

Aus dem
Institut für Lasermedizin der Heinrich-Heine-Universität Düsseldorf

Direktor: Prof. Dr. med. Rainer Bayer

**Computerassistierte Erkennung von Absorptionsspektren
für die labormedizinische Diagnostik**

Dissertation

zur Erlangung des Grades eines Doktors der
Medizin

Der Medizinischen Fakultät der Heinrich-Heine-Universität
Düsseldorf

vorgelegt von

Andreas Hans-Hermann Fischer

2005

Als Inauguraldissertation gedruckt mit Genehmigung der Medizinischen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Gez.: Univ.-Prof. Dr. med. dent. Wolfgang H.-M. Raab
Dekan

Referent: Univ.-Prof. Dr. Bayer
Korreferent: Prof. Dr. Godehardt

Inhaltsverzeichnis

| | | |
|-------|--|----|
| 1 | Einleitung: Optische Spektrometrie in der Medizin..... | 5 |
| 1.1 | Probenanalyse im Großlabor | 9 |
| 1.2 | Probenanalyse in der Arztpraxis..... | 11 |
| 1.3 | Aufgabenstellung..... | 12 |
| 2 | Material und Methoden | 14 |
| 2.1 | Der Analysand: Probe abzüglich Leerwert..... | 23 |
| 2.2 | Der Analysand: bereinigte Probe verkettet mit Leerwert..... | 25 |
| 3 | Physikalisch-chemische Grundlagen..... | 27 |
| 3.1 | Das Lambert-Beersche Gesetz | 28 |
| 3.1.1 | Der Leerwert..... | 29 |
| 3.1.2 | Quantitative Bestimmung von Konzentrationen | 31 |
| 3.1.3 | Die Extinktionsfunktion $E_{\lambda}(c)$ | 31 |
| 3.2 | Das Spektrum als Charakteristikum einer Probe: der „Fingerprint“ | 32 |
| 4 | Ergebnisse | 34 |
| 4.1 | Mathematische Formalisierung | 34 |
| 4.1.1 | Das Spektrum einer Probe als Vektor des $(\mathbb{R}^+)^n$ | 34 |
| 4.1.2 | Spektren einer Substanz als Parameterfunktion der Konzentration | 34 |
| 4.2 | Abstandsmessungen im $(\mathbb{R}^+)^n$ | 36 |
| 4.2.1 | Punkt zu „beliebigem“ Polygon : Euklidischer Abstand..... | 37 |
| 4.2.2 | Punkt zu „monotonem“ Polygon : Abstand als Bogenmaß..... | 40 |
| 4.2.3 | Punkt zu „linearem“ Polygon : einfaches Skalarprodukt | 42 |
| 4.3 | Die Parameterfunktion..... | 43 |
| 4.3.1 | Darstellung durch Interpolation..... | 43 |
| 4.3.2 | Darstellung durch Approximation | 44 |
| 4.4 | Drei Lösungsansätze..... | 46 |
| 4.4.1 | Interpolationsmethode (Lösungsansatz 1)..... | 47 |
| 4.4.2 | Winkelmethode (Lösungsansatz 2) | 48 |
| 4.4.3 | Approximationsmethode (Lösungsansatz 3) | 49 |
| 4.5 | Numerische Anwendung der Teilverfahren auf die Spektren | 50 |
| 4.5.1 | Erkennung der Referenzsubstanzen | 50 |
| 4.5.2 | Anwendung der Algorithmen zur Testerkennung | 54 |
| 4.6 | Trenngüte..... | 57 |
| 4.7 | Vergleich der Verfahren | 58 |
| 4.7.1 | Analysand : Probe abzüglich Leerwert..... | 58 |
| 4.7.2 | Analysand: bereinigte Probe verkettet mit Leerwert..... | 60 |
| 5 | Diskussion | 62 |
| 6 | Zusammenfassung | 66 |
| 7 | Programm | 67 |
| 7.1 | Blockdiagramm der beteiligten Units..... | 67 |
| 7.2 | Kommentar zu den einzelnen Units | 67 |
| 7.3 | Laufzeit- und Speicherbedarfsanalyse..... | 72 |
| 7.4 | Quelltext im Pascal-Quellcode | 73 |
| 8 | Tabellen-, Abbildungs- und Beispielverzeichnis | 74 |
| 9 | Anhang | 75 |
| 9.1 | Mathematisch | 75 |
| 9.2 | Sonstige | 76 |

| | | |
|----|--|----|
| 10 | Literatur- und Abkürzungsverzeichnis | 78 |
| 11 | Danksagung | 81 |
| 12 | Lebenslauf | 82 |

1 Einleitung: Optische Spektrometrie in der Medizin

„Spektrometer dienen der spektralen Zerlegung und Messung von Strahlung mit dem Ziel, aus den Spektren Informationen über Emissions- und Absorptionsvermögen von Materie...zu gewinnen...(und)...Wechselwirkungen in der Materie zu untersuchen, die zu Änderungen der Spektren führen...“ (17). Den Spezialfall der Messung von sichtbarem Licht als elektromagnetische Strahlung bezeichnet man als Photometrie. Den grundlegenden Aufbau eines Spektrometers zeigt die folgende Skizze :

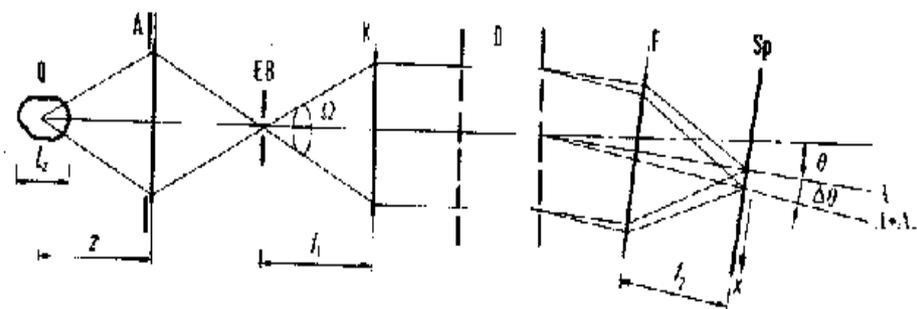


Abbildung 1-1 : Schematischer Aufbau eines Spektrometers (17)

Licht aus einer Strahlungsquelle Q fällt durch die Eintrittsblende EB in das Spektrometer. Nach Parallelisierung durch den Kollimator K fallen die Strahlen in das so genannte „dispersierende“ Element D; hierbei handelt es sich um ein Prisma, (Reflexions-)gitter oder Interferometer. Strahlung verschiedener Wellenlängen λ und $\Delta\lambda$ tritt in verschiedenen Richtungen θ und $\Delta\theta$ aus (Dispersion). Physikalische Grundlage der Dispersion ist die konstruktive Interferenz von Teilstrahlen mit ganzzahligem Gangunterschied (17). Eine fokussierende Optik F erzeugt dann das Spektrum Sp, in dem benachbarte Wellenlängen getrennt sind (17).

Die Registrierung des Spektrums kann durch eine Photoplatte erfolgen (Spektrograph) (17); hierbei wird eine lichtempfindliche Platte je nach Intensität des einfallenden Lichtes an von der Wellenlänge abhängiger Lokalisation geschwärzt. „Durch die integrierende Funktion der photoempfindlichen Schicht können bei langen Messzeiten auch schwache Spektren erfasst werden. Die quantitative Auswertung der Platten

(Densitometrie) ist aufwendig, da die...Schwärzung der Platte als Funktion der...Strahlungsleistung bekannt sein...muss.“ (17)

Moderner ist die photoelektrische Messung des Strahlenflusses; je nach Position des photoelektrischen Detektors zum dispergierenden Element wird zunächst nur die Intensität von Licht eines engen Wellenlängenbereiches – im Idealfall *einer* Wellenlänge – gemessen (Monochromator) (17). Um die Intensität mehrerer Wellenlängen mittels eines Detektors zu messen, wird das Spektrum durch Drehen des dispergierenden Elementes über den Austrittsspalt verfahren. „Mit rotierenden Gittern und Prismen ist es möglich, große Spektralbereiche schnell über den Austrittsspalt zu fahren“ (17). Mittels einer solchen „...Kurzzeitregistrierung kann die zeitliche Entwicklung der Emission einer Lichtquelle gemessen werden...“ (17). Zuverlässiger ist die Abtastung des Spektrums durch schrittmachermotorgetriebenes Verfahren des Empfängers (17).

Die folgende Abbildung zeigt den schematischen Aufbau eines Prismenspektrometers in Verbindung mit einer Photoplatte (a) und eines (Reflexions-)gitterspektrometers mit einkanaliger photoelektrischer Registrierung (b).

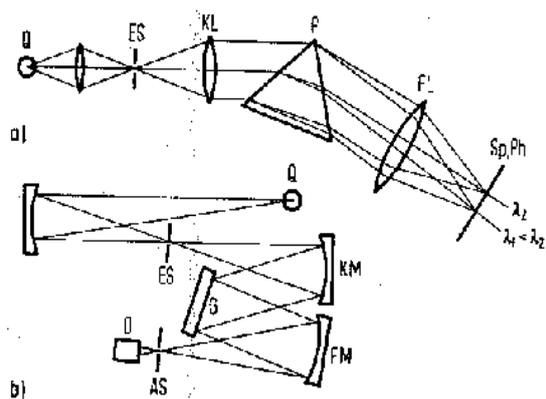


Fig. 5.80
a) Prismenspektrometer mit Linsenoptik und photographischer Registrierung des Spektrums
b) Gitterspektrometer mit Spiegeloptik und photoelektrischer Registrierung hinter dem Austrittsspalt AS

Abbildung 1-2 : Prismen- und Gitterspektrometer aus (17)

Umgangen wird diese mechanische Problematik durch vielkanalige spektrale photoelektrische Erfassung mit Photodiodenarrays, die – ein- oder zweidimensional – in der Ebene des Spektrums aufgestellt werden. „Die Sperrschichtkapazitäten der

einzelnen Photodioden wirken z.B. als Speicherkondensatoren für die photoelektrisch freigesetzten Ladungsträger, die elektronisch ...ausgelesen werden können.“ „Die Breite ... der einzelnen Photodioden begrenzt die Ortsauflösung...“ (17).

Im Unterschied zur vorgenannten dispersiven verwendet man bei der nichtdispersiven Spektrometrie optische Filter. Man unterscheidet dabei Farbgläser und Interferenzfilter (17). Sie „...dienen zur Selektion von Spektral-...linien sowie ausgesuchter Bereiche aus kontinuierlichen Spektren...“ (17). Nach Auswahl des interessierenden Wellenlängenbereiches erfolgt der Einschub eines geeigneten Filters zwischen Lichtquelle und Spektrometer.

„Spektrometrische Untersuchungsmethoden finden Anwendung in vielen Bereichen der Physik sowie unter anderem in Chemie, Biologie und Industrie“ (17), so auch in der Medizin (22) und Umwelttechnik (21). Durch Positionierung einer Flüssigkeitsprobe in einer lichtdurchlässigen Küvette detektionsseitig – aufgrund des geringeren Energieeffektes *hinter* dem dispergierenden Element im dann weitgehend monochromatischen Strahlengang – wird das Ausmaß der Absorption zu Aussagen über die Konzentrationen der in der Probe enthaltenen Testsubstanz herangezogen.

In der modernen Laboratoriumsdiagnostik werden dazu sowohl die dispersive als auch die nichtdispersive Spektrometrie als etablierte Verfahren zur qualitativen und quantitativen Bestimmung von klinisch relevanten Blut- und Blutserumbestandteilen benutzt. So umfaßt die Routineblutdiagnostik heute in Deutschland in einem Krankenhaus der Grundversorgung etwa in einer internistischen Abteilung ca. 40 verschiedene Bluttests aus den Funktionsgruppen Elektrolyte, Nierenwerte, Leber- u. Herzenzyme, Fette, großes Blutbild, kleine Gerinnung, Blutzucker, Schilddrüsenhormone u.v.m. (5). Darüberhinaus ist die Analyse anderer Körperflüssigkeiten häufig richtungsweisend für die Diagnosestellung, so z.B. die Proteinkonzentration im Urin (27) oder Liquor. Nach dem gleichen physikalischen Ansatz dient die transkutane Bilirubinometrie als Screeningmethode für einen beginnenden Kerninkterus bei Neugeborenen poststationär und zur Therapiekontrolle

(23)(7)(11)(4); die entsprechenden Geräte zeichnen sich durch Handlichkeit und Nichtinvasivität aus, z.B. BiliCheck^R. Diese Tests erfolgen dispersiv. Aber auch nichtdispersive Anwendungen in vivo sind beschrieben: am bekanntesten ist die sogenannte Pulsoximetrie, die über die transkutane Absorption an lediglich zwei Wellenlängen Aussagen über den Anteil an oxygeniertem Hämoglobin im Kapillarblut machen kann (6). Auch eine alternative Bilirubinmessmethode ist bekannt, Colormate III^R (4). Unterschiede bestehen nur in der Anordnung von Sender und Empfänger : bei der – reflexiven – Bilirubinometrie wird das von der Haut reflektierte Licht gemessen, bei der Pulsoximetrie befinden sich Sender und Empfänger vor und hinter dem durchstrahlten Gewebe. Experimentellen Charakter hingegen haben die spektroskopische Erfassung von oxgeniertem Hämoglobin zur nichtinvasiven Erfassung funktionell aktiver Gehirnareale beim Erwachsenen (13) und die Messung des zerebralen Blutvolumens bei Neugeborenen (28).

Die schnelle Bestimmung einiger Blutbestandteile ist von elementarer Bedeutung für die Einleitung spezifischer Notfallmaßnahmen. So sind die schnelle Diagnosestellung einer Anämie und die konsekutive Transfusion von Erythrozytenkonzentraten nach dem Ausmaß des Hämoglobinabfalles bei traumatischen und nichtraumatischen – in der Regel gastrointestinalen – Blutungen lebenswichtig.

In westlichen Industrieländern sind aber von den oben genannten Laborparametern für den niedergelassenen Arzt je nach indikationsbedingt angefordertem Profil nur zwei! bis zwölf bedeutsam(5), etwa mit der Fragestellung kardiovaskuläre Risikofaktoren nur Glucose im Serum und Cholesterin. Eine flächendeckende Erhebung der gesamten Routineparameter ist selten erforderlich. Neben Anamnese und körperlicher Untersuchung dienen diese wenigen Parameter zur Diagnosesicherung, Abgrenzung von Differentialdiagnosen und zur Verlaufskontrolle; oder sie sind als Notfallparameter bei der Indikationsstellung zur raschen Klinikeinweisung hilfreich. Wird die Bestimmung jedoch nicht vor Ort sondern in einem kooperierenden Labor durchgeführt, vergehen zwischen Probenentnahme und Ergebnismitteilung Tage, oft Wochen.

In ländlichen Regionen mit schlechter medizinischer Infrastruktur oder in Entwicklungsländern stellt sich dieses Problem aggraviert dar: Zentrallaboratorien oder Kliniken der Grundversorgung sind teilweise so weit entfernt, dass eine angemessene, zeitnahe Patientenbetreuung ohne lokal verfügbares Labor ambulant unmöglich ist.

1.1 Probenanalyse im Großlabor

Für Anwendungen in Großlabors mit mehreren tausend Analysen pro Tag stellen Anbieter wie Olympus^R vollautomatische Anlagen (z.B. AU640) zur Verfügung. Pro Stunde werden neben bis zu 400 Ionenkonzentrationsbestimmungen über Potentialdifferenzen in der ISE zeitgleich bis zu 800 Blutproben (20) nach nur geringfügiger manueller Beteiligung durch den Bediener mit den entsprechenden Reagenzien versetzt und durch Extinktionsmessung Konzentrationen bestimmt.

Physikalisch-technisch wird hier zwar die Gitterspektrometrie angewandt: das aus der Lichtquelle ausgesandte Gemisch verschiedener Wellenlängen durchdringt die Probe und trifft danach auf ein Reflexionsgitter. Je nach Wellenlänge erfahren die Lichtstrahlen bei der Reflexion eine bestimmte Ablenkung; ein im reflektierten Strahlengang positioniertes Photodiodenarray empfängt Licht definierter Wellenlängen(20). Verarbeitet wird aber im Weiteren nur das Signal *einer* Photodiode resp. einer Wellenlänge. Der Informationsgehalt der an anderen Wellenlängen gemessenen Lichtintensität wird nicht genutzt.

Der Automat ist dabei so programmiert, dass beispielsweise im Arbeitsschritt „Bestimmung der Bilirubinkonzentration“ das erforderliche Reagenz aus dem entsprechenden Behälter geladen und mit der Blutserumprobe versetzt wird. In Abhängigkeit von der aktuell zu bestimmenden Substanz erfolgt Auslesung der Lichtintensität aus dem Photodiodenarray an den für den Test charakteristischen Wellenlängen des Bilirubins, 546 nm und 560 nm, und die Extinktionsmessung. Die Ausgabe der gemessenen Konzentration erfolgt nach jeweiliger Multiplikation mit einer Konstanten, hier 63 mg/dl und 81 mg/dl und gegebenenfalls Mittelwertbildung.

Die Konstanten setzen sich aus den bekannten Größen Extinktionskoeffizient, Verdünnung und Schichtdicke der Küvette zusammen.

Die Qualitätskontrolle erfolgt durch eine breit gefächerte Anzahl von Qualitätskontrolle-Techniken (20); eine *spektrale* Analyse erfolgt nicht. Der Betrieb der Anlage setzt neben einem für den Bediener gut zugänglichen Stellplatz von mehreren Quadratmetern (20) auch eine genaue Schulung und Lagerungsmöglichkeiten für die Reagenzienkanister voraus. Der Kosten-Nutzenaufwand ist nur für sehr hohe Testzahlen pro Zeiteinheit gut.

1.2 Probenanalyse in der Arztpraxis

Durch Rationalisierung und Zentralisierung im Gesundheitswesen ist ein zunehmender Trend auch zur Zentralisation teurer Ressourcen wie Laborgeräten auszumachen. Dies impliziert häufig lange Wege von der Probe zum Labor und der Information vom Labor zu der weiter entfernt liegenden Station. Zur Bestimmung elementarer Laborinformationen auf Intensivstationen werden deshalb auch oft kleinere dezentrale Kombinationen aus Elektroden, Barometer und Photometer verwendet, die in einem Arbeitsgang unter anderem Blutgasparameter, Hämoglobin-Wert und Kaliumspiegel messen (z.B. Radiometer Kopenhagen ABL 500^R).

Für die Analyse kleiner Probenstückzahlen bei begrenztem Platzangebot – etwa in Arztpraxen und Ambulanzen – werden heute oft einfache, preiswerte „Photometer“ eingesetzt. Hier wird im Gegensatz zum Großlabor die *nicht*dispersive Spektroskopie angewandt, d.h. Messung von Absorption bzw. Extinktion in einem engen Wellenlängenbereich. Gewählt wird je nach Probe derjenige Wellenlängenbereich, in dem eine lineare Beziehung zwischen Konzentration der Probe und Extinktion besteht und so über den bekannten Extinktionskoeffizienten aus der Extinktion die Konzentration berechnet werden kann. Nach Auswahl der testspezifischen Wellenlänge erfolgt der manuelle Einschub eines entsprechenden Filters (Bandpass-, Interferenzfilter) zwischen Weißlichtquelle und Probe. Das so erzeugte – mehr oder weniger – monochromatische Licht durchdringt die Probe (18). Eine Positionierung der Probe *vor* dem Filter (emissionsseitig) ist aufgrund der zusätzlich in die Probe eingestrahlten Energie der Weißlichtquelle mit konsekutiven verfälschenden Effekten weniger günstig, aber möglich. Zuletzt wird die Intensität des Lichtes der selektierten Wellenlänge von einem Detektor gemessen. Die rechnerische Anwendung des Lambert-Beer'schen Gesetzes ergibt die Konzentration. Eine solche Anlage hat exklusive der angeschlossenen EDV etwa die Größe eines Schuhkartons; auch der finanzielle Aufwand liegt für die Betreiber im akzeptablen Bereich.

Die Bestückung des „Photometers“ mit der Probe erfolgt jedoch manuell. Dem Messgerät ist deshalb zunächst unbekannt, *welche* Substrate gerade bestimmt werden

sollen. Bedienerfehler sind zu befürchten: neben der fehlerhaften Wahl der Wellenlänge (Filter) könnte zur Bestimmung des Serumparameter A irrtümlicherweise mit dem Reagenz für Serumparameter B versetzt werden. Auch die bedienerseitig durchgeführte rechnerische Anwendung des Lambert-Beer'schen Gesetzes durch Multiplikation der Extinktion mit dem richtigen Extinktionskoeffizienten, Schichtdicke etc. ist fehlerträchtig. Ein falscher Messwert für Serumparameter A würde erhoben werden.

1.3 Aufgabenstellung

Die herkömmlichen spektralphotometrischen Ansätze können daher dahingehend zusammengefasst werden, dass

- 1) nach Erzeugung von monochromatischem Licht durch ein dispergierendes Element durch „Verfahren“ von Gittermonochromator/Prisma gegen Austrittsspalt/einkanaligem Detektor (oder umgekehrt) das Spektrum einschließlich der interessierenden Wellenlänge sukzessive gescannt wird oder
- 2) ein durch Dispersion an
 - a) Prisma oder
 - b) Reflexionsgitter
 erzeugtes Spektrum parallel von einem Photodiodenarray eingelesen wird oder
- 3) durch einen Filter erzeugtes, quasi monochromatisches Licht einkanalig gemessen wird.

Gemeinsames Kennzeichen sowohl der dispersiven als auch der nichtdispersiven Photometrie ist weiterhin, dass jeweils nur die Extinktion *einer* vorgegebenen Wellenlänge resp. eines vorgegebenen, engen Wellenlängenbereiches genutzt wird. Der Informationsgehalt der Extinktionen der Probe an anderen Wellenlängen wird ignoriert.

Die unter 1.) genannte Methode ist zum einen kostenintensiv und hat zum zweiten den Nachteil, dass das sukzessive Messen bei den verschiedenen Wellenlängen relativ viel

Zeit benötigt; bei der Methode 2.) liegt das Spektrum binnen Bruchteilen einer Sekunde an. Während jedoch der Aufbau von Prismenspektrometern (2a) noch recht aufwendig ist, können heute mittels moderner Mikrospritzgusstechnik Reflexionsspektrometer (2b) äußerst preisgünstig hergestellt werden.

Es liegt daher nahe, die unter 3) genannte einfache „Photometrie“ durch Spektrometrie unter Verwendung eines Mikrospektrometers zu ersetzen. Der zusätzliche Informationsgehalt eines Spektrums kann dazu genutzt werden, die Präzision der Messung zu erhöhen, den Test automatisch zu erkennen und Fehler zu identifizieren; der manuelle Einschub eines Filters (bzw. die Positionierung eines Detektors) ist überflüssig. Dies ist nur unter Verwendung der dispersiven Spektrometrie möglich.

Es sollen also mathematische Modelle für ein Mikrospektrometer entwickelt werden, die anhand der spektralen Verteilung von Absorptionsspektren

- Tests (*Testsubstanzen*) identifizieren. Dies geschieht durch Auswahl anhand ihrer Absorptionsspektren aus einer endlichen, bekannten Menge von möglichen *Referenzsubstanzen*, deren Spektrum jeweils in mindestens einer Konzentration in der Datenbank vorliegt.
- automatisch zwischen dem Spektrum der eigentlichen Probe und dem Spektrum des dazugehörigen Leerwertes unterscheiden
- die Konzentration der identifizierten Testsubstanz messen und ausgeben
- grob die Fehleranfälligkeit des o.g. Systems abschätzen und Hilfe bei der Fehleridentifikation geben.

2 Material und Methoden

Um die theoretisch entwickelten Algorithmen zu testen, wurden konzentrationsabhängige Spektren von folgenden sieben Referenzsubstanzen (Fa. Lange, Berlin) entsprechend den Angaben des Herstellers erhoben:

| Lfd. Nr. | Referenz- substanz | Substrat Methode | empfohlene Messwellen- längen [nm] für Extinktion | Extinktions- koeffizient | Einheit |
|----------|--------------------|---------------------------------|---|--------------------------|------------------------------|
| 1 | Bilirubin | Secutrol DPD | 546 | 0.007353 | $(mm \cdot mg/dl)^{-1}$ |
| 2 | Calcium | Secutrol o-CPC | 560-590 | Vergleich mit Standard | |
| 3 | Cholesterin | Secutrol CHOD-PAP | 520 | 0.000306 | $(mm \cdot mg/dl)^{-1}$ |
| 4 | Eisen | Secutrol Ferrozin | 578 | 0.000067 | $(mm \cdot \mu g/dl)^{-1}$ |
| 5 | Erythrozyten | Vollblut Photometrische Trübung | 500-600 | 0.037037 | $(mm \cdot 10^6/\mu l)^{-1}$ |
| 6 | Gesamteiweiß | Secutrol Biuret | 546 | 0.005263 | $(mm \cdot g/dl)^{-1}$ |
| 7 | Hämoglobin | Vollblut Hämiglobincyanid | 520-560 | 0.002331 | $(mm \cdot g/dl)^{-1}$ |

Zu Vergleichszwecken wurden zwei weitere Spektrenscharen nach folgenden Vorschriften erzeugt :

| | | | |
|---|--|------|-----------------------|
| 8 | Synthetisch A: $E_A = c \cdot \sin \lambda + e^{-\lambda}$ | alle | $(mm)^{-1} \cdot 0,1$ |
| 9 | Synthetisch B: $E_B = c \cdot \sin \lambda \cdot e^{-\lambda}$ | alle | $(mm)^{-1} \cdot 0,1$ |

Tabelle 2-1 : Referenzsubstanzen mit Testverfahren

Secutrol U ist ein lyophilisiertes Kontrollserum tierischen Ursprungs (Secutrol U^R LCQ 010 03/98 von Fa. Lange, Berlin), das Vollblut entstammt einer weiblichen gesunden Probandin von 28 Jahren; die Bestimmung der Erythrozytenzahl erfolgte über die Zählkammermethode, die Hämoglobinkonzentration durch photometrische Bestimmung nach Lyse der Erythrozyten mit einem cyanidfreien Detergenz (SLS-Methode).

Die unter 8. und 9. aufgeführten Modellspektren – die Funktionsangabe oben erfolgt ohne modifizierende Konstanten, die Funktionsvorschrift mit den verwendeten

Konstanten findet sich im Anhang in Form des PASCAL-Quelltextes – wurden numerisch zu Testzwecken erzeugt.

Durch Verwendung steigender Mengen des Secutrols bzw. Vollblut (und Variierung des Parameters c für die Konzentration in den zwei o.g. Formeln) wurden steigende Konzentrationen der Referenzsubstanzen erzeugt. Die vorgenommenen Mischungen und dabei entstehenden Konzentrationen ergeben sich aus der Tabelle 2-2:

| AnalySAT | Substrat | Konz. AnalySAT in Substrat | Herstellung der Küvettenlösung | | | Endkon- zentration AnalySAT in Küvette | Konzentration normierte Probe zurück- gerechnet | | | | |
|---------------------|----------|-------------------------------------|-----------------------------------|----------------------|----------------------|--|---|-----------|-------|--------|------|
| | | | FR/ul | Sub- strat /ul | Aqua Dest. /ul | | | | | | |
| Bilirubin | Secutrol | /[mg/dl] | FR/ul | /ul | /ul | /[mg/dl] | /[mg/dl] | | | | |
| | | | 1.98 | 1000 | 50 | | | 50 | 0.09 | 0.99 | |
| | | | 1.98 | 1000 | 100 | | | 0 | 0.18 | 1.98 | |
| | | | 1.98 | 1000 | 200 | | | 0 | 0.33 | 3.63 | |
| | | | 1.98 | 1000 | 400 | 0 | 0.57 | 6.22 | | | |
| Calcium | Secutrol | /[mmol/l] | RM/ul | /ul | /ul | /ul | /[mmol/l] | /[mmol/l] | | | |
| | | | 2.52 | 500 | 5 | 5 | | | 0 | 0.02 | 1.26 |
| | | | 2.52 | 500 | 0 | 0 | | | 10 | 0.05 | 2.52 |
| | | | 2.52 | 500 | 10 | 0 | | | 10 | 0.10 | 4.94 |
| | | | 2.52 | 500 | 10 | 0 | | | 20 | 0.14 | 7.27 |
| | | | 2.52 | 500 | 10 | 0 | 30 | 0.19 | 9.52 | | |
| Cholesterin | Secutrol | /[mg/dl] | FR/ul | /ul | /ul | /[mg/dl] | /[mg/dl] | | | | |
| | | | 150 | 500 | 5 | | | 5 | 1.47 | 75.00 | |
| | | | 150 | 500 | 10 | | | 0 | 2.94 | 150.00 | |
| | | | 150 | 500 | 20 | | | 0 | 5.77 | 294.23 | |
| | | | 150 | 500 | 40 | 0 | 11.11 | 566.67 | | | |
| Eisen | Secutrol | /[ug/dl] | FR/ul | /ul | /ul | /[ug/dl] | /[ug/dl] | | | | |
| | | | 152 | 500 | 50 | | | 50 | 12.67 | 76.00 | |
| | | | 152 | 500 | 100 | | | 0 | 25.33 | 152.00 | |
| | | | 152 | 500 | 200 | | | 0 | 43.43 | 260.57 | |
| | | | 152 | 500 | 400 | 0 | 67.56 | 405.33 | | | |
| Erythrozyten | Vollblut | /[10 ⁶ /ul] | GO/ml | /ml | /ml | /[10 ⁶ /ul] | /[10 ⁶ /ul] | | | | |
| | | | 4 | 5 | 0.01 | | | 0 | 0.01 | 2.00 | |
| | | | 4 | 5 | 0.02 | | | 0 | 0.02 | 4.00 | |
| | | | 4 | 5 | 0.04 | | | 0 | 0.03 | 7.97 | |
| | | | 4 | 5 | 0.08 | 0 | 0.06 | 15.81 | | | |
| Gesamteiweiß | Secutrol | /[g/dl] | BI/ul | /ul | /ul | /[g/dl] | /[g/dl] | | | | |
| | | | 7.52 | 1000 | 10 | | | 10 | 0.07 | 3.76 | |
| | | | 7.52 | 1000 | 20 | | | 0 | 0.15 | 7.52 | |
| | | | 7.52 | 1000 | 40 | | | 0 | 0.29 | 14.75 | |
| | | | 7.52 | 1000 | 80 | 0 | 0.56 | 28.41 | | | |
| Hämoglobin | Vollblut | /[g/dl] | TR/ml | /ml | /ml | /[g/dl] | /[g/dl] | | | | |
| | | | 11 | 5 | 0.01 | | | 0 | 0.02 | 5.51 | |
| | | | 11 | 5 | 0.02 | | | 0 | 0.04 | 11.00 | |
| | | | 11 | 5 | 0.08 | 0 | 0.17 | 43.48 | | | |

Tabelle 2-2 : Referenzspektren¹ : Berechnung der Konzentrationen der einzelnen Tests

¹ FR = Farbreagenz, RM = Reagenzmischung, GO = Gowers'sche Lösung, BI = Biuretreaktion, TR = Transformationslösung

In der Tabelle 2-3 sind die Konzentrationen der zu Testzwecken numerisch erzeugten Spektren aufgeführt:

| Analysat | Konzentration |
|----------------------|---|
| A Synthetisch | 0.83 |
| | 0.91 |
| | $E_A = c \cdot \sin \lambda + e^{-\lambda}$ |
| | 1.00 |
| | 1.10 |
| B Synthetisch | 1.21 |
| | 0.25 |
| | 0.91 |
| | $E_B = c \cdot \sin \lambda \cdot e^{-\lambda}$ |
| | 1.00 |
| | 2.00 |
| | 4.00 |

Tabelle 2-3 : Synthetische Spektren² : gewählte Konzentrationen

Die Funktionen E_A und E_B beschreiben die Abhängigkeit der Extinktion von der Konzentration und Wellenlänge. Dabei wird gefordert, dass die Extinktion über den gesamten Wellenlängenbereich bis auf eine additive Konstante proportional zur Konzentration ist. Die synthetischen Spektren sind in Abbildung 2-8 und Abbildung 2-9 dargestellt. E_B verhält sich damit exakt nach dem Lambert-Beerschen Gesetz. E_A wurde erzeugt, um das später verwandte Verfahren der linearen Regression, welches auch auf nichtverschwindende Achsenabschnitte anwendbar ist, zu testen.

Nach Pipettierung in Halbmikro-Rechteck-Küvetten der Schichtdicke 10 mm erfolgte die Extinktionsmessung im Bereich 400 nm bis 800 nm und einer Auflösung von 1 nm mit einem Spektralphotometer der Marke Perkin-Elmer Lambda 19 Near IR/Vis/UV bei Raumtemperatur (siehe Spektren der Abbildung 2-1 bis Abbildung 2-7). Proben und zugehörige Leerwerte lagen als ASCII-Textdateien vor; nach Abzug der jeweiligen Leerwerte erfolgte die Speicherung zur weiteren Verarbeitung mit einem einfachen PC:

² E_A : Extinktion des Spektrums A. E_B : Extinktion des Spektrums B. c: Konzentration. λ : Wellenlänge. sin: Sinusfunktion. e: Exponentialfunktion. Modifizierende Konstanten in den Termen wurden für die bessere Lesbarkeit weggelassen. Sie finden sich konkret im Anhang.

```

Calcium      //Substanzname
7.274716981 //Konzentration
800 0.00612 //Extinktion bei 800 nm
799 0.00673
798 0.00639
.
.
404 0.01964
403 -0.00493
402 -0.02313
401 -0.0116
400 -0.01422 //Extinktion bei 400 nm

```

Beispiel 1 : ASCII-Testdatei der Extinktionen von Calcium bei 7,27 mmol/l (// Kommentar)

Eine Steuerdatei enthält dann weitere Daten über die Referenzsubstanz:

```

Calcium      // Substanzname
J           // J : ein Standard muss angefordert werden
560 570 580 590 // Messwellenlängen lt. Testverfahren
mmol/l     // Ausgabeeinheit
1.26 9.52  // Minimale und maximal vorliegende Konzentration
C:\prom_asc\calcium\Ca_00_05.txt // Die einzelnen Spektren, nach steigenden
C:\prom_asc\calcium\Ca_10_00.txt // Konzentrationen geordnet
C:\prom_asc\calcium\Ca_10_10.txt
C:\prom_asc\calcium\Ca_20_10.txt
C:\prom_asc\calcium\Ca_30_10.txt

```

Beispiel 2 : ASCII-Steuerdatei von Calcium, Anforderung eines Standards (//Kommentar)

oder :

```

Erythrozyten
N           // N : es braucht kein Standard angefordert zu werden
500 520 560 585 600 534 // Messwellenlängen lt. Testverfahren
2.7 2.7 2.7 2.7 2.7 2.7 *10^6/ul // Faktoren lt. Testverfahren, mir denen
2.003992016 15.81102362 // die Extinktionen an den o.g. Messwellenlängen
C:\prom_asc\ery\ery05.txt // multipliziert werden müssen mit Einheit
C:\prom_asc\ery\ery1.txt
C:\prom_asc\ery\ery2.txt
C:\prom_asc\ery\ery4.txt

```

Beispiel 3 : ASCII-Steuerdatei von Erythrozyten, direkte Berechnung der Konzentration (//Kommentar)

Diese Daten wurden in das mit Turbo Pascal 6.0 erstellte Programm eingelesen und waren Datenbasis der Algorithmen. Diese selbsterklärende Programmiersprache wurde verwendet, um eine optimale Lesbarkeit des Programmkerns sicherzustellen.

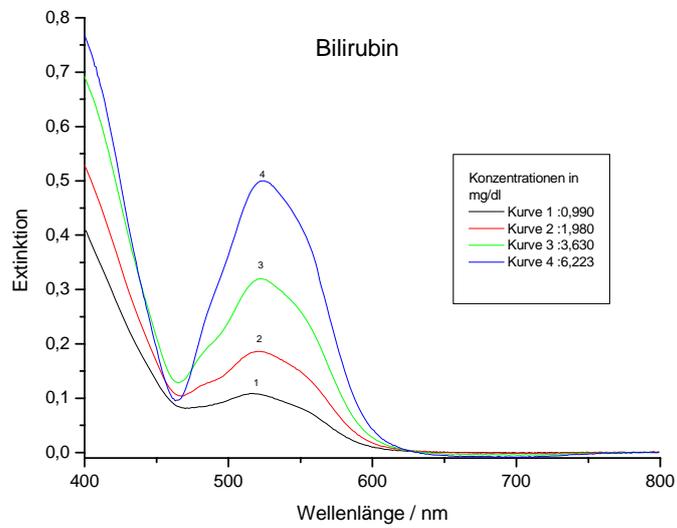


Abbildung 2-1 : Bilirubin

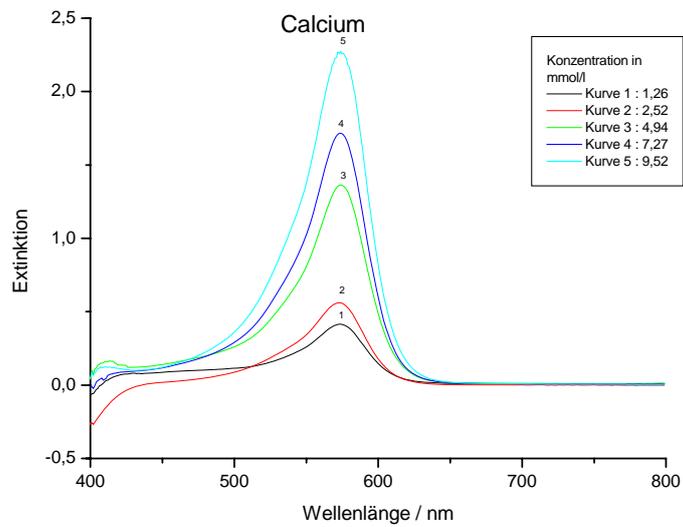


Abbildung 2-2 : Calcium

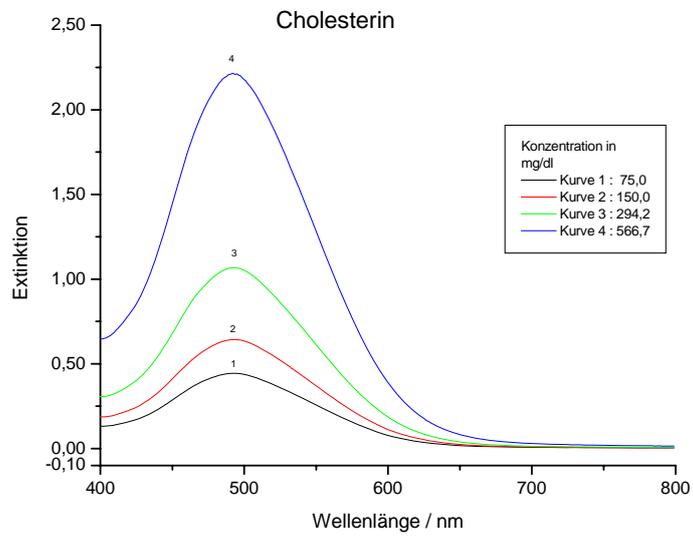


Abbildung 2-3 : Cholesterin

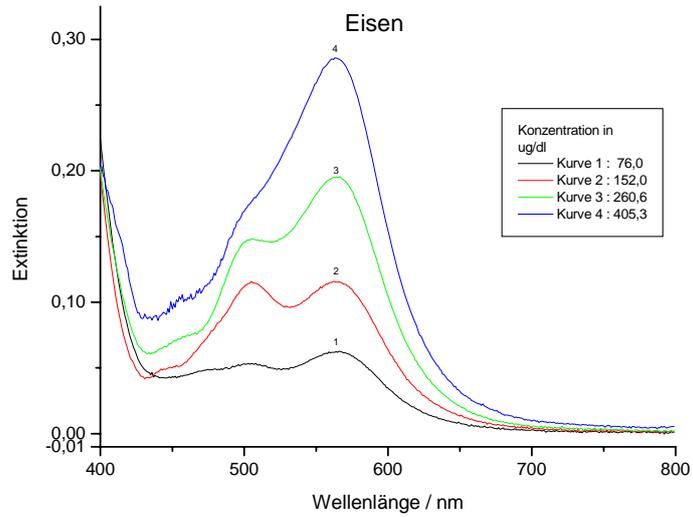


Abbildung 2-4 : Eisen

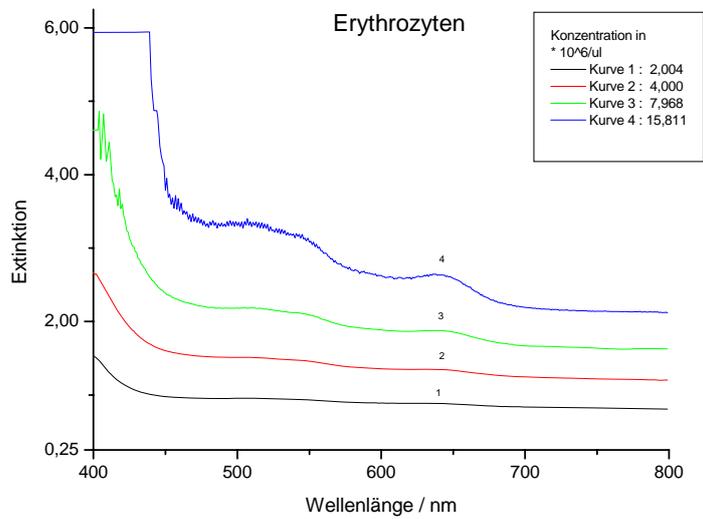


Abbildung 2-5 : Erythrozyten

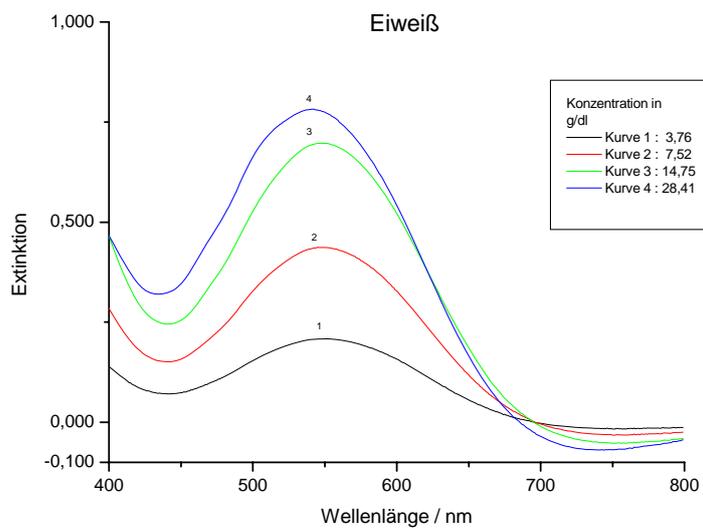


Abbildung 2-6 : Eiweiß

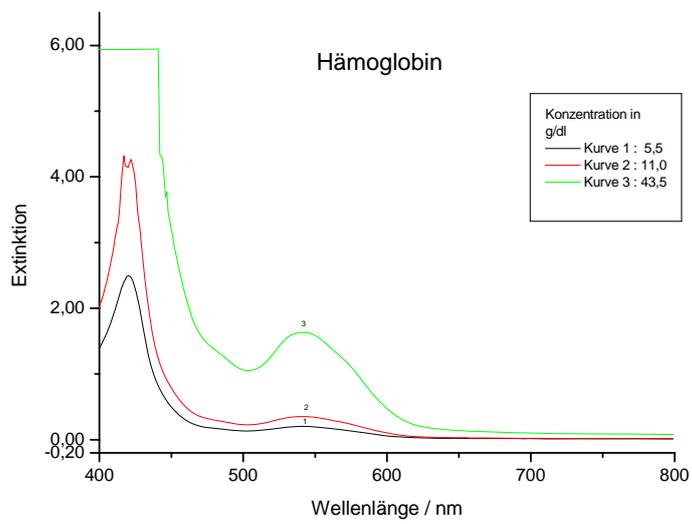


Abbildung 2-7 : Hämoglobin

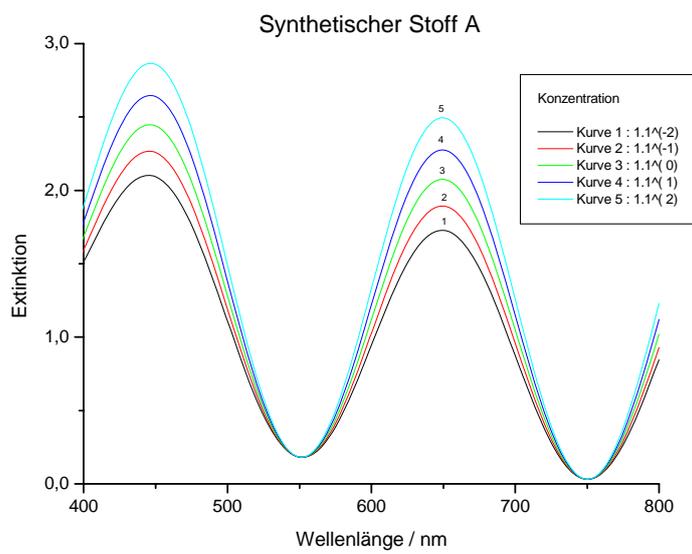


Abbildung 2-8 : Synthetisches Spektrum A

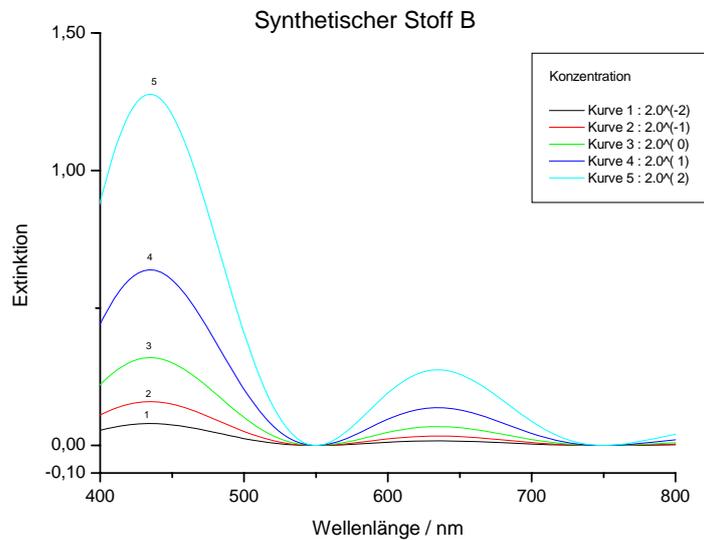


Abbildung 2-9 : Synthetisches Spektrum B

Es galt nun, Algorithmen zu entwickeln, die die bei der Messung einer Probe und ihres dazugehörigen Leerwertes anfallenden Daten weiterverarbeiten und jeweils einer der Scharen (Abbildung 2-1 bis Abbildung 2-9) möglichst „sicher“ zuordnen. „Sicher“ bedeutet, dass die Zuordnung mit einer hohen Trenngüte gegenüber der „zweitbesten“ Kurvenschar erfolgt und damit unempfindlich gegenüber messfehlerbedingten Fehlinterpretationen.

Prinzipiell ist darüber hinaus denkbar, dass der zusätzliche Informationsgehalt, den der Leerwert bietet, die Identifizierung erleichtert.

2.1 Der Analysand: Probe abzüglich Leerwert

Absorptionsspektren lassen sich mathematisch als Tupel³ reeller Zahlen modellieren. Daher diene zunächst das Spektrum der um ihren Leerwert bereinigten Testsubstanz als Eingabevektor (bereinigte Probe). Die folgende Abbildung zeigt die verwendeten Testsubstanzen 1.-7. nach Differenzbildung mit dem dazugehörigen

³ Zu den Begriffen Tupel und Vektor : Ein Tupel stellt eine Zusammenfassung voneinander verschiedener oder gleicher mathematischer Objekte dar. Ein Punkt des dreidimensionalen Raumes ist beispielsweise ein 3-Tupel reeller Zahlen. Sind auf diesen gleichartigen Komponenten des Tupels die üblichen algebraischen Operationen (Addition, Multiplikation) etc. möglich, spricht man präziser von einem Vektor. Zur genauen Definition eines Vektorraumes vergleiche (12).

Leerwert. Sie lagen aus einer anderen Messreihe in physiologischen Konzentrationen vor (siehe Abbildung 2-10) und sollten den jeweiligen Scharen möglichst fehlerfrei zugeordnet werden.

Wie bei der visuellen Zuordnung sollte auch die von den Algorithmen zu gebende Antwort lauten:

„Kurve Nr. 4 (Cholesterin) in Abbildung 2-10 gehört am ehesten zu der Kurvenschar in Abbildung 2-3 (Kurvenschar für Cholesterin)“.

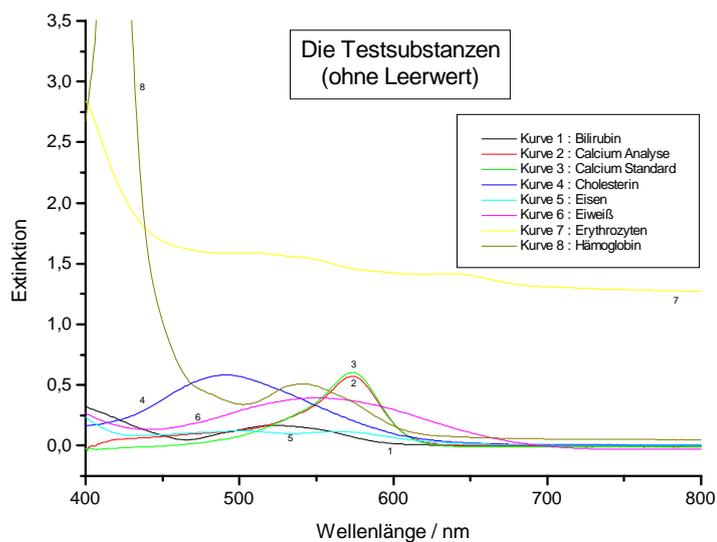


Abbildung 2-10 : Die Testsubstanzen nach Abzug der Leerwerte

Die meisten Spektren lassen sich intuitiv leicht zuordnen, da sie typische spektrale Charakteristika haben wie z.B. Eiweiß, andere hingegen wie z.B. Eisen schwerer.

2.2 Der Analysand: bereinigte Probe verkettet mit Leerwert

Zu prüfen war weiterhin, ob die Berücksichtigung der absoluten Extinktion der Leerwerte in Verbindung mit der Differenz Probe – Leerwert die Identifizierung erleichtert. Hierfür wurden eine Verkettung zwischen den Vektoren der bereinigten Probe und des Leerwertes gebildet.

Es werden dabei „neben“ den jeweiligen Differenzspektren aus Abbildung 2-1 bis Abbildung 2-9 die entsprechenden Leerwertspektren vorgehalten. Zur Identifikation werden dazu auch die Leerwerte (siehe linker Teil der Abbildung 2-11) der jeweiligen Tests (siehe rechter Teil der Abbildung 2-11) benötigt.

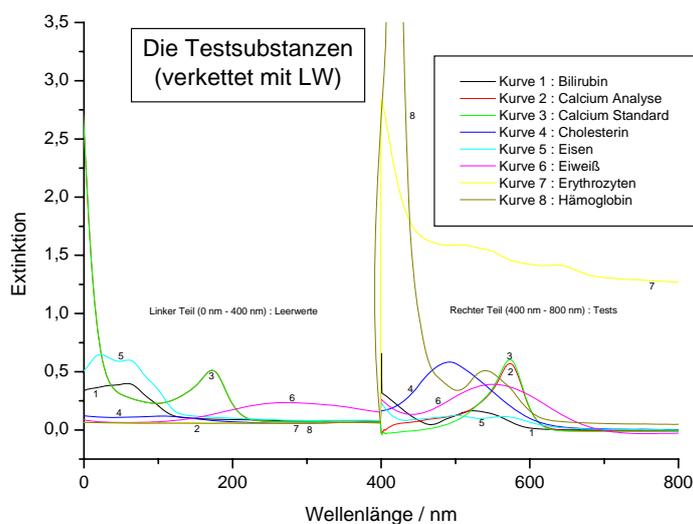


Abbildung 2-11 : Die Testsubstanzen verkettet mit dem Leerwert

Zur Vereinfachung der Darstellung wurden die Leerwerte um 400 nm nach links verschoben; auch die Speicherung findet so statt.

Mithilfe der zu entwickelnden Algorithmen sollten Angaben gemacht werden über die

1. Abweichung der Spektren der jeweiligen Einzelsubstanz von ihrer „ähnlichsten“ Schar
2. geschätzte Konzentration einer Substanz anhand des *gesamten* Spektrums
3. berechnete Konzentration der Einzelsubstanz unter Berücksichtigung lediglich der nach Hersteller empfohlenen Wellenlängen
4. Abweichungen von den anderen Kurvenscharen

In einem weiteren Schritt sollten dann Trenngüte und Aussagefähigkeit der verschiedenen Algorithmen unter Variation der folgenden Parameter miteinander verglichen werden:

1. Wellenlängenbereich
2. Wellenlängendichte

Damit sollte der Einfluss einer Reduktion des Wellenlängenbereiches oder der spektralen Auflösung auf die Identifikationsgenauigkeit geprüft werden.

3 Physikalisch-chemische Grundlagen

Bestimmte chemische Substanzen, bzw. einzelne Atomgruppen oder Bindungen sind in der Lage, Licht bestimmter Wellenlängen zu absorbieren. Dies beruht im Wesentlichen auf der Wechselwirkung des Lichts als elektromagnetischer Strahlung mit Atomen und Molekülen, so genannte induzierte Absorption(14). Wie viel des eingestrahlteten Lichtes nun von einer Probe absorbiert wird, hängt im Wesentlichen von der eingesetzten Substanz und von der Zahl der bestrahlten Moleküle ab. Je nach Wellenlänge des eingestrahlteten Lichtes ergibt sich in der Regel eine andere Absorption, so dass das Absorptionsspektrum, also die Absorption in Abhängigkeit von der Wellenlänge, für die Probe charakteristisch sein kann. Das Lösungsmittel selbst erzeugt dabei ebenfalls eine Absorption, die als so genannter Leerwert von der eigentlichen Absorption abgezogen werden muss, siehe Abbildung 3-1. Zum additiven Beitrag der Extinktion von Lösungsmittel und Substanz zur Gesamtextinktion vergleiche (26).

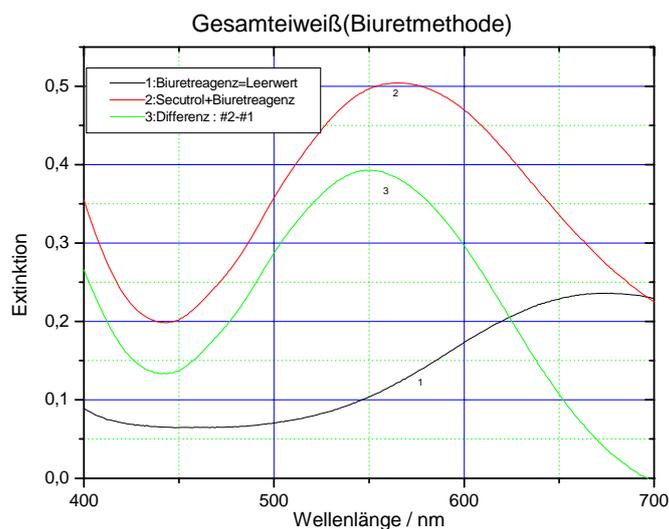


Abbildung 3-1 : Eiweißspektrum mit LW und Differenz

3.1 Das Lambert-Beersche Gesetz

Der funktionale Zusammenhang zwischen der Wellenlänge des eingestrahnten Lichtes und der Absorption wird beschrieben durch das, unabhängig von Pierre Bouguer 1729 zunächst experimentell entdeckte, von Heinrich Lambert 1760 und August Beer 1852 theoretisch begründete Lambert-Beersche Gesetz (25).

Aus einfachen Annahmen über das Absorptionsverhalten einzelner gleichverteilter Moleküle in Lösungen lässt sich eine Differentialgleichung aufstellen (3), deren Lösung das Lambert-Beersches Gesetz darstellt (9). Dieses hält einer empirischen Prüfung stand.

„Lambert-Beersches Gesetz“ :

$$E(\lambda) := - \ln \frac{I(\lambda)}{I_0(\lambda)} = \varepsilon(\lambda, x) \cdot c \cdot d$$

$$E(\lambda) : [\lambda_1; \lambda_2] \rightarrow]0; +\infty[^4$$

Dabei bezeichnen die prinzipiell messbaren bzw. abgeleiteten Parameter :

- x : im Lösungsmittel gelöste Substanz
- λ : Wellenlänge [nm]
- d : Durchmesser [cm] der Meßküvette
- $I(\lambda)$: durchgelassene Lichtintensität [lux] der Wellenlänge λ
- $I_0(\lambda)$: eingestrahlte Lichtintensität [lux] der Wellenlänge λ
- ln : „Logarithmus naturalis“ zur Basis „e“ = 2,71..., d.h. $\ln(e)=1$
- λ_1, λ_2 : Wellenlängen, zwischen denen die Absorption gemessen wird
- ε : Extinktionskoeffizient [l/mol/cm], ε ist abhängig von der Substanz x und der betrachteten Wellenlänge λ

⁴ Zum Wertebereich der Absorption und Extinktion: Aufgrund physikalisch-chemischer Gegebenheiten kann I alle Werte des Intervalls $]0; I_0[$ annehmen; dabei bedeutet der Zustand $I = I_0$ komplette „Durchlässigkeit“ der Probe, d.h. $A=0$, was aufgrund von natürlichen Verunreinigungen in natura nicht auftritt und $I=0$ komplette Absorption des eingestrahnten Lichtes, d.h. $A=1$, was ebenfalls nicht eintritt. Wegen der strengen Monotonie der Logarithmusfunktion nimmt dann der Term $E = - \ln(I/I_0)$ die Werte in $\mathfrak{R}^+ :=]0; +\infty[$ an.

$E(\lambda)$: dimensionslos, so genannte „Extinktion“ also Auslöschung des eingestrahlten Lichtes der Wellenlänge λ durch die Substanz x in der Konzentration c

und der unbekannte Parameter

c : Konzentration der Substanz x [mol/l] in der Meßküvette

In diesem Zusammenhang werden in der Literatur auch die Begriffe

T := I/I_0 , „Transmission“, also Durchlässigkeit und

A := $1-T$, „Absorption“

verwendet.

3.1.1 Der Leerwert

Der so genannte Leerwert bezeichnet den Anteil der Gesamtabsorption, welche durch die Küvette selbst und die Lösungsmittel *ohne* die Testsubstanz verursacht wird. Denn da die Absorptionsmaxima der Testsubstanz x häufig im durch das Spektralphotometer nicht messbaren Bereich des Lichtes liegen, ist eine Verschiebung in den spektralphotometrisch erfassbaren Bereich des Lichtes erforderlich. Diese wird durch chemische Reaktion mit einem Reagenz erzeugt, welche ein hinsichtlich Stoffmenge proportionales Produkt erzeugt. Da Reagenz und Lösungsmittel selber absorbieren, muss ihre Absorption zusammen mit der Absorption der Küvette als so genannter „Leerwert“ von den Spektren von Testsubstanz (und ggfls. Standard) abgezogen werden. Da die Absorptionsverstärkung durch die (normierte) Küvette sowohl in die Absorptionsmessung des Leerwertes als auch der Testsubstanz als Summand eingeht und nur die Differenz Testsubstanz - Leerwert in die Konzentrations-Bestimmung eingeht, braucht sie – einen Küvettenwechsel ausgenommen – nicht berücksichtigt zu werden.

Das Spektrum der Küvette selbst – siehe Abbildung 3-2 – und der additive Zusammenhang erklären, warum Wellenlängen um 300 nm für Messungen nicht geeignet sind: Die Absorption der leeren Küvette ist in diesem Wellenlängenbereich

so hoch, dass die betragsmäßig deutlich geringeren Absorptionen des Küvetteninhalts aufgrund der Überlagerung nicht mehr messbar sind.

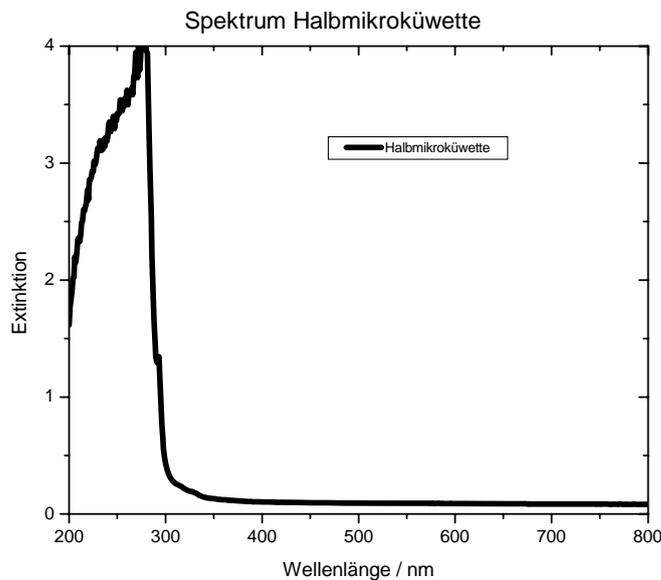


Abbildung 3-2 : Spektrum Halbmikro-Küvette

Nach den oben genannten Ausführungen ist die Absorption von gelöster Testsubstanz an jeder Wellenlänge größer als die Absorption des Leerwertes, i.e. Lösungsmittels. Damit sind – um Fehlinterpretationen durch Messfehler bei einzelnen Wellenlängen zu vermeiden – die Integrale unter Testsubstanz und Leerwert geeignet, zwischen den beiden Proben zu unterscheiden. Wir nutzen diesen Zusammenhang später bei der Implementierung in der EDV.

Regulär ist die Extinktion des Leerwertes von der unbereinigten Probe unabhängig, da gegen die Reagenzmischung als Leerwert bzw. gegen destilliertes Wasser gemessen wird. Von den untersuchten Substanzen sind lediglich die Leerwerte bei der Bilirubin- und der Eisenmessung von den jeweiligen Bilirubin- bzw. Eisenkonzentrationen abhängig.

3.1.2 Quantitative Bestimmung von Konzentrationen

Der im Lambert-Beerschen Gesetz genannte Zusammenhang wird traditionell auf zwei verschiedene Arten verwendet (16):

- Es wird bei einer festen Wellenlänge λ die Extinktion E gemessen und unter Kenntnis des Extinktionskoeffizienten ε die Konzentration c bestimmt nach Formel 3.1.
- Ohne Zugriff auf den bekannten Extinktionskoeffizienten ε wird die Proportionalität $E \sim c$ ausgenutzt: es wird die Meßküvette mit unbekannter Konzentration c_1 und Extinktion E_1 bei fester Wellenlänge λ gegen einen so genannten Standard mit bekannter Konzentration c_2 und bekannter Extinktion E_2 gemessen (ebenfalls bei Wellenlänge λ). Die gefragte Konzentration c_1 wird anhand $\frac{c_1}{c_2} = \frac{E_1}{E_2}$ ermittelt. Dies ist z.B. das übliche Verfahren bei der Bestimmung von Ca^{2+} -Konzentrationen. Vorteil dieses Verfahrens ist darüber hinaus die Unabhängigkeit von der Basis des gewählten Logarithmus, da sich dieser in Form einer Proportionalitätskonstante wegekürzt. Als Nachteil ist jedoch die jeweilig zusätzlich erforderliche Absorptionsmessung des Standards zu erwähnen.

Die jeweilig geeigneten Wellenlängen und Umrechnungsfaktoren für die konkrete Konzentrationsbestimmung sind den Anleitungen der Hersteller der Reagenzien zu entnehmen.

3.1.3 Die Extinktionsfunktion $E_\lambda(c)$

Die Extinktion bei der Wellenlänge λ in Abhängigkeit von der Konzentration c , kurz $E = E_\lambda(c)$ hat folgende Eigenschaften :

- a) E ist an jeder Wellenlänge monoton steigend in $c \in \mathfrak{R}^{\geq 0}$. Dies modelliert die Erfahrung, dass bei steigenden Konzentrationen in der Regel auch die Absorptionen, also die „Trübung“ der Probe zunehmen und folgt aus der Additivitätseigenschaft der Extinktion.
- b) E ist an mindestens einer Wellenlänge λ und damit wegen der Stetigkeit der Funktion E approximativ auch in der Umgebung von λ linear. Dies modelliert die empirische Erfahrung, dass sich einige E_λ – nämlich die Extinktionsfunktionen an den so genannten Messwellenlängen – strikt nach dem Lambert-Beerschen Gesetz verhalten. Sie eignen sich gut zur Konzentrationsmessung, so z.B. für Calcium im Wellenlängenbereich von 560 bis 590 nm. Theoretisch besteht auch bei den anderen Wellenlängen Linearität.
- c) E ist auf \mathfrak{R}^+ für alle relevanten Werte der Wellenlänge mindestens einmal stetig differenzierbar. Obwohl es sich bei der Absorption um ein per se „unstetiges“ Quantenphänomen handelt, kann dies als vereinfachende, empirisch gesicherte Annahme betrachtet werden.

Weitere Eigenschaften von E sind aus den physikalisch-chemischen Gegebenheiten nicht prima vista ableitbar.

3.2 Das Spektrum als Charakteristikum einer Probe: der „Fingerprint“

Die Tatsache, dass je nach bestrahltem Molekül an bestimmten Wellenlängen Auslöschungen des einstrahlenden Lichtes stattfinden, wird in der analytischen Chemie seit langem genutzt:

Im für das menschliche Auge sichtbaren Spektralbereich des Lichtes kann der erfahrene Chemiker mit den eigenen Augen als „Spektroskop“ Substanzen anhand ihrer Farbmischung identifizieren.

Aber auch im infraroten Bereich des Lichtes kann die Identifizierung erfolgen. „Da die im Spektrum enthaltenen Schwingungsbanden charakteristisch für bestimmte

Substanzen sind, lassen sich diese durch IR-Spektren identifizieren, wie durch einen Fingerabdruck“(19)(3). Bisher unbekannte chemische Verbindungen lassen sich durch Einsatz neuronaler Netze anhand der charakteristischen Spektren einzelner Atomgruppen im Einzelfall identifizieren (15). Diese Analyse erfolgt aber bisher weitgehend noch durch hierfür geschultes Personal.

Durch den enormen Zuwachs an Rechengeschwindigkeit, Speicherkapazität und Vernetzung bisher getrennter Datenbestände in den letzten zwanzig Jahren bei gleichzeitig fallenden Preisen stehen nunmehr auch kleineren Institutionen Hochleistungsrechner zur Verfügung.

Man bemüht sich daher seit einiger Zeit Algorithmen zu entwickeln, die dieses so genannte Expertenwissen nachbilden. „So gesehen bildet die Kombination von Struktur/Spektren-Tabellen mit den Augen eines erfahrenen Spektroskopikers ein Vorhersagesystem, an dessen Interpretationsfähigkeiten sich computergestützte Systeme messen lassen müssen.“(24)

4 Ergebnisse

Zur Verkürzung der späteren Notationen und zur Implementierung in die EDV ist eine angepasste mathematische Formalisierung erforderlich.

4.1 Mathematische Formalisierung

4.1.1 Das Spektrum einer Probe als Vektor des $(\mathfrak{R}^+)^n$

Die Funktion $E(\lambda)$ liegt in der Praxis nicht mit dem Definitionsbereich $[\lambda_1; \lambda_2] \subset \mathfrak{R}^+$ vor, sondern je nach Messintervall und Abtastgenauigkeit des Spektralphotometers als ein Tupel von n Messwerten. In den hier betrachteten Messungen wird die Extinktion gemessen von $\lambda_1 = 400$ nm bis $\lambda_2 = 800$ nm mit einem Abstand von $\Delta\lambda = 1$ nm, also $n = 1 + \frac{\lambda_2 - \lambda_1}{\Delta\lambda} = 401$.

Man kann deshalb das Spektrum einer Probe P auch auffassen als Vektor bzw. Punkt

$$\vec{x}_p := (x_1, \dots, x_i, \dots, x_n)^T \in (\mathfrak{R}^+)^n \text{ mit } x_i = E(\lambda_1 + (i-1) \cdot \Delta\lambda)$$

4.1.2 Spektren einer Substanz als Parameterfunktion der Konzentration

In Abhängigkeit von der Konzentration c ergeben sich verschiedene Spektren für jedes $c \in \mathfrak{R}^{\geq 0}$, so dass die Spektren einer Substanz formal als Parameterfunktion oder anschaulich als „Parameterzug“ zu betrachten sind:

$$\vec{F}^* = \vec{F}^*(c) : \mathfrak{R}^{\geq 0} \rightarrow (\mathfrak{R}^+)^n.$$

Von dieser Funktion sind aber nur n Stützstellen bekannt.

Liegen jetzt m Spektren \bar{x}_i einer einzigen Referenzsubstanz bei den Konzentrationen $c_i \in \mathfrak{R}^{\geq 0}$, $i = 1 \dots m$ aufsteigend geordnet, also $c_i < c_{i+1}$, vor, so stellt sich die Frage, welchen Punkt im $(\mathfrak{R}^+)^n$ das Spektrum bei einer beliebigen Konzentration $c \in \mathfrak{R}^{\geq 0}$ annimmt.

Es ist also gefragt nach einer Funktion

$$\bar{F} : \mathfrak{R}^{\geq 0} \rightarrow (\mathfrak{R}^+)^n \text{ mit } \bar{F}(c_i) := (f_1(c_i), \dots, f_n(c_i))^T = \bar{x}_i, i = 1 \dots m$$

bzw.

$$\bar{F} : \mathfrak{R}^{\geq 0} \rightarrow (\mathfrak{R}^+)^n \text{ mit } \bar{F}(c_i) := (f_1(c_i), \dots, f_n(c_i))^T \approx \bar{x}_i, i = 1 \dots m,$$

welche den realen Verlauf von \bar{F}^* unter Berücksichtigung der bekannten Eigenschaften der Extinktionsfunktion möglichst genau wiedergibt.

Im Hinblick auf das Folgende sollte darüber hinaus eine „Abstandsmessung“ von \bar{F} zu einem beliebigen Punkt des $(\mathfrak{R}^+)^n$ mit vertretbarem numerischen Aufwand möglich sein. Schließlich soll die Auswertung der Spektren durch die EDV in Echtzeit geschehen.

Die Bedeutung der näherungsweise Gleichheit „ \approx “ an den Stützstellen wird später präzisiert.

4.2 Abstandsmessungen im $(\mathfrak{R}^+)^n$

Im Hinblick auf die Aufgabenstellung ist nun danach gefragt, ob ein Punkt in einer – kleinen – „Umgebung“ einer der Punkte des Parameterzuges P liegt, also nach dem Abstand eines Punktes von einem Parameterzug P. Liegt also ein Punkt „nahe“ bei einem Parameterzug P, so ist anschaulich dieser Punkt als Testsubstanz der Referenzsubstanz als Parameterzug „ähnlich“.

Sei also jetzt $\bar{F}(t)$ eine – beliebige – Parameterdarstellung eines Parameterzuges P. Die Frage nach dem minimalen Abstand von $P \subseteq (\mathfrak{R}^+)^n$ zu einem Punkt $\bar{x} \in (\mathfrak{R}^+)^n$ lautet demnach

Abstand $(\bar{x}, \bar{F}(c)) \rightarrow \text{Minimum} !$

Prinzipiell ist als Abstandsmessung im $(\mathfrak{R}^+)^n$ natürlich auch jede Norm zu verwenden. (Zur mathematischen Definition einer Norm vergleiche etwa(2)). Die Maximumnorm etwa würde sich in der Praxis aber insofern problematisch gestalten, als dass ein großer „Ausreißer“ im Sinne einer Extinktionsfehlmessung den Abstand bestimmen würde. Die so genannte Summennorm als der andere Extremfall der möglichen Normen in endlich-dimensionalen Vektorräumen lässt sich mathematisch als Betragsfunktion nur umständlich handhaben. Da in endlichen Vektorräumen alle Normen äquivalent sind (2), entscheidet man sich deshalb für die *euklidische* Norm.

Diese lässt sich durch das kanonische Skalarprodukt darstellen und ist deshalb für die weitere mathematische Behandlung günstig. Also:

Abstand $(\bar{x}, \bar{F}(c)) := \sqrt{(\bar{x} - \bar{F}(c))^2} \rightarrow \text{Minimum} !$

Dieses Problem lässt sich mathematisch nach der Extremalwertrechnung in der Regel auf die Bestimmung der Nullstellen der Funktion $d\sqrt{(\bar{x} - \bar{F}(c))^2} / dc$ zurückführen

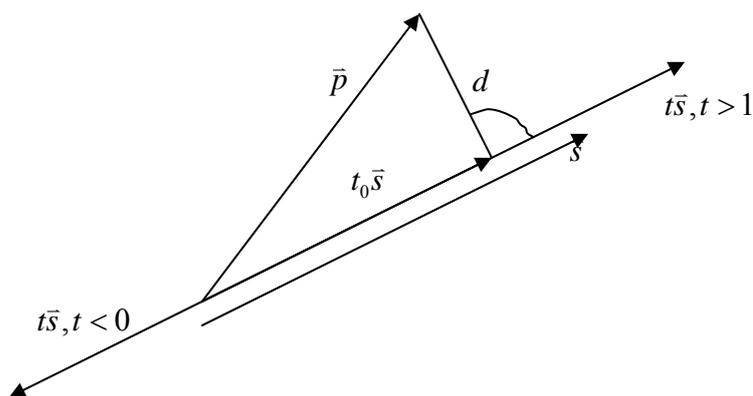
oder numerisch durch „Durchprobieren“ jedes möglichen Wertes des Parameters t und Bestimmung des Ausdruckes $\sqrt{(\bar{x} - \bar{F}(c))^2}$.

Wie weiter unten ausgeführt, wird der Parameterzug $\bar{F}(c)$ jedoch durch Polygonzüge interpoliert bzw. durch eine Gerade approximiert. Wir betrachten daher nur die folgenden Abstandsmessungen.

4.2.1 Punkt zu „beliebigem“ Polygon : Euklidischer Abstand

Da jedes Polygon aus Strecken bzw. einem Strahl besteht, reduziert sich das Problem auf die Berechnung des minimalen Abstandes eines Punktes \bar{x} von einer durch den Anfangspunkt \bar{x}_1 und den Endpunkt (Zweitpunkt) \bar{x}_2 gegebenen Strecke (Strahl). Ohne Beschränkung der Allgemeinheit sei $\bar{s} = \bar{x}_2 - \bar{x}_1$ und $\bar{p} = \bar{x} - \bar{x}_1$, also Verschiebung des Punktes \bar{x}_1 auf $\bar{0}$.

Die entsprechende Parameterdarstellung der Strecke lautet dann $t\bar{s}$, $t \in [0; 1]$ für die Strecke und $t \in [0; \infty[$ für den Strahl, siehe Skizze:



Die Frage ist also, für welches t der Term $\sqrt{(\bar{p} - t_0\bar{s})^2}$ ein Minimum annimmt.

Aufgrund der streng steigenden Monotonie der Wurzelfunktion ist dies genau dort der Fall, wo die quadratische Funktion $g(t) = (\bar{p} - t\bar{s})^2 = t^2\bar{s}^2 - 2t\bar{p}\bar{s} + \bar{p}^2$ auf $[0; 1]$ bzw.

$[0; \infty[$ ein Minimum annimmt. Wegen $\bar{s}^2 > 0$ für $\bar{s} \neq \bar{0}$ handelt es sich um eine nach

oben geöffnete Parabel (mit maximal einer Nullstelle wie die Berechnung der Diskriminanten unter Verwendung der Cauchy-Schwarz'schen Ungleichung zeigt). Die Parabel nimmt deshalb ein Minimum auf \mathfrak{R} am Scheitelpunkt t_0 mit $g'(t_0)=0$ an, also

$$g'(t_0) = 2t_0\bar{s}^2 - 2\bar{p}\bar{s} = 0 \Leftrightarrow$$

$$t_0 = \frac{\bar{p}\bar{s}}{\bar{s}^2} \quad ^5$$

Wie man sich anschaulich⁶ klar macht, nimmt $g(t)$ ein Minimum an im Fall

der Strecke auf $[0;1]$ für :

$$t_0 < 0 \quad \text{bei} \quad t = 0 \quad \text{mit} \quad \sqrt{g(t)} = |\bar{p}|$$

$$t_0 > 1 \quad \text{bei} \quad t = 1 \quad \text{mit} \quad \sqrt{g(t)} = |\bar{p} - \bar{s}|$$

$$t_0 \in [0;1] \quad \text{bei} \quad t = \frac{\bar{p}\bar{s}}{\bar{s}^2} \quad \text{mit} \quad \sqrt{g(t)} = \sqrt{(\bar{p} - t_0\bar{s})^2}$$

des Strahles auf $[0;\infty[$ für :

$$t_0 < 0 \quad \text{bei} \quad t = 0 \quad \text{mit} \quad \sqrt{g(t)} = |\bar{p}|$$

$$t_0 \geq 0 \quad \text{bei} \quad t = \frac{\bar{p}\bar{s}}{\bar{s}^2} \quad \text{mit} \quad \sqrt{g(t)} = \sqrt{(\bar{p} - t_0\bar{s})^2}$$

Für spätere Zwecke in der Programmerstellung erhält man durch Einsetzen von t_0 in den mittleren Summanden:

$$\sqrt{(\bar{p} - t_0\bar{s})^2} = \sqrt{\bar{p}^2 - 2t_0\bar{s}\bar{p} + t_0^2\bar{s}^2} = \sqrt{\bar{p}^2 - t_0^2\bar{s}^2}$$

⁵ Geometrisch ist t_0 die anteilige Länge der Projektion von p auf s , im Modell ist $t_0 = (c - c_{i-1}) / (c_i - c_{i-1})$

⁶ Die von der Anschauung unabhängige Herleitung der o.g. Beziehungen ohne Rückgriff auf den \mathfrak{R}^2 sind vom Autor durchgeführt und können bei Bedarf abgerufen werden.

Es steht uns damit ein numerisch relativ effizientes Verfahren zur Verfügung, um den euklidischen Abstand eines Punktes von einer Strecke bzw. einem Strahl zu bestimmen.

Das Verfahren zur Abstandsbestimmung eines Polygons P mit den Stützstellen \bar{x}_i , $i = 1 \dots m$ von einem Punkt \bar{x} lautet dann:

Beginn Algorithmus

Durchlaufe i von 1 bis $m-1$

Bestimme gemäß 4.2.1 den Abstand \bar{x} zu der Strecke $\bar{x}_{i-1} - \bar{x}_i$;

Merke den Abstand als d_i

Ende Durchlaufe i ;

Bestimme gemäß der eben vorgestellten Abstandsmessung den Abstand \bar{x} zu dem Strahl $\bar{x}_{m-1} - \bar{x}_m \rightarrow$;

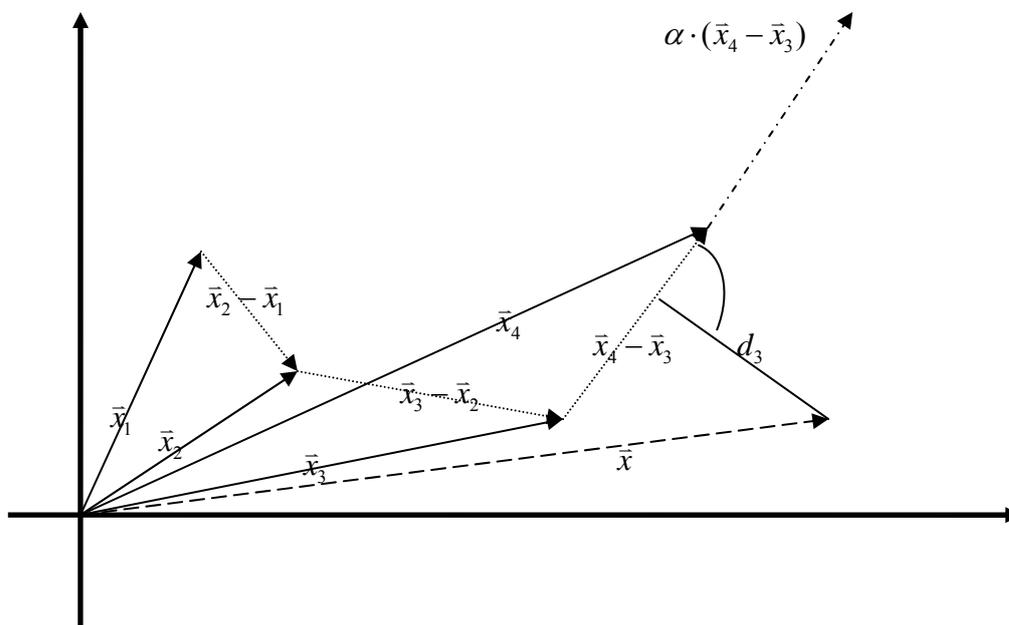
Merke den Abstand als d_m ;

Ermittle das Minimum von $d_1 \dots d_m$ als d ;

d ist der Abstand des Punktes \bar{x} zum Polygon P

Ende Algorithmus.

Skizze :



4.2.2 Punkt zu „monotonem“ Polygon : Abstand als Bogenmaß

Liegt jetzt ein Polygon P mit den Stützstellen \bar{x}_i , $i = 1 \dots m$ mit der Eigenschaft $|\bar{x}_i| < |\bar{x}_k|$ für $i < k$ – dies ist nach den Eigenschaften der Extinktionsfunktion der Fall – vor, so lässt sich dem Punkt \bar{x} *eindeutig* eine Strecke $\bar{x}_{i-1} - \bar{x}_i$ bzw. ein Strahl $\bar{x}_{m-1} - \bar{x}_m \rightarrow$ zuordnen mit

$$|\bar{x}_{i-1}| \leq |\bar{x}| < |\bar{x}_i| \text{ für ein } i \text{ aus } 1.. m-1 \text{ bzw. } |\bar{x}_{m-1}| \leq |\bar{x}|.^7$$

Wie oben sei $\bar{x}_0 := \vec{0}$. Sei also i – eindeutig – so gewählt, dass $|\bar{x}_{i-1}| \leq |\bar{x}| < |\bar{x}_i|$.

Man fragt also nach $\alpha \in [0; 1[$ mit $|\bar{x}| = |\bar{x}_{i-1} + \alpha \cdot (\bar{x}_i - \bar{x}_{i-1})|$. Wie im Anhang näher ausgeführt, erhalten wir genau eine Lösung für α , nämlich

$$\alpha_0 = \frac{-\bar{x}_{i-1}(\bar{x}_i - \bar{x}_{i-1})}{(\bar{x}_i - \bar{x}_{i-1})^2} + \sqrt{\frac{r^2 - \bar{s}^2}{(\bar{x}_i - \bar{x}_{i-1})^2}} \text{ mit}$$

$$\bar{s}^2 = \bar{x}_{i-1}^2 - \frac{(\bar{x}_{i-1}(\bar{x}_i - \bar{x}_{i-1}))^2}{(\bar{x}_i - \bar{x}_{i-1})^2} \text{ und}$$

$$r^2 = \bar{x}^2$$

Als Abstand des Punktes $\bar{x}_{i-1} + \alpha_0 \cdot (\bar{x}_i - \bar{x}_{i-1})$ vom dem Punkt \bar{x} bietet sich wegen der gleichen Beträge der Abstand auf einer gedachten Kugel mit dem Radius $|\bar{x}| = r$ an. Da der Winkel zwischen zwei Vektoren invariant gegenüber Drehungen der von den Vektoren aufgespannten Ebene ist, erhalten wir – nach $\vec{a} \cdot \vec{b} = |\vec{a}| \cdot |\vec{b}| \cdot \cos \gamma$ mit γ als von den Vektoren \vec{a} und \vec{b} eingeschlossenen Winkel – für den Abstand (in Bogenmaß)

$$d = \arccos\left(\frac{(\bar{x}_{i-1} + \alpha_0(\bar{x}_i - \bar{x}_{i-1}))\bar{x}}{r^2}\right)$$

⁷ Der dem Punkt x am nächsten liegende Punkt auf dem Polygonzug muss nicht auf dieser Strecke liegen.

Das Verfahren zur Abstandsbestimmung eines Polygons P mit den Stützstellen \bar{x}_i , $i = 1 \dots m$ von einem Punkt \bar{x} lautet dann (auch hier ist definitionsgemäß $\bar{x}_0 = \bar{0}$):

Beginn Algorithmus

Durchlaufe i von 1 bis $m-1$ bis ($|\bar{x}_{i-1}| \leq |\bar{x}| < |\bar{x}_i|$) oder $|\bar{x}| \geq |\bar{x}_{m-1}|$;

Wenn $|\bar{x}_{i-1}| \leq |\bar{x}| < |\bar{x}_i|$ dann

Bestimme gemäß der eben vorgestellten Abstandsmessung den Abstand d von \bar{x} zu der Strecke $\bar{x}_{i-1} - \bar{x}_i$;

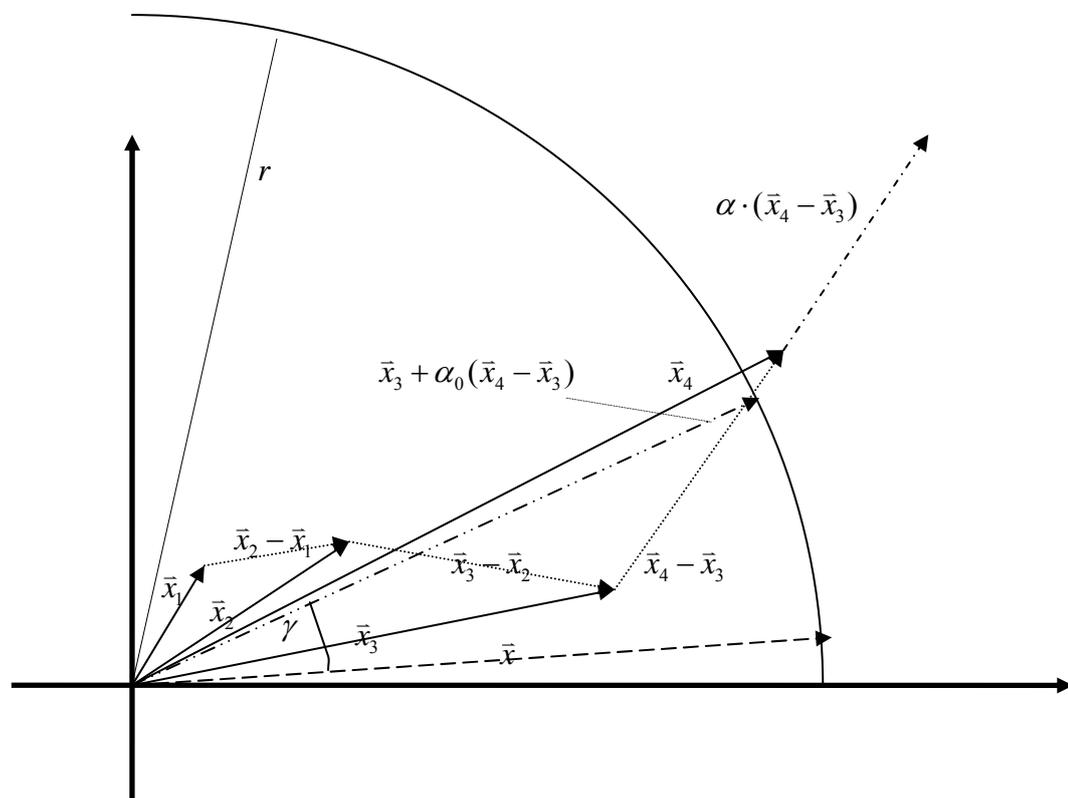
sonst

Bestimme gemäß der eben vorgestellten Abstandsmessung den Abstand d von \bar{x} zu dem Strahl $\bar{x}_{m-1} - \bar{x}_m \rightarrow$;

d ist der Abstand des Punktes \bar{x} zum Polygon P

Ende Algorithmus.

Skizze:



Günstig an diesem Verfahren ist, dass d nur Werte zwischen 0 und π annimmt, also beschränkt ist und so eine „prozentuale“ Distanzangabe möglich ist.

4.2.3 Punkt zu „linearem“ Polygon : einfaches Skalarprodukt

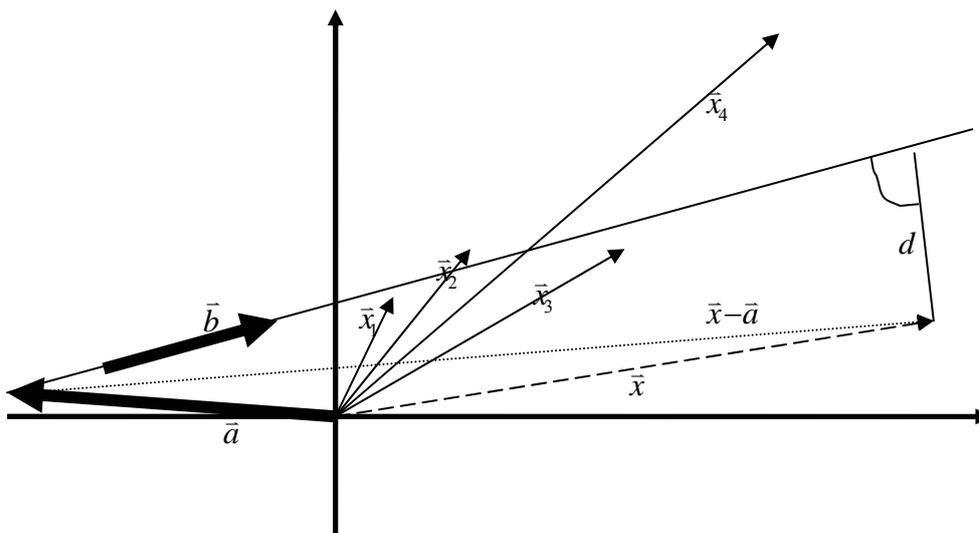
Liegt jetzt ein lineares Polygon P, also eine Gerade vor, so lautet die Frage:
Was ist der Abstand des Punktes \vec{x} von der Gerade $\vec{a} + c \cdot \vec{b}$? Analog zu der
Teilproblemstellung oben erhalten wir mit

$$t_0 = \frac{(\vec{x} - \vec{a}) \cdot \vec{b}}{\vec{b} \cdot \vec{b}}$$

für den Abstand

$$d = \sqrt{(\vec{x} - \vec{a})^2 - t_0^2 \vec{b}^2}$$

Skizze:



Die Angabe eines Algorithmus entfällt aufgrund der Einfachheit der Berechnung.

Welche mathematischen Verfahren und Funktionen eignen sich nun zur Modellierung der Funktion \vec{F}^* bzw. ihrer Komponentenfunktionen f_k^* ?

Im Folgenden bezeichne „Stützstelle“ die Konzentration, bei der eine Absorption bekannt ist, und „Stützwert“ ein Tupel $(c, \vec{F}^*(c))$.

4.3 Die Parameterfunktion

4.3.1 Darstellung durch Interpolation

Unter Berücksichtigung der o.g. Eigenschaften der Extinktionsfunktion

$$\bar{F} : \mathfrak{R}^{\geq 0} \rightarrow (\mathfrak{R}^+)^n \text{ mit}$$

$$\bar{F}(c_i) := (f_1(c_i), \dots, f_n(c_i))^T = \bar{x}_i, i = 1 \dots m$$

versucht man, durch Auswahl einer Funktion aus einer bestimmten Funktionenmenge mit günstigen numerischen Eigenschaften diejenige zu ermitteln, die genau an den Stützstellen $c_1 \dots c_m$ die bekannten Absorptionen annimmt. Von solchen Funktionenmengen stehen im Wesentlichen die trigonometrischen Funktionen und Polynomfunktionen zur Verfügung sowie stückweise definierte, insbesondere lineare Funktionen. Wie oben ausgeführt verhält sich jedoch die Absorption in Abhängigkeit von der Konzentration nicht periodisch, so dass nur Polynomfunktionen in Frage kommen:

Konkret führt die Interpolation durch ein Polynom grad $p_k = k$ bei m Stützstellen auf ein lineares Gleichungssystem (LGS) mit k Variablen:

- Für $k < m$ ist dieses LGS im Allgemeinen nicht lösbar.
- Für $k = m$ ist das LGS eindeutig lösbar ist; problematisch ist allerdings, dass Polynome vom Grad ≥ 2 ohne Zusatzbedingungen nicht zwingend monoton sind, so z.B. das Polynom $p_2(x) = x^2$. Damit wäre die für unsere Modellierung essentielle Eigenschaft der Monotonie möglicherweise nicht erfüllt.
- Für $k > m$ ist das LGS im Allgemeinen nicht ohne Zusatzbedingungen eindeutig lösbar. Es wäre zwar denkbar, durch die Forderung $p'_k \geq 0$ (Monotonie) mit $k = m+1$ diese Eigenschaft der Monotonie zu erzwingen. Meines Erachtens lässt sich aber dieser Ansatz nicht weiter mathematisch fortführen, weil eine Methode zur Erzeugung eines solchen Polynoms nicht bekannt ist.

Über der gesamten Menge $\mathfrak{R}^{\geq 0}$ definierte Polynome sind daher ungeeignet, selbst wenn die Eigenschaft der beliebigen Differenzierbarkeit kanonisch erfüllt wäre.

Es ist jedoch möglich, jeweils zwischen den Stützstellen eine Interpolation mit monoton steigenden Polynomen 1. Grades, also Geraden vorzunehmen⁸. Hierbei ist die Eigenschaft der Monotonie gesichert. Auf die Eigenschaft auch nur der einmaligen Differenzierbarkeit muss verzichtet werden: es ist leicht einzusehen, dass etwa die Verwendung von B-Splines – ein Interpolationsverfahren aus der numerischen Mathematik mit beliebig wählbarer Differenzierbarkeit an den Stützstellen – zur Sicherung der Differenzierbarkeit in Polynomen von höherem Grad als 1 mündet und damit die Monotonie nicht mehr gesichert wäre.

Wir wählen deshalb zur Interpolation die stückweise Verbindung der Stützpunkte durch Polynome 1. Grades; das im Folgenden „Interpolationsmethode“ genannte Verfahren ist Gegenstand eines der folgenden Abschnitte.

Ein unter Ausnutzung der Monotonie vereinfachter Lösungsansatz, im Folgenden „Winkelmethode“ genannt, ist Gegenstand eines der folgenden Abschnitte.

4.3.2 Darstellung durch Approximation

Verlangt man aus Gründen der Vereinfachung nicht, dass die Funktion \bar{F} die Stützstellen von \bar{F}^* genau annimmt, sondern sich lediglich den Stützstellen möglichst „gut annähert“, also

$$\bar{F}(c_i) = (f_1(c_i), \dots, f_n(c_i))^T \approx \bar{x}_i, \quad i = 1..m,$$

⁸ Man kann zeigen (vereinfachte Version des Approximationssatzes von Weierstraß), dass sich jede über einem Intervall stetige Funktion durch solche stückweise affinen Funktionen im Sinne gleichmäßiger Konvergenz annähern lässt (1).

dann besteht die Möglichkeit, in jeder Komponente f_k , $k \in [1 \dots n]$ eine Funktion anzugeben, die den Verlauf von f_k mit einer an den Stützstellen minimalen summierten quadratischen Abweichung wiedergibt, so genanntes „Least-square“-Problem (8). Wie später ausgeführt, bieten sich dafür Geraden an.

Eine Approximation mit Polynomen höheren Grades ist mit Hilfe der Extremalwertbestimmung (8) natürlich ebenfalls möglich, erhöht aber erheblich den numerischen Aufwand: so ist zunächst die Bestimmung der Approximationsfunktion erheblich aufwändiger. Aber auch die Abstandsmessung zu einem Punkt des $(\mathbb{R}^+)^n$ würde die Nullstellenbestimmung eines Polynoms des Grades $2 \cdot n - 1$ nach sich ziehen:

$$\sqrt{((p_{n,1}(c)-x_1)^2 + \dots + (p_{n,\lambda}(c)-x_\lambda)^2 \dots + (p_{n,401}(c)-x_{401})^2)} \rightarrow \text{Minimum} \Rightarrow$$

$$d(\sqrt{((p_{n,1}(c)-x_1)^2 + \dots + (p_{n,\lambda}(c)-x_\lambda)^2 \dots + (p_{n,401}(c)-x_{401})^2)})/dc = 0$$

Hierbei bezeichne $p_{n,\lambda}$ das Approximationspolynom vom Grad n an der Wellenlänge λ .

Gerade diese Abstandsbestimmungen sollen jedoch in Echtzeit durchgeführt werden. Schließlich ist wie oben ausgeführt auch hier bei Polynomen vom Grad > 1 die Monotonie nicht sichergestellt.

Darüber hinaus lassen sich Polynome 1. Grades günstig mit Mitteln der linearen Algebra beherrschen, was Vorteile bei der Behandlung zeitabhängiger Absorptionen und Analyse von Substanzmischungen ergibt.

Wir wählen deshalb als Approximationsfunktion eine Gerade; das im Folgenden „Approximationsmethode“ genannte Verfahren ist Gegenstand eines der folgenden Abschnitte.

4.4 Drei Lösungsansätze

Stehen nun r Referenzsubstanzen mit je k Konzentrationen zur Verfügung und eine Testsubstanz, so fassen wir die r Referenzsubstanzen als r Polygonzüge (mit k Eckpunkten) und die Testsubstanz als einen Punkt im $(\mathcal{R}^+)^n$ auf.

Die Frage nach der Identifikation der Testsubstanz lautet dann:

„Zu welchem der r Polygonzüge hat die Testsubstanz den geringsten *Abstand*?“

Der entsprechende Algorithmus lautet:

- (1) Lies zwei Spektren \vec{a} und \vec{b} ein;
- (2) Wenn $|\vec{a}| < |\vec{b}|$ dann nehme als Leerwert \vec{a} und als Messung \vec{b} , sonst umgekehrt (denn : die Probe ist insgesamt immer „trüber“ als der dazugehörige Leerwert);
- (3) Bilde die Differenz Testsubstanz := Messung – Leerwert;
- (4) Ermittle für die Testsubstanz nach den unten⁹ genannten alternativen Verfahren das „beste“ Polygon, i.e. die beste Referenzsubstanz;
- (5) Bestimme die Konzentration an den spezifischen Wellenlängen;
- (6) Gib den Namen der identifizierten Substanz und die ermittelte Konzentration aus.

Der Basisalgorithmus steht damit fest; wir führen lediglich noch aus, in welcher Form nach den verschiedenen Interpolations- bzw. Approximationsverfahren die Aufbereitung der Referenzspektren erfolgt und welche zusätzlichen Informationen aus den oben genannten Verfahren erzielbar sind.

⁹ Im Rahmen dieser Schrift werden alle drei Verfahren geprüft.

4.4.1 Interpolationsmethode (Lösungsansatz 1)

Es wird gefordert $\bar{F}(c_i) := (f_1(c_i), \dots, f_n(c_i))^T = \bar{x}_i$

Den Verlauf von F zwischen den Stützstellen c_i gibt man durch Geraden wieder. Unter Erhalt der ersten drei, oben genannten Eigenschaften der Extinktionsfunktion geben wir eine stückweise Approximation durch Polynome 1. Grades mittels folgender Definition an: Es sei

$$\bar{F}(c) := \bar{x}_{i-1} + \alpha \cdot (\bar{x}_i - \bar{x}_{i-1}) \text{ mit } \alpha := (c - c_{i-1}) / (c_i - c_{i-1}) \text{ für } c_{i-1} \leq c < c_i$$

$$\bar{F}(c) := \bar{x}_m + \alpha \cdot (\bar{x}_m - \bar{x}_{m-1}) \text{ mit } \alpha := (c - c_m) / (c_m - c_{m-1}) \text{ für } c \geq c_m$$

mit $\bar{x}_0 := \bar{0}$ und $c_0 := 0$.

Hierbei bezeichne $\bar{0}$ den Nullvektor, die Vektoroperationen sollen in üblicher Weise verstanden werden.

Man beachte, dass für die Fälle $c < c_1$ und $c > c_m$ Extrapolation vorliegt: man nimmt für das Spektrum bei $c = 0$ den Vektor $\bar{x}_0 = \bar{0}$ an. D.h., jede um den Leerwert bereinigte Probe hat bei der Konzentration $c = 0$ keine Absorption mehr. Weiter gehe ich davon aus, dass sich F oberhalb c_m genauso linear wie von \bar{x}_{m-1} bis \bar{x}_m verhält. Insofern ist also auch die Wahl des Faktors α für $c \geq c_m$ willkürlich. Dies ist jedoch kein inhaltlicher Fehler, da ja aufgrund der begrenzten Messwerte für diese Konzentrationen in der Tat keine Spektren mehr vorliegen. Es ist aber nötig, die Extrapolation anzubieten, da auch Substanzmessungen außerhalb der Stützstellen prinzipiell vorkommen können. Durch o.g. Definition erhalten wir anschaulich einen (an einer Seite ins Unendliche fortgesetzten) Polygonzug im $(\mathfrak{R}^+)^n$ mit Start im Nullpunkt und streng linearem Verlauf ab Punkt \bar{x}_m . Dieser Polygonzug besteht aus m Strecken und einem Strahl.

Für die weitere Behandlung ist allerdings die Parameterdarstellung der Funktion F zunächst unerheblich. Zum Durchlaufen des Polygonzuges ist letztendlich jede Funktion mit dem Wertebereich \mathfrak{R}^+ geeignet (so z.B. $2t$, t^2 u.v.m.). Insofern stellt die o.g. Definition nur eine Existenzaussage dar. *Wie genau* durch Veränderung des Parameters das Polygon durchlaufen wird, ist im Folgenden bedeutungslos. Die letztendliche Konzentrationsbestimmung wird nämlich über lediglich einige Wellenlängen vorgenommen. Die hier dargestellten Schritte dienen lediglich der Identifikation der Testsubstanz.

Die Distanzmessung erfolgt über den oben vorgestellten Algorithmus durch euklidische Abstandsmessung.

Als „Nebenprodukt“ fällt die anhand der *ganzen* Kurve geschätzte Konzentration c durch die Formel

$t_0 = \frac{\bar{p}\bar{s}}{\bar{s}^2}$ aus der euklidischen Abstandsmessung in Verbindung mit

$$\bar{F}(c) := \bar{x}_{i-1} + \alpha \cdot (\bar{x}_i - \bar{x}_{i-1}) \text{ mit } \alpha := (c - c_{i-1}) / (c_i - c_{i-1}) \text{ für } c_{i-1} \leq c < c_i$$

$$\bar{F}(c) := \bar{x}_m + \alpha \cdot (\bar{x}_m - \bar{x}_{m-1}) \text{ mit } \alpha := (c - c_m) / (c_m - c_{m-1}) \text{ für } c \geq c_m \text{ an.}$$

Die Bestimmung des Faktors c kann aber durchaus als Plausibilitätskontrolle dienen. Liegt nämlich die geschätzte Konzentration c „weit“ außerhalb der in der Natur vorkommenden Konzentrationen, d.h. $c \ll c_1$ oder $c \gg c_m$, so ist die Identifikation zumindest anzuzweifeln, Näheres dazu unten.

4.4.2 Winkelmethode (Lösungsansatz 2)

Auch hier wird gefordert $\bar{F}(c_i) := (f_1(c_i), \dots, f_n(c_i))^T = \bar{x}_i$.

Der Verlauf von F zwischen den Stützstellen c_i wird analog zur Interpolationsmethode durch Geraden wiedergegeben.

Nur die Distanzmessung erfolgt abweichend, nämlich über den in oben vorgestellten Algorithmus per Bogenmaß.

Auch hier fällt eine geschätzte Konzentration anhand von

$$\alpha_0 = \frac{-\bar{x}_{i-1}(\bar{x}_i - \bar{x}_{i-1})}{(\bar{x}_i - \bar{x}_{i-1})^2} + \sqrt{\frac{r^2 - \bar{s}^2}{(\bar{x}_i - \bar{x}_{i-1})^2}} \quad \text{in Verbindung mit}$$

$$\bar{F}(c) := \bar{x}_{i-1} + \alpha \cdot (\bar{x}_i - \bar{x}_{i-1}) \quad \text{mit } \alpha := (c - c_{i-1}) / (c_i - c_{i-1}) \quad \text{für } c_{i-1} \leq c < c_i$$

$$\bar{F}(c) := \bar{x}_m + \alpha \cdot (\bar{x}_m - \bar{x}_{m-1}) \quad \text{mit } \alpha := (c - c_m) / (c_m - c_{m-1}) \quad \text{für } c \geq c_m$$

an.

4.4.3 Approximationsmethode (Lösungsansatz 3)

Hier wird lediglich gefordert: $\bar{F}(c_i) := (f_1(c_i), \dots, f_n(c_i))^T \approx \bar{x}_i, i=1..m$

Nach der so genannten „Methode der kleinsten Quadrate (Least-Square-Verfahren)“

(8) lässt sich also für jedes f_k eine Näherungsfunktion g_k bzw. Konstanten $a_k, b_k \in \mathfrak{R}$ angeben mit

$$g_k(c) := a_k + c \cdot b_k \quad \text{mit}$$

$$f_k \approx g_k \quad \text{in dem Sinne, dass } \sum_{i=1}^m (f_k(c_i) - g_k(c_i))^2 \rightarrow \text{Minimum}$$

In Vektorschreibweise erhalten wir eine Gerade:

$$\bar{F}(c) \approx \bar{a} + c \cdot \bar{b}$$

Die Distanzmessung erfolgt über die kanonische, oben vorgestellte Formel als einfaches Skalarprodukt; eine geschätzte Konzentration ergibt sich durch

$$t_0 = \frac{(\bar{x} - \bar{a})\bar{b}}{\bar{b}^2}.$$

4.5 Numerische Anwendung der Teilverfahren auf die Spektren

Dem Hauptalgorithmus folgend wurden die drei verschiedenen Algorithmen in Turbo-Pascal implementiert.

4.5.1 Erkennung der Referenzsubstanzen

Zunächst wurden die Referenzsubstanzen selbst der Erkennungsanalyse unterworfen. Erwartungsgemäß erfolgte für alle Referenzsubstanzen in jeder vorliegenden Konzentration bei Interpolationsmethode (= Lösungsansatz 1) und Winkelmethode (= Lösungsansatz 2) die sichere Erkennung mit einer Distanz von 0 und der richtigen geschätzten Konzentration. Anschaulich liegen die Spektren nämlich *auf* dem Polygonzug.

Die Tabelle 4-1 zeigt ein typisches Ergebnis der Analyse einer mit 0,99 mg/dl angesetzten Bilirubinprobe (siehe auch Kurve 1 in Abbildung 2-1) unter Berücksichtigung des Wellenlängenbereiches von 400 nm bis 800 nm und einer Abtastweite von 1 nm durch die Interpolationsmethode.

| Nr. | Distanz | Konz. geschätzt | Konz. Minimum | Konz. Maximum | Referenzsubstanz |
|-----|---------|-----------------|---------------|---------------|------------------|
| 1 | 0.000 | 0.990 | 0.990 | 6.223 | Bilirubin |
| 7 | 0.901 | 0.901 | 5.511 | 43.480 | Haemoglobin |
| 9 | 1.215 | 0.697 | 0.250 | 4.000 | Bsynthetisch |
| 4 | 1.396 | 81.769 | 76.000 | 405.333 | Eisen |
| 3 | 1.660 | 26.496 | 75.000 | 566.667 | Cholesterin |
| 8 | 1.661 | 0.050 | 0.826 | 1.210 | Asynthetisch |
| 5 | 1.667 | 0.160 | 2.004 | 15.811 | Erythrozyten |
| 6 | 1.759 | 2.227 | 3.760 | 28.409 | Gesamteiweiss |
| 2 | 2.038 | 0.405 | 1.260 | 9.520 | Calcium |

Die beste Substanz ist Nr. 1 : Bilirubin mit
 GESCHAETZT **0.990** mg/dl
 GEMESSEN **1.173** mg/dl bei 546 nm im arith. Mittel
 // da die Distanz 0.000 minimal ist
 // nach Interpolationsmethode
 // unter Verwendung des Extinktions-
 // koeffizienten bei 546 nm
 //(Tabelle 2-1): 1.173 mg/dl=
 // 0.08626*13,6 mg/dl(konst.)

Tabelle 4-1 : Erkennung einer Bilirubinprobe von 0,99 mg/dl, sortiert nach Distanz, und Berechnung der Konzentration

Die Spalte „Nr.“ in Tabelle 4-1 enthält die laufende Nummer des ganz rechts stehenden Identifiers der Zeile, der „Referenzsubstanz“. Diese reichen von Bilirubin, Nr.1, bis Bsynthetisch, Nr. 9. „Distanz“ gibt den Abstand der Bilirubinprobe zu der ganz rechts stehenden Referenzsubstanz wieder. „Konz. geschätzt“ nennt die nach dem Interpolationsverfahren bezüglich der ganz rechts stehenden Referenzsubstanz geschätzte Konzentration. Die Spalten „Konz. Minimum“ und „Konz. Maximum“ bezeichnen *hier* die minimal und maximal vorliegenden Konzentrationen der jeweiligen Referenzsubstanz, siehe Tabelle 2-2.

Dieses Beispiel zeigt, dass die Berechnung der Konzentration anhand *einer* Extinktionswellenlänge ungenauer als die Berechnung der Konzentration anhand des gesamten Spektrums sein kann: die geschätzte Konzentration ist mit 0,990 mg/dl dem realen Wert von 0,99 mg/dl wesentlich näher als die bei einer Wellenlänge von 546 nm gemessene Konzentration von 1.173 mg/dl.

Die Tabelle 4-2 zeigt ein typisches Ergebnis der Analyse einer mit 150mg/dl angesetzten Cholesterinprobe (siehe Kurve 2 in Abbildung 2-3) unter Berücksichtigung des Wellenlängenbereiches von 470 nm bis 620 nm und einer Abtastweite von 1 nm durch die Winkelmethode¹⁰:

| Nr. | Distanz | Konz. geschätzt | Konz. Minimum | Konz. Maximum | Referenzsubstanz |
|-----|---------|-----------------|---------------|---------------|------------------|
| 3 | 0.000 | 150.000 | 75.000 | 566.667 | Cholesterin |
| 7 | 20.315 | 17.142 | 5.511 | 43.480 | Haemoglobin |
| 1 | 24.885 | 8.778 | 0.990 | 6.223 | Bilirubin |
| 5 | 26.279 | 0.940 | 2.004 | 15.811 | Erythrozyten |
| 6 | 30.810 | 9.911 | 3.760 | 28.409 | Gesamteiweiss |
| 8 | 35.990 | 0.384 | 0.826 | 1.210 | Asynthetisch |
| 9 | 38.430 | 5.051 | 0.250 | 4.000 | Bsynthetisch |
| 4 | 38.661 | 963.087 | 76.000 | 405.333 | Eisen |
| 2 | 52.480 | 3.388 | 1.260 | 9.520 | Calcium |

Die beste Substanz ist Nr. 3 : Cholesterin mit
 GESCHAETZT **150.000** mg/dl
 GEMESSEN 179.543 mg/dl bei 520 nm im arith. Mittel

Tabelle 4-2 : Erkennung einer Cholesterinprobe von 150 mg/dl, sortiert nach Distanz, und Berechnung der Konzentration

¹⁰ Wie weiter oben ausgeführt, ist die berechnete Distanz als Bogenmaß, nämlich \arccos eines positiven Wertes, beschränkt auf Werte zwischen 0 und $\pi/2$. Der besseren Anschaulichkeit halber wurde die Distanz daher auf Werte zwischen 0° und 90° normiert.

Dagegen zeigte sich, dass die Approximationsmethode (= Lösungsansatz 3) nicht ohne weiteres anwendbar ist: bei Berücksichtigung aller Wellenlängen des Bereiches von 400 nm bis 800 nm (Auflösung 1 nm) wurde das eine oder andere Spektrum einer Referenzsubstanz nicht richtig erkannt. Der Grund dafür liegt darin, dass das für die Approximationsmethode entworfene Modell eine lineare Beziehung von Konzentration und Extinktion voraussetzt. Soll die Approximationsmethode angewendet werden, muss geprüft werden, ob und inwieweit eine lineare Abhängigkeit zwischen Konzentration und Extinktion besteht.

Die Abbildung 4-1 zeigt eine Untersuchung dieses Zusammenhanges für 400 nm bis 800 nm.

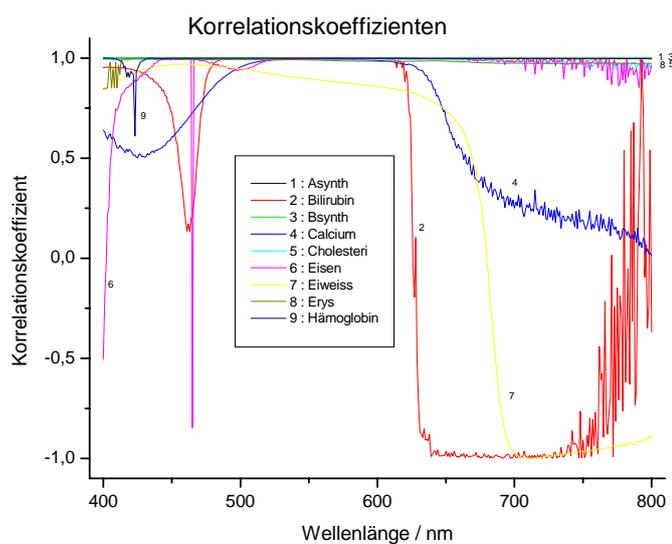


Abbildung 4-1¹¹: Prüfung der linearen Abhängigkeit von Extinktion und Konzentration in Abhängigkeit von der Wellenlänge

¹¹ Man erkennt in der Graphik, dass die Korrelationen für Asynth und Bsynth jeweils über dem gesamten Wellenlängenbereich bei 1.0 liegen. Dies verwundert nicht, da die Vorschriften für beide Funktionen strikt linear abhängig von der Konzentration gewählt wurden.

Trägt man den so genannten Korrelationskoeffizienten(8) r als Maß der linearen Korrelation (Korrelation von $r = \pm 1.0$ bezeichnet eine perfekte Linearität) zwischen Konzentration und Extinktion gegen die Wellenlänge auf, so erkennt man, dass $r > 0,7$ – diese Konstante ist *willkürlich* gewählt – für *alle* Referenzsubstanzen maximal im Wellenlängenbereich von 470 nm bis 620 nm erfüllt ist, oberhalb und unterhalb davon nicht.

Für die Summe der hier gestellten Tests war damit die Approximationsmethode nur von 470 nm bis 620 nm anwendbar. Durch entsprechende Einschränkung erfolgte die Analyse der Referenzsubstanzen dann auch für *alle* Konzentrationsspektren *aller* Referenzsubstanzen korrekt.

Ein Beispiel dafür ist in Tabelle 4-3 wiedergegeben (typisches Ergebnis der Analyse einer Erythrozytenlösung von $4,0 \cdot 10^6/\text{ul}$ – bestimmt mit Zählkammermethode – unter Berücksichtigung des Wellenlängenbereiches von 470 nm bis 620 nm und einer Abtastweite von 1 nm durch die Approximationsmethode, siehe Kurve 2 in Abbildung 2-5).

| Nr. | Distanz | Konz. geschätzt | Konz. Minimum | Konz. Maximum | Referenzsubstanz |
|-----|---------|-----------------|---------------|---------------|------------------|
| 5 | 0.099 | 4.585 | 2.004 | 15.811 | Erythrozyten |
| 6 | 2.871 | 67.380 | 3.760 | 28.409 | Gesamteiweiss |
| 7 | 5.735 | 49.966 | 5.511 | 43.480 | Haemoglobin |
| 4 | 6.205 | 2852.362 | 76.000 | 405.333 | Eisen |
| 3 | 7.610 | 503.239 | 75.000 | 566.667 | Cholesterin |
| 1 | 9.069 | 25.385 | 0.990 | 6.223 | Bilirubin |
| 8 | 9.697 | 1.058 | 0.826 | 1.210 | Asynthetisch |
| 2 | 10.628 | 9.272 | 1.260 | 9.520 | Calcium |
| 9 | 12.592 | 11.887 | 0.250 | 4.000 | Bsynthetisch |

Die beste Substanz ist Nr. 5 : Erythrozyten mit

GESCHAETZT $4.585 \cdot 10^6/\text{ul}$

GEMESSEN $3.892 \cdot 10^6/\text{ul}$ bei 500 520 560 585 600 534 nm im arith. Mittel

Tabelle 4-3 : Erkennung einer Erythrozytenprobe von $4,0 \cdot 10^6/\text{ul}$, sortiert nach Distanz, und Berechnung der Konzentration

4.5.2 Anwendung der Algorithmen zur Testerkennung

Hierfür wurden die Spektren für Leerwert und unbereinigte Probe als zwei Vektoren eingelesen.

Da es vorstellbar ist, dass bei der Messung die übliche Reihenfolge – erst Leerwert, dann Probe – nicht eingehalten wurde, entscheidet ein Algorithmus dann anhand der Beträge der beiden Vektoren, welcher Vektor Leerwert und welcher Probe ist¹².

Eine exemplarische Übersicht über die vorgenannten Beträge gibt die Tabelle 4-4.

| Beträge der zu er- kennenden Spektren | 400 nm bis 800 nm | | 470 nm bis 620 nm | |
|--|--------------------|------------|--------------------|----------|
| | unbereinigte Probe | Leerwert | unbereinigte Probe | Leerwert |
| Bilirubin | 5.6 | 3.9 | 3.2 | 2.0 |
| Calcium Analyse | 11.5 | 9.4 | 7.5 | 3.9 |
| Calcium Standard | 11.3 | 9.4 | 7.5 | 3.9 |
| Cholesterin | 7.0 | 1.7 | 6.0 | 1.2 |
| Eisen | 7.4 | 6.0 | 4.2 | 3.1 |
| Erythrozyten | 32.1 | 1.2 | 19.4 | 0.7 |
| Eiweiß | 6.6 | 3.2 | 5.4 | 1.5 |
| Hämoglobin | 29.1 | 1.2 | 5.3 | 0.7 |

Tabelle 4-4: Beträge der unbereinigten Proben im Vergleich zu ihren Leerwerten

Wie theoretisch bereits abgeleitet wurde, haben alle unbereinigten Proben über die Wellenlängenbereiche in Tabelle 4-4 eine höhere Extinktion als die zugeordneten Leerwerte. Das Beispiel 4 für Hämoglobin zeigt, dass sich ein hoher Betrag von 29,1 für den ersten Vektor ergibt. Damit handelt es sich um die Probe, bei dem zweiten Vektor mit einem niedrigeren Betrag von 1,2 um den Leerwert.

¹² Diese Zuordnung ist auch dann erforderlich, wenn der Leerwert mit in die Analyse eingeht, da die Verkettung von Probe und Leerwert in der definierten Reihenfolge geschehen muss.

| Probe Hämoglobin | | Leerwert Hämoglobin | |
|------------------|-------------------------------|---------------------|-------------------------------|
| 800 | 0.110 * 0.110 = 0.0120 | 800 | 0.061 * 0.061 = 0.0037 |
| | + | | + |
| 799 | 0.109 * 0.109 = 0.0118 | 799 | 0.060 * 0.060 = 0.0037 |
| | + | | + |
| . | . * . = . | . | . * . = . |
| | + | | + |
| 401 | 2.858 * 2.858 = 8.1671 | 401 | 0.066 * 0.066 = 0.0043 |
| | + | | + |
| 400 | 2.739 * 2.739 = <u>7.5000</u> | 400 | 0.065 * 0.065 = <u>0.0043</u> |
| | 29,1 | | 1,2 |

Beispiel 4 : Berechnung der Beträge von Probe und Leerwert für den Hämoglobintest

Dies ermöglicht, die Differenz unbereinigte Probe – Leerwert zu bilden und entsprechend den drei Algorithmen den Abstand zu jeder der unter Abbildung 2-1 bis Abbildung 2-9 abgebildeten Referenzsubstanzen zu bestimmen und die Konzentration zu berechnen. Die Tabelle 4-5 zeigt dies beispielhaft für die Erkennung des Hämoglobintestes durch die Interpolationsmethode unter Berücksichtigung des Wellenlängenbereiches von 400 nm bis 800 nm und einer Abtastweite von 1 nm.

| Nr. | Distanz | Konz. geschätzt | Konz. Minimum | Konz. Maximum | Referenzsubstanz |
|-----|---------|-----------------|---------------|---------------|------------------|
| 7 | 6.358 | 17.923 | 5.511 | 43.480 | Haemoglobin |
| 9 | 17.989 | 8.249 | 0.250 | 4.000 | Bsynthetisch |
| 8 | 23.033 | 0.570 | 0.826 | 1.210 | Asynthetisch |
| 5 | 24.042 | 1.656 | 2.004 | 15.811 | Erythrozyten |
| 1 | 24.231 | 14.557 | 0.990 | 6.223 | Bilirubin |
| 6 | 25.767 | 66.903 | 3.760 | 28.409 | Gesamteiweiss |
| 3 | 25.893 | 345.800 | 75.000 | 566.667 | Cholesterin |
| 4 | 27.626 | 266.979 | 76.000 | 405.333 | Eisen |
| 2 | 27.761 | 4.943 | 1.260 | 9.520 | Calcium |

Die beste Substanz ist Nr. 7 : Haemoglobin mit

// da der Abstand der Testsubstanz

// zu der Referenzsubstanz

// **Hämoglobin** minimal ist

GESCHAETZT 17.92339 g/dl

// Konz.-Schätzung durch Algorithmus

GEMESSEN 17.78291 g/dl bei 520nm im arith. Mittel

// Unter Verwendung des

// Extinktionskoeffizienten bei 520 nm

Tabelle 4-5: Erkennung einer Hämoglobinprobe, sortiert nach Distanz, und Berechnung der Konzentration

Ein Vergleich der Distanzen zu den restlichen Referenzsubstanzen im obigen Beispiel zeigt, dass alle anderen Referenzsubstanzen mindestens um das Zwei- bis Dreifache

weiter entfernt liegen. Die zweitnächste Referenzsubstanz, Bsynthetisch, liegt nämlich mit 17.989 etwa dreimal soweit entfernt wie Hämoglobin, die nächste nichtsynthetische, Erythrozyten, etwa viermal so weit wie Hämoglobin. Es ist damit anzunehmen, dass kleinere Messungenauigkeiten der spektralen Daten die Identifikation wenig erschweren.

Sollten aber in die Erkennung sehr viel mehr Tests eingezogen werden, so könnte sich die Identifikation schwieriger gestalten. Für diesen Fall muss ein Gütekriterium für die Identifikationssicherheit gefunden werden, das man als „Trenngüte“ bezeichnen kann.

4.6 Trenngüte

Neben anderen aufwendigen statistischen Verfahren wie der Fehlerquadratmethode wird hier ein einfaches Vorgehen wie folgt vorgeschlagen.

Es sei x die Testsubstanz und d_1 bis d_m die Abstände der Substanz x zu den Referenzsubstanzen P_1 bis P_m und die Testsubstanz x richtig als zugehörig zu der Referenzsubstanz P_b erkannt, d.h. $d_b := \min \{ d_i, i = 1..m \}$.

Dann sei $d_z := \min \{ d_i, i = 1..m, i \neq b \}$, also der Abstand zur "zweitbesten" Referenzsubstanz. Wir definieren dann die Trenngüte als

$$t := 1 - d_b/d_z.$$

Bei t nahe 1 liegt offenbar eine sehr gute Trenngüte vor, bei t um 0 eine sehr schlechte. Liegt nämlich ein Vektor genau zwischen zwei Polygonzügen ist klar, dass dann eine Zuordnung zumindest nach der Interpolations- oder Approximationsmethode alleine nicht eindeutig erfolgen kann. Unter zusätzlicher Berücksichtigung der Winkelmethode gelingt dies, siehe unten. Die Abstände d_b und d_z sind dann gleich und die Trenngüte ergibt sich zu $t = 1 - 1 = 0$. Liegt die Testsubstanz jedoch *auf* der besten Referenzsubstanz, so ist $d_b = 0$ und man erhält $t = 1 - 0/d_z = 1$.

Im vorgenannten Beispiel mit Hämoglobin ist $t = 1 - 6.358/17.989 = 65\%$. Das bedeutet nicht, dass die Wahrscheinlichkeit einer Fehlerkennung bei 35% läge. Es sagt vielmehr aus, dass eine Fehlerkennung erst dann auftreten könnte, wenn sich durch eine extrem unwahrscheinliche Reihe von Fehlmessungen der Extinktionen der Testsubstanz diese der zweitbesten Referenzsubstanz auf unter 35% des wirklichen Wertes annähern würde. Eine Trenngüte von 65% stellt damit einen sehr *guten* Wert dar.

Um einen Überblick über die zu erwartenden Trenngüte zu erhalten, wurde zunächst die Interpolationsmethode als anzunehmender „Goldstandard“ unter „Verengung“ des

Abtastintervalls getestet; dabei wurde der Bereich so lange reduziert, bis mindestens eine Testsubstanz nicht mehr richtig erkannt wurde. Bei den folgend aufgeführten Bereichen erfolgte noch eine richtige Identifikation:

| Verfahren | W.- von | länge bis | $\Delta\lambda$ | Bilirub | Ca.-Ar | Ca.-St | Chole | Eisen | Erythr | Eiweiß | Haem | Arith. Mittel ü. Verfahren |
|--------------------------------------|------------|--------------|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------------------------|
| Interpolation | 400 | 630 | 1 | 0.346 | 0.610 | 0.700 | 0.970 | 0.069 | <u>0.961</u> | 0.961 | 0.641 | 0.657 |
| Interpolation | 400 | 650 | 1 | 0.388 | 0.620 | 0.704 | 0.972 | 0.277 | 0.960 | <u>0.972</u> | 0.644 | 0.692 |
| Interpolation | 400 | 700 | 1 | <u>0.389</u> | <u>0.623</u> | <u>0.788</u> | <u>0.973</u> | 0.449 | 0.957 | 0.969 | <u>0.647</u> | 0.724 |
| Interpolation | 400 | 800 | 1 | 0.389 | 0.618 | 0.699 | 0.973 | <u>0.457</u> | 0.961 | 0.969 | 0.647 | 0.714 |
| Arith. Mittel ü. Substanz | | | | 0.378 | 0.618 | 0.723 | 0.972 | 0.313 | 0.960 | 0.968 | 0.644 | |

Tabelle 4-6 : Trenngüten bei verschiedenem Wellenlängenbereich

Hierbei bedeutet der Eintrag 0.388 in der Zeile „Interpolation 400 650 1“ und der Spalte „Bilirubin“:

„Der Algorithmus Interpolationsmethode erkennt im Bereich 400 nm bis 650 nm bei einer Abtastweite von 1 nm, d.h. also mit $(650 - 400) / 1 = 150$ Datenpunkten die Testsubstanz Bilirubin mit einer Trenngüte von 0,38.“

Der Unterstrich markiert die beste Trenngüte einer Testsubstanz über den Block.

4.7 Vergleich der Verfahren

4.7.1 Analysand : Probe abzüglich Leerwert

Um die verschiedenen Verfahren mit einander vergleichen zu können, wurden wegen der Nichtlinearität zwischen Konzentration und Extinktion einiger Substanzen im gesamten Wellenlängenbereich nur noch der Wellenlängenbereich 470 nm bis 620 nm berücksichtigt. Es wurde geprüft, ob die Wellenlängenauflösung (1 nm bis 50 nm) einen wesentlichen Einfluss auf die Präzision des Verfahrens hat. Analog zu Tabelle 4-6 wurde diesmal die Wellenlängenauflösung solange verdoppelt, bis mindestens eine der sieben Testspektren falsch zugeordnet wurde.

| Verfahren | W.- von | länge bis | $\Delta\lambda$ | Biliruk | Ca.-Ar | Ca.-St | Chole | Eisen | Erythr | Eiweiß | Haem | Arith. Mittel ü. Verfahren |
|------------------------------|------------|--------------|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------------------------|
| Interpolation | 470 | 620 | 25 | 0.524 | 0.692 | 0.886 | 0.961 | 0.365 | <u>0.991</u> | <u>0.931</u> | 0.904 | 0.782 |
| Interpolation | 470 | 620 | 16 | 0.576 | 0.700 | 0.886 | 0.964 | 0.308 | 0.991 | 0.924 | 0.911 | 0.783 |
| Interpolation | 470 | 620 | 8 | 0.586 | 0.712 | 0.892 | 0.965 | 0.313 | 0.991 | 0.928 | 0.913 | 0.787 |
| Interpolation | 470 | 620 | 4 | 0.592 | 0.724 | 0.893 | 0.966 | 0.392 | 0.991 | 0.929 | <u>0.914</u> | 0.800 |
| Interpolation | 470 | 620 | 2 | 0.598 | <u>0.813</u> | <u>0.894</u> | 0.966 | 0.423 | 0.991 | 0.931 | 0.912 | 0.816 |
| Interpolation | 470 | 620 | 1 | <u>0.603</u> | 0.732 | 0.894 | <u>0.967</u> | <u>0.431</u> | 0.991 | 0.930 | 0.913 | 0.808 |
| Arith. Mittel ü. Substanz | | | | 0.580 | 0.729 | 0.891 | 0.965 | 0.372 | 0.991 | 0.929 | 0.911 | |
| Winkel | 470 | 620 | 25 | 0.542 | 0.695 | 0.887 | 0.961 | 0.334 | <u>1.000</u> | 0.930 | 0.906 | 0.782 |
| Winkel | 470 | 620 | 16 | 0.587 | 0.701 | 0.888 | 0.964 | 0.279 | 1.000 | 0.927 | 0.913 | 0.782 |
| Winkel | 470 | 620 | 8 | 0.591 | 0.713 | 0.892 | <u>0.973</u> | 0.279 | 1.000 | 0.928 | <u>0.915</u> | 0.786 |
| Winkel | 470 | 620 | 4 | 0.599 | 0.725 | 0.894 | 0.967 | 0.362 | 1.000 | 0.929 | 0.915 | 0.799 |
| Winkel | 470 | 620 | 2 | 0.605 | 0.731 | <u>0.895</u> | 0.967 | 0.399 | 1.000 | <u>0.941</u> | 0.914 | 0.807 |
| Winkel | 470 | 620 | 1 | <u>0.608</u> | <u>0.734</u> | 0.895 | 0.968 | <u>0.403</u> | 1.000 | 0.931 | 0.914 | 0.807 |
| Arith. Mittel ü. Substanz | | | | 0.589 | 0.717 | 0.892 | 0.967 | 0.343 | 1.000 | 0.931 | 0.913 | |
| Approx. | 470 | 620 | 50 | 0.397 | 0.715 | 0.892 | 0.938 | 0.026 | <u>0.992</u> | <u>0.937</u> | 0.844 | 0.718 |
| Approx. | 470 | 620 | 25 | 0.397 | 0.715 | 0.892 | 0.938 | 0.026 | 0.992 | 0.937 | 0.844 | 0.718 |
| Approx. | 470 | 620 | 16 | 0.576 | 0.700 | 0.886 | 0.964 | 0.308 | 0.991 | 0.924 | 0.911 | 0.783 |
| Approx. | 470 | 620 | 8 | 0.586 | 0.712 | 0.892 | 0.965 | 0.313 | 0.991 | 0.928 | 0.913 | 0.787 |
| Approx. | 470 | 620 | 4 | 0.592 | 0.724 | 0.893 | 0.966 | 0.392 | 0.991 | 0.929 | <u>0.914</u> | 0.800 |
| Approx. | 470 | 620 | 2 | 0.598 | 0.729 | <u>0.894</u> | 0.966 | 0.423 | 0.991 | 0.931 | 0.912 | 0.806 |
| Approx. | 470 | 620 | 1 | <u>0.603</u> | <u>0.732</u> | 0.894 | <u>0.967</u> | <u>0.431</u> | 0.991 | 0.930 | 0.913 | 0.808 |
| Arith. Mittel ü. Substanz | | | | 0.536 | 0.718 | 0.892 | 0.958 | 0.274 | 0.991 | 0.931 | 0.893 | |

Tabelle 4-7 : Trenngüten bei verschiedener Wellenlängenauflösung

Die Tabelle 4-7 zeigt, dass eine bessere Auflösung eine höhere Trenngüte nach sich zieht.

4.7.2 Analysand: bereinigte Probe verkettet mit Leerwert

Berücksichtigte man Probe UND Leerwert, so war die Winkelmethode für JEDEN Wellenlängenbereich fehlerhaft. Die Interpolationsmethode und die Approximationsmethode ergaben erst durch Dehnung des Wellenlängenbereiches auf Werte wie unten dargestellt ein fehlerfreies Ergebnis.

| Verfahren | W.- von | länge bis | $\Delta\lambda$ | Bilirub | Ca.-Ar | Ca.-St | Chole | Eisen | Erythr | Eiweiß | Haem | Arith. Mittel ü. Verfahren |
|---------------|------------|--------------|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------------------------|
| Interpolation | 400 | 750 | 1 | 0.033 | 0.007 | 0.011 | 0.046 | 0.442 | 0.815 | 0.442 | 0.815 | 0.326 |
| Approx. | 400 | 800 | 1 | 0.207 | 0.926 | 0.885 | 0.921 | 0.814 | 0.897 | 0.969 | 0.637 | 0.782 |
| Arith. Mittel | | | | | | | | | | | | |
| ü. Substanz | | | | 0.120 | 0.467 | 0.448 | 0.483 | 0.628 | 0.856 | 0.705 | 0.726 | |

Tabelle 4-8 : Trenngüten bei Berücksichtigung des Leerwertes

Auch die Interpolations- und Winkelmethode waren hinsichtlich der Trenngüte weniger befriedigend (siehe Tabelle 4-8) als bei Verwendung des Differenzspektrums (siehe Tabelle 4-7).

Insgesamt lassen Tabelle 4-6 bis Tabelle 4-8 folgende Aussagen zu:

Im linearen Wellenlängenbereich (470 nm bis 620 nm) – siehe Tabelle 4-6 – ist die Trenngüte für alle drei Lösungsansätze bei 1 nm Abtastweite mit im Mittel 81% gleich. Bei allen drei Lösungsansätzen führt eine Reduktion der spektralen Auflösung *nicht* zu einer wesentlich schlechteren Erkennung. Selbst bei lediglich (620 nm – 470 nm) / 25 nm = 6 Messpunkten – also Reduktion der Anzahl der berücksichtigten Extinktionen auf 1/25! – reduziert sich die mittlere Trenngüte von 81% auf lediglich je 78% für die Interpolations- und Winkelmethode und auf 72% für die Approximationsmethode. Weiterhin ist die Trenngüte einer Testsubstanz konstant über alle Lösungsansätze, siehe etwa für Erythrozyten mit rund 100% in jedem Block.

Bezüglich der Auswahl des richtigen Spektral**bereiches** kann festgehalten werden, dass eine Spreizung des Wellenlängenbereiches auf größtmögliche Werte wie in Tabelle 4-6 gezeigt, *nicht* die Trenngüte verbessert: so liegt die durchschnittliche Trenngüte der Interpolationsmethode bei einem Wellenlängenbereich von 570 nm bis 620 nm – 1 nm Abtastweite – mit 81% deutlich über der besten durchschnittlichen Trenngüte aus Tabelle 4-6 in Höhe von 72%.

Eine noch schlechtere Trenngüte ergibt sich durchgehend bei Einbeziehung der Leerwerte: wie aus Tabelle 4-8 ersichtlich, funktioniert die Winkelmethode gar nicht im Sinne eine über alle Testsubstanzen fehlerfreien Erkennung. Interpolations- und Approximationsmethode liefern bei Einbeziehung der Leerwerte nur deutlich schlechtere Trenngüten.

5 Diskussion

Ziel dieser Arbeit war es, Algorithmen zu entwerfen, die eine sichere Identifikation von Spektren aus medizinisch-diagnostischen Tests gestatten. Hierzu wurden exemplarisch Spektren von sieben Laborbestimmungen mit Hilfe eines handelsüblichen Spektrometers mit einer Wellenlängenauflösung von 1 nm über den Bereich von 400 nm bis 800 nm erfasst und nach Etablierung der Algorithmen der Versuch unternommen, die geprüften Testsubstanzen zu erkennen und ihre Konzentrationen zu bestimmen.

Die Referenzspektren selbst wurden durch jeweils höhere Konzentrationen von Secutrol bzw. Vollblut erzeugt. Dies ist insofern nur ein Modell der in der Realität anfallenden Proben, weil in der Regel nicht *alle* Serumparameter ansteigen, wenn ein Wert pathologisch verändert ist. Da aber spezifische Reagenzien benutzt wurden, erscheint diese Form der Konzentrationserhöhung als vertretbar.

Im Rahmen der mathematischen Formalisierung wurden die Spektren als Vektoren betrachtet. Es wurden drei Algorithmen auf Basis der Vektorenrechnung entwickelt. Zusätzlich wurde geprüft, ob die Berücksichtigung der Testleerwerte in Kombination mit den Tests die Erkennung des Testansatzes verbessert.

So ließen sich die Spektren interpolatorisch darstellen (siehe 4.4.1), indem vorausgesetzt wurde, dass sich Spektren zwischen zwei benachbarten Konzentrationen linear verhalten. Bei der Winkelmethode (siehe 4.4.2) wird insbesondere das Integral unter den Spektren berücksichtigt, die Distanzmessung erfolgt über einen Winkel. Der dritte Lösungsansatz, die „Approximationsmethode“ (siehe 4.4.3), bedient sich der üblichen Methode der kleinsten Fehlerquadrate, wobei die Spektren als lineare Näherungsfunktion beschrieben werden. Die Distanzmessung erfolgt wie bei der Interpolationsmethode über ein einfaches Skalarprodukt.

Alle vorgeschlagenen Algorithmen können Testspektren erkennen.

Die Approximationsmethode (siehe 4.4.3) konnte aber nur dann die Tests anhand ihrer Spektren identifizieren, wenn sichergestellt war, dass für den gesamten Spektralbereich das Lambert-Beersche Gesetz uneingeschränkt gilt. Da dies nicht für alle hier verwendeten Tests zutraf, konnte mit diesem Algorithmus nur dann eine eindeutige Zuordnung erreicht werden, wenn der Wellenlängenbereich entsprechend eingeschränkt wurde, hier auf 470 bis 620 nm, siehe Abbildung 4-1. Im Hinblick auf eine Ergänzung durch weitere Tests müsste für die Verwendung dieses Algorithmus jeweils der Wellenlängenbereich identifiziert werden, der eine lineare Beziehung zwischen Konzentration und Extinktion aller eingeschlossenen Testsubstanzen ergibt.

Der Algorithmus, der auf der Winkelmethode beruht (siehe 4.4.2), ist von der unbedingten Einhaltung des Lambert-Beerschen Gesetzes nicht betroffen. Andererseits setzt die Winkelmethode voraus, dass das Integral unter dem Spektrum mit der Konzentration ansteigt. Schwierigkeiten sind für den Fall vorherzusehen, wenn beispielsweise der testspezifische Peak ansteigt, an anderer Stelle aber beispielsweise durch Substratverbrauch die Absorption abnimmt.

Die Interpolationsmethode (siehe 4.4.1) ist vergleichsweise rechenintensiv, wird aber weder durch den Wellenlängenbereich noch durch das Integral unter den Spektren gestört. Es bietet sich daher an, diesen Algorithmus bevorzugt zur Identifizierung und Konzentrationsmessung zu nutzen, insbesondere dann, wenn man eine größere Anzahl neuer spektroskopischer Tests einführen möchte.

Die Einbeziehung der Leerwertdaten in die Algorithmenstruktur brachte in keinem Fall eine Verbesserung und barg darüber hinaus die Gefahr der fehlerhaften Testidentifikation. Dies ist leicht dadurch erklärbar, dass viele Tests gleiches oder ähnliches Leerwertmedium verwenden.

Zur Abschätzung der Güte der Erkennung der beiden nächstliegenden Referenzspektren wurde mit der Trenngüte ein Maß etabliert, das etwas über die Treffergenauigkeit aussagt. Es entspricht nicht der Berechnung der Irrtumswahrscheinlichkeit nach statistischen Methoden. Es benutzt ganz ähnlich wie

im Least-Square-Verfahren die Summe der Abweichungen, wobei die hier verwendeten Algorithmen – insbesondere bei Interpolationsmethode und Approximationsmethode – den Abstand als Wurzel aus der Summe der Quadrate der Absorptionsunterschiede darstellen. Die Abschätzung der Trenngüte wurde in dieser Arbeit aber nicht dazu genutzt, die Irrtumswahrscheinlichkeit zu bestimmen. Mit Hilfe dieses Parameters sollte vielmehr die Frage beantwortet werden, ob und wie weit die Reduktion der spektralen Daten – Spektralbereich und spektrale Auflösung – sich hinsichtlich der Detektionssicherheit auswirkt. Dies erschien deshalb wichtig, weil heute vielfach die Verwendung kleiner Reflexionsmikrospektrometer die Wellenlängenauflösung nur 4 nm bis 12 nm beträgt. Unter Benutzung der Interpolationsmethode konnte hier gezeigt werden, dass (siehe Tabelle 4-7) selbst bei einer spektralen Auflösung von 25 nm die Tests mit großer Sicherheit detektiert werden können.

Durch Anwendung des Algorithmus aus der Interpolationsmethode ist es möglich, Spektren mit großer Sicherheit zu erkennen und Substratkonzentrationen zu bestimmen. Zur Ermittlung der Konzentration einer Testsubstanz bieten alle Verfahren die Berechnung der Konzentration anhand der Extinktionskoeffizienten an bestimmten Wellenlängen und anhand des gesamten Spektrums an. Weicht die durch Einzelmessung hervorgegangene Konzentration stark von der durch den Algorithmus geschätzten Konzentration ab, kann eine alternative Wellenlänge zur Einzelkonzentrationsmessung herangezogen werden oder der Nutzer zur Wiederholung der Messung aufgefordert werden. Bei den meisten Verfahren ist die Einzelkonzentrationsmessung an mehreren Wellenlängen möglich. Voraussetzung hierfür ist das Vorliegen von verlässlichen Datensätzen, die die zu erwartenden Konzentrationsbereiche voll abdecken.

Eine sinnvolle Anwendung dieser Algorithmen ist erst dann gegeben, wenn neben den sieben hier exemplarisch verwendeten Tests eine große Zahl von Labortests eingeht, die die üblichen medizinischen Labortests weitgehend abdecken.

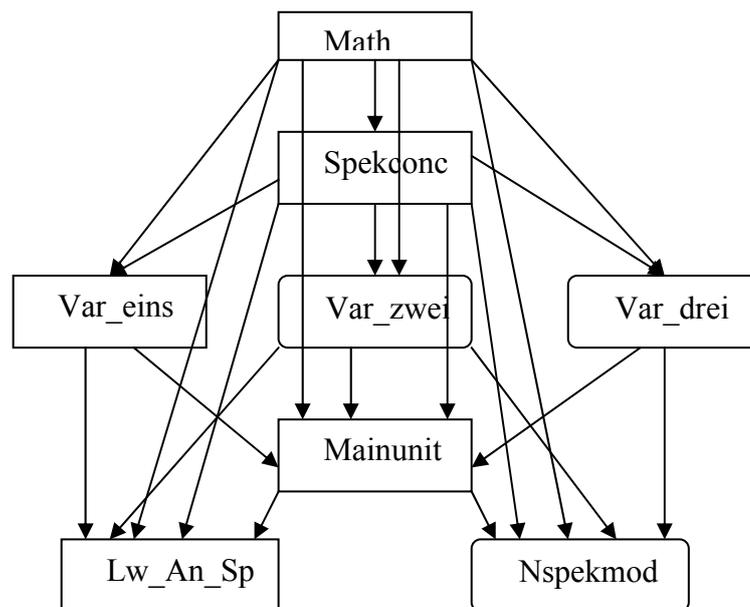
Auf diese Weise wird es möglich sein, mit Hilfe eines Spektrometers oder Mikrospektrometers nur durch sukzessive Messung von Leerwert und Test Substanzen einfach zu erkennen. Auf der anderen Seite ist es denkbar, dass diese Form der Spektrenerkennung dazu genutzt wird, in Großlaboratorien im Zuge der Qualitätssicherung fehlerhafte Proben anhand ihrer atypischen Spektren zu identifizieren.

6 Zusammenfassung

- I. Anhand der vorliegenden Arbeit sollte der Versuch unternommen werden, Spektren bekannter Labortests mittels mathematischer Verfahren zu beschreiben, so dass in einem zweiten Schritt Spektren mit unbekannter Substanzkonzentration qualitativ den bekannten Spektren zugeordnet werden können und deren Konzentrationen im Testansatz festgestellt werden können.
- II. Die Daten von sieben bekannten medizinischen Labortests wurden für einen Wellenlängenbereich von 400 nm bis 800 nm in verschiedenen Konzentrationen aufgenommen.
- III. Es wurden drei mathematische Modelle verwendet. Das Erste setzt lediglich einen linearen Verlauf zwischen Spektren benachbarter Konzentrationen voraus. Das Zweite verlangt eine monoton steigende Beziehung zwischen dem Integral eines Spektrums und der zugrunde liegenden Konzentration. Das Dritte geht von einer linearen Beziehung zwischen Spektren und Konzentrationen über den gesamten Konzentrationsbereich aus.
- IV. Alle Algorithmen sind in der Lage, die Tests zu erkennen, wobei der erste Algorithmus, die Interpolationsmethode, die stabilsten Ergebnisse liefert.
- V. Unter Verwendung dieses Algorithmus können die Tests sicher erkannt werden, auch dann, wenn die Wellenlängenauflösung gering ist oder die Beziehung zwischen Extinktion und Konzentration nicht linear ist.
- VI. Zur Ermittlung der Konzentration einer Testsubstanz bietet das Verfahren die Berechnung der Konzentration anhand der Extinktionskoeffizienten an bestimmten Wellenlängen und anhand des gesamten Spektrums an.
- VII. Vorausgesetzt, dass das vorhandene Testkontingent sich erheblich erweitern lässt, erscheint die Identifikation des Spektrums geeignet, den Messablauf in einem Praxislabor einfacher und sicherer zu gestalten.
- VIII. Eine Anwendung dieser Verfahren im Großlabor kann erheblich zur Güte der Qualitätskontrolle beitragen.

7 Programm

7.1 Blockdiagramm der beteiligten Units



7.2 Kommentar zu den einzelnen Units

MATH.PAS :

Die Unit enthält die wesentlichen MATHEMATISCHEN Prozeduren und Funktionen. Weiterhin exportiert sie einige Konstanten (die maximal gegebenen Wellenlängen) und Typen wie tVektor, die von den übergeordneten Units gebraucht werden.

Da das gesamte Programm Methoden der linearen Algebra benutzt, finden sich auch im Wesentlichen die entsprechenden Operationen wieder. Die Prozeduren SummeVektoren, DifferenzVektoren, SkalarMalVektor, Skalarprodukt, BetragVektor bedürfen keiner weiteren Erläuterung.

Die Prozedur LGS löst ein lineares Gleichungssystem mit 2 Variablen der Form:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Man benötigt diese Operation im Rahmen des Least-Square-Verfahren.

Die Funktionen HaevySide, NaheNull und Wurzel wurden implementiert, weil aufgrund von Rundungsfehlern im Modul Var_dei.pas z.B. zu der Anforderung einer Wurzel aus einer sehr kleinen, negativen Zahl kam. Der fehlerbedingte Abbruch wird dadurch umgangen, dass einem solchen Ausdruck dann der Wert 0 zugeordnet wird.

SPEKCONC.PAS :

In dieser Unit werden aus Vektoren abgeleitete Strukturen definiert und exportiert wie der Typ tPolygon: Zunächst wird ein Polygon durch eine endliche Anzahl von Vektoren definiert.

Aus dem Polygonzug leitet sich der Typ tSubstanz, also eine Referenzsubstanz, ab: Er beinhaltet außer dem eigentlichen Polygonzug (Spektrum) auch die dazugehörigen Substanzkonzentrationen (KonzentrationNr) und den Substanznamen (Substanzname). Weiter enthält tSubstanz auch Informationen über den Modus der eigentlichen Konzentrationsmessung:

Ist ein Standard erforderlich (Standard_anford) , an welchen Wellenlängen findet später die Konzentrationsmessung statt (Mess_wellenlge) , in welcher Einheit erfolgt die Konzentrationsangabe (Einheit) und mit welchem Faktor erfolgt bei den einzelnen Wellenlängen die Umrechnung von Extinktionswert zu Konzentration.

Die Prozedur TauscheZuMessungLeerwert nutzt den im Text hergeleiteten Zusammenhang, dass die "trübere" Mischung, i.e. diejenige MIT der Probe an allen –

bis auf Messfehler bedingte Ausnahmen – Wellenlängen größere Extinktion hat als die Mischung OHNE Probe, i.e. der Leerwert.

Da der Nutzer auch von der Zuordnung – was ist Probe, was ist Leerwert – entlastet werden soll, sortiert die Prozedur die beiden Spektren also Vektoren so, dass an nach Ablauf der Prozedur an erster Stelle die eigentliche Messung und an zweiter Stelle der Leerwert steht.

Die Prozedur `BekanntesSpektrumEinlesen` schließlich liest ein File mit dem Namen `inTextFileName` und folgendem Aufbau ein:

```
Bilirubin  
6.222857143  
800 5.3E-4  
799 0.00107
```

```
403 0.74663  
402 0.75575  
401 0.7606  
400 0.76901
```

und erzeugt einen Vektor `outVektor` mit den Extinktionen als Einträgen, hier

0.76901 bis 5.3E-4 und übergibt weiter, welche Konzentration diesem Spektrum entspricht (`outKonz`), hier 6.222857143.

Exportiert in die übergeordneten Module wird aber nur die abgeleitete Prozedur `NeueSubstanzEinlesen`:

Sie liest ein Textfile des Typs

```

Bilirubin          /* Substanzname
N                  /* Standard_anford
546 560            /* Mess_wellenlge
13.6 17.4 mg/dl    /* Faktor und Einheit
0.99 6.222857143  /* Minimale und Maximale Konzentration der
                  /* Referenzsubstanz
c:\prom_asc\bili\bili050.txt /* Pfad und Name zu den einzelnen
c:\prom_asc\bili\bili100.txt /* bekannten Referenzspektren
c:\prom_asc\bili\bili200.txt /* Diese müssen nach aufsteigenden
                  /* Konzentrationen
c:\prom_asc\bili\bili400.txt /* geordnet sein.

```

und erzeugt daraus eine Variable des Typs TSubstanz.

Var_eins.PAS (alternativ eingebunden: Var_zwei.pas oder Var_drei.pas):

Die Prozedur RedundanteInformationenSubstanzErzeugen liest einen Vektor inSubstanz als Referenzsubstanz ein und verwandelt die darin enthaltenen Informationen in einen an den Lösungsansatz angepassten Datentyp; im Fall der Approximationsmethode wäre dies etwa eine Regressionsgerade für jede Referenzsubstanz

Eine Schlüsselrolle kommt der Prozedur AbstandundKonzentrationPunktSubstanz zu. Sie berechnet je nach vorliegender Methode den Abstand eines Punktes inPunkt zu einer Referenzsubstanz inSubstanz und gibt die Distanz als outMinDistanz und die geschätzte Konzentration als outEstimatedConcentration aus.

MAINUNIT.PAS :

Dieses Unit enthält u.a. die Prozedur AusgabeDistanzen: Sie ermittelt die Abstände einer Testsubstanz zu allen Referenzsubstanzen und gibt diese Abstände aus; die beste Referenzsubstanz wird vermerkt, siehe auch Ausgabebeispiel unter 4.5.

LW_AN_SP.PAS (alternativer Aufruf : NSPEKMOD.PAS) :

Diese Hauptprogramme haben folgende Aufgabe: NSPEKMOD prüft alle einzelnen Spektren der Referenzsubstanzen auf ihre Abstände und speichert das Ergebnis in der ASCII-Datei `c:\prom_asc\tpu&out\TestRefS.txt`. Dieser Schritt dient also lediglich als Plausibilitätskontrolle.

LW_AN_SP schließlich ist das eigentliche Identifikationsprogramm. Es liest anhand der Steuerdatei `c:\prom_asc\ascii\steueren.txt` paarweise jeweils zwei Vektoren ein, ordnet diese als Probe und Leerwert zu und bestimmt dann die am besten passende Referenzsubstanz. Die Ausgabe erfolgt in die Datei `c:\prom_asc\tpu&out\AusDist.txt`.

7.3 Laufzeit- und Speicherbedarfsanalyse

Sei n die Anzahl der berücksichtigten Wellenlängen, m die durchschnittliche Anzahl der Spektren pro Referenzsubstanz und k die Anzahl der Referenzspektren.

Speicherbedarf für Interpolations- und Winkelmethode:

Legt man einen Speicherbedarf von l für einen Extinktionswert zugrunde, ergibt sich ein benötigter Festplattenspeicher $S = O(n*m*k*l)$. Zur O-Notation im Sinne von Größenordnung vergleiche (10).

In unserem Fall mit $n=400$, $m=5$, $k=9$ und $l=4$ Byte erhalten wir einen Speicherbedarf von rund 72000 Byte, also lediglich 70 KByte. Allerdings steigt damit der Speicherbedarf linear mit der Anzahl der Spektren pro Referenzsubstanz m , was bei langer Laufzeit im Sinne einer Speicherung aller bis dahin erhobenen Messungen problematisch werden könnte.

Speicherbedarf für Approximationsmethode:

Hier ergibt sich lediglich ein Speicherbedarf von $O(2*n*k*l) = O(n*k*l)$, in unserem Fall 28800 Byte = 28 KByte. Dies wäre auch nach Anpassung durch die während der Laufzeit erhobenen Messungen unproblematisch.

Laufzeit für Interpolations- und Winkelmethode:

Legt man einen Rechenaufwand von l für die Multiplikation zweier Gleitkommazahlen zugrunde und vernachlässigt die Additionen und Subtraktionen so ergibt sich ein Aufwand für die beiden Algorithmen von $O(l*n*m*k)$.

Laufzeit für Approximationsmethode:

Hier ergibt sich lediglich ein Rechenaufwand von $O(l*n*k)$.

Auch der Rechenaufwand würde also mit feinerer Interpolation für die Interpolations- und Winkelmethode linear steigen und die Approximationsmethode unberührt lassen.

7.4 Quelltext im Pascal-Quellcode

Der Quelltext umfasst etwa 1000 Zeilen Programmcode und kann vom Autor angefordert werden.

8 Tabellen-, Abbildungs- und Beispielverzeichnis

| | |
|--|----|
| Abbildung 1-1 : Schematischer Aufbau eines Spektrometers (17) | 5 |
| Abbildung 1-2 : Prismen- und Gitterspektrometer aus (17)..... | 6 |
| Abbildung 2-1 : Bilirubin | 19 |
| Abbildung 2-2 : Calcium..... | 19 |
| Abbildung 2-3 : Cholesterin | 20 |
| Abbildung 2-4 : Eisen..... | 20 |
| Abbildung 2-5 : Erythrozyten..... | 21 |
| Abbildung 2-6 : Eiweiß | 21 |
| Abbildung 2-7 : Hämoglobin | 22 |
| Abbildung 2-8 : Synthetisches Spektrum A | 22 |
| Abbildung 2-9 : Synthetisches Spektrum B | 23 |
| Abbildung 2-10: Die Testsubstanzen nach Abzug der Leerwerte..... | 24 |
| Abbildung 2-11: Die Testsubstanzen verkettet mit dem Leerwert..... | 25 |
| Abbildung 3-1 : Eiweißspektrum mit LW und Differenz | 27 |
| Abbildung 3-2 : Spektrum Halbmikro-Küvette..... | 30 |
| Abbildung 4-1 : Prüfung der linearen Abhängigkeit von E zu c..... | 52 |
| | |
| Tabelle 2-1 : Referenzsubstanzen mit Testverfahren | 14 |
| Tabelle 2-2 : Referenzspektren : Berechnung der Konz. der einzelnen Tests..... | 16 |
| Tabelle 2-3 : Synthetische Spektren : gewählte Konzentrationen..... | 17 |
| Tabelle 4-1 : Erkennung einer Bilirubinprobe von 0,99 mg/dl, sortiert nach Distanz, 50 | |
| Tabelle 4-2 : Erkennung einer Chol.-probe von 150 mg/dl, sortiert nach Distanz, | 51 |
| Tabelle 4-3 : Erkennung einer Ery.-probe von $4,0 \cdot 10^6$ /ul, sortiert nach Distanz | 53 |
| Tabelle 4-4 : Beträge der unbereinigten Proben im Vergleich zu ihren Leerwerten... 54 | |
| Tabelle 4-5 : Erkennung einer Hämoglobinprobe, sortiert nach Distanz..... | 55 |
| Tabelle 4-6 : Trenngüten bei versch. Wellenlängenbereich..... | 58 |
| Tabelle 4-7 : Trenngüten bei verschiedener Wellenlängenauflösung | 59 |
| Tabelle 4-8 : Trenngüten bei Berücksichtigung des Leerwertes | 60 |
| Tabelle 10-1: Abkürzungsverzeichnis..... | 81 |
| | |
| Beispiel 1 : ASCII-Testdatei der Extinktionen von Calcium bei 7,27 mmol/l..... | 18 |
| Beispiel 2 : ASCII-Steuerdatei von Calcium, Anforderung eines Standards..... | 18 |
| Beispiel 3 : ASCII-Steuerdatei von Erythrozyten, Berechnung der Konzentration..... | 18 |
| Beispiel 4 : Berechnung der Beträge von Probe und LW für den Hämoglobintest | 55 |

9 Anhang

9.1 Mathematisch

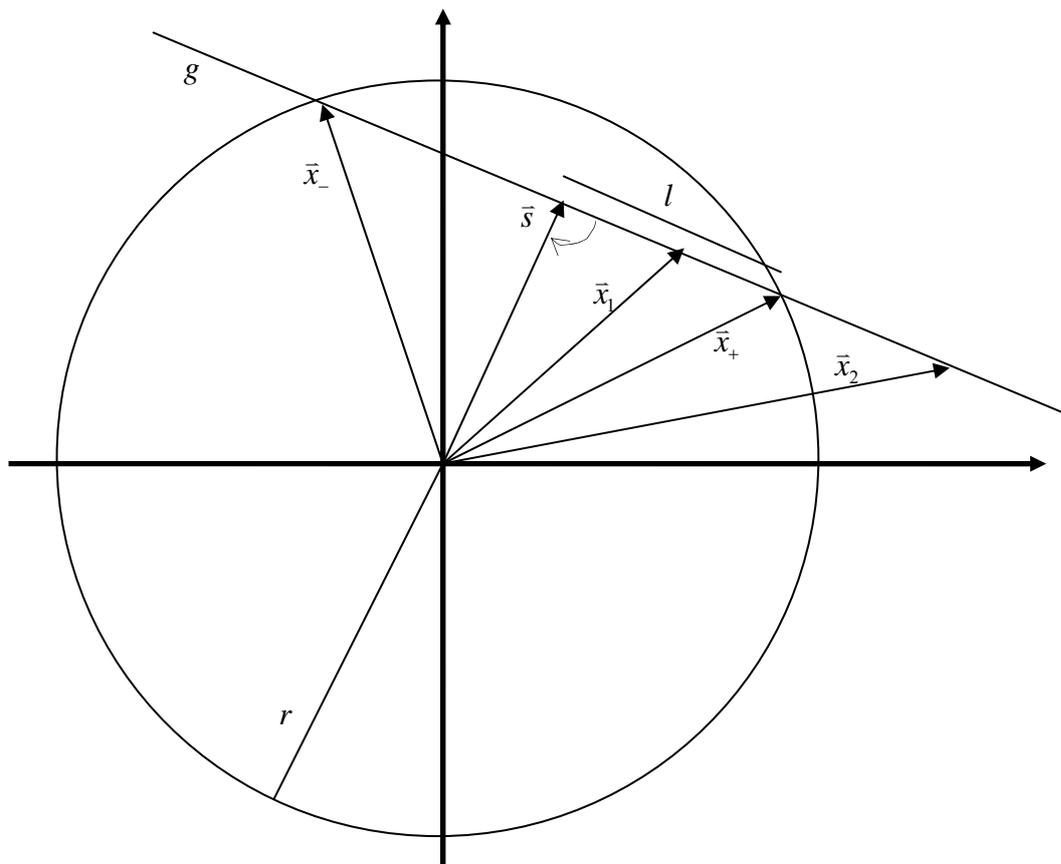
Seien $\bar{x}_1, \bar{x}_2 \in \mathfrak{R}^n$, $0 < |\bar{x}_1| \leq r < |\bar{x}_2| \Leftrightarrow \bar{x}_1^2 \leq r^2 < \bar{x}_2^2$ und $|\bar{x}| := \sqrt{\bar{x}^2}$ mit dem kanonischen Skalarprodukt auf \mathfrak{R}^n .

Behauptung :

$\exists \alpha_+ \in [0;1[$ mit $|\bar{x}| = r$, wobei $\bar{x} := \bar{x}_1 + \alpha_+(\bar{x}_2 - \bar{x}_1) \wedge$

$\exists \alpha_- \leq 0$ mit $|\bar{x}| = r$, wobei $\bar{x} := \bar{x}_1 + \alpha_-(\bar{x}_2 - \bar{x}_1)$

Folgende Skizze



legt als Lösungen von $|\bar{x}| = r$ die Lösungen

$$\bar{x}_{\pm} = \bar{s} \pm l \frac{x_2 - x_1}{|x_2 - x_1|} \text{ nah.}$$

Wie vom Autor ausgerechnet, ergibt sich

$$\alpha_0 = \frac{-\bar{x}_{i-1}(\bar{x}_i - \bar{x}_{i-1})}{(\bar{x}_i - \bar{x}_{i-1})^2} + \sqrt{\frac{r^2 - \bar{s}^2}{(\bar{x}_i - \bar{x}_{i-1})^2}}$$

Die genaue Herleitung kann vom Autor angefordert werden.

9.2 Sonstige

```

program erzeugeSpektrenA (input , output) ;
const
    lambdaA      = 200 ;
    lambdaB      = 800 ;
    periode      = 3.0 ;
    basiskonzentration = 1.1 ;
    faktorbasis  =      4 ;
    daempfung    = 0.005 ;
var
    lambda ,
    exponent          : integer ;
    pi ,
    basis ,
    nulleins ,
    extinktion ,
    schwingung ,
    konzentration    : real ;
    dummy            : string ;
    textfile         : text ;
begin
    pi := 4 * arctan (1) ;
    for exponent := -2 to 2 do
    begin
        konzentration := exp (ln (basiskonzentration) * exponent) ;
        str (exponent + 2 : 1, dummy) ;
        assign (textfile , 'a:asynt'+ dummy + '.txt') ;
        rewrite (textfile) ;
        writeln (textfile , 'Synthetisch A : E = c * sin (lambda) +
            exp (- lambda)') ;
        writeln (textfile , konzentration) ;
        for lambda := lambdaA to lambdaB do
        begin
            nulleins := (lambda - lambdaA) / (lambdaB - lambdaA) ;
            basis := faktorbasis * exp ( ln (daempfung ) *
                nulleins) ;
            schwingung := 1 + sin (2 * pi * nulleins * periode) ;
            extinktion := konzentration * schwingung + basis ;
            writeln (textfile , lambda : 3 , ' ' , extinktion)
        end ;
        close (textfile)
    end
end.

```

```

program erzeugeSpektrenB (input , output) ;
const
  lambdaA    = 200 ;
  lambdaB    = 800 ;
  periode    = 3.0 ;
  basiskonzentration = 2.0 ;
  faktorbasis =      1.025 ;
  daempfung  = 0.01 ;
var
  lambda ,
  exponent          : integer ;
  pi ,
  basis ,
  nulleins ,
  extinktion ,
  schwingung ,
  konzentration    : real ;
  dummy            : string ;
  textfile         : text ;
begin
  pi := 4 * arctan (1) ;
  for exponent := -2 to 2 do
  begin
    konzentration := exp (ln (basiskonzentration) * exponent) ;
    str (exponent + 2 : 1, dummy) ;
    assign (textfile , 'a:bsynth'+ dummy + '.txt') ;
    rewrite (textfile) ;
    writeln (textfile , 'Synthetisch B : E = c * sin (lambda) *
      exp (-lambda)');
    writeln (textfile , konzentration) ;
    for lambda := lambdaA to lambdaB do
    begin
      nulleins := (lambda - lambdaA) / (lambdaB - lambdaA) ;
      basis    := faktorbasis * exp ( ln (daempfung ) *
        nulleins) ;
      schwingung := 1 + sin (2 * pi * nulleins * periode) ;
      extinktion := konzentration * schwingung * basis ;
      writeln (textfile , lambda : 3 , ' ' , extinktion)
    end ;
    close (textfile)
  end
end
end.

```

10 Literatur- und Abkürzungsverzeichnis

- 1 Barner, Martin; Friedrich Flohr. 1983. Analysis I. Berlin; New York: de Gruyter. S. 328.
- 2 Barner, Martin; Friedrich Flohr. 1996. Analysis II. Berlin; New York: de Gruyter. S. 15, S. 10.
- 3 Barrow, Gordon M.. 1982. Physikalische Chemie: Lehrbuch für Chemiker, Verfahrenstechniker, Physiker ab 3. Sem.. Wien: Bohmann; Braunschweig/Wiesbaden: Vieweg. S. 196-200.
- 4 Bertini G, Rubaltelli FF. Non-invasive bilirubinometry in neonatal jaundice. Semin Neonatol. 2002 Apr;7(2):129-33.
- 5 Bott, M. Schriftliche Mitteilung Laborpraxis Dres. Eveld/Gyenes Nienkampstraße 1 45326 Essen vom 03.08.04
- 6 Brecht, J. „Monitoring : Pulsoximetrie“ <http://home.t-online.de/home/Dr.J.Brecht/SKRIPTE/Monitoring.htm#_Toc50632019>(01.08.2004)
- 7 Briscoe L, Clark S, Yoxall CW. Can transcutaneous bilirubinometry reduce the need for blood tests in jaundiced full term babies? Arch Dis Child Fetal Neonatal Ed 2002; 86: F190-F192.
- 8 Bronstein, Il'ja Nikolaevic; K.A. Semendjajew. 1981. Taschenbuch der Mathematik; Thun, Frankfurt/Main: Verlag Harri Deutsch. S. 693, S. 287.
- 9 Encyclopædia Britannica from Encyclopædia Britannica Premium Service. „spectrophotometry“ <<http://www.britannica.com/eb/article?eu=70823>>(03.02.2004)

- 10 Engesser, Hermann; Volker Claus; Andreas Schwill. 1988. Duden Informatik. Mannheim; Wien; Zürich: Dudenverlag. S. 425.
- 11 Engle WD, Jackson GL, Sendelbach D, Manning D, Frawley WH. Assessment of a transcutaneous device in the evaluation of neonatal hyperbilirubinemia in a primarily Hispanic population. Pediatrics 2002; 110: 61-67.
- 12 Fischer, Gerd. 1986. Lineare Algebra. Braunschweig; Wiesbaden: Vieweg. S. 36.
- 13 Hirth, Christina. 1999. „NICHTINVASIVES OPTISCHES MAPPING UND SPEKTROSKOPIE ZUR FUNKTIONELLEN UNTERSUCHUNG DES GEHIRNS Räumliche, zeitliche und physiologische Aspekte lokaler Veränderungen der Blutoxygenierung bei funktioneller Aktivierung“ <<http://dohost.rz.hu-berlin.de/dissertationen/medizin/hirth-christina/HTML/hirth.html>> (22.05.2004)
- 14 Hollas, J.Michael. 1995. Moderne Methoden in der Spektroskopie; Braunschweig; Wiesbaden: Vieweg. S. 27.
- 15 Klötzer, Hans-Martin 11.07.1992. „Methoden zur Bibliothekssuche im IR-Bereich“ <<http://www.hmkq.de/download/Dipl.pdf>> (23.05.2004)
- 16 Köchli, Hans P.; Brigitte Walz; Richard W. Köchli. 02.02.2001. „Absorptionsphotometrie“ <<http://www.biorama.ch/biblio/b40laban/b24meth/m04abs/abs310.htm>> (02.02.2004)
- 17 Kohlrausch, Friedrich. 1985. Praktische Physik : zum Gebrauch für Unterricht, Forschung u. Technik. Bd.1. 23.,neubearb. U. erw. Auflage. Stuttgart : Teubner. S. 601,602,604,605,612,614,615
- 18 Krause, Christian J.. 2003. „Das Verfahren der Photometrie“ <<http://www.unki.de/schulcd/physik/photom.htm>> (22.05.2004)

- 19 Maiwald, Michael. 25.07.2000. „Instrumentelle Analytik am ITT“
<<http://www.itt.uni-stuttgart.de/~maiwald/interessen/analytik.html>>
(02.02.2004)
- 20 Olympus Diagnostica GmbH. 2003. AU 640 Users Guide V6.
S. H31ff., S. H27ff., S. H41ff., S. H3.
- 21 Orłowski, Reiner; Joachim Kaczmarek. 1999. „Diodenspektrometer zur
Schadstoffanalyse in Flüssigkeiten“ <[http://www.fh-
luebeck.de/et/impulse/themen/diospekt.html](http://www.fh-luebeck.de/et/impulse/themen/diospekt.html)> (06/2001)
- 22 Radiometer A/S. Radiometer Kopenhagen ABL 500, Referenzhandbuch.
S. 1.6.5..
- 23 Rubaltelli FF, Gourley GR, Loskamp N, Modi N, Roth-Kleiner M, Sender A,
Vert P. Transcutaneous bilirubin measurement: a multicenter evaluation of a
new device. Pediatrics 2001; 107: 1264-1271.
- 24 Selzer, Paul. 1998. „Vorhersage von Infrarotspektren mittels neuronaler Netze
zur Identifikation organischer Verbindungen“ <[http://www2.chemie.uni-
erlangen.de/services/dissonline/data/dissertation/Paul_Selzer/html/einleit.doc1.
html](http://www2.chemie.uni-erlangen.de/services/dissonline/data/dissertation/Paul_Selzer/html/einleit.doc1.html)> (02.02.2004)
- 25 Sigrist-Photometer AG. 30.01.2004. „ABC der Prozessphotometrie“
<http://www.photometer.com/ge/abc/abc_061.htm> (02.02.2004)
- 26 Tissue, Brian M.. 2000. „Beer-Lambert Law“
<<http://www.chem.vt.edu/chem-ed/spec/beerslaw.html>> (02.02.2004)
- 27 Wenzel, Ulrich; Peter Dominiak; Hans-Hellmut Neumayer; Gunter Wolf.
Dtsch Arztebl 2003;100:A 2072-2079[Heft 31-32]

- 28 Wyatt JS, Cope M, Delphy DT, Richardson CE, Edwards AD, Wray S, Reynolds EO. Quantitation of cerebral blood volume in human infants by near-infrared spectroscopy. J Appl Physiol. 1990 Mar; 68(3):1086-91

Abkürzungen:

| | |
|--------|---------------------------------|
| ^ | Textzeichen für mathem. "hoch" |
| bzw. | beziehungsweise |
| EDV | Elektronische Datenverarbeitung |
| i.e. | id est |
| ISE | Ion Selectiv Elektrod |
| LGS | Lineares Gleichungssystem |
| LW | Leerwert |
| o.g. | oben genannt |
| OD | Optical Density |
| resp. | respektive |
| s. | siehe |
| u.a. | unter anderem |
| u.ä. | und ähnliches |
| u.s.w. | und so weiter |
| u.v.m. | und verschiedenes mehr |
| z.B. | zum Beispiel |

Tabelle 10-1 : Abkürzungsverzeichnis

11 Danksagung

Herrn Prof. Dr. med. Bayer danke ich für die Überlassung des Themas und den Enthusiasmus, mit dem er mein Interesse für dieses Sujet zu vertiefen wusste. Frau Roth hat mir nicht nur die Daten zur Verfügung gestellt, sondern auch den richtigen Umgang mit der notwendigen Auswertungssoftware vermittelt. Für sein Korrekturlesen und die humorvollen konstruktiven Anmerkungen danke ich Herrn Prof. Dr. rer. pol. Joachim Henkel.

12 Lebenslauf

Angaben zur Person:

Name: Andreas Fischer
 Wohnort: 40237 Düsseldorf
 Schumannstr. 73
 Geburtstag/-ort: 15.04.1965 in Mülheim an der Ruhr
 Nationalität: deutsch

Ausbildungsdaten:

1971 - 1975 Städt. Grundschule in Mülheim an der Ruhr, Trooststraße
 1975 - 1984 Staatl. Gymnasium Otto-Pankok in Mülheim an der Ruhr
 1984 - 1993 Studium der Humanmedizin an der Heinrich-Heine
 Universität in Düsseldorf und an der Medizinischen Universität
 zu Lübeck. Vorlesungen der Mathematik und Physik an der
 Heinrich Heine Universität in Düsseldorf
 November 1993 Staatsexamen in Humanmedizin an der
 Heinrich-Heine-Universität in Düsseldorf
 1994 - 1996 Arzt im Praktikum in der Abteilung für innere Medizin
 St. Vinzenz Hospital Düsseldorf
 1996 - 1999 Assistenzarzt in der Herz- und Thoraxchirurgie Herzzentrum
 Duisburg
 1999 - 2003 Assistenzarzt für Innere Medizin in der Medizinischen Klinik
 des Städtischen Krankenhauses Düsseldorf-Gerresheim
 2003 - 2005 Assistenzarzt Kardiologie in der Klinik für Kardiologie
 St.-Vinzenz-Krankenhaus Essen
 Januar 2005 Facharztprüfung Innere Medizin

COMPUTERASSISTIERTER ERKENNUNG VON ABSORPTIONSSPEKTREN FÜR DIE LABORMEDIZINISCHE DIAGNOSTIK

ABSTRACT

HINTERGRUND: Anhand der vorliegenden Arbeit sollte der Versuch unternommen werden, Spektren bekannter Labortests mittels mathematischer Verfahren zu beschreiben, so dass in einem zweiten Schritt Spektren mit unbekannter Substanzkonzentration qualitativ den bekannten Spektren zugeordnet werden können und deren Konzentrationen im Testansatz festgestellt werden können. **METHODE:** Die Daten von sieben bekannten medizinischen Labortests wurden für einen Wellenlängenbereich von 400 nm bis 800 nm in verschiedenen Konzentrationen aufgenommen. Es wurden drei mathematische Modelle verwendet. Das Erste setzt lediglich einen linearen Verlauf zwischen Spektren benachbarter Konzentrationen voraus. Das Zweite verlangt eine monoton steigende Beziehung zwischen dem Integral eines Spektrums und der zugrunde liegenden Konzentration. Das Dritte geht von einer linearen Beziehung zwischen Spektren und Konzentration über den gesamten Konzentrationsbereich aus. **RESULTATE:** Alle Algorithmen sind in der Lage, die Tests zu erkennen, wobei der erste Algorithmus, die Interpolationsmethode, die stabilsten Ergebnisse liefert. Unter Verwendung dieses Algorithmus können die Tests sicher erkannt werden, auch dann, wenn die Wellenlängenauflösung gering ist oder die Beziehung zwischen Extinktion und Konzentration nicht linear ist. Zur Ermittlung der Konzentration einer Testsubstanz bietet das Verfahren die Berechnung der Konzentration anhand der Extinktionskoeffizienten an bestimmten Wellenlängen und anhand des gesamten Spektrums an. **SCHLUSSFOLGERUNG:** Vorausgesetzt, dass das vorhandene Testkontingent sich erheblich erweitern lässt, erscheint die Identifikation des Spektrums geeignet, den Messablauf in einem Praxislabor einfacher und sicherer zu gestalten. Eine Anwendung dieser Verfahren im Großlabor kann erheblich zur Güte der Qualitätskontrolle beitragen