₂-Abspaltung) und CZ7-CZ8 (Scrambling) zusammengefaßt. Stellvertretend für die erste Sequenz steht CZ4 und für die zweite CZ7.
6. Reaktionen, bei denen Moleküle gespalten werden, wurden als unidirektional angenommen, wenn die Rückreaktion aus thermodynamischen Gründen als unwahrscheinlich angesehen wird. Dies ist im Einzelfall zu diskutieren. Beispielsweise kommt es im anaplerotischen Abschnitt zu einer CO₂-Refixierung, wogegen der oxidative Pentosephosphatweg PPW1 als irreversibel angesehen wird.
7. Ein zellexterner CO₂-Pool im Medium wurde hinzugenommen, um die Refixierung adäquat beschreiben zu können.

12.1.2 Freie Flüsse und extrazelluläre Flußdaten

Es liegt nahe, den Glucose-Aufnahmeschritt Infl, die Gesamt-Lysin-Produktion über beide Wege und alle Biomasseabflüsse als freie Flüsse zu definieren, da alle diese Flüsse direkt gemessen werden können. In Bezug auf die Nettoflüsse verbleiben dann drei Freiheitsgrade, die durch intrazelluläre Flüsse repräsentiert werden müssen (vgl. Definition 6.3). Dies sind die für die Stoffflußanalyse eigentlich interessanten Flüsse. Folgende Netto-Flüsse erweisen sich als linear unabhängig von den bereits angenommenen freien Flüssen:

Fluß	Meßwert [%]	Geschätzter Wert [%]	Angenommener Meßfehler	Relative Abweichung
Substrataufnahme:				
<i>infl</i>	100.0	98.7	2.000	-0.63
Biomasseabflüsse:				
<i>glu6peffl</i>	1.3	1.3	0.026	0.04
<i>fru6peffl</i>	0.5	0.5	0.010	0.01
<i>gapeffl</i>	0.9	0.9	0.018	0.02
<i>pyreffl₁</i>	18.0	18.0	0.360	0.37
<i>pyreffl₂</i>	23.0	23.1	0.460	0.15
<i>e4peffl</i>	1.8	1.8	0.036	0.06
<i>ri5peffl₁</i>	1.0	1.0	0.020	0.05
<i>ri5peffl₂</i>	4.9	4.9	0.098	0.21
<i>akgeffl₁</i>	7.0	7.0	0.140	0.07
<i>akgeffl₂</i>	1.2	1.2	0.024	0.01
<i>oaapeffl</i>	11.6	11.6	0.232	0.17
Produktbildung:				
<i>lyseffl</i>	18.3	18.5	0.370	0.40
<i>co2effl</i>	275.1	239.0	13.800	-2.62

Tabelle 12.1: Extrazelluläre Flüsse, gemessen in einer kontinuierlichen Kultur von *C. glutamicum* MH20-22B. Alle Werte wurden auf eine Substrataufnahmerate von 100% entsprechend 1.49 mmol/(gTM · h) normiert. Die Flüsse *pyreffl₂*, *e4peffl* und *lyseffl₁* sind an eine CO₂-Bildung gekoppelt, während *ri5peffl₁* und *ri5peffl₂* CO₂ reflexieren. Die Metabolit-Abkürzungen sind aus Tabelle 3.1 zu entnehmen. Alle Meßfehler wurden mit 2% angenommen, bis auf den CO₂-Meßwert mit 5%. Rechts zum Vergleich das Resultat der Parameterschätzung und die Abweichung des Meßwerts vom Schätzer relativ zum angenommenen Meßfehler.

der Pentosephosphatweg-Netto-Zufluß ppw_1^{net} ,
 der Netto-Fluß des zweiten Lysin-Produktionsschritts as_{1b}^{net} ,
 der Glyoxylatzyklus-Netto-Fluß goz_1^{net} .

Austauschflüsse werden angenommen für die meisten Reaktionen der Glykolyse und alle Reaktionen des Pentosephosphatwegs mit Ausnahme des CO₂-Abspaltungsschritts PPW1 des oxidativen Pentosephosphatwegs. Ebenso ist der zusammengefaßte anaplerotische Schritt bidirektional, während im Citratzyklus nur der nicht an eine CO₂-Abspaltung gekoppelte Schritt einen Austausch hat. Zusammenfassend sind die freien Austauschflüsse gegeben durch:

die Glykolyse-Austauschflüsse $gly_2^{xch[0,1]}$, $gly_5^{xch[0,1]}$,
 die Pentosephosphatweg-Austauschflüsse $ppw_4^{xch[0,1]}$, $ppw_5^{xch[0,1]}$, $ppw_6^{xch[0,1]}$,
 den anaplerotischen Austauschfluß $ar_1^{xch[0,1]}$,
 den Citratzyklus-Austauschfluß $cz_7^{xch[0,1]}$.

Das Experiment wurde unter für die Lysin-Produktion günstigen Bedingungen bei einer Durchflußrate von 0.1 h⁻¹ durchgeführt. Das Substrat war [1-¹³C]-markierte Glucose. Unter diesen Bedingungen betrug die Substrataufnahmerate 1.49 mmol/(g · h) (bezogen auf die Biotrockenmasse). Dieser Wert wird nachfolgend auf 100% gesetzt, um die Interpretation der Daten zu erleichtern. Die einzelnen Biomasseabflüsse wurden auf der Grundlage der Biomassezusammensetzung nach [NIS90] berechnet (vgl. dazu die Ausführungen in Abschnitt 2.1). Die Flußmessungen werden als unabhängig mit den in Tabelle 12.1 angegebenen Standardabweichungen an-

genommen (vgl. dazu auch die Ausführungen in Abschnitt 6.2). Die statistischen Daten in dieser Tabelle werden in einem nachfolgenden Abschnitt diskutiert.

C-Atom	Meßwert [%]	Geschätzter Wert [%]	Angenommener Meßfehler [%]	Relative Abweichung
<i>ri5p</i> ₁	12.6	12.5	0.2	-0.14
<i>ri5p</i> ₂	2.9	2.8	0.5	-0.28
<i>ri5p</i> ₃	2.1	2.3	2.0	0.11
<i>ri5p</i> ₄	1.5	2.1	0.5	1.28
<i>ri5p</i> ₅	19.3	18.4	1.3	-0.67
<i>e4p</i> ₁	2.0	3.0	1.0	1.01
<i>e4p</i> ₂	3.6	2.1	1.0	-1.49
<i>e4p</i> ₃	2.0	1.9	1.0	-0.08
<i>e4p</i> ₄	16.7	15.5	2.0	-0.59
<i>gap</i> ₁	2.9	2.9	0.2	-0.11
<i>gap</i> ₂	2.6	2.5	0.1	-0.60
<i>gap</i> ₃	26.7	26.4	0.2	-1.65
<i>pyr</i> ₂	3.0	2.7	1.0	-0.33
<i>pyr</i> ₃	26.4	26.3	0.5	-0.20
<i>akg</i> ₂	24.1	22.5	0.3	-5.10
<i>akg</i> ₃	11.1	9.7	0.5	-2.78
<i>akg</i> ₄	28.1	26.3	0.6	-2.95
<i>oaa</i> ₂	7.6	9.7	2.0	1.05
<i>oaa</i> ₃	20.9	22.5	2.0	0.81
<i>oaa</i> ₄	16.8	17.3	2.7	0.17
<i>lys</i> ₂	6.8	7.1	0.2	1.36
<i>lys</i> ₃	21.9	23.9	0.3	6.78
<i>lys</i> ₄	18.9	17.3	1.0	-1.63
<i>lys</i> ₅	22.2	24.9	1.0	2.69
<i>lys</i> ₆	5.6	5.3	0.3	-1.00
<i>co2</i> ₁	23.0	21.6	0.4	-4.16

Tabelle 12.2: Aus Proteinhydrolysat bzw. RNA/DNA bestimmte Markierungsanteile in einzelnen C-Atom-Positionen von *C. glutamicum*. Der angenommene Meßfehler hängt von der Qualität der erzeugten NMR-Spektren ab. Markiertes CO₂ wurde massenspektroskopisch gemessen. Rechts zum Vergleich das Resultat der Parameterschätzung und die Abweichung des Meßwerts vom Schätzer relativ zum angenommenen Meßfehler.

12.1.3 Markierungsdaten

Die Biomasse wurde nach drei Verweilzeiten (also 30 h) abgeerntet. Für die Auswaschkorrektur gemäß Satz 2.1 muß die Zeit τ_1 bis zum Erreichen des isotopisch stationären Zustands in den intrazellulären Intermediaten abgeschätzt werden. Dabei ist zu bedenken, daß in *C. glutamicum* ein extrem großer Glutamat-Pool mit Konzentrationen von über 100 mmol/l vorliegt [Sch95]. Dieser verzögert aufgrund seiner Pufferwirkung das Erreichen des isotopisch stationären Zustands. Aufgrund der in Abbildung 12.6 dargestellten instationären Meßdaten wird (pessimistisch) die Zeitdauer bis zum Erreichen des isotopisch stationären Zustands mit $\tau_1 = 4\text{h}$ an-

```

{
  ...
  NET_FLUXES :=
  {
    ...

    MEASUREMENTS :=
    {
      Glu6P_Efflux:= 1.3 +- 0.09;
      Fru6P_Efflux:= 0.5 +- 0.04;
      GA3P_Efflux := 0.9 +- 0.02;
      ...
    };
  };
  ...

  POOLS :=
  {
    Ru5P :=
    {
      ...

      MEASUREMENT :=
      {
        LABEL :=
        {
          #1 := 12.6 +- 0.2;
          #2 := 2.9 +- 0.5;
          #3 := 2.1 +- 2.0;
          #4 := 1.5 +- 0.5;
          #5 := 19.3 +- 1.3;
        }
      } ;
    };
    ...
  };
};

```

Abbildung 12.2: Erweiterung der in Abbildung 9.2 dargestellten Simulatoreingabe für die Bereitstellung von Meßdaten und Standardabweichungen zur Parameterschätzung. Dargestellt sind exemplarisch die neu hinzukommenden Datenfelder in der hierarchischen Datenstruktur.

genommen. Der entsprechende Korrekturfaktor liegt dann gemäß Satz 2.1 zwischen 1.05 und 1.07. Die Korrektur führt also einen *relativen* Meßfehler von maximal 2% ein.

Die aus dem Proteinhydrolysat und der RNA/DNA-Fraktion gewonnenen Metabolite erlauben die Erzeugung qualitativ hochwertiger NMR-Spektren mit einer (durch den NMR-Experten geschätzten) Standardabweichung unterhalb eines Prozentes des Markierungsanteils. Die in Tabelle 12.2 wiedergegebenen Werte stellen die noch nicht mit dem Faktor 1.05 korrigierten Rohdaten dar. Diese Korrektur kann auf viel einfachere Weise berücksichtigt werden, indem die zugefütterte markierte Glucose als nicht 100-prozentig markiert angenommen wird.¹ Stattdessen wurde eine 95%-Markierung eingesetzt.

Offenbar liegen alle Markierungsanteile unterhalb von 30%, einige weit darunter. Der durch die Auswaschkorrektur eingeführte zusätzliche Fehler beträgt also maximal $0.02 \cdot 30\% = 0.6\%$. Er kann also nur in wenigen Fällen in der Größenordnung der Genauigkeit der NMR-Messung liegen. Die statistischen Daten der Parameteranpassung werden in einem nachfolgenden Abschnitt diskutiert.

Dem Mikrobiologen fallen auf den ersten Blick die zum Teil großen Markierungsanreicherungen im Pentosephosphatweg auf. Dies ist deshalb ungewöhnlich, weil das erste C-Atom der Glucose — also das markierte C-Atom — beim Übergang von Glucose-6-phosphat in den Pentosephosphatweg als CO_2 abgespalten wird. Würden alle Reaktionsschritte unidirektional ablaufen, so könnte daher überhaupt keine Markierung in den Pentosephosphatweg gelangen. Dies beweist bereits ohne weitere Analyse der Daten die Existenz bidirektionaler Reaktionsschritte.

12.1.4 Parameterschätzung

Bereits in Kapitel 9.2 wurde ein Textformat vorgestellt, daß die Eingabe eines metabolischen Netzwerks und der entsprechenden Simulationsparameter gestattet. Dieses Format ermöglicht

¹Es sei darauf hingewiesen, daß auch die übrigen C-Atom-Positionen stets mit der natürlichen Häufigkeit des ^{13}C -Isotops von 1.13% angereichert sind, was entsprechend berücksichtigt wurde. Vollkommen unmarkierte C-Atom-Pools kommen also real nicht vor.

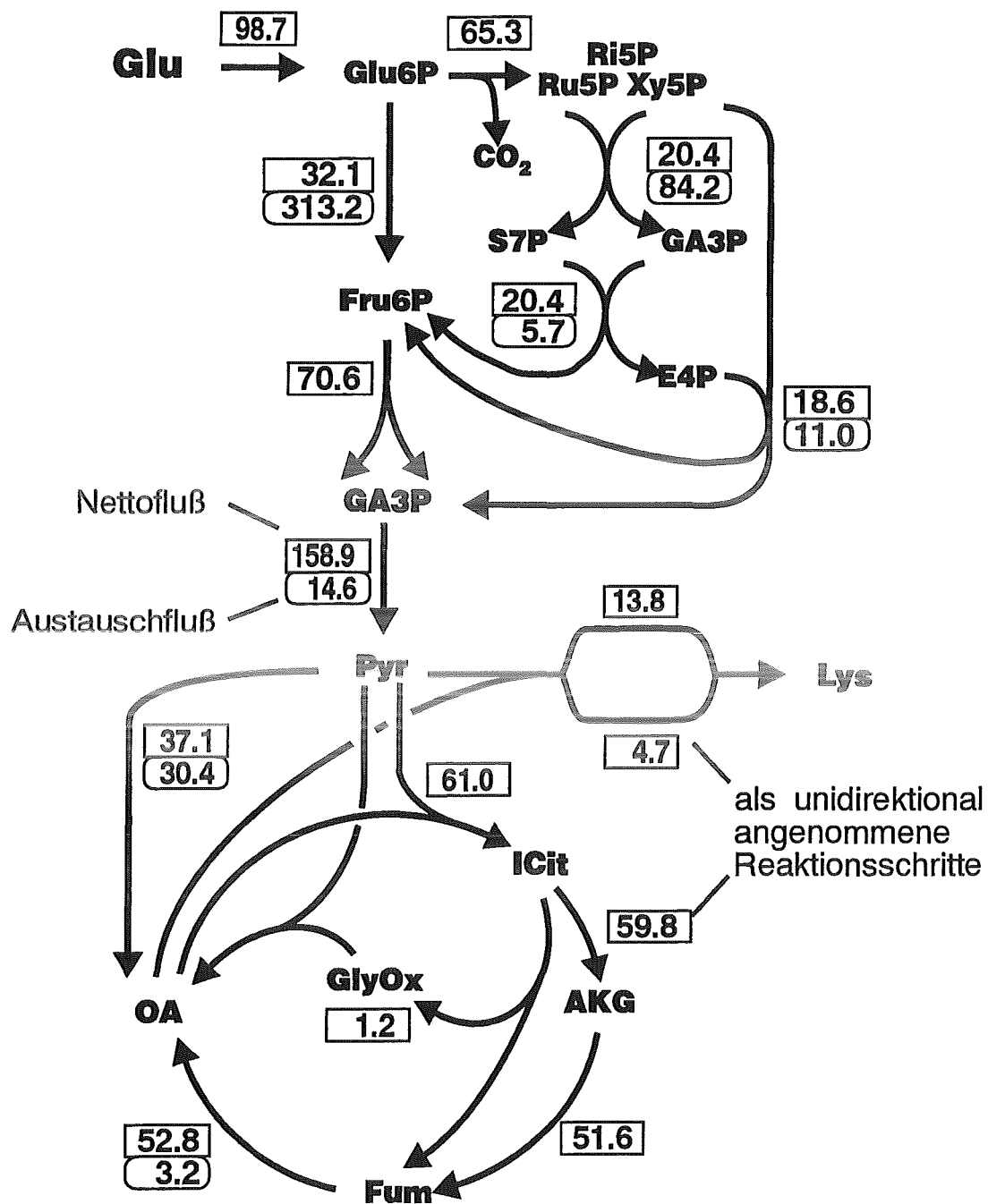


Abbildung 12.3: Geschätzte intrazelluläre Stoffflüsse in *C. glutamicum*. Wurden Austauschflüsse angenommen, so sind die geschätzten Werte jeweils im unteren Rahmen angegeben.

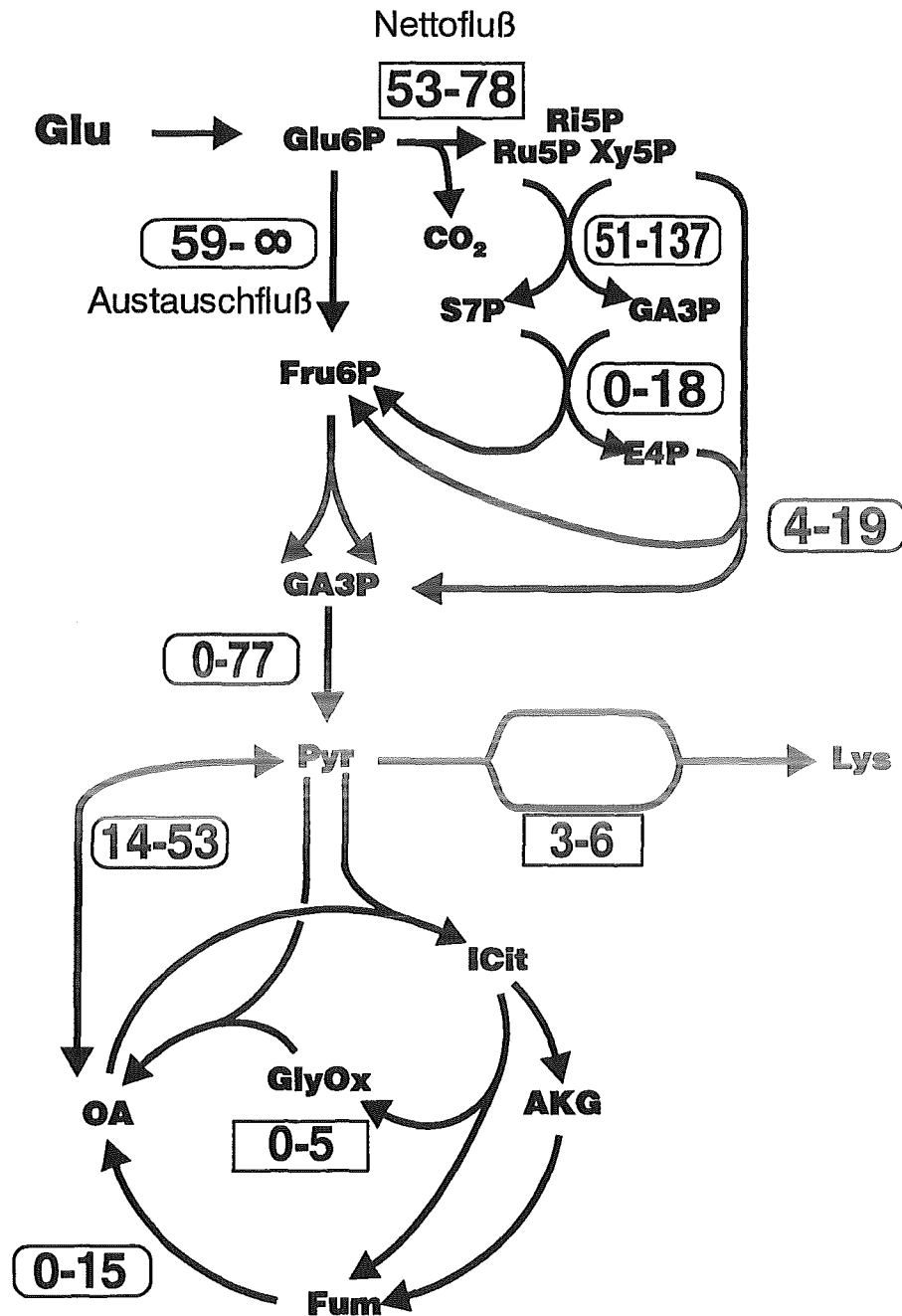


Abbildung 12.4: Geschätzte Standard-Vertrauensbereiche für die geschätzten freien Flüsse in *C. glutamicum*. Bei den Austauschflüssen treten unsymmetrische Intervalle auf (vgl. Abb. 12.3).

auch die Eingabe der gemessenen Werte als Grundlage für die Parameterschätzung. Abbildung 12.2 zeigt exemplarisch wie die Meßdaten in das hierarchische Eingabeformat eingeordnet werden.

Auf der Grundlage der so bereitgestellten Meßdaten findet der in Abschnitt 10.3 beschriebene Optimierungsalgorithmus ein Optimum der Fehlerquadratsumme nach etwa 10-20 Schritten [Sie96]. Wiederholtes Neustarten mit verschiedenen Ausgangswerten liefert reproduzierbare Ergebnisse, d.h. es kann von der Auffindung eines globalen Optimums ausgegangen werden. Ferner wurde ausgehend vom gefundenen Optimum eine Einzelvariation jedes Parameters durchgeführt, die jeweils ein deutlich ausgeprägtes, wenn auch oft breites, Optimum zeigte [Sie96].

Die aufgrund des Datensatzes geschätzten Nettoflüsse und Austauschflüsse zeigen die Tabellen 12.1 (für die direkt meßbaren Flüsse) und 12.3 (für die übrigen freien Flüsse) sowie die Abbildung 12.3. Hervorzuheben ist die gute Reproduktion aller Markierungsanteile in Glykolyse und Pentosephosphatweg, die weder mit dem Modell unidirektionaler Flüsse im Pentosephosphatweg noch mit einem schnellen Markierungsgleichgewicht bzgl. der Reaktionen PPW4, PPW5 und PPW6 erklärt werden kann (vgl. dazu die Ausführungen in Abschnitt 9.2, insbesondere Abbildung 9.5).

12.2 Statistische Analyse

Zur statistischen Analyse des Anpassungsergebnisses werden nun die in Kapitel 11 beschriebenen statistischen Methoden (insbesondere der nichtlineare Reskalierungsansatz aus Abbildung 11.6) angewendet. Das Konfidenzniveau für die nachfolgenden Analysen ist stets das für Regressionsanalysen mit mittleren Anforderungen an die Signifikanz der Ergebnisse üblicherweise gewählte 90%-Niveau (vgl. auch die Ausführungen in der Einleitung des Abschnitts 11.1).

Fluß	Geschätzter Wert [%]	Verhältnis zum Nettofluß	Geschätztes Standardintervall
ppw_1^{net}	65.3		[53, 78]
goz_1^{net}	1.2		[0, 5]
as_{1b}^{net}	4.7		[3, 6]
gly_2^{xch}	313.2	9.9	[59, ∞]
gly_5^{xch}	14.6	0.1	[0, 77]
ppw_4^{xch}	84.2	4.4	[51, 137]
ppw_5^{xch}	5.7	0.3	[0, 18]
ppw_6^{xch}	11.0	0.6	[4, 19]
ar_1^{xch}	30.4	0.8	[14, 53]
cz_7^{xch}	3.2	0.1	[0, 15]

Tabelle 12.3: Geschätzte nicht direkt meßbare freie Netto- und Austauschflüsse (relativ zur Substrataufnahme von 100%) und deren aus dem $(v^{net}, v^{xch[0,1]})$ -Raum rücktransformierte Standardabweichungen. Zur Einschätzung der Größenordnung ist zudem das Verhältnis zum jeweiligen Nettofluß angegeben.

12.2.1 Vertrauensbereiche für Netto- und Austauschflüsse

Man kann zeigen, daß bei einem direkt meßbaren Nettofluß die geschätzte Standardabweichung des geschätzten Flusses nicht größer sein kann als die vorgegebene Standardabweichung des

Meßfehlers. Dies ist unmittelbar einsichtig, denn durch weitere Meßwerte kann die Information über diesen direkt meßbaren Fluß höchstens verbessert werden. Aus diesem Grunde sind in Tabelle 12.1 die geschätzten Standardabweichungen nicht mit angegeben, denn sie decken sich bei fast allen Flüssen genau mit den vorgegebenen Meßfehler-Standardabweichungen oder sind sogar geringfügig kleiner. Die Analyse der Parameter-Korrelationsmatrix und der Hauptachsen der Kovarianzmatrix (hier nicht in vollem Umfang reproduziert) zeigt ferner, daß die direkt gemessenen Flüsse völlig unabhängig von den übrigen freien Flüssen geschätzt werden.

Tabelle 12.3 und Abbildung 12.3 stellt nun die auf der Basis der Linearisierung des Modells berechneten Vertrauensintervalle für die nicht direkt meßbaren Netto- und Austauschflüsse zusammen. Bei diesen im wesentlichen durch die Markierungsmessungen bestimmten Flüssen zeigt sich eine deutlich größere Unbestimmtheit wie schon aufgrund der Beispielstudien in Abschnitt 11.3 zu erwarten war (vgl. Abb. 11.5).

Zur Bestimmung eines 90%-Konfidenzintervalls für einzelne Flüsse sind die tabellierten Werte gemäß Satz 11.1 mit $\sqrt{\chi_1^2(0.9)} = 1.64$ zu multiplizieren. Soll jedoch ein simultanes 90%-Konfidenzgebiet für alle Flüsse *zugleich* bestimmt werden, so ist stattdessen der Faktor $\sqrt{\chi_{16}^2(0.9)} = 4.8$ zu wählen. Dies zeigt, daß die Anwendung strenger statistischer Gütemaße zu extrem großen Vertrauensbereichen führt, die auf der Basis von Simulationsstudien nicht realistisch erscheinen. Da dies vor allem die schlecht bestimmten Flüsse betrifft, erscheint eine separate Analyse dieser Flüsse sinnvoller zu sein, denn bei sehr großen Konfidenzintervallen kann nicht mehr davon ausgegangen werden, daß der Linearisierungsansatz adäquat ist (vgl. Abschnitt 11.1).

Die geschätzten intrazellulären Austauschflüsse werden nach der in Abschnitt 11.3 beschriebenen Reskalierungsmethode durch Linearisierung im $(v^{net}, v^{xch[0,1]})$ -Raum und anschließende Rücktransformation in den (v^{net}, v^{xch}) -Raum bestimmt (vgl. auch Abb. 11.6). Dabei entstehen unsymmetrische Vertrauensbereiche, die in Tabelle 12.3 und Abbildung 12.4 dargestellt sind. Bei der Interpretation der Werte hilft das Verhältnis zum jeweiligen Nettofluß.

Auch hier sind nur die Standardintervalle wiedergegeben, die für eine 90%-Konfidenz noch mit etwa dem Faktor zwei vergrößert werden müssen. Trotzdem zeigt sich, daß die Austauschflüsse innerhalb weiter Grenzen bestimmt sind, die zumindest eine Aussage über ihre Größenordnung relativ zum entsprechenden Nettofluß zulassen. Die Auswertung zeigt, daß der Fluß ppw_4^{xch} einen sehr hohen Austausch hat, während die Austauschflüsse ppw_5^{xch} , ppw_6^{xch} und ar_1^{xch} den Nettoflüssen vergleichbar sind und cz_7^{xch} nahezu verschwindet. Hervorzuheben ist, daß damit erstmals eine quantitative Aussage über Austauschflüsse im Pentosephosphatweg gemacht werden konnte.

Über die Glykolyseflüsse gly_2^{xch} und gly_5^{xch} kann aufgrund der großen Vertrauensbereiche nur eine ungefähre quantitative Aussage gemacht werden. Diese Vertrauensbereiche sind so groß, daß die Brauchbarkeit der Linearisierung zweifelhaft ist. Diese Flüsse sollten daher als unbestimmt angesehen werden. Zur untersuchen ist nun, in wie weit sich diese Unbestimmtheit auf die übrigen Flußschätzungen auswirkt.

12.2.2 Korrelationsanalyse

Mit dem Ziel der Analyse von Abhängigkeiten zwischen den geschätzten Parametern wird nun die Parameter-Korrelationsmatrix aus Satz 11.1 herangezogen. Diese (23×23) -Matrix kann hier aus Platzgründen nicht vollständig reproduziert werden. Stattdessen beschränkt sich die Analyse auf schlecht bestimmte Netto- und Austauschflüsse. Alle übrigen (zur besseren Übersicht hier nicht aufgeführten) Korrelationen liegen unterhalb von 0.2, sind also unbedeutend. Die derart reduzierte Korrelationsmatrix zeigt Tabelle 12.4. Folgende Feststellungen können getroffen

	ppw_1^{net}	$gly_2^{xch[0,1]}$	$ppw_4^{xch[0,1]}$	$ppw_5^{xch[0,1]}$	$ppw_6^{xch[0,1]}$	goz_1^{net}	$ar_1^{xch[0,1]}$	$gly_5^{xch[0,1]}$	cz_7
ppw_1^{net}	1.00	0.98	.	-0.30	0.50	.	.	.	.
$gly_2^{xch[0,1]}$	0.98	1.00	.	-0.26	0.44	.	.	.	.
$ppw_4^{xch[0,1]}$	.	.	1.00	-0.36	0.26	.	.	.	.
$ppw_5^{xch[0,1]}$	-0.30	-0.26	-0.36	1.00	-0.97	.	.	.	-0.20
$ppw_6^{xch[0,1]}$	0.50	0.44	0.26	-0.97	1.00	.	.	.	.
goz_1^{net}	.	.	.	.	.	1.00	0.43	-0.48	-0.39
$ar_1^{xch[0,1]}$	.	.	.	.	.	0.43	1.00	-0.95	0.39
$gly_5^{xch[0,1]}$	.	.	.	.	.	-0.48	-0.95	1.00	-0.38
cz_7	.	.	.	-0.20	.	-0.39	0.39	-0.38	1.00

Tabelle 12.4: Korrelationsmatrix für freie intrazelluläre Netto- und Austauschflüsse. Wiedergegeben sind aus Gründen der Übersichtlichkeit nur Korrelationen größer als 0.2.

werden:

- Es sind deutlich drei Gruppen von Flüssen unterscheidbar, derart daß innerhalb der Gruppen deutlich von Null verschiedene Korrelationen auftreten, während die Gruppen untereinander nahezu unkorreliert sind. Jede Gruppe besteht aus Flüssen, die im Stoffwechselnetzwerk eng benachbarten sind:
 - Einstieg in den Pentosephosphatweg $ppw_1^{net}, gly_2^{xch[0,1]}$
 - Austauschflüsse im Pentosephosphatweg $ppw_4^{xch[0,1]}, ppw_5^{xch[0,1]}, ppw_6^{xch[0,1]}$
 - Reaktionen nahe dem oder im Citratzyklus $goz_1^{net}, ar_1^{xch[0,1]}, gly_5^{xch[0,1]}, cz_7$
- Mit drei Ausnahmen treten keine Korrelationen auf, die auf eine mögliche lineare Abhängigkeit zwischen einzelnen Flußschätzungen hinweisen. Ausnahmen bilden die Korrelationen zwischen ppw_1^{net} und $gly_2^{xch[0,1]}$ (0.98), zwischen $ppw_5^{xch[0,1]}$ und $ppw_6^{xch[0,1]}$ (-0.97) sowie zwischen $ar_1^{xch[0,1]}$ und $gly_5^{xch[0,1]}$ (-0.95). Jedoch ist keine dieser Korrelationen exakt gleich 1.00 oder -1.00, so daß grundsätzlich keine modellbedingten Nichtidentifizierbarkeiten von Flüssen vorliegen.
- Eine hohe Korrelation sagt zunächst nichts über die Bestimmtheit von Parametern aus, denn erst in Verbindung mit der Größe des jeweiligen Vertrauensbereichs wird die Größenskala für den Konfidenzellipse festgelegt. Dieser kann zwar stark entartet sein, d.h. einen linearen Unterraum approximieren, jedoch spielt dies nur bei großen Ellipsoiden eine Rolle. Kritisch sind in dieser Hinsicht also nur die unbestimmten Austauschflüsse $gly_2^{xch[0,1]}$ und $gly_5^{xch[0,1]}$.

12.2.3 Hauptachsenanalyse

Wie in Abschnitt 11.2 erläutert, ist die Korrelationsmatrix für eine tiefergehende Analyse der Bestimmtheit von Parametern wenig geeignet, da sie gegenseitige Abhängigkeiten von mehr als zwei Parametern nicht zuverlässig aufdeckt. Insbesondere betrifft dies die wichtige Frage nach den mit den unbestimmten Austauschflüssen $gly_2^{xch[0,1]}$ und $gly_5^{xch[0,1]}$ stark korrelierten Flüssen ppw_1^{net} und $ar_1^{xch[0,1]}$. Denn die Unbestimmtheit der ersteren wirkt sich über die Korrelation auch auf die beiden letztgenannten Flüsse aus. Dies kann nun durch eine Hauptachsenanalyse genauer quantifiziert werden. Die längsten auftretenden Hauptachsen des Standard-Konfidenzellipse zeigt die Tabelle 12.5. Folgendes ist feststellbar:

Fluß	Achsenlänge							
	0.256	0.206	0.085	0.058	0.032	0.019	0.016	0.001
$gly_2^{xch[0,1]}$	0.94	.	.	.	.	.	.	.
ppw_1^{net}	0.30	.	.	.	-0.27	.	-0.44	.
$ppw_4^{xch[0,1]}$	.	.	0.79	-0.61	.	.	.	.
$ppw_5^{xch[0,1]}$	.	.	-0.53	-0.66	.	.	.	.
$ppw_6^{xch[0,1]}$	.	.	0.28	0.40	.	.	.	.
goz_1^{net}	.	.	.	.	-0.24	.	.	-0.31
$ar_1^{xch[0,1]}$	.	0.34	.	.	.	-0.87	0.27	.
$gly_5^{xch[0,1]}$	.	-0.93	.	.	.	-0.30	.	.
$cz_7^{xch[0,1]}$	.	.	.	.	0.91	.	-0.23	.
$inff^{net}$	.	.	.	.	.	-0.32	-0.75	.
as_{1b}^{net}	.	.	.	.	.	.	.	0.92

Tabelle 12.5: Die längsten Hauptachsen des Parameter-Konfidenzellipsoids skaliert auf die einfache Standardabweichung. Zur besseren Übersicht sind jeweils nur die längsten Hauptachsen wiedergegeben. Zudem sind die Gruppe aus Tabelle 12.4 hervorgehoben.

1. In der Tat zeigt sich die normierte längste Hauptachse des Konfidenzellipsoids bereits zu 99% durch die beiden Einträge zum Austauschfluß $gly_2^{xch[0,1]}$ (nämlich 0.94) und zum Nettofluß ppw_1^{net} (nämlich 0.30) bestimmt. Die Unbestimmtheit von $gly_2^{xch[0,1]}$ überträgt sich also auch auf ppw_1^{net} , allerdings stark abgeschwächt. Dies erklärt die im Vergleich zu den anderen freien Nettoflüssen weniger gute Bestimmtheit von ppw_1^{net} . Trotzdem ist dieser Fluß in akzeptablen Grenzen festgelegt.
2. Ein völlig analoger Effekt ergibt sich bei den Flüssen $gly_5^{xch[0,1]}$ und $ar_1^{xch[0,1]}$. Hier ist die zweitlängste Achse dafür verantwortlich, daß sich die Unbestimmtheit des ersten Flusses abgeschwächt auf den zweiten überträgt. Wiederum ist dies ausreichend für eine akzeptable Bestimmung von $ar_1^{xch[0,1]}$.
3. Alle weiteren Hauptachsen sind bereits um mindestens den Faktor vier kürzer als die beiden längsten Achsen, so daß keine großen Unbestimmtheiten mehr durch sie hervorgerufen werden können. Deutlich zu erkennen ist, daß die dritte und vierte Achse einen linseförmigen Konfidenzbereich in dem durch $ppw_4^{xch[0,1]}$, $ppw_5^{xch[0,1]}$ und $ppw_6^{xch[0,1]}$ erzeugten Raum aufspannen. Analoges gilt für die dritte Gruppe von korrelierten Flüssen.

Der Nettofluß ppw_1^{net} und der Austauschfluß $ar_1^{xch[0,1]}$ sind also durchaus aus den Daten schätzbar. Jedoch ist die hierfür zur Verfügung stehende Information äußerst gering, woraus die großen Konfidenzintervalle resultieren. Aufgrund von Simulationsstudien, die das globale nichtlineare Verhalten des Systems untersuchen sollten, zeigte sich ferner, daß der gemessene Markierungszustand nur mit einem hohen Austauschfluß $gly_2^{xch[0,1]}$ erklärt werden kann. Da $gly_2^{xch[0,1]}$ im geschätzten Zustand des Systems zudem nur einen äußerst geringen Einfluß auf den Markierungszustand des Systems hat, ist zumindest der Nettofluß ppw_1^{net} mit akzeptabler Güte schätzbar.

Meßwert	Flußschätzer						
	$gly_2^{xch[0,1]}$	ppw_1^{net}	$gly_5^{xch[0,1]}$	$ar_1^{xch[0,1]}$	$ppw_4^{xch[0,1]}$	$ppw_5^{xch[0,1]}$	$ppw_6^{xch[0,1]}$
$ri5p_2$	0.32	0.29	.	.	.	.	.
$ri5p_5$	.	.	.	.	0.42	0.11	.
$e4p_4$	.	.	.	.	.	0.21	0.23
gap_1	0.19	0.20	.	.	.	0.22	0.25
pyr_2	.	.	0.10	0.10	.	.	.
akg_3	.	.	0.12	0.14	.	.	.
lys_6	.	.	0.10	.	.	.	.
Gesamt-Sensit.	0.59	0.19	0.71	0.27	0.16	0.17	0.10

Tabelle 12.6: Parameter-Sensitivitäten für die Flüsse mit den größten geschätzten Vertrauensbereichen. Alle Sensitivitäten sind gemäß Gleichung (11.2) durch die Standardabweichungen der Meßwerte skaliert und sodann zum besseren Vergleich auf die Gesamtsumme 1.0 pro Fluß normiert. Die absolute Gesamtsensitivität der Spalte ist darunter angegeben. Ferner wurden zur besseren Übersicht nur Werte größer als 0.1 aufgeführt.

12.2.4 Sensitivitätsanalyse

Ein Ziel zukünftiger Experimente muß sein, die korrelierten Schätzwerte $gly_2^{xch[0,1]}$ und ppw_1^{net} sowie $gly_5^{xch[0,1]}$ und $ar_1^{xch[0,1]}$ durch Bereitstellung weiterer Meßinformation zu entkoppeln und damit die Vertrauensbereiche zu verkleinern. An dieser Stelle sei zunächst analysiert, welche Meßwerte die Flüsse mit den größten Vertrauensbereichen bestimmen. Hierzu wird die Parameter-Sensitivitätsmatrix herangezogen, die in Tabelle 12.6 wiedergegeben ist.

Folgendes kann festgestellt werden:

1. Die stark korrelierten Flüsse sind im wesentlichen durch die gleichen intrazellulären Markierungsanteile bestimmt. Konkret gilt:

$gly_2^{xch[0,1]}, ppw_1^{net}$	sind stark bestimmt durch	$ri5p_2, gap_1$
$gly_5^{xch[0,1]}, ar_1^{xch[0,1]}$	sind stark bestimmt durch	pyr_2, akg_3, lys_6
$ppw_4^{xch[0,1]}, ppw_5^{xch[0,1]}, ppw_6^{xch[0,1]}$	sind stark bestimmt durch	$ri5p_5, e4p_4, gap_1$

2. Bei $gly_2^{xch[0,1]}$ und ppw_1^{net} , $gly_5^{xch[0,1]}$ und $ar_1^{xch[0,1]}$ sowie $ppw_5^{xch[0,1]}$ und $ppw_6^{xch[0,1]}$ sind die jeweiligen skalierten Sensitivitäten annähernd gleich. Dies erklärt die entsprechenden hohen Korrelation in Tabelle 12.4, denn die Meßwerte enthalten jeweils annähernd die gleiche Information über die korrelierten Parameter.

Das Datenmaterial enthält also nur sehr wenig Information zur Unterscheidung der stark korrelierten Flüsse. Ein Vorschlag für ein Experiment zur Verbesserung der Schätzgüte wird im übernächsten Abschnitt diskutiert.

12.2.5 Anpassungsgüte und Ausreißertest

Der χ^2 -Test für die Beschreibbarkeit der Daten durch das Modell kommt bei einem 90%-Konfidenzniveau zu einem negativen Ergebnis. Die Quadratsumme ist 137, während der Test lediglich

$\chi^2_{17}(0.9) = 25$ zuläßt (vgl. Tabelle 12.2). Hierzu können verschiedene Erklärungsansätze vorgeschlagen werden:

1. Mit Ausnahme der CO_2 -Messung stimmen die Schätzwerte für die direkt gemessenen Netto-Flüsse fast genau mit den Meßwerten überein. Große Abweichungen bei der CO_2 -Messung jedoch sind nichts ungewöhnliches, da die CO_2 -Messung empfindlich vom Kohlen-säuregleichgewicht im Medium abhängt (vgl. auch [ZS95]).
2. Dagegen weisen die extrem großen Ausreißer in einzelnen C-Atom-Positionen von Lysin und α -Ketoglutarat möglicherweise auf systematische Fehler hin, zumal die übrigen Markierungsanteile der C-Atome in diesen Stoffen besser reproduziert werden. Ein Blick auf Tabelle 12.2 zeigt, daß der weitaus größte Beitrag zur Quadratsumme durch die Meßwerte von lys_3 (mit 46.2), akg_2 (mit 26.0) und co_2_1 (mit 17.6) gegeben ist. Durch Streichen dieser Werte würde die Quadratsumme bereits auf 47.2 reduziert.
3. Ganz allgemein ist die Anpassungsgüte des Modells im Bereich des Pentosephosphatweges und der Glykolyse weitaus besser als die im Citratzyklus. Bei Beschränkung auf die ersten beiden Wege wäre der χ^2 -Test (mit einem entsprechend geringeren Freiheitsgrad) erfolgreich. Dies wirft die Frage nach systematischen Fehlern im Citratzyklus-Bereich auf. Eine erste Hypothese wäre die, daß die Auswaschkorrektur (Satz 2.1) falsch gewählt wurde, denn die Markierung im Zyklus wird während der isotopisch instationären Phase von der Glutamat-Markierung dominiert (vgl. auch Abschnitt 12.3 und Abb. 12.6). Insbesondere ist der α -Ketoglutarat-Pool unmittelbar mit dem Glutamat-Pool verbunden. Dagegen ist die Rückwirkung des Glutamat-Pools auf den "entfernten" Zuckerstoffwechsel gering. Dies legt eine für verschiedene Stoffwechselwege unterschiedlich gewählte Auswaschkorrektur nahe.
4. Die Genauigkeit für die Lysin-Messung ist möglicherweise zu eng vorgegeben. Hier kommt es zu einer starken Peak-Überlagerung der Lys#3-, Lys#4- und Lys#5-Peaks, die die Auswertung des Spektrums erschwert [WMWG95, Möl95].
5. Eine generelle Ursache für die zu große Quadratsumme kann in einer Überschätzung der Güte der Meßdaten liegen. Bereits eine Korrektur der Meßfehler-Standardabweichungen in Tabelle 12.1 und 12.2 um den Faktor 1.5 würde die Quadratsumme um den Faktor $1.5^2 = 2.25$ verringern. Werden zusätzlich die Ausreißer gestrichen, so ist der χ^2 -Test bereits passiert. Dann sind jedoch auch die Standardabweichungen aller geschätzten Parameter um den Faktor 1.5 zu vergrößern.

An dieser Stelle wird klar, warum exakte statistische Verfahren bei den oben diskutierten Problemen nicht weiterhelfen können. Einzelne der Hypothesen könnten zwar eventuell mit dem vorhandenen Datenmaterial statistisch getestet werden, nicht aber alle zugleich, was die Resultate für die Praxis fragwürdig machen würde. Die oben diskutierten möglichen Fehlerquellen sollten vielmehr als Anregungen für eine nachfolgende experimentelle Überprüfung verstanden werden.

12.2.6 Versuchsplanung

Im Anschluß an die oben durchgeführte Sensitivitätsanalyse wird deutlich, daß eine verbesserte Information über der Pentosephosphatweg-Fluß ppw_1^{net} wünschenswert wäre. Die Bestimmung weiterer Markierungsdaten ist hier nicht mehr denkbar, da bereits alle Vorläufermetabolite für

Komponenten der Biomasse ausgeschöpft wurden. Rückwirkend betrachtet, hat jedoch die Zunahme der Markierungen des Ru5P-Pools im Vergleich zu [MGWE⁺96] bereits die statistische Qualität des Ergebnisses entscheidend verbessert.

Zur experimentellen Gewinnung weiterer Information kann jedoch ein anderer Weg beschritten werden. Hierfür ist die im PPW1-Schritt erfolgende CO₂-Abspaltung nutzbar. Offenbar reicht dafür jedoch die ¹³C-Information nicht aus. Es kann aber versucht werden, mit Hilfe weiterer Tracer eine Diskriminierung zu erreichen (vgl. dazu auch [RKBY94]).

Hierfür bietet sich beispielsweise eine ¹⁴C-Markierung an, um eine Unterscheidung des im Pentosephosphatweg freigesetzten CO₂ vom übrigen CO₂ zu erreichen. Da eine Radioaktivitätsmessung um Größenordnungen sensitiver ist als eine NMR-Messung, genügt die Zugabe von ¹⁴C-Glucose in äußerst geringen Mengen. Die Markierung sollte allerdings auf einem der C-Atome 4-6 liegen, denn dieser Teil des Substrats passiert den Pentosephosphatweg auch im Falle von Austauschreaktionen völlig unverändert (vgl. Abb. 9.5). Das freigesetzte ¹⁴CO₂ muß dann aus einem der anderen Stoffwechselwege stammen. Damit kann der PPW1-Schritt möglicherweise von den übrigen Schritten diskriminiert werden.

Eine weitere Idee besteht darin, einen neuen Meßwert in der Glykolyse einzubringen. Dies gelingt z.B., wenn Fructose-6-Phosphat als Vorläufer-Metabolit der Zellwandsynthese vermessen wird. Dazu ist eine neue analytische Methode zur präparativen Gewinnung von Fructose-6-Phosphat zu entwickeln, was derzeit bearbeitet wird.

12.3 Ein instationäres Markierungsexperiment

Isotopisch instationäre Markierungsexperimente werden in dieser Arbeit nur am Rande berücksichtigt. Die Diskussion eines derartigen Experiments erfolgt daher an dieser Stelle nur unter dem Aspekt einer groben Abschätzung der so erzielbaren Informationsmenge im Vergleich zum oben durchgeführten stationären Ansatz.

12.3.1 Experimentelle Bedingungen

Die in Abbildung 12.6 dargestellten Daten wurden in dem in Jülich entwickelten Hydrozyklonreaktor mit Biomasse-Verweilzeitentkopplung gemessen [Har95]. Dabei wird die Biomasse mit Hilfe eines Filtrationsmoduls im Reaktor zurückgehalten. Lediglich über einen kleinen "Bleed-Strom" kann sie den Reaktor verlassen (vgl. dazu Abb. 12.5). Der Hydrozyklonreaktor erlaubt den Betrieb einer Hochzelllichtfermentation innerhalb des Magnetkerns eines NMR-Gerätes (vgl. auch Abb. 2.8, die den Vorläufer des wesentlich aufwendigeren Hydrozyklonreaktors zeigt), also eine wirkliche In-vivo-Beobachtung des Systems. So können stabile physiologische Bedingungen bei gleichzeitiger Verstärkung des NMR-Signals erreicht werden.

Das Experiment wurde mit dem bereits oben beschriebenen Stamm *C. glutamicum* MH20-22B durchgeführt. Allerdings ist die Wachstumsrate der Biomasse jetzt um den Faktor 10 kleiner. Als Substrat wurde diesmal [6-¹³C]-Glucose verwendet. Da bei einer verweilzeitentkoppelten Kultur der erforderliche Durchfluß des Substrats um ein vielfaches höher ist als der Durchfluß der Biomasse, müssen Substratmengen eingesetzt werden, die einem zehnfach größeren Reaktor entsprechen. Die damit verbundenen Kosten führten bei dem diskutierten Experiment zu einer Beendigung der Zufuhr markierter Glucose noch vor dem Erreichen eines isotopisch stationären Zustands (vgl. Abb. 12.6).

Die in Abbildung 12.6 gezeigten zeitlichen Verläufe geben alle unter diesen Bedingungen im NMR-Gerät detektierbaren Markierungen wieder. Es zeigt sich, daß selbst bei hohen Zelldichten

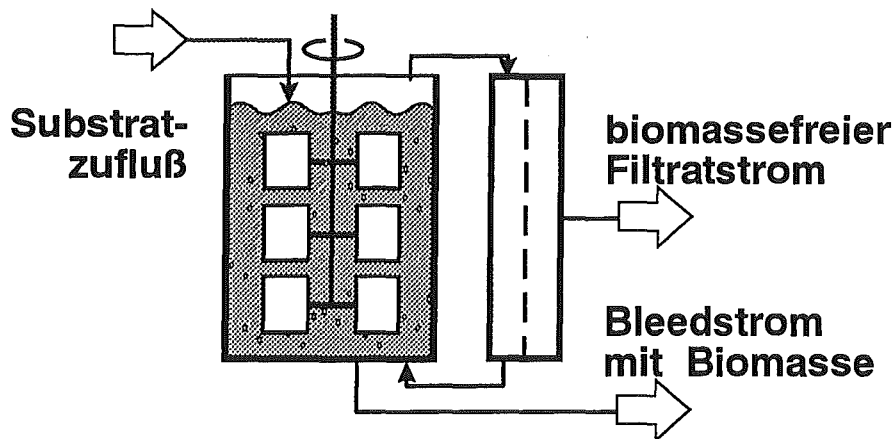


Abbildung 12.5: Schemazeichnung eines Bioreaktors mit Verweilzeitentkopplung. Die Biomasse wird mittels eines Filtrationsmoduls im Reaktor zurückgehalten. Ihre Verweilzeit wird durch den Bleed-Strom bestimmt, während Substrate und Produkte als Filtrat abgetrennt werden.

ausschließlich die stark konzentrierten Metabolite einer Beobachtung zugänglich sind. In diesem Fall sind dies Lactat, Glutamat, Succinat und Lysin, wobei lediglich der Glutamat-Pool ein rein intrazellulärer Pool ist. Alle übrigen Pools setzen sich aus intrazellulären Konzentrationen und aus im Medium akkumulierten Stoffen zusammen, da die Versuchsanordnung die Messung der Konzentrationen dieser Stoffe nur für den gesamten Reaktor erlaubt.

12.3.2 Abschätzung der Zeitkonstanten

Es soll nun eine grobe Abschätzung des Verhältnisses zwischen dem Substratzufluß in die Zellen und den intrazellulären Poolgrößen in Glykolyse und Pentosephosphatweg durchgeführt werden. Die während des Experiments erreichte Biomassekonzentration betrug 25 g/l. Die mit einer Konzentration von 125 g/l zugeführte Glucose wurde nahezu vollständig verbraucht. Dabei betrug die Verweilzeit der Biomasse 100 h, die des Substrats 20 h. Folglich setzten 25 g Biomasse innerhalb einer Stunde etwa 6 g Glucose (entsprechend 30 mmol) um.

Da mit dem NMR-Gerät erst ab einer gewissen Poolgröße ein auswertbares Signal erzeugt werden kann, ist davon auszugehen, daß alle intrazellulären Pools (außer den oben genannten) in einer Konzentration von weniger als 40 mmol/l vorliegen. Wahrscheinlich sind aber wesentlich kleinere Werte. Die nachfolgenden Überlegungen gelten zunächst für den Zuckerstoffwechsel, für den Poolgrößen von 1 mmol/l angenommen werden. Das intrazelluläre Volumen der Zellen beträgt etwa 0.002 l/g [Sch95], d.h. nur etwa 5% des Reaktorvolumens sind mit Biomasse gefüllt. Die intrazellulären Poolgrößen bezogen auf den Gesamtreaktor betragen also schätzungsweise 0.05 mmol/l. Diese Pools werden nun pro Stunde von den oben errechneten 30 mmol Substrat durchflossen.

Das Verhältnis $30/0.05 = 600$ entspricht recht genau dem Verhältnis bei der in Kapitel 9.2 vorgestellten instationären Simulation (vgl. Abb. 1.4). Demgemäß wäre zu erwarten, daß der intrazelluläre isotopisch stationäre Zustand (zumindest im Zuckerstoffwechsel) bereits nach einigen Minuten erreicht ist. Dagegen zeigen die gemessenen Daten des instationären Experiments, daß der isotopisch stationäre Zustand des Gesamtsystems frühestens nach 2 Stunden zu erwarten ist. Dies steht im Widerspruch zu der soeben ermittelten Zeitkonstante.

Zur Erklärung dieser Diskrepanz muß bedacht werden, daß das Medium im Vergleich zum intrazellulären Raum der Zellen eine 20-fach höhere Kapazität hat. Die Zeitkonstante der Mar-

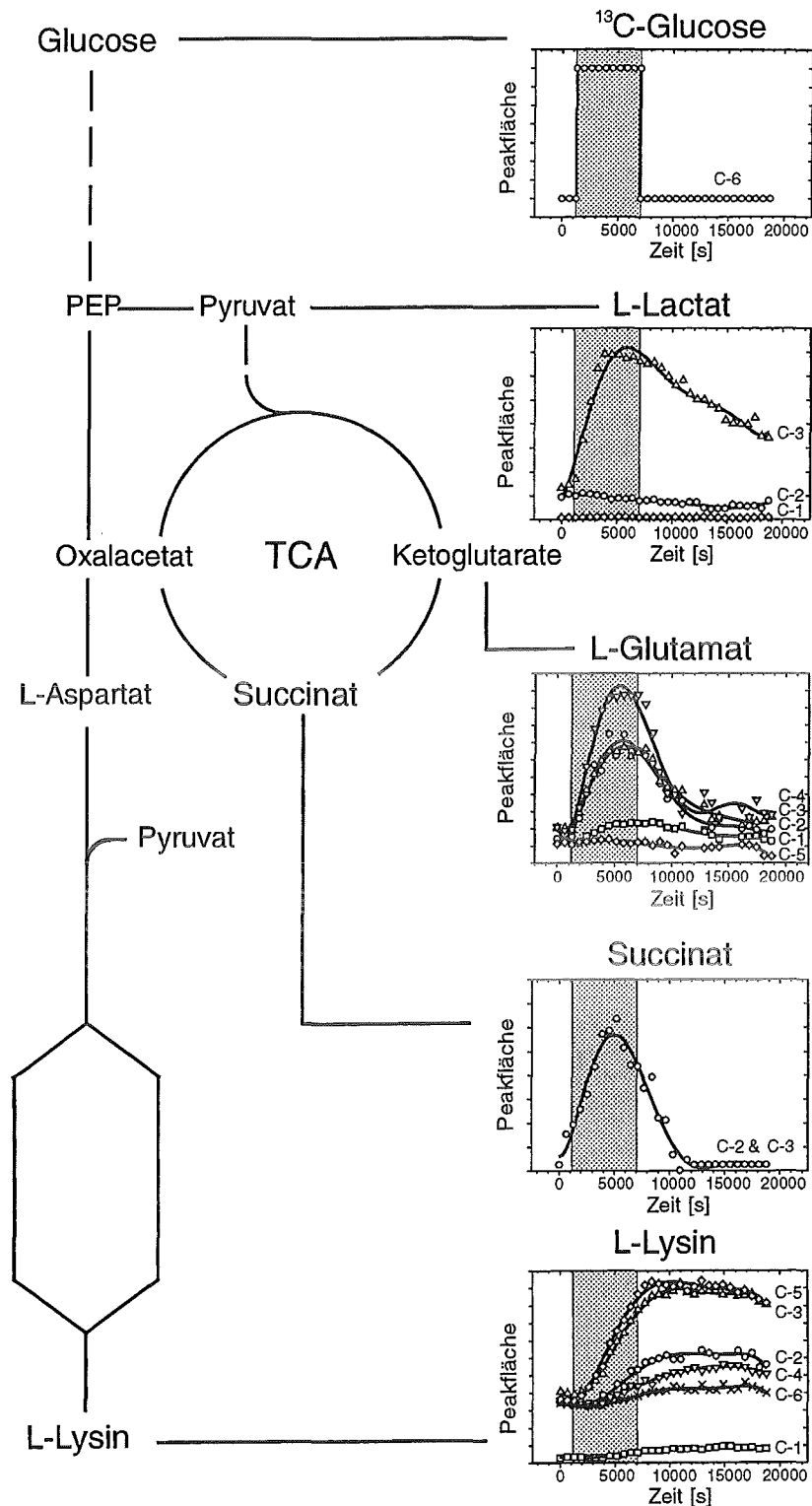


Abbildung 12.6: Zeitverläufe des Markierungsaufbaus bei einem instationären In-vivo-Markierungsexperiment, durchgeführt innerhalb eines NMR-Reaktors mit Verweilzeitentkopplung (aus [Har95]). Während des gekennzeichneten Zeitraums wurde $[6-^{13}\text{C}]$ -Glucose zugeführt.

kierungsanreicherung im Medium ist also etwa um diesen Faktor größer als die entsprechende Zeitkonstante für das intrazelluläre System. Das in Abbildung 12.6 dargestellte Verhalten des Systems beruht also in erster Linie auf der Akkumulation des markierten Materials im Medium. Diese Erklärung ist mit den gemessenen Daten gut vereinbar. Eine inzwischen durchgeführte Simulationsstudie mit angenommenen intrazellulären Flüssen konnte dieses Ergebnis genau bestätigen.

Die obigen Überlegungen treffen auf den Citratzyklus nur bedingt zu, denn der große intrazelluläre Glutamat-Pool (160 mmol/l intrazellulär, entsprechend 8 mmol/l bezogen auf den Reaktorinhalt) verzögert das Erreichen des intrazellulären isotopisch stationären Zustands. Das Glutamat-Signal ist also im wesentlichen die einzige auswertbare Information über das intrazelluläre instationäre Markierungsverhalten. Aus den übrigen Signalen läßt sich dagegen im Grunde nur der isotopisch stationäre intrazelluläre Zustand rekonstruieren.

12.4 Vergleich der experimentellen Ansätze

Das instationäre In-vivo-Experiment im Magnetkern des NMR-Geräts zeigt deutlich die Problematik solcher Markierungsexperimente bei schnell wachsenden Mikroorganismen. Eine Beobachtung des instationären Zustands in einer größeren Anzahl intrazellulärer Pools im NMR-Gerät bei einer ^{13}C -Markierung ist nach dem derzeitigen Stand der Technik wohl nur mit aufkonzentrierten Zellen möglich, wobei gemäß den Ausführungen in Kapitel 2.3 ein wohldefinierter physiologischer Zustand nicht gewährleistet werden kann.²

Die auf den isotopisch stationären Zustand angewendete Entkopplungsmethode (Definition 2.6) liefert dagegen eine ungleich höhere Informationsmenge, wobei zusätzlich auf die Bestimmung (oder Annahme) der intrazellulären Poolgrößen verzichtet werden kann. Insbesondere ist mit diesem Ansatz erstmals auch der Markierungszustand schwach konzentrierter intrazellulärer Pools indirekt einer Messung zugänglich geworden, wodurch — wie oben gezeigt — auch die bisher nicht mögliche Quantifizierung von Austauschflüssen im Pentosephosphatweg gelingt. Dieser wurde bei den meisten bisher durchgeführten Markierungsstudien zu einem einzigen unidirektionalen Schritt zusammengefaßt [PSMC93, STM⁺94].

²Dies stellt nicht die Anwendung eines In-vivo-NMR-Reaktors für metabolisch instationäre ^{31}P -Experimente in Frage [GSPW⁺91, WG96a].

Kapitel 13

Schlußdiskussion und Ausblicke

Diese Arbeit schließt mit einer kurzen Diskussion der erzielten Ergebnisse in Hinblick auf die allgemeine Problematik der Modellbildung in den biologischen Disziplinen. Dabei wird darauf verzichtet, eine detaillierte Zusammenfassung der Einzelergebnisse bzgl. der Modellierung, Analyse, Simulation und Auswertung von Markierungsexperimenten zu bieten — diese werden in den einzelnen Kapiteln ohnehin ausführlich dokumentiert. Vielmehr soll gefragt werden, welchen Beitrag diese Arbeit zur Theoretischen Biologie leistet, d.h. welche Fortschritte die hier entwickelten theoretischen Konzepte für die praktische biologische Anwendung erbringen können.

13.1 Biologische Resultate

Schon die exemplarische Anwendung der entwickelten Methoden in Kapitel 12 hat gezeigt, daß die neu entwickelte experimentelle Methode in Kombination mit den vorgestellten Analysewerkzeugen zu wichtigen stoffwechselphysiologischen Erkenntnissen führen kann, die derzeit mit anderen Methoden nicht erzielbar sind. Eine umfassende und detaillierte Analyse der biologischen Resultate des beschriebenen stationären Experiments ist nicht die Aufgabe dieser Arbeit. Hierzu sei verwiesen auf [Mar94, MGWE⁺96]. An dieser Stelle sollen nur die wichtigsten Ergebnisse skizziert werden.

Aufgrund der Ausführungen in Abschnitt 12.2 kann davon ausgegangen werden, daß die geschätzten Werte jeweils im Rahmen der ermittelten Genauigkeit ein korrektes Bild der intrazellulären Verhältnisse geben. Selbst unter der Voraussetzung, daß die berechneten Werte zum Teil große Ungenauigkeiten aufweisen, sind die nachfolgenden Schlüsse zulässig:

1. Die Bestimmung der Nettoflüsse im Zentralstoffwechsel von *C. glutamicum* gelang anhand von Markierungs- und extrazellulären Flußdaten auch ohne Rückgriff auf Energiebilanzen. Zusätzlich konnte die Existenz von Austauschflüssen gezeigt und deren Größenordnung geschätzt werden. Insbesondere beim Pentosephosphatweg zeigte sich, daß hinsichtlich der Austauschflüsse weder von Unidirektionalität noch von einem schnellen Gleichgewicht (vgl. dazu auch [FCA⁺93]) gesprochen werden kann. Dies folgt im übrigen bereits aus der Tatsache, daß nach Abbildung 9.5 in diesem Fall die Markierungen in Ri5P#1 und Ri5P#2 übereinstimmen müßten, was offenbar nicht zutrifft.
2. Die sich aus der Flußbestimmung ergebende NADPH-Bilanz zeigt, daß dieses Reduktionsäquivalent im Überschuß produziert wird [MGWE⁺96]. Es folgt, daß ein bisher unbekannter Verbraucher existieren muß. Im Gegensatz dazu wurde für die Flußbestimmung in [Val91] eine ausgewogene NADPH-Bilanz dem Modell zugrundegelegt.

3. Wie bereits in anderen Arbeiten gezeigt [HBS81, CMC⁺94], deutet der Austauschfluß im anaplerotischen Abschnitt auf einen Futile Cycle hin. Die Bedeutung solcher energieverbrauchenden Schritte ist bis heute nicht hinreichend geklärt.

13.2 Austauschflüsse

Besonderes Gewicht wurde in dieser Arbeit auf die modellmäßige Beschreibung von Austauschflüssen und die Analyse ihrer Auswirkungen auf das metabolische Markierungssystem gelegt. Das ausgewertete Experiment zeigt deutlich die Relevanz dieses Konzeptes auf. Obwohl ein Austauschfluß ohne weitere Untersuchungen auf enzymkinetischer und thermodynamischer Ebene nicht genauer verstanden werden kann, wird bereits jetzt ein verändertes Verständnis der zellulären Regulation nahegelegt.

Offenbar kann man sich den Metabolismus nicht als ein hochgradig gerichtetes System von Stoffflüssen vorstellen. Vielmehr sind viele Reaktionsschritte in beide Richtungen durchlässig, so daß eine Stoffverteilung über benachbarte Pools mit vergleichbarem Energieniveau stattfindet. Diese Vorstellung ist aus thermodynamischer Sicht nicht neu, jedoch gehen Modellansätze zur Beschreibung der zellulären Regulation eher von einer "hydrodynamischen" Vorstellung eines Flusses aus.

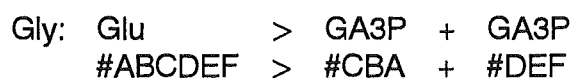
Inwieweit sich aus einer veränderten Sichtweise der zellulären Regulation auch alternative Modellansätze entwickeln lassen, kann hier nicht mehr diskutiert werden (vgl. auch [WD87]). Weitere Modellierungsstudien sollen in der Folge untersuchen, welche Zusammenhänge zwischen Austauschflüssen und enzymkinetischen Mechanismen bestehen.

13.3 Problematik der Modellbildung

Ein zentraler Gegenstand dieser Arbeit war die Entwicklung eines allgemeinen Modellierungsansatzes, der insbesondere die Auswirkungen von Austauschflüssen in vollem Umfang berücksichtigt. Besonderes Gewicht wurde dabei auf eine sorgfältige Dokumentation der Modellannahmen gelegt. Hat sich das daraus für die Stoffflußanalyse bei *C. glutamicum* resultierende Modell auch weitgehend bei der Analyse der Experimente bewährt, so traten doch auch viele Fragen auf, die aus einer Differenz zwischen Modell und Meßwerten resultieren. Anhand eines Vergleichs mit einer einfacheren Methode zur Flußbestimmung wird im folgenden deutlich, daß es geradezu ein Anliegen einer Modellierung sein muß, solche Unstimmigkeiten hervorzurufen.

Die in der Literatur stark dominierende Methode zur Flußbestimmung aus Markierungsdaten beruht auf der Lösung von Markierungs-Bilanzgleichungen für einzelne C-Atom-Pools [WHK⁺82, WK84, Szy95]. Dabei werden – methodenbedingt – starke Vereinfachungen des metabolischen Netzwerks vorgenommen. Insbesondere können nicht in großem Umfang Austauschflüsse berücksichtigt werden, es sei denn, es werden schnelle Gleichgewichte angenommen.

Als Beispiel sei eine häufig verwendete Vorgehensweise zur Bestimmung des Pentosephosphatweg-Flusses ppw_1^{net} anhand der GA3P#3-Markierung herangezogen. Dabei wird die gesamte Glykolyse zu einem unidirektionalen Reaktionsschritt



zusammengefaßt und ebenso der Pentosephosphatweg zu einem einzigen unidirektionalen Schritt [PSMC93, STM⁺94] (der einen Glykolyse-schritt hinzunimmt):



Unter Zugrundelegung dieses Modells bleibt der Pentosephosphatweg bei einer [$1\text{-}^{13}\text{C}$]-Markierung völlig unmarkiert und die Markierung aus Glu#1 fließt zu 100% nach GA3P#3. Allerdings kommt mit jedem markierten GA3P-Molekül auch ein unmarkiertes hinzu, da die Reaktion Gly4 den Metaboliten Fru16BP in zwei GA3P-Moleküle aufspaltet, von denen nur das eine eine Markierung trägt. Die Bilanz für den GA3P#3-Pool ist folglich

$$1 \cdot gly^{\text{net}} = ga3p_3 \cdot (2 gly^{\text{net}} + 5 ppw^{\text{net}}) \implies \frac{ppw^{\text{net}}}{gly^{\text{net}}} = \frac{1 - 2 ga3p_3}{5 ga3p_3}, \quad (13.1)$$

so daß das Flußverhältnis unmittelbar aus der GA3P#3-Markierung bestimmt werden kann. GA3P#3 wird darüber hinaus häufig erst über den Umweg des Glutamat-Pools unter weiteren Unidirektionalitätsannahmen gemessen.

Bezüglich des vorliegenden Datensatzes ist zu beachten, daß der obige Fluß ppw^{net} aufgrund der trimolekularen Reaktionsgleichung mit dem gedrittelten Fluß von ppw_1^{net} im vorangehenden Modell verglichen werden muß. Ferner muß berücksichtigt werden, daß aufgrund der Auswaschungskorrektur nur eine 95-prozentig markierte Glucose zugrunde gelegt wird (vgl. die Ausführungen in Abschnitt 12.1). Dann folgt aus den oben präsentierten Meßdaten:

$$\frac{ppw_1^{\text{net}}}{gly^{\text{net}}} = \frac{95 - 2 \cdot 26.7}{5/3 \cdot 26.7} = 0.93$$

Der mit dem in dieser Arbeit entwickelten, komplexen Modell ermittelte Wert ist jedoch 2.08. Damit wird deutlich, in welchem Umfang die Modellannahmen bei identischem Datensatz das Ergebnis der Flußanalyse beeinflussen können.

Erstaunlich ist, daß bzgl. des Ergebnisses im Falle der Rechnung mit Hilfe von Gleichung 13.1 keinerlei statistische Probleme auftreten, denn bei Messung nur einer Größe können Unstimmigkeiten zwischen Meßdaten nicht auftreten. Das Resultat hat daher — trotz der in großem Umfang vorgenommenen Modellvereinfachungen — den Anschein der Zuverlässigkeit, wogegen der Einsatz komplexer Modelle — wie oben gezeigt — viele Fragen aufwirft.

Im einem wissenschaftstheoretischen Kontext kann dies nur positiv bewertet werden, denn das komplexe Modell ermöglicht die kritische Diskussion der berechneten Ergebnisse einerseits und der Meß- und Modellfehler andererseits. Zudem sind die erzielten Resultate aufgrund redundanter Information in einem gewissen Ausmaß überprüfbar. Dies entspricht einer auch die Möglichkeit einer Modellfalsifikation berücksichtigenden Vorgehensweise [Pop84].

13.4 Einsatz mathematischer Methoden

Die in dieser Arbeit bereitgestellten mathematischen Werkzeuge zur Analyse, Simulation und Auswertung von Markierungsexperimenten können vom biologischen Anwender in aller Regel nicht mehr in allen Einzelheiten nachvollzogen werden. Jedoch sollte dies den Einsatz fortgeschrittener mathematischer Methoden nicht ausschließen. Vielmehr hat diese Arbeit gezeigt, daß die Entwicklung zuverlässiger numerischer Verfahren die Voraussetzung für die umfassende, methodenbewußte und statistisch zuverlässige (soweit für biologische Systeme sinnvoll) Auswertung der Experimente darstellt.

Das Jülicher Stoffflußanalyseprojekt legt dabei besonderen Wert auf eine adäquate Behandlung von Austauschflüssen. Dies wurde erst durch die in dieser Arbeit vorgestellten numerischen Methoden, die auch große Austauschflüsse numerisch stabil behandeln können, ermöglicht. Insbesondere zeigt das Anwendungsbeispiel, daß dies bei Markierungsexperimenten eine Notwendigkeit darstellt. Die hier dargestellten numerischen Methoden sind erstmalig in der Lage, die C-Atom-Bilanzgleichungen einschließlich der verschiedenen Nebenbedingungen allgemein zu behandeln.

Die zukünftige Herausforderung liegt in der numerischen Behandlung von Isotopomeren-Systemen zur Simulation und Auswertung stationärer Experimente mit Multipllett-Messungen. Die ersten Grundlagen hierfür wurden bereits in Kapitel 8 gelegt. Wie dort gezeigt wurde, besteht der entscheidende Unterschied zwischen den C-Atom- und Isotopomeren-Bilanzgleichungssystemen darin, daß letztere nichtlinear bzgl. der Markierungsvariablen sind. Daher lassen sich die Isotopomeren-Bilanzgleichungen nicht mehr mit Methoden der linearen Algebra lösen. Vielmehr sind iterative Verfahren zur Lösung nichtlinearer Gleichungssysteme erforderlich. Das in Abschnitt 8.1 beschriebene einfache Euler-Verfahren ist hierfür nur bedingt einsetzbar. Insbesondere im Falle großer Austauschflüsse treten die analogen Probleme wie beim linearen C-Atom-Bilanzgleichungssystem auf (vgl. Abschnitt 8.2), d.h. das entsprechende nichtlineare Differentialgleichungssystem (4.34) wird dann steif. Folglich benötigt jedes an einem Differentialgleichungslöser angelehnte Verfahren unvertretbar lange Rechenzeiten bis zum Erreichen des stationären Punktes.

Es ist abzusehen, daß sich die in Abschnitt 8.2 eingeführte Vorkonditionierung des linearen Gleichungssystems auf Isotopomeren-Systeme verallgemeinern läßt. Schließlich beruht die Vorkonditionierung im Grunde auf einer Linearkombination von Bilanzgleichungen (vgl. das Beispiel in Abschnitt 8.2). Die Vorkonditionierungsoperation, also die Multiplikation der Systemmatrizen mit einer Transformationsmatrix von links, läßt sich daher unmittelbar auch auf die Isotopomeren-Transitionsmatrizen übertragen. Es ist zu erwarten, daß die derart vorkonditionierten Isotopomeren-Bilanzgleichungssysteme problemlos numerisch lösbar sind. Zur Berechnung der Ableitungen nach den Flüssen gemäß Satz 8.2 muß dann ein sehr großes lineares Gleichungssystem gelöst werden, das ebenfalls analog zu Satz 8.3 vorkonditionierbar ist.

Schwieriger ist die Erweiterung des in Kapitel 10 beschriebenen Optimierungsalgorithmus zur Parameteranpassung. Zunächst kann das Optimierungsproblem aus Definition 10.2 vollkommen analog auf Isotopomeren-Systeme übertragen werden, indem darin Gleichung (4.29) durch Gleichung (4.30) ersetzt wird. Die Vereinfachung des Problems gemäß Satz 10.1 ist aber nicht mehr ohne weiteres möglich, da eine explizite Lösung der Isotopomeren-Bilanzen, etwa in Form der Abbildung Γ aus Satz 7.3, nicht möglich ist. Trotzdem kann Γ und dessen Ableitungen in jedem Schritt des Optimierungsalgorithmus wie oben beschrieben iterativ berechnet werden, was allerdings zu hohen Rechenzeiten führen wird.

Daher sollte ein alternativer Ansatz entwickelt werden, der das Problem aus Definition 10.2 von vorn herein als ein nichtlineares Optimierungsproblem unter nichtlinearen Nebenbedingungen auffaßt. Zu seiner Lösung können erneut die Konzepte der Sequentiellen Quadratischen Programmierung (vgl. Abschnitt 10.2) herangezogen werden. Zur Behandlung der nichtlinearen Gleichheits-Nebenbedingungen, also der Bilanzgleichungen, muß die oben beschriebene Vorkonditionierung geeignet in das Verfahren integriert werden. Andernfalls ist das Optimierungsproblem ebenso schlecht konditioniert wie das in ihm enthaltene nichtlineare Gleichungssystem. Von großem Interesse wird sein, ob Isotopomeren-Markierungsexperimente genau wie C-Atom-Markierungsexperimente mit einer vertretbaren Rechenzeit auswertbar sind, so daß ein interaktives Arbeiten ermöglicht wird.

13.5 Versuchsplanung

Wie in den Kapiteln 11 und 12 gezeigt, ermöglichen statistische Methoden eine detaillierte Analyse des Informationsgehaltes von Meßdaten in Bezug auf Bestimmung von intrazellulären Flüssen. Diese Methoden können auch zum Entwurf solcher Experimente eingesetzt werden, die eine optimale Parameterinformation liefern. Da das Experiment bei vorgegebenen Flüssen nur wenig beeinflusst werden kann, gibt es hierzu grundsätzlich nur die beiden nachfolgenden Möglichkeiten. In jedem Fall muß a priori eine Flußverteilung angenommen werden, in der Hoffnung, daß die später im Experiment auftretende Verteilung nicht allzu stark hiervon abweicht:

1. Die Markierung des Substrates kann vom Experimentator frei gewählt werden (vgl. Abb. 1.3). Dabei sind nicht nur Substrate denkbar, die zu nahezu 100% aus einem einzigen Isotopomer bestehen, sondern auch Mischungen aus mehreren solchen Substraten. Mit Methoden der nichtlinearen Versuchsplanung kann nun versucht werden, eine Zusammensetzung des Substrates so zu berechnen, daß ein kleiner Parameter-Konfidenzellipsoid entsteht. Schließlich kann auch eine Versuchsplanung für Experimente durchgeführt werden, bei denen zugleich ^{13}C - und ^{14}C -Isotope eingesetzt werden.
2. Unter dem Gesichtspunkt der Arbeitersparnis ist auch die Art und Anzahl der durchzuführenden Messungen von Bedeutung. Zwar erfordert die Vermessung der Aminosäuren im wesentlichen nur einen Arbeitsgang, die Gewinnung weiterer Markierungsinformation aus RNA/DNA oder Zellwandbestandteilen erfordert dagegen die Entwicklung spezifischer quantitativer Aufreinigungsmethoden. Hier kann aufgrund einer Sensitivitätsanalyse im voraus entschieden werden, ob sich die Aufwand für die Gewinnung weiterer Markierungsanteile lohnt.

Erneut ist die Frage von großem Interesse, ob sich generell durch die Hinzunahme von Isotopomeren-Meßdaten ein wesentlicher Informationszuwachs einstellt. Grundsätzlich muß dies nicht gelten, denn in dieser Arbeit wurde sowohl ein Beispiel für einen Informationszuwachs (Abb. 1.7) als auch eines für gleichbleibende Information (Abb. 3.1) vorgestellt. Realistisch komplexe Systeme sind allerdings kaum noch algebraisch wie in Kapitel 4 behandelbar, sondern müssen auf numerischem Wege mit Hilfe von Simulationsstudien untersucht werden.

Die in der Vergangenheit publizierten Beispiele haben hier wenig Aussagekraft, da sie meistens auf nur auf der Vermessung eines einzigen Metabolit-Pools beruhen. Mit der in Jülich entwickelten experimentellen Methode steht jedoch erstmals eine derartige Menge von Markierungsdaten zur Verfügung, daß erst die nächsten Jahren zeigen werden, wann Isotopomeren-Meßdaten zusätzliche Flüsse bestimmbar machen oder zumindest die statistische Qualität der Flußschätzung deutlich verbessern. Ist dies der Fall, so stellt sich das Problem der optimalen Versuchsplanung erneut für Isotopomerensysteme.

Literaturverzeichnis

- [AAC⁺91] R. Altenburger, S. Abarzua, R. Callies, L.H. Grimme, A. Mayer, and D. Leibfritz. Ammonia rhythm in *microcystis firma* studied by in vivo ¹⁵N and ³¹P NMR spectroscopy. *Arch. Microbiol.*, 156:471–476, 1991.
- [AC74] A.C. Atkinson and D.R. Cox. Planning experiments for discriminating between models. *J.R.Statist.Soc.*, 3:321–348, 1974.
- [AHRT93] K.M. Anstreicher, D. den Hertog, C. Roos, and T. Terlaky. A long-step barrier method for convex quadratic programming. *Algorithmica*, 10(5):365–382, 1993.
- [ALS77] G. Adam, P. Luger, and G. Stark. *Physikalische Chemie und Biophysik*. Springer, 1977.
- [And83] D.H. Anderson. *Compartmental Modelling and Tracer Kinetics*, volume 50 of *Lecture Notes in Biomathematics*. Springer, 1983.
- [Arn90] S.F. Arnold. *Mathematical Statistics*. Ellis Horwood, 1990.
- [ASU86] A.V. Aho, R. Sethi, and J.D. Ullman. *Compilers — Principles, Techniques and Tools*. Addison Wesley, 1986.
- [AW92] K.R. Albe and B.E. Wright. Systems analysis of the tricarboxylic acid cycle in *dictyostelium discoideum*: II. control analysis. *J. Biol. Chem.*, 267:3106–3114, 1992.
- [Bai91] J.E. Bailey. Towards a science of metabolic engineering. *Science*, 252:1668–1674, 1991.
- [Baz83] M. Bazin, editor. *Mathematics in Biotechnology*. Academic Press, 1983.
- [BC85] S. Barnett and R.G. Cameron. *Introduction to Mathematical Control Theory*. Oxford Applied Mathematics and Computing Science Series. Clarendon, 1985.
- [BD85] G.C. Barrett and J.S. Davies. *Nuclear Magnetic Resonance Spectra of Amino Acids and their Derivatives*, chapter 18. Chapman and Hall, 1985.
- [BD93] D. Birkes and Y. Dodge. *Alternative Methods of Regression*. Wiley, 1993.
- [Ber83] F. Bergter. *Wachstum von Mikroorganismen*. Verlag Chemie, second edition, 1983.
- [Bet87] J. Betten. *Tensorrechnung fur Ingenieure*, volume 64 of *Leitfaden der angewandten Mathematik und Mechanik*. Teubner, 1987.

- [BG68] R.L. Bishop and S.I. Goldberg. *Tensor Analysis on Manifolds*. Dover Publications, 1968.
- [Bjo96] A. Bjorck. *Numerical Methods for Least Squares Problems*. SIAM, 1996.
- [BKW80] D.A. Belsley, E. Kuh, and R.E. Welsch. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Probability and Mathematical Statistics. Wiley, 1980.
- [BL94] D. Bray and S. Lay. Rapid numerical integration algorithm for finding the equilibrium state of a system of coupled binding reactions. *CABIOS*, 10(5):471–476, 1994.
- [BM88a] T.D. Brock and M.T. Madigan. *Biology of Microorganisms*. Prentice Hall, 5th edition, 1988.
- [BM88b] T.D. Brock and M.T. Madigan. *Biology of Microorganisms*. Prentice Hall, 5th edition, 1988.
- [BO86] J.E. Bailey and D.F. Ollis. *Biochemical Engineering Fundamentals*. McGraw-Hill Chemical Engineering Series. McGraw Hill, second edition, 1986.
- [BS82] J.J. Blum and R.B. Stein. On the analysis of metabolic networks. In R.F. Goldberger, editor, *Biological Regulation and Development*, volume 3A, chapter 3, pages 99–124. Plenum Press, 1982.
- [BT95] P.T. Boggs and J.W. Tolle. Sequential quadratic programming. *Acta Numerica*, pages 1–51, 1995.
- [BW88] D.M. Bates and D.G. Watts. *Nonlinear Regression Analysis and its Applications*. Probability and Mathematical Statistics. Applied Probability and Statistics. Wiley, 1988.
- [BW93] T. Becker and V. Weispfennig. *Gröbner Bases — A computational Approach to Commutative Algebra*, volume 141 of *Graduate Texts in Mathematics*. Springer, 1993.
- [CB83] J.M. Crawford and J.J. Blum. Quantitative analysis of flux along the gluconeogenic, glycolytic and pentose phosphate pathways under reducing conditions in hepatocytes isolated from fed rats. *Biochem. J.*, 212:595–598, 1983.
- [CB93] R. Chen and J.E. Bailey. Observations of aerobic, growing *escherichia coli* metabolism using an on-line nuclear magnetic resonance spectroscopy system. *Biotech. Bioeng.*, 42:215–221, 1993.
- [CB94] M. Coccagn-Bousquet. Growth rate dependent modulation of carbon flux through central metabolism and the kinetic consequences for glucose-limited chemostat cultures of *corynebacterium glutamicum*. *Appl. Env. Microbiol.*, 1994.
- [CB95] M. Cohen and R.N. Bergman. Estimation of TCA cycle flux, aminotransferase flux, and anaplerosis in heart: validation with syntactic model. *Am. J. Physiol.*, 268:E397–E409, 1995.

- [CF80] J.-P. Chambost and D.G. Fraenkel. The use of 6-labeled glucose to assess futile cycling in *escherichia coli*. *J. Biol. Chem.*, 255:2867–2869, 1980.
- [CFG95] J.C. Chatham, J.R. Forder, J.D. Glickson, and E.M. Chance. Calculation of absolute metabolic fluxes and the elucidation of the pathways of glutamate labeling in perfused rat heart by ^{13}C NMR spectroscopy and nonlinear least squares analysis. *J. Biol. Chem.*, 270:7999–8008, 1995.
- [CGST93] A.R. Conn, N. Gould, A. Sartener, and P.L. Toint. Global convergence of a class of trust region algorithms for optimization using inexact projections on convex constraints. *SIAM J. Optim.*, 3(1):164–221, 1993.
- [CH88] S. Chatterjee and A.S. Hadi. *Sensitivity Analysis in Linear Regression*. Probability and Mathematical Statistics. Wiley, 1988.
- [Chr75] N. Christofides. *Graph Theory: An Algorithmic Approach*. Computer Science and Applied Mathematics. Academic Press, 1975.
- [CKS90] S. Cerdan, B. Künnecke, and J. Seelig. Cerebral metabolism of $[1,2-^{13}\text{C}_2]$ acetate as detected by *in vivo* and *in vitro* ^{13}C NMR. *J. Biol. Chem.*, 265:12916–12926, 1990.
- [CL94] Y.-P. Chao and J.C Liao. Metabolic responses to substrate futile cycling in *escherichia coli*. *J. Biol. Chem.*, 269:5122–5126, 1994.
- [CLO92] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms — An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics. Springer, 1992.
- [CMCM⁺94] M.-F. Chauvin, F. Megnin-Chanet, G. Martin, J.-M. Lhoste, and G. Baverel. The rabbit kidney tubule utilizes glucose for glutamine synthesis. *J. Biol. Chem.*, 269:26025–26033, 1994.
- [Coh83] S.M. Cohen. Simultaneous ^{13}C and ^{31}P NMR studies of perfused rat liver. *J. Biol. Chem.*, 258:14294–14308, 1983.
- [Cop65] W.A. Coppel. *Stability and asymptotic behaviour of differential equations*. D.C. Heath, 1965.
- [CRSK81] S.M. Cohen, R. Rognstad, R.G. Shulman, and J. Katz. A comparison of ^{13}C nuclear magnetic resonance and ^{14}C tracer studies of hepatic metabolism. *J. Biol. Chem.*, 256:3428–3432, 1981.
- [CS94] J.L. Connell and L. Shafer. *Object-Oriented Rapid Prototyping*. Prentice Hall, 1994.
- [CSKW83] E.M. Chance, S.H. Seeholzer, K. Kobayashi, and J.R. Williamson. Mathematical analysis of isotope labeling in the citric acid cycle with applications to ^{13}C NMR studies in perfused rat hearts. *J. Biol. Chem.*, 258:13785–13794, 1983.
- [DAM88] Dori, Arato, and Maga. Determination of unresolved peaks via a new non linear regression procedure. *Annali di chimica*, 78:529–542, 1988.

- [Del87] H. Dellweg. *Biotechnologie - Grundlagen und Verfahren*. Verlag Chemie Weinheim, 1987.
- [Deu95] P. Deuffhard. *Newton Methods for Highly Nonlinear Problems: Theory, Algorithms, Code*. Academic Press, 1995.
- [DF95] J. Dolezal and J. Fidler, editors. *System Modelling and Optimization*. Chapman and Hall, 1995.
- [DHB94] P. Deuffhard, A. Hohmann, and F. Bornemann. *Numerische Mathematik I+II*. de Gruyter, 1993 + 1994.
- [DLC⁺84] M.M. Domach, S.K. Leung, R.E. Cahn, C.G. Cocks, and M.L. Shuler. Computer model for glucose-limited growth of a single cell of *eschericia coli* b/r-a. *Biotechn.Bioeng.*, 26:203–216, 1984.
- [Dom93] H. Dominguez. Modified carbon flux during oxygen limited growth of *corynebacterium glutamicum* and the consequences for amino acid overproduction. *Biotech. Letters*, 15:449–454, 1993.
- [Don79] J.J. Dongarra. *LINPACK, User's guide*. SIAM, Philadelphia, 1979.
- [DRM⁺93] L. Di Donato, C. Des Rosiers, J.A. Montgomery, F. David, M. Garneau, and H. Brunengraber. Rates of gluconeogenesis and citric acid cycle in perfused livers, assessed from the mass spectrometric assay of the ¹³C labeling pattern of glutamate. *J. Biol. Chem.*, 268(6):4170–4180, 1993.
- [EG96] L. Eggeling and A.A. de Graaf. Quantifying and directing metabolite flux: Application to amino acid overproduction. *Adv. Biochem. Eng. Biotechnol.*, 1996.
- [EGL89] H.D. Ehrich, M. Gogolla, and U. Lipeck. *Algebraische Spezifikation abstrakter Datentypen*. Teubner, 1989.
- [Eik93] B.J. Eikmanns. Molecular aspects of lysine threonine, and isoleucine biosynthesis in *corynebacterium glutamicum*. *A. v. Leeuwenhoek*, 64:145–163, 1993.
- [ES90] M.A. Ellis and B. Stroustrup. *The Annotated C++ Reference Manual*. Addison Wesley, 1990.
- [ESS83] I. Ekiel, I.C.P. Smith, and G.D. Sprott. Biosynthetic pathways in *methanospirillum hungatei* as determined by ¹³C nuclear magnetic resonance. *J. Bacteriol.*, 156:316–326, 1983.
- [ESS⁺93] C. Evans, B. Sumegi, P.A. Srere, A.D. Sherry, and C.R. Malloy. [¹³C]propionate oxidation in wild-type and citrate synthase mutant *escherichia coli*: evidence for multiple pathways of propionate utilization. *Biochem. J.*, 291:927–932, 1993.
- [ESW⁺93] W. Eisenreich, G. Strauss, U. Werz, G. Fuchs, and A. Bacher. Retrobiosynthetic analysis of carbon fixation in the phototrophic eubacterium *chloroflexus aurantiacus*. *Eur. J. Biochem.*, 215:619–632, 1993.

- [FCA⁺93] I. Flanigan, J.G. Collins, K.K. Arora, J.K. MacLeod, and J.F. Williams. Exchange reactions catalysed by group-transferring enzymes oppose quantitation and the unravelling of the identity of the pentose phosphate pathway. *Eur.J.Biochem.*, 213:477–485, 1993.
- [FHBS90] S.M. Fitzpatrick, H.P. Hetherington, K.L. Behar, and R.G. Shulman. The flux from glucose to glutamate in the rat brain in vivo as determined by ¹H-observed, ¹³C-edited NMR spectroscopy. *J. Cerebr. Blood Flow Metab.*, 10:170–179, 1990.
- [Fie84] A. Fiechter. Physical and chemical parameters in microbial growth. *Adv. Biochem. Eng. Biotechnol.*, 30:7–60, 1984.
- [Fle87] R. Fletcher. *Practical Methods of Optimization (Volume 1+2)*. Wiley-Interscience. Wiley, second edition, 1987.
- [FR95] C.A. Fernandez and C. Des Rosiers. Modeling of liver citric acid cycle and gluconeogenesis based on ¹³C mass isotopomer distribution analysis of intermediates. *J. Biol. Chem.*, 270:10037–10042, 1995.
- [Fri86] H. Friebolin. *Ein- und zweidimensionale NMR-Spektroskopie*. Verlag Chemie, 1986.
- [Gad82] D.G. Gadian. *Nuclear Magnetic Resonance and its Application to Living Systems*. Oxford University Press, 1982.
- [Gea71] G.W. Gear. *Numerical Initial Value problems in Ordinary Differential Equations*. Prentice-Hall, 1971.
- [GFJA93] A. Goel, J. Ferrance, J. Jeong, and M.M. Atai. Analysis of metabolic fluxes in batch and continuous cultures of *bacillus subtilis*. *Biotechnol. Bioeng.*, 42:686–696, 1993.
- [GHJV95] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison Wesley, 1995.
- [GI83] D. Goldfarb and A. Idnani. A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming*, 27(1):1–33, 1983.
- [GR89] A. Goldberg and D. Robson. *Smalltalk-80, The Language*. Addison Wesley, 1989.
- [Gro93] A. Groetschel. *Inverse Probleme*. Vieweg, 1993.
- [GS91] G. Gielen and W. Sansen. *Symbolic Analysis for Automated Design of Analog Integrated Circuits*. Kluwer Academic Publishers, 1991.
- [GSB90] J.L. Galazzo and J.V. Shanks and J.E. Bailey. Comparison of intracellular sugar-phosphate levels from ³¹P NMR spectroscopy of intact cells and cell-free extracts. *Biotech. Bioeng.*, 35:1164–1168, 1990.

- [GSPW⁺91] A.A. de Graaf, S.M. Schoberth, U. Probst, R.M. Wittig, J. Strohhäcker, and H. Sahm. Charakterisierung des Zuckerstoffwechsels im Bakterium *zymomonas mobilis* mit Hilfe der In-vivo-³¹P-NMR-Spektroskopie. *Chem. Ing. Techn.*, 63(6):628–631, 1991.
- [GWPS⁺92] A.A. de Graaf, R.M. Wittig, U. Probst, J. Strohhäcker, S.M. Schoberth, and H. Sahm. Continuous-flow NMR bioreactor for *in vivo* studies of microbial cell suspensions with low biomass concentrations. *J. Magn. Res.*, 98:654–659, 1992.
- [HA83] H-Amman. *Gewöhnliche Differentialgleichungen*. de Gruyter, 1983.
- [Hac93] W. Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme*. Number 69 in LAMM. Teubner Verlag, 1993.
- [Hag88] W.W. Hager. *Applied Numerical Linear Algebra*. Prentice Hall, 1988.
- [Har64] H.O. Hartley. Exact confidence regions for parameters in non-linear regression laws. *Biometrika*, 51(3,4):347–353, 1964.
- [Har95] A. Hartbrich. *Verfahrenstechnische Charakterisierung von Zyklonreaktoren in der Biotechnologie*. PhD thesis, RWTH Aachen, 1995.
- [HBS81] J.A. den Hollander, K.L. Behar, and R.G. Shulman. ¹³C NMR study of transamination during acetate utilization by *saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.*, 78:2693–2697, 1981.
- [Hea91] A.C. Hearn. *REDUCE User's Manual*, volume CP78. RAND, 1991.
- [HHHR⁺94] R.T.J.M. van Heijden, J.J. Heijnen, C. Hellinga, B. Romein, and K.C.A.M. Luyben. Linear constraint relations in biochemical reaction systems: I. classification of the calculability and the balanceability of conversion rates. *Biotechnol. Bioeng.*, 43:3–10, 1994.
- [Him78] D.M. Himmelblau. *Fault Detection in Chemical and Petrochemical Processes*. Number 8 in Chemical Engineering Monographs. Elsevier Scientific Publishing Company, 1978.
- [Him89] M. Himsolt. Graph^{ed}: An interactive graph editor. In B. Monien and R. Lori, editors, *STACS89, Proceedings of the 6th Annual Symposium on Theoretical Aspects of Computer Science, Paderborn, FRG, 1989*, volume 349 of *Lecture Notes in Computer Science*, pages 532–533. Springer, 1989.
- [HJ85] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [HNW87] E. Hairer, S.P. Norsett, and G. Wanner. *Solving Ordinary Differential Equations I — Nonstiff Problems*, volume 8 of *Series in Computational Mathematics*. Springer, 1987.
- [Hof86] J.H.S. Hofmeyr. Steady-state modelling of metabolic pathways: a guide for the prospective simulator. *CABIOS*, 2(1):5–11, 1986.

- [Hof93] R. Hofestädt. Grammatical formalization of metabolic processes. In Hunter et al. [HSS93], pages 181–189.
- [Hol85] H.G. Holzhütter. Compartmental analysis: Theoretical aspects and applications. *Biomed.Biochim.Acta*, 44:863–873, 1985.
- [Hol86] W.H. Holms. The central metabolic pathways of escherichia coli: relationship between flux and control at a branch point, efficiency of conversion to biomass, and excretion of acetate. *Curr. Topics Cell. Regul.*, 28:69–104, 1986.
- [Hol94] J.A. Den Hollander. Potential metabolic limitations in lysine production by corynebacterium glutamicum as revealed by metabolic network analysis. *Appl. Microbiol. Biotechnol.*, 42:508–515, 1994.
- [HRHH⁺94] R.T.J.M. van Heijden, B. Romein, J.J. Heijnen, C. Hellinga, and K.C.A.M. Luyben. Linear constraint relations in biochemical reaction systems: II. diagnosis and estimation of gross errors. *Biotechnol. Bioeng.*, 43:11–20, 1994.
- [HS93] H.G. Holzhütter and A. Schwendel. Estimation of enzymatic flux rates from kinetic isotope experiments. In Schuster et al. [SROM93].
- [HSG⁺93] H.-G. Holzhütter, A. Schwendel, T. Grune, J. Quedenau, and W. Siems. Estimation of steady-state flux rates in metabolic systems by computer simulations of radioactive tracer experiments. *CABIOS*, 9(5):573–580, 1993.
- [HSS93] L. Hunter, D. Searls, and J. Shavlik, editors. *ISMB-93, Proceedings of the First International Conference on Intelligent Systems for Molecular Biology*. AAAI Press, 1993.
- [Hub81] P.J. Huber. *Robust Statistics*. Wiley, 1981.
- [HUBB⁺86] J.A. den Hollander, K. Ugurbil, T.R. Brown, M. Bednar, C.Redfield, and R.G. Shulman. Studies of anaerobic and aerobic glycolysis in *saccharomyces cerevisiae*. *Biochemistry*, 25:203–211, 1986.
- [HW91] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II —Stiff and Differential-Algebraic Problems*, volume 14 of *Series in Computational Mathematics*. Springer, 1991.
- [IL87] L. Inbar and A. Lapidot. ¹³C-NMR, ¹H-NMR and gas-chromatography mass-spectrometry studies of the biosynthesis of ¹³C-enriched L-lysine by *brevibacterium flavum*. *Eur. J. Biochem.*, 162:621–633, 1987.
- [ISNY⁺91] S. Ishino, J. Shimomura-Nishimuta, K. Yamaguchi, K. Shirahata, and K. Arai. ¹³C nuclear magnetic resonance studies of glucose metabolism in L-glutamic acid and L-lysine fermentation by *corynebacterium glutamicum*. *J. Gen. Appl. Microbiol.*, 37:157–165, 1991.
- [Jor95] H. Jorgensen. Metabolic flux distributions in penicillium chrysogenum during fed-batch cultivations. *Biotech. Bioeng.*, 46:117–131, 1995.

- [JRMS91] F.M.H. Jeffrey, A. Rajagopal, C.R. Malloy, and A.D. Sherry. ^{13}C -NMR: a simple yet comprehensive method for analysis of intermediary metabolism. *TIBS*, 16:5–10, 1991.
- [JS87] D.W. Jordan and P. Smith. *Nonlinear Ordinary Differential Equations*. Oxford Applied Mathematics and Computing Science Series. Clarendon, 1987.
- [JSJ⁺93] J.G. Jones, A.D. Sherry, F.M.H. Jeffrey, C.J. Storey, and C.R. Malloy. Sources of acetyl-CoA entering the tricarboxylic acid cycle as determined by analysis of succinate ^{13}C isotopomers. *Biochemistry*, 32:12240–12244, 1993.
- [JWBL89] A.W.H. Jans, C. Winkel, L. Buitenhuis, and J. Lugtenburg. ^{13}C -n.m.r. study of citrate metabolism in rabbit renal proximal-tubule cells. *Biochem. J.*, 257:425–429, 1989.
- [KA93] H. Kacser and L. Acerenza. A universal method for achieving increases in metabolite production. *Eur. J. Biochem.*, 216:361–367, 1993.
- [Kac88] H. Kacser. Regulation and control of metabolic pathways. In M.J. Bazin and J.I. Prosser, editors, *Physiological Models in Microbiology*, volume 2 of *CRC Series in Mathematical Models in Microbiology*, chapter 1, pages 1–23. CRC Press, 1988.
- [KC83] P.W. Kuchel and B.E. Chapman. NMR spin exchange kinetics at equilibrium in membrane transport and enzyme systems. *J.Theor.Biol.*, 105:569–589, 1983.
- [KCS93] B. Künnecke, S. Cerdan, and J. Seelig. Cerebral metabolism of (1,2- $^{13}\text{C}_2$)glucose and (U- $^{13}\text{C}_4$)3-hydroxybutyrate in rat brain as detected by ^{13}C NMR spectroscopy. *NMR in Biomedicine*, 6:264–277, 1993.
- [KKW79] P. Kelly, J.K. Kelleher, and B.E. Wright. The tricarboxylic acid cycle in *dictyostelium discoideum* — metabolite concentrations, oxygen uptake and ^{14}C -labelled amino acid labelling patterns. *Biochem. J.*, 184:581–588, 1979.
- [Kof92] M. Kofler. *Mathematica —Einführung undLeitfaden für den Praktiker*. Addison Wesley, 1992.
- [Kof94] M. Kofler. *Maple V, Release 3*. Addison Wesley, 1994.
- [KR79] P. Kaps and P. Rentrop. Generalized runge-kutta methods of order four with step-size control for stiff ordinary differential equations. *Numer. Math.*, 33:55–68, 1979.
- [KW86] D.B. Kell and H.V. Westerhoff. Metabolic control theory: its role in microbiology and biotechnology. *FEMS Microbiology Reviews*, 39:305–320, 1986.
- [KWL93] J. Katz, P. Wals, and W.-N.P. Lee. Isotopomer studies of gluconeogenesis and the krebs cycle with ^{13}C -labeled lactate. *J. Biol. Chem.*, 268:25509–25521, 1993.
- [KWR88] K. Kanamori, R.L. Weiss, and D. Roberts. Glutamate biosynthesis in *bacillus azotofixans* — ^{15}N NMR and enzymatic studies. *J. Biol. Chem.*, 263(6):2871–2823, 1988.

- [Lan85] S. Lang. *Complex Analysis*, volume 103 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 1985.
- [LC80] A. Lent and Y. Censor. Extensions of hildreth's row-action method for quadratic programming. *SIAM Journal Contr. Opt.*, 18(4):444–454, 1980.
- [Lee93] W.-N.P. Lee. Analysis of tricarboxylic acid cycle using mass isotopomer ratios. *J. Biol. Chem.*, 268:25522–25526, 1993.
- [LH74] C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems*. Prentice Hall, 1974.
- [LHHV90] P. Lundberg, E. Harmsen, C. Ho, and H.J. Vogel. Nuclear magnetic resonance studies of cellular metabolism. *Anal. Biochem.*, 191:193–222, 1990.
- [Lon88] R.E. London. ^{13}C labeling in studies of metabolic regulation. *Prog. NMR Spectroscopy*, 20:337–383, 1988.
- [Lon92] R.E. London. In vivo ^2H NMR studies of cellular metabolism. In L.J. Berliner and J. Reuben, editors, *Biological Magnetic Resonance*, volume 11, chapter 6, pages 277–306. Plenum Press, 1992.
- [LPdPB88] M.E. Ligthelm, B.A. Prior, J.C. du Preez, and V. Brandt. An investigation of d-[1- ^{13}C] xylose metabolism in pichia stipitis under aerobic and anaerobic conditions. *Appl. Microbiol. Biotechnol.*, 28:293–296, 1988.
- [LR83] R.M. Lambrecht and A. Rescigno, editors. *Tracer Kinetics and Physiological Modelling*. Number 48 in *Lecture Notes in Biomathematics*. Springer, 1983.
- [LR87] Y. Lecourtier and A. Raksanyi. The testing of structural properties through symbolic computation. In Walter [Wal87], chapter 7, pages 75–84.
- [LVHHV95] E.M. Lohmeier-Vogel, B. Hahn-Hägerdal, and H.J. Vogel. Phosphorus-31 and carbon-13 nuclear magnetic resonance studies of glucose and xylose metabolism in *candida tropicalis* cell suspensions. *Appl. Env. Microbiol.*, 61:1414–1419, 1995.
- [Mar63] D.W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *J.Soc.Indust.Appl.Math.*, 11(2):431–441, 1963.
- [Mar94] A. Marx. Charakterisierung des Zentralstoffwechsels bei *corynebacterium glutamicum* mittels Metabolitbilanzierung und computergestützter Analyse von ^{13}C -NMR-Markierungsdaten. Diplomarbeit, Universität Bonn, 1994.
- [Mat82] N.A. Matwiyoff. Recent applications of ^{13}C NMR spectroscopy to biological systems. In Schmidt et al. [SFH82], pages 573–585.
- [Mat93] C.K. Mathews. The cell: Bag of enzymes or network of channels? *J. Bacteriol.*, 175(20):6377–6381, 1993.
- [Mav93] M.L. Mavrovouniotis. Identification of localized and distributed bottlenecks in metabolic pathways. In Hunter et al. [HSS93], pages 275–283.

- [MBW94] W.J. Malaisse, M. Biesemans, and R. Willem. Mathematical modelling for the generation of L-[3-²H, 3-¹³C]lactic acid isotopomers by erythrocytes exposed to either D-[1-¹³C]glucose or D-[6-¹³C]glucose in the presence of ²H₂O. *Mol. Cell. Biochem.*, 130:129–136, 1994.
- [MCDB94] G. Martin, M.-F. Chauvin, S. Dugelay, and G. Baverel. Non-steady state model applicable to NMR studies for calculating flux rates in glycolysis, gluconeogenesis, and citric acid cycle. *J. Biol. Chem.*, 269:26034–26039, 1994.
- [Mes93] G. Meszéna. Path analysis in general control theory. In Schuster et al. [SROM93].
- [MGWE⁺96] A. Marx, A.A. de Graaf, W. Wiechert, L. Eggeling, and H. Sahm. Determination of the fluxes in central metabolism of *corynebacterium glutamicum* by NMR spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.*, 49:111–129, 1996.
- [Mif79] R. Mifflin. A stable method for solving certain constrained least squares problems. *Mathematical Programming*, 16:141–158, 1979.
- [MMN89] H. Melenk, H.M. Möller, and W. Neun. Symbolic solution of large stationary chemical kinetic problems. *Impact of Computing in Science and Engineering*, 1(2), 1989.
- [MN91] C.A. Marinov and P. Neittanmäki. *Mathematical Models in Circuit Theory: Theory and Applications*. Kluwer Academic Publishers, 1991.
- [Möl95] M. Möllney. Zeiteffiziente Parameteranpassung bei NMR-Gemischspektren. Diplomarbeit, Universität Bonn, 1995.
- [Mor78] J.J. Moré. The levenberg-marquardt algorithm: Implementation and theory. In A. Dold and B. Eckmann, editors, *Numerical Analysis, Proceedings of the Biennial Conference, Held at Dundee, June 28-July 1, 1977*, number 630 in Lecture Notes in Mathematics, pages 105–116. Springer, 1978.
- [MS83] J.J. More and D.C. Sorensen. Computing a trust region step. *SIAM J. Sci.-Statist. Comput.*, 4(3):553–572, 1983.
- [MSJ88] C.R. Malloy, A.D. Sherry, and F.M.H. Jeffrey. Evaluation of carbon flux and substrate selection through alternate pathways involving the citric acid cycle of the heart by ¹³C NMR spectroscopy. *J. Biol. Chem.*, 263(15):6964–6971, 1988.
- [MSS90] M.L. Mavrovouniotis, G. Stephanopoulos, and G. Stephanopoulos. Computer-aided synthesis of biochemical pathways. *Biotechnol. Bioeng.*, 36:1119–1132, 1990.
- [MST⁺94] A.M. Mancuso, S.T. Sharfstein, S.N. Tucker, D.S. Clark, and H.W. Blanch. Examination of primary metabolic pathways in a murine hybridoma with carbon-13 nuclear magnetic resonance spectroscopy. *Biotechnol. Bioeng.*, 44:563–585, 1994.

- [MTBK89] L.M. McIntyre, D.R. Thorburn, W.A. Bubb, and P.W. Kuchel. Comparison of computer simulations of the F-type and L-type non-oxidative hexose monophosphate shunts with ^{31}P experimental data from human erythrocytes. *Eur.J.Biochem.*, 180:399–420, 1989.
- [MTG94] D. Massiot, H. Thiele, and A. Germanus. Winfit – a windows-based program for lineshape analysis. *Bruker Report*, pages 43–46, 1994.
- [Mul90] M. Mullin. *Rapid Prototyping for Object-Oriented Systems*. Addison Wesley, 1990.
- [Mun89] A. Munack. Optimal feeding strategy for identification of monod-type models by fed-batch experiments. In N.M. Fish and R.I. Fox, editors, *Proceedings of the 4th international Congress on Computer Applications in Fermentation Technology: Modelling and Control of biotechnical Processes, Cambridge, September 25-29, 1988*. Society of Chemical Industry, Pergamon Press, 1989.
- [Ni93] Q. Ni. General large-scale nonlinear programming using sequential quadratic programming methods. *Bayreuther-Mathematische-Schriften*, 45:133–236, 1993.
- [NM94] O.M. Neijssel and T. de Mattos. The energetics of bacterial growth: A reassessment. *Mol. Microbiol.*, 13:179–182, 1994.
- [NIS90] F.C. Neidhardt, J.L. Ingraham, and M. Schaechter. *Physiology of the Bacterial Cell — A Molecular Approach*. Sinauer Associates, 1990.
- [NS83] Wang N.S. and G. Stephanopoulos. Application of macroscopic balances to the identification of gross measurement errors. *Biotech. Bioeng.*, 25:2177–2208, 1983.
- [NS89] F. Ni and H.A. Scheraga. Constrained iterative spectral deconvolution with applications in NMR spectroscopy. *J. Magn. Res.*, 82:413–418, 1989.
- [Oku94] K. Okuno. A new cultivation system operated under a super high magnetic field. *J. Ferment. Bioeng.*, 77:453–456, 1994.
- [OLea82] M.H. O’Leary. Heavy-atom isotope effects on enzyme-catalyzed reactions. In Schmidt et al. [SFH82], pages 67–75.
- [Pat92] R. Patnaik. Stimulation of glucose catabolism in escherichia coli by a potential futile cycle. *J. Bac.*, 174:7527–7532, 1992.
- [Pau93] F.N. Paulisch. *The Design of an Extendible Graph Editor*, volume 704 of *Lecture Notes in Computer Science*. Springer, 1993.
- [Páz86] A. Pázman. *Foundations of Optimal Experimental Design*. Kluwer Academics, 1986.
- [Páz93] A. Pázman. *Nonlinear Statistical Models*. Number 254 in *Mathematics and its Applications*. Kluwer Academic Publishing, 1993.

- [Per93] L. Perko. *Differential Equations and Dynamical Systems*, volume 7 of *Texts in Applied Mathematics*. Springer, 1993.
- [PFTV88] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C - The Art of Scientific Computing*. Cambridge University Press, 1988.
- [PK92] R. Pilette and D.T. Kincaid. First flow-thru analysis in ecosystem studies. *Ecol.Modelling*, 1992.
- [Pop84] K.R. Popper. *Objektive Erkenntnis - Ein evolutionärer Entwurf*. Hoffmann und Campe, 4th edition, 1984.
- [PSMC93] J.-C. Portais, R. Schuster, M. Merle, and P. Canioni. Metabolic flux determination in C6 glioma cells using carbon-13 distribution upon $[1-^{13}\text{C}]$ glucose incubation. *Eur. J. Biochem.*, 217:457–468, 1993.
- [PT95] C. Posten and B. Tibken. Application of algebraic methods to the calculation of steady states in continuous cultures. *Contr. Eng. Pract.*, 3(7), 1995.
- [PWK94] M.W. Pickett, M.P. Williamson, and D.J. Kelly. An enzyme and ^{13}C -NMR study of carbon metabolism in heliobacteria. *Photosynth. Res.*, 41:75–88, 1994.
- [RB85] M. Rabkin and J.J. Blum. Quantitative analysis of intermediary metabolism in hepatocytes incubated in the presence and absence of glucagon with a substrate mixture containing glucose, ribose, fructose, alanine and acetate. *Biochem. J.*, 225:761–786, 1985.
- [RBP⁺91] J. Rumbaugh, M. Blaha, W. Premerlani, F. Eddy, and W. Lorensen. *Object-Oriented Modeling and Design*. Prentice Hall, 1991.
- [RCRL90] M.F. Roberts, B.-S. Choi, D.E. Robertson, and S. Lesage. Free amino acid turnover in methanogens measured by ^{15}N NMR spectroscopy. *J. Biol. Chem.*, 265(30):18207–18212, 1990.
- [RDC⁺95] C. Des Rosiers, L. Di Donatos, B. Comte, A. Laplante, C. Marcoux, F. David, C.A. Fernandez, and H. Brunengraber. Isotopomer analysis of citric acid cycle and gluconeogenesis in rat liver — reversibility of isocitrate dehydrogenase and involvement of ATP-citrate lyase in gluconeogenesis. *J. Biol. Chem.*, 270:10027–10036, 1995.
- [Red88] C. Reder. Metabolic control theory: A structural approach. *J.Theor.Biol.*, 135:175–201, 1988.
- [RKBY94] B.D. Ross, P.B. Kingsley, and O. Ben-Yoseph. Measurement of pentose phosphate-pathway activity in a single incubation with $[1,6-^{13}\text{C}_2, 6,6-^2\text{H}_2]$ glucose. *Biochem.J.*, 302:31–38, 1994.
- [RMGGK95] C. Rollin, V. Morgant, A. Guyonvarch, and J.-L. Guerquin-Kern. ^{13}C -NMR studies of *corynebacterium melassecola* metabolic pathways. *Eur. J. Biochem.*, 227:488–493, 1995.
- [RML93] V.N. Reddy, M.L. Mavrovouniotis, and M.N. Liebman. Petri net representations in metabolic pathways. In Hunter et al. [HSS93], pages 328–336.

- [Roe83] J.A. Roels. *Energetics and Kinetics in Biotechnology*. Elsevier Biomedical Press, 1983.
- [RR93] H. Ratschek and J. Rokne. Experiments using interval analysis for solving a circuit design problem. *J. Globl. Opt.*, 3(4):501–518, 1993.
- [RR95] J. Ratschek and J. Rokne. *New Computer Methods for Global Optimization*. Statistics, Operational Research and Computational Mathematics. Ellis Horwood, 1995.
- [RTQ⁺96] M. Rizzi, U. Theobald, E. Querfurth, T. Rohrhirsch, M. Baltes, and M. Reuss. In vivo investigations of glucose transport in *saccharomyces cerevisiae*. *Biotechnol. Bioeng.*, 49:316–327, 1996.
- [SB73] J. Stoer and R. Bulirsch. *Einführung in die numerische Mathematik (Band 1+2)*. Heidelberger Taschenbücher. Springer, 1973.
- [SB79] R.B. Stein and J.J. Blum. Quantitative analysis of intermediary metabolism in *tetrahymena* — cells grown in proteose-peptone and resuspended in a defined nutrient-rich medium. *J. Biol. Chem.*, 254:10385–10395, 1979.
- [SB88] J. Vanni Shanks and E. Bailey. Estimation of intracellular sugar phosphate concentrations in *saccharomyces cerevisiae* using ³¹P nuclear magnetic resonance spectroscopy. *Biotechnol. Bioeng.*, 32:1138–1152, 1988.
- [Sch91] K. Schügerl, editor. *Biotechnology, Vol.4: Measuring, Modelling and Control*. Verlag Chemie, 2 edition, 1991.
- [Sch95] B. Schruppf. *Lysinbildung in Corynebacterium glutamicum: Analyse des Metabolitflusses durch Bestimmung zellinterner Aminosäurekonzentrationen und enzymatische Untersuchungen*. PhD thesis, Universität Düsseldorf, 1995.
- [Sch96] V. Schwingenheuer. Redundanzanalyse bei metabolischen Kohlenstoff-Markierungssystemen. Diplomarbeit, Universität Bonn, 1996. In Vorbereitung.
- [Sco92] D.W. Scott. *Multivariate Density Estimation — Theory, Practice, and Visualization*. Probability and Mathematical Statistics. Wiley, 1992.
- [SEBF92] G. Strauss, W. Eisenreich, A. Bacher, and G. Fuchs. ¹³C-NMR study of autotrophic CO₂ fixation pathways in the sulfur-reducing archaebacterium *thermoproteus neutrophilus* and in the phototrophic eubacterium *chloroflexus aurantiacus*. *Eur. J. Biochem.*, 205:853–866, 1992.
- [SEGS93] K. Sonntag, L. Eggeling, A.A. de Graaf, and H. Sahm. Flux partitioning in the split pathway of lysine synthesis in *corynebacterium glutamicum* — quantification by ¹³C- and ¹H-NMR spectroscopy. *Eur. J. Biochem.*, 213:1325–1331, 1993.
- [Seg75] I.H. Segel. *Enzyme Kinetics: Behaviour and Analysis of Rapid-Equilibrium and Steady-State Enzyme Systems*. Wiley, 1975.
- [SFH82] H.-L. Schmidt, H. Förstel, and K. Heinzinger, editors. *Stable Isotopes, Proceedings of the 4th International Conference, Jülich, March 23-26, 1981*, volume 11 of *Analytical Chemistry Symposia Series*. Elsevier, 1982.

- [SGSW⁺93] J. Strohhäcker, A.A. de Graaf, S.M. Schoberth, R.M. Wittig, and H. Sahm. ³¹P nuclear magnetic resonance studies of ethanol inhibition in *zymomonas mobilis*. *Arch. Microbiol.*, 159:484–490, 1993.
- [Sie96] C. Siefke. Ein numerisches Verfahren zur Flußschätzung bei metabolischen ¹³C-Markierungsexperimenten. Diplomarbeit, Universität Bonn, 1996.
- [SM95] K. Schügerl and A. Munack, editors. *3rd IFAC Symposium on Modelling and Control of Biotechnical Processes, Garmisch-Partenkirchen, 1995*. Pergamon Press, 1995.
- [Spe86] P. Spellucci. *Sequential quadratic programming: theory, implementation, problems*, volume 53 of *Methods Oper. Res.*, pages 183–213. Athenäum/Hain/Hanstein, Königstein/Ts., 1986.
- [Spe93] P. Spelluci. *Numerische Verfahren der nichtlinearen Optimierung*. Internationale Schriftenreihe zur Numerischen Mathematik, Lehrbuch. Birkhäuser, 1993.
- [Sre90] P.A. Srere. Citric acid cycle redux. *TIBS*, 15:411–412, 1990.
- [SROM93] S. Schuster, M. Rigoulet, R. Ouhabi, and J.-P. Mazat, editors. *Modern Trends in Biothermokinetics — Proc. 5th Int. Meeting on Biothermokinetics, Sept. 23-26, Bordeaux, France*. Plenum, 1993.
- [SS79] K. Schittkowski and J. Stoer. A factorization method for the solution of constrained linear least squares problems allowing subsequent data changes. *Numerische Mathematik*, 31:431–463, 1979.
- [SS93] G. Stephanopoulos and A.J. Sinskey. Metabolic engineering - methodologies and future prospects. *TibTech*, 11:392–396, 1993.
- [SSGM⁺95] K. Sonntag, J. Schwinde, A.A. de Graaf, A. Marx, B.J. Eikmanns, W. Wiechert, and H. Sahm. ¹³C NMR studies of the fluxes in the central metabolism of *corynebacterium glutamicum* during growth and overproduction of amino acids in batch cultures. *Appl. Microbiol. Biotech.*, 1995.
- [SSH92] R. Schuster, S. Schuster, and H.-G. Holzhütter. Simplification of complex kinetic models used for the quantitative analysis of nuclear magnetic resonance or radioactive tracer studies. *J.Chem.Soc. Faraday Trans.*, 88(19):2837–2844, 1992.
- [SSK⁺91] B. Schrupf, A. Schwarzer, J. Kalinkowski, A. Pühler, L. Eggeling, and H. Sahm. A functionally split pathway for lysine synthesis in *c. glutamicum*. *J. Bacteriol.*, 173(14):4510–4516, 1991.
- [SSMS93] B. Sumegi, A.D. Sherry, C.R. Malloy, and P.A. Srere. Evidence for orientation-conserved transfer in the TCA cycle in *saccharomyces cerevisiae*: ¹³C NMR studies. *Biochemistry*, 32:12725–12729, 1993.
- [STM⁺94] S.T. Sharfstein, S.N. Tucker, A. Mancuso, H.W. Blanch, and D.S. Clark. Quantitative in vivo nuclear magnetic resonance studies of hybridoma metabolism. *Biotechnol. Bioeng.*, 43:1059–1074, 1994.

- [Sto71] J. Stoer. On the numerical solution of constrained least squares problems. *SIAM Journal on Numerical Analysis*, 8(2):382–411, 1971.
- [Str88] L. Stryer. *Biochemistry*. W.H. Freeman Company, third edition, 1988.
- [Sun78] N. Sundararaj. A method for confidence regions for nonlinear models. *Austral. J. Statist.*, 20(3):270–274, 1978.
- [SV91] G. Stephanopoulos and J.J. Vallino. Network rigidity and metabolic engineering in metabolite overproduction. *Science*, 252:1675–1681, 1991.
- [SW89] G.A.F. Seber and C.J. Wild. *Nonlinear Regression*. Wiley, 1989.
- [Syk88] P. Sykes. *Reaktionsmechanismen der Organischen Chemie*. Verlag Chemie, 9th edition, 1988.
- [Szy95] T. Szyperski. Biosynthetically directed fractional ^{13}C -labeling of proteinogenic amino acids — an efficient analytical tool to investigate intermediary metabolism. *Eur.J.Biochem.*, 232:433–448, 1995.
- [TDHW91] S. Tran-Dinh, M. Herve, and J. Wietzerbin. Determination of flux through different metabolite pathways in *saccharomyces cerevisiae* by ^1H -NMR and ^{13}C -NMR spectroscopy. *Eur. J. Biochem.*, 201:715–721, 1991.
- [TMT79] O. Tosaka, H. Morioka, and K. Takinami. The role of biotin dependent pyruvate carboxylase in l-lysine production. *Agric.Biol.Chem.*, 43:1513–1519, 1979.
- [TWWB96] R. Takors, W. Wiechert, and D. Weuster-Botz. Experimental design for identification of macrokinetic models and model discrimination. 1996. Eingereicht.
- [Val91] J.J. Vallino. *Identification of Branch-Point Restrictions in Microbial Metabolism through Metabolic Flux Analysis and local Network Perturbations*. PhD thesis, Massachusetts Institute of Technology, 1991.
- [VP94a] A. Varma and B.O. Palsson. Metabolic flux balancing: basic concepts, scientific and practical use. *Bio/Technol.*, 12:994–998, 1994.
- [VP94b] A. Varma and B.O. Palsson. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type escherichia coli w3110. *Appl. Env. Microbiol.*, 60:3724–3731, 1994.
- [VS93] J.J. Vallino and G. Stephanopoulos. Metabolic flux distribution in *corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnol. Bioeng.*, 41:633–646, 1993.
- [Wal87] E. Walter, editor. *Identifiability of Parametric Models*. Pergamon, 1987.
- [WBA92] B.E. Wright, M.H. Butler, and K.R. Albe. Systems analysis of the tricarboxylic acid cycle in *dictyostelium discoideum*: I. the basis for model construction. *J. Biol. Chem.*, 267:3101–3105, 1992.
- [WD87] H.V. Westerhoff and K. van Dam. *Mosaic Nonequilibrium Thermodynamics and Control of biological Free-Energy Transduction*. Elsevier Amsterdam, 1987.

- [WG93] W. Wiechert and A.A. de Graaf. Modelling of ^{13}C labelling in metabolic pathways for in vivo flux analysis using NMR. In V. Bales, editor, *Modelling for improved Bioreactor Performance, EFB Working Party on Bioreactor Performance*, 27.-28. September 1993, Papiernicka, Slovakia, pages 19–24. Malé Centrum, Publisher & Bookshop, Bratislava, 1993.
- [WG96a] D. Weuster and A.A. de Graaf. Reaction engineering methods to study intracellular metabolite concentrations. *Adv.Biochem.Eng.Biotechnol.*, 54:76–108, 1996.
- [WG96b] W. Wiechert and A.A. de Graaf. Bidirectional reaction steps in metabolic networks. part i: Modelling and simulation of carbon isotope labelling experiments. 1996. Eingereicht.
- [WG96c] W. Wiechert and A.A. de Graaf. In vivo stationary flux analysis by ^{13}C labelling experiments. *Adv.Biochem.Eng.Biotechnol.*, 54:109–154, 1996.
- [WGKF⁺92] R.G. Weiss, S.T. Gloth, R. Kalil-Filho, V.P. Chacko, M.D. Stern, and G. Gerstenblith. Indexing tricarboxylic acid cycle flux in intact hearts by carbon-13 nuclear magnetic resonance. *Circ. Res.*, 70:392–408, 1992.
- [WGM95] W. Wiechert, A.A. de Graaf, and A. Marx. In vivo stationary flux determination using ^{13}C NMR isotope labelling experiments. In Schügerl and Munack [SM95].
- [WHK⁺82] T.E. Walker, C.H. Han, V.H. Kollman, R.E. London, and N.A. Matwiyoff. ^{13}C nuclear magnetic resonance studies of the biosynthesis by *microbacterium ammoniaphilum* of L-glutamate selectively enriched with carbon-13. *J. Biol. Chem.*, 257:1189–1195, 1982.
- [Wie90] W. Wiechert. *Interaktive Datenanalyse bei biotechnischen Prozeßdaten*. PhD thesis, Universität Bonn, 1990.
- [Wie94a] W. Wiechert. Design of a software framework for flux determination by ^{13}C NMR isotope labelling experiments. In E. Gnaiger, F.N. Gellerich, and M. Wyss, editors, *What is Controlling Life?*, volume 3 of *Modern Trends in BioThermoKinetics*. Innsbruck University Press, 1994.
- [Wie94b] W. Wiechert. Diskrete Analyse von Stoffwechselnetzwerken. In F. Haubensack and J. Sühnel, editors, *Bioinformatik—Computereinsatz in den Biowissenschaften*, pages 100–111. Institut für Molekulare Biotechnologie, Jena, Available on Internet: ftp.imb-jena.de, 1994.
- [Wie95a] W. Wiechert. Algebraic methods for the analysis of redundancy and identifiability in metabolic ^{13}C labelling systems. In D. Schomburg and U. Lessel, editors, *Bioinformatics; From Nucleic Acids and Proteins to Cell Metabolism*, pages 169–184. Verlag Chemie, 1995.
- [Wie95b] W. Wiechert. C++ im technisch-naturwissenschaftlichen Bereich. *Objekt Spektrum*, 7/8, 1995.
- [Wie96a] W. Wiechert. Metabolic flux determination by stationary ^{13}C tracer experiments: Analysis of sensitivity, identifiability and redundancy. In Dolezal and Fidler [DF95], pages 128–135.

- [Wie96b] W. Wiechert. Objektorientierte Entwurfsmuster für die Wiederverwendung numerischer Software-Komponenten. In W. Mackens, editor, *Workshop Software Engineering im Scientific Computing, Hamburg 1995*, 1996. In Vorbereitung.
- [WJW⁺95] W. Wiechert, B. Joksche, R. Wittig, A. Hartbrich, T. Höner, and M. Möllney. Object oriented programming for the biosciences. *CABIOS*, 11(8), 1995.
- [WK84] K. Walsh and D.E. Koshland. Determination of flux through the branch point of two metabolic cycles — the tricarboxylic acid cycle and the glyoxylate shunt. *J. Biol. Chem.*, 259:9646–9654, 1984.
- [WKKS82] F.J. Winkler, H. Kexel, C. Kranz, and H.-L. Schmidt. Parameters affecting the $^{13}\text{CO}_2/^{12}\text{CO}_2$ discrimination of the ribulose-1,5-bisphosphate carboxylase reaction. In Schmidt et al. [SFH82], pages 83–89.
- [WMWG95] R. Wittig, M. Möllney, W. Wiechert, and A.A. de Graaf. Interactive evaluation of nmr spectra from in vivo isotope labelling experiments. In Schügerl and Munack [SM95].
- [Wol82] M. Wolfsberg. The theoretical analysis of isotope effects. In Schmidt et al. [SFH82], pages 3–14.
- [Woo85] T. Wood. *The Pentose Phosphate Pathway*. Academic Press, 1985.
- [WP89] S.S. Walther and R.L. Peskin. Strategies for scientific prototyping in smalltalk. In N. Meyrowitz, editor, *Proc. OOPSLA 89*, pages 159–168. ACM Press, 1989.
- [WSGM96] W. Wiechert, C. Siefke, A.A. de Graaf, and A. Marx. Bidirectional reaction steps in metabolic networks. part ii: Flux estimation and statistical analysis. 1996. Eingereicht.
- [You89] E. Yourdon. *Modern Structured Analysis*. Prentice Hall, 1989.
- [You91] D.A. Young. *The X Windows System, OSF/Motif Edition*. Prentice Hall, 1991.
- [Yua95] Y. Yuan. On the convergence of a new trust-region algorithm. *Num. Math.*, 70(4), 1995.
- [ZED⁺95] M. Zhang, C. Eddy, K. Deanda, M. Finkelstein, and S. Picataggio. Metabolic engineering of a pentose metabolism pathway in ethanologenic *zymomonas mobilis*. *Science*, 267:240–243, 1995.
- [Zou91] G. Zoutendijk. *Mathematical Programming Methods*. Nort Holland, 1991.
- [ZS94] C. Zupke and G. Stephanopoulos. Modeling of isotope distributions and intracellular fluxes in metabolic networks using atom mapping matrices. *Biotechnol. Prog.*, 10:489–498, 1994.
- [ZS95] C. Zupke and G. Stephanopoulos. Intracellular flux analysis in hybridomas using mass balances and in vitro ^{13}C NMR. *Biotechn. Bioeng.*, 45:292–303, 1995.

Index

- ^{13}C -NMR, 23
- ^1H -NMR, 23
- 3-D-Input-Isotopomeren-Transitionsmatrix, 77
- 3-D-Isotopomeren-Transitionsmatrix, 77
- 3-D-Matrix, 73
- 4-D-Isotopomeren-Transitionsmatrix, 77
- 4-D-Matrix, 73
- äußeres Produkt, 74
- aktive Nebenbedingung, 160
- Aminosäuren, 36
- Anwendungs-Flußkoordinatensystem, 83, 91
- Approximationsfehler, 86
- Architektur
 - generische, *siehe* generische Architektur
- asymptotische Grenzlinie, 71
- ATP, 33
- Austauschfluß, 81, 82
 - Nebenbedingungsmatrix, 94
 - Wertevektor, 94
 - vektor, 90
- Auswaschungskorrektur, 40
- Batch-Kultur, 30
- Beispielauswertung, 192
- bidirektionaler Reaktionsschritt, 16, 34, 81
- Bilanzgleichung, 19, 61
 - C-Atom-, *siehe* C-Atom-Bilanzgleichung
 - instationäre, *siehe* instationäre Bilanzgleichung
 - Isotopomeren-, *siehe* Isotopomeren-Bilanzgleichung
 - Metabolit-, *siehe* Metabolit-Bilanzgleichung
- bimolekularer Reaktionsschritt, 66
- biochemisches Netzwerk, 29, 31
- Biomasse
 - abfluß, 20, 55
 - zusammensetzung, 19, 55
- C-Atom
 - Bilanzgleichung, 66, 72, 77
 - Markierungsvektor, 64
 - Netzwerk, 47, 50
 - Notation, 48
 - Pool, 13, 39, 48
 - Transition, 48, 76
 - Transitionsmatrix, 76
 - extrazelluläres, *siehe* extrazelluläres C-Atom
 - Input-, *siehe* Input-C-Atom
 - intrazelluläres, *siehe* intrazelluläres C-Atom
 - Output-, *siehe* Output-C-Atom
 - C-Atom-Metabolit-Zugehörigkeitsmatrix, 78
 - chemische Summennotation, *siehe* Reaktionsschritt-Notation
- Chi-Quadrat
 - Test, 174, 204
 - Verteilung, 175, 201
- Computer-Algebra, 68
- Corynebacterium glutamicum, 15, 22, 33, 57, 137, 192
- Cosubstrat, 56
- Datenflußdiagramm, 147
- differenzierbare Approximation, 85
- Direktionalitätsannahme, 34
- dissipative Matrix, 114
- dissipatives System, 113
- DNA, 36
- Doublett, 24
- dynamische Modellierung, 41
- elliptischer Konfidenzbereich, *siehe* Konfidenzellipsoid
- Energie
 - bilanzierung, 33
 - stoffwechsel, 33
- Entkopplung von Messung und Experiment, 36, 40
- Enzymkinetik, 42
- Erhaltungs

- gleichung, *siehe* Markierungs-Erhaltungsgleichung
- stoffwechsel, 33
- Escherichia coli, 30
- Eukaryont, 30
- explizite Lösung, 68
- extrazellulär, 52
- extrazellulärer
 - Fluß, 6, 19
 - Metabolit, 52
 - Reaktionsschritt, 53
- extrazelluläres
 - C-Atom, 53
 - Isotopomer, 53
- Falle, 108
- Fallenfreiheit, 108
- Fed-Batch-Kultur, 30
- Fehlerquadratsumme, 156
- Fisher-Informationsmatrix, 182
- Fluß
 - Meßmatrix, 98
 - Meßvektor, 98
 - Austausch-, *siehe* Austauschflußextrazellulärer, *siehe* extrazellulärer Fluß
 - freier, *siehe* freier Fluß
 - geglätteter Austausch-, *siehe* geglätteter Austauschfluß
 - großer Austausch-, *siehe* großer Austauschfluß
 - intrazellulärer, *siehe* intrazellulärer Fluß
 - kompaktifizierter Austausch-, *siehe* kompaktifizierter Austauschfluß
 - Netto-, *siehe* Nettofluß
 - zulässiger, *siehe* zulässiger Fluß
- Flußkoordinatensystem, 82
 - Anwendungs-, *siehe* Anwendungs-Flußkoordinatensystem
 - kompaktifiziertes, *siehe* kompaktifiziertes Flußkoordinatensystem
 - natürliches, *siehe* natürliches Flußkoordinatensystem
- Flußkoordinatentransformationen, *siehe* Koordinatentransformation
- Flußschätzung, 70, 153
- Flußvektor
 - Austausch-, *siehe* Austauschflußvektor
 - Gesamt-, *siehe* Gesamtflußvektor
- Hin-, *siehe* Hinflußvektor
- kompaktifizierter Austausch-, *siehe* kompaktifizierter Austauschfluß
- Netto-, *siehe* Nettoflußvektor
- Rück-, *siehe* Rückflußvektor
- formale Reaktionsschritt-Notation, *siehe* Reaktionsschritt-Notation
- freie Reaktionsenergie, 33
- freier Fluß, 68, 96
- Futile Cycle, 33, 211
- geglättete
 - Maximumfunktion, 85
 - Minimumfunktion, 85
- geglätteter Austauschfluß, 85
- Gemischanalyse, 23, 25, 28
- Gemischspektrum, *siehe* NMR-Gemischspektrum
- generische Architektur, 144
- Gesamt
 - Nebenbedingungsmatrix, 96
 - Ungleichungs-Nebenbedingungsmatrix, 97
 - Ungleichungs-Wertevektor, 97
 - Wertevektor, 96
 - flußvektor, 63
- gewichtete Norm, 155
- Glättungsparameter, 85
- globale Stabilität, 111
- Grenzgleichung, 132
- Grenzlinie
 - asymptotische, *siehe* asymptotische Grenzlinie
 - realisierbare, *siehe* realisierbare Grenzlinie
- großer Austauschfluß, 87
- Grundannahmen, 29
- Hilfsintermediat, 56
- Hinflußvektor, 63
- hochauflösendes NMR-Spektrum, 25
- Homogenität, 29
- Identifizierbarkeitsproblem, 70
- In-vivo NMR, 11, 23, 27
- Inaktivierung, 32
- inneres Produkt, 74
- Input
 - C-Atom, 53
 - C-Atom-Markierungsvektor, 64

- C-Atom-Transitionsmatrix, 76
- Isotopomer, 53
- Markierungsvektor, *siehe* Markierungsvektor
- Metabolit, 53
- Reaktionsschritt, 54
- Insel, 108
- instationäre Bilanzgleichung, 67, 79
- Intermediat, 52, *siehe* intrazellulärer Metabolit
- intrazellulär, 52
- intrazellulärer
 - Fluß, 6
 - Metabolit, 6, 52
 - Reaktionsschritt, 53
- intrazelluläres
 - C-Atom, 53
 - Isotopomer, 53
- inverses Problem, 153
- irreversibler Reaktionsschritt, 34
- Isotopen
 - Anreicherung, 7
 - Masseneffekte, 29, 31
- isotopische Stationarität, 10, 38
- Isotopomer, 10, 24
 - extrazelluläres, *siehe* extrazelluläres Isotopomer
- Input-, *siehe* Input-Isotopomer
- intrazelluläres, *siehe* intrazelluläres Isotopomer
- Output-, *siehe* Output-Isotopomer
- Isotopomeren
 - Anteil, 10
 - Bilanzgleichung, 66, 72, 77
 - Input-Markierungsvektor, 64
 - Markierungsvektor, 64
 - Netzwerk, 47, 50, 51
 - Notation, 48
 - Pool, 13, 48
- Isotopomeren-C-Atom-Zugehörigkeitsmatrix, 78
- Isotopomeren-Metabolit-Zugehörigkeitsmatrix, 78
- Kernresonanz, 22
- Kleinste-Quadrate-Schätzung, 156, 171
- Kohlenstoff-Markierungssystem, 47
- kompaktifizierte Flußkoordinaten, 91
- kompaktifizierter Austauschfluß, 89, 90
- kompaktifiziertes Flußkoordinatensystem, 89
- Kompaktifizierung, 89
- Kompaktifizierungs-Abbildung, 89
- kompartimentelle Matrix, 108
- Konfidenz
 - bereich, 175
 - ellipsoid, 175
- kontinuierliche Kultur, 19, 27, 30, 36
- Konzentrationsmessung, 11, 19, 42
- Koordinatentransformation, 90, 91
- Korrelationsmatrix
 - Parameter-, *siehe* Parameter-Korrelationsmatrix
- Kovarianzmatrix
 - Meßfehler-, *siehe* Meßfehler-Kovarianzmatrix
 - Parameter-, *siehe* Parameter-Kovarianzmatrix
- KQ-Schätzung, 156
- Kultur
 - Batch-, *siehe* Batch-Kultur
 - Fed-Batch-, *siehe* Fed-Batch-Kultur
 - kontinuierliche, *siehe* kontinuierliche Kultur
- Lösung
 - numerische, *siehe* numerische Lösung
 - zulässige, *siehe* zulässige Lösung
- Lösung von Bilanzgleichungen
 - explizite, *siehe* explizite Lösung
- lineares Regressionsmodell, 173
- Lumping, *siehe* Pool-Lumping
- markierte Substanz, 8
- Markierung, 8
- Markierungs
 - Erhaltungsgleichung, 183
 - Meßmatrix, 98
 - Meßvektor, 98
 - Scrambling, 57
- Markierungsanteil
 - Isotopomeren-, 10
 - Positions-, *siehe* Positions-Markierungsanteil
- Markierungsexperiment, 7, 8
- Markierungsinput
 - zulässiger, *siehe* zulässiger Markierungsinput

- Markierungssystem, *siehe* Kohlenstoff-Markierungssystem
- Markierungsvektor
- C-Atom-, *siehe* C-Atom-Markierungsvektor
 - C-Atom-Input-, *siehe* C-Atom-Input-Markierungsvektor
 - Isotopomeren-, *siehe* Isotopomeren-Markierungsvektor
 - Isotopomeren-Input-, *siehe* Input-Isotopomeren-Markierungsvektor
- Markierungszustand
- realisierbarer, *siehe* realisierbarer Markierungszustand
- Massenspektroskopie, 11, 21, 26
- Matrix
- 3-D-, *siehe* 3-D-Matrix
 - 4-D-, *siehe* 4-D-Matrix
 - dissipative, *siehe* dissipative Matrix
 - Fisher-Informations-, *siehe* Fisher-Informationsmatrix
 - kompartimentelle, *siehe* kompartimentelle Matrix
 - mehrdimensionale, *siehe* mehrdimensionale Matrix
 - Nebenbedingungs-, *siehe* Nebenbedingungsmatrix
 - Sensitivitäts-, *siehe* Sensitivitätsmatrix
 - stöchiometrische, *siehe* stöchiometrische Matrix
 - Transitions-, *siehe* Transitionsmatrix
 - Zugehörigkeits-, *siehe* Zugehörigkeitsmatrix
- Matrix-Matrix-Produkt, 75
- Matrix-Vektor-Produkt, 74
- Meßfehler, 40, 197
- Kovarianzmatrix, 98
 - vektor, 98
- Meßmatrix, 98
- Fluß-, *siehe* Fluß-Meßmatrix
 - Markierungs-, *siehe* Markierungs-Meßmatrix
 - Poolgrößen-, *siehe* Poolgrößen-Meßmatrix
- Meßvektor, 98
- Fluß-, *siehe* Fluß-Meßvektor
 - Markierungs-, *siehe* Markierungs-Meßvektor
 - Poolgrößen-, *siehe* Poolgrößen-Meßvektor
- mehrdimensionale Matrix, 72
- Metabolic Engineering, 4, 6
- metabolische Stationarität, 4, 19, 29, 30
- metabolischer Zyklus, 50
- Metabolit
- Bilanzgleichung, 65, 72, 75
 - Bilanzierungstechnik, 7, 15
 - Netzwerk, 47, 50
 - Notation, 48
 - Pool, 13, 34, 48
 - extrazellulärer, *siehe* extrazellulärer Metabolit, *siehe* extrazellulärer Metabolit
 - Input-, *siehe* Input-Metabolit
 - intrazellulärer, *siehe* intrazellulärer Metabolit, *siehe* intrazellulärer Metabolit
 - Output-, *siehe* Output-Metabolit
- Mitochondrien, 30
- monomolekularer Reaktionsschritt, 66
- multimolekularer Reaktionsschritt, 56
- Multiplett, 24
- analyse, 24, 25
- NADH, 33
- NADPH, 33, 210
- natürliches Flußkoordinatensystem, 63, 81, 91
- Nebenbedingungsmatrix
- Austauschfluß-, *siehe* Austauschfluß-Nebenbedingungsmatrix
 - Gesamt-, *siehe* Gesamt-Nebenbedingungsmatrix
 - Gesamt-Ungleichungs-, *siehe* Gesamt-Ungleichungs-Nebenbedingungsmatrix
 - Nettofluß-, *siehe* Nettofluß-Nebenbedingungsmatrix
 - Ungleichungs-, *siehe* Ungleichungs-Nebenbedingungsmatrix
 - zusätzliche, *siehe* zusätzliche Nebenbedingungsmatrix
- Nettofluß, 81, 82
- Nebenbedingungsmatrix, 94
 - Wertevektor, 94
 - vektor, 90
- Netzwerk
- struktur, 47
 - synthese, 147

- vereinfachung, 51
- biochemisches, *siehe* biochemisches Netzwerk
- C-Atom-, *siehe* C-Atom-Netzwerk
- Isotopomeren-, *siehe* Isotopomeren-Netzwerk
- Metabolit-, *siehe* Metabolit-Netzwerk
- Netzwerkspezifikation, 47
 - zulässige, *siehe* zulässige Netzwerkspezifikation
- Nichtlineares Programm, 159
- nichtlineares Regressionsmodell, 157
- NMR, 11, 21
 - In-vivo, *siehe* In-vivo NMR
- NMR-Gemischspektrum, 28
- NMR-Spektrum, 22
 - hochauflösendes, *siehe* hochauflösendes NMR-Spektrum
- Notation
 - C-Atom-, *siehe* C-Atom-Notation
 - Isotopomeren, *siehe* Isotopomeren-Notation
 - Metabolit-, *siehe* Metabolit-Notation
 - Reaktionsschritt-, *siehe* Reaktionsschritt-Notation
- numerische Lösung, 124
- Objektorientierung, 149
- Output
 - C-Atom, 53
 - Isotopomer, 53
 - Metabolit, 53
 - Reaktionsschritt, 54
 - Sensitivitätsmatrix, 176, 194
- oxidative Phosphorylierung, 33
- P/O-Quotient, 33
- Parameter
 - Effekt-Krümmung, 188
 - Korrelationsmatrix, 174, 201
 - Kovarianzmatrix, 174, 201
 - Sensitivitätsmatrix, 176, 204
- Pentose-5-phosphat-Pools, 34
- Pentosephosphatweg, 34
- Pichia stipitis, 135
- Pool
 - Lumping, 17, 34, 42
 - größe, 40
- C-Atom-, *siehe* C-Atom-Pool, *siehe* C-Atom-Pool
- Isotopomeren-, *siehe* Isotopomeren-Pool, *siehe* Isotopomeren-Pool
- Metabolit-, *siehe* Metabolit-Pool, *siehe* Metabolit-Pool
- Poolgrößen
 - Meßmatrix, 98
 - Meßvektor, 98
 - vektor, 65
- Positions-Markierungsanteil, 8
- positive Invarianz, 113
- Produktbildungsrate, 19
- Produkt, *siehe* Output-Metabolit
 - äußeres, *siehe* äußeres Produkt
 - inneres, *siehe* inneres Produkt
- Matrix-Matrix-, *siehe* Matrix-Matrix-Produkt
- Matrix-Vektor, *siehe* Matrix-Vektor-Produkt
- Produktbildungsschritt, 54
- Programmkomponenten, 146
- Proteine, 36
- Quadratisches Programm, 164
- Quartett, 24
- Rückflußvektor, 63
- Radioaktivitäts-Messung, 11
- Reaktionsgleichung, *siehe* Reaktionsschritt-Notation
- Reaktionsschritt
 - Notation, 47
 - bidirektionaler, *siehe* bidirektionaler Reaktionsschritt
 - bimolekularer, *siehe* bimolekularer Reaktionsschritt
 - extrazellulärer, *siehe* extrazellulärer Reaktionsschritt
 - Input-, *siehe* Input-Reaktionsschritt
 - intrazellulärer, *siehe* intrazellulärer Reaktionsschritt
 - irreversibler, *siehe* irreversibler Reaktionsschritt
 - monomolekularer, *siehe* monomolekularer Reaktionsschritt
 - multimolekularer, *siehe* multimolekularer Reaktionsschritt
- Output-, *siehe* Output-Reaktionsschritt

- unidirektionaler, *siehe* unidirektionaler Reaktionsschritt
- realisierbare Grenzlinie, 71
- realisierbarer Markierungszustand, 71
- Redundanz
 - beziehung, 69
 - matrix, 180
- Regressionsmodell
 - lineares, *siehe* lineares Regressionsmodell
 - nichtlineares, *siehe* nichtlineares Regressionsmodell
- Resonanz
 - entkopplungstechnik, 23, 25
 - frequenz, 23
- RNA, 36
- schnelles Gleichgewicht, 85
- Sensitivitäts
 - analyse, 204
 - maße, 176
- Sensitivitätsmatrix
 - Output-, *siehe* Output-Sensitivitätsmatrix
 - Parameter-, *siehe* Parameter-Sensitivitätsmatrix
- Shift-Experiment, 41
- Simplex, 120
- Simulation, 134
- Singulärwerte, 179
- Singulärwertzerlegung, 179
- skalare Spin-Spin-Kopplung, 24
- Srambling-Reaktion, *siehe* Markierungs-Scrambling
- stöchiometrische Gleichung, *siehe* Metabolit-Bilanzgleichung
 - erweiterte, 94
- stöchiometrische Matrix, 75
- stöchiometrischer Koeffizient, 55
- Stabilität
 - globale, *siehe* globale Stabilität
- Stabilitätsanalyse, 107
- Standardmeßtechnik, 19, 81
- stationäre Stoffflußanalyse, 4, 43
- Stationarität
 - isotopische, *siehe* isotopische Stationarität
- statistische Auswertung, 171
- Stoffflußanalyse, *siehe* stationäre Stoffflußanalyse
- Stoffwechselzyklus, 16
- streng dissipatives System, 114
- strukturelles Modell, 62
- Substrat
 - aufnahme, 19
 - aufnahmeschritt, 54
 - Input-, *siehe* Input-Metabolit
- Thermodynamik, 12, 16, 33, 34
- Transitions-Matrix
 - 3D-Input-Isotopomeren-, *siehe* 3D-Input-Isotopomeren-Transitionsmatrix
 - 3D-Isotopomeren-, *siehe* 3D-Isotopomeren-Transitionsmatrix
 - 4D-Isotopomeren-, *siehe* 4D-Isotopomeren-Transitionsmatrix
 - C-Atom-, *siehe* C-Atom-Transitionsmatrix
 - Input-C-Atom-, *siehe* Input-C-Atom-Transitionsmatrix
- Triplett, 24
- Ungleichungs
 - Nebenbedingungsmatrix, 97
 - Wertevektor, 97
- unidirektionaler Reaktionsschritt, 34, 55, 82
- unverzerrte Messung, 32
- Versuchsplanung, 214
- Verweilzeit, 39
- verzerrungsfreie Messung, 29, 36
- Verzweigung, 15
- Visualisierung, 149
- Vorläufermetabolit, 20, 36, 39
- Vorteilsfunktion, 164
- Wachstumsrate, 19, 55
- Wertevektor
 - Austauschfluß-, *siehe* Austauschfluß-Wertevektor
 - Gesamt-, *siehe* Gesamt-Wertevektor
 - Gesamt-Ungleichungs-, *siehe* Gesamt-Ungleichungs-Wertevektor
 - Nettofluß-, *siehe* Nettofluß-Wertevektor
 - Ungleichungs-, *siehe* Ungleichungs-Wertevektor
 - zusätzlicher, *siehe* zusätzlicher Wertevektor

Wertsetzung, 94

zusätzliche, *siehe* zusätzliche Wertsetzung
wohldefinierte physiologische Bedingungen,
29, 36

Zellkomponenten, 36

Zentralstoffwechsel, 6, 58

zerstörungsfreie Messung, 29, 31

Zuckerstoffwechsel, 10

Zugehörigkeitsmatrix

C-Atom-Metabolit-, *siehe* C-Atom-Metabolit-
Zugehörigkeitsmatrix

Isotopomeren-C-Atom-, *siehe* Isotopomeren-
C-Atom-Zugehörigkeitsmatrix

Isotopomeren-Metabolit-, *siehe* Isotopomeren-
Metabolit-Zugehörigkeitsmatrix

zulässige

Netzwerkspezifikation, 54

zulässige Lösung, 119

zulässiger

Fluß, 98

Markierungsinput, 119

Punkt, 160

zusätzliche

Nebenbedingungsmatrix, 95

Wertsetzung, 95

zusätzlicher Wertevektor, 95

Zymomonas mobilis, 135

Jül-3301
November 1996
ISSN 0944-2952