

## Fortsetzung Optimierungsverfahren

Letzte Woche Gradientenbasierte Verfahren,  
jetzt:

### Methoden basierend auf Newton Verfahren

Bislang haben wir nur die Ableitung erster  
Ordnung zur Richtungsbestimmung genutzt,  
jetzt eine Ordnung höher (Taylorentwicklung)

$$f(x) = f(x_0) - \underset{\substack{\uparrow \\ \text{negative} \\ \text{Gradient}}}{g} \cdot (x - x_0) + \frac{1}{2} (x - x_0) \cdot \underset{\substack{\sim \\ \text{Hesse'sche} \\ \text{Matrix}}}{H} \cdot (x - x_0)$$

Die erste Ableitung muss bei Minimum  
verschwinden, Approximation ableiten  
und 0 setzen.

$$\begin{aligned} f'(x) \Big|_{x=x_0} &= -g + \frac{1}{2} H \cdot (x - x_0) + \frac{1}{2} (x - x_0) \cdot H \\ &\quad H \text{ symmetrisch} \\ &= -g + \frac{1}{2} H \cdot (x - x_0) = 0 \end{aligned}$$

$$\Rightarrow \frac{1}{2} H \cdot (x - x_0) = g$$

$$\Leftrightarrow (x - x_0) = \frac{1}{2} H^{-1} \cdot g \quad (\text{vorausgesetzt } H^{-1} \text{ existiert})$$

Diese Gleichung  $\Delta x$  erlaubt die Bestimmung der Richtung  
 $x - x_0$  zu zum Minimum.

1) Zur Illustration, wird  $\Delta x$  in das Koordinatensystem  
der Eigenvektoren zu  $H$  überführt

$$\Delta x' = (\Delta x'_1, \dots, \Delta x'_i, \dots, \Delta x'_n)$$

$$\Delta x_i' = \frac{f_i}{\epsilon_i} \leftarrow \begin{array}{l} \text{Projektion des Gradienten entlang des} \\ \text{Eigenvektors} \end{array}$$

Positiv in der Nähe ein Minimum.

Falls ein negativer Wert dabei  $\Rightarrow$  Sattelpunkt

2) Nähe ein Minimum kein Eigenwert klein werden...  
 $\Rightarrow$  Mehrere Schritte!

In dem Fall wird teilweise ein Vertrauensradius  
 für die Stepsize vorgegeben, falls der Step zu groß  
 $\Rightarrow$  Line search.

3) Die Schrittweite sollte den Bereich der  
 Gültigkeit der Taylornäherung nicht verlassen.

Falls Hessian negative Eigenwert hat, hilft ein  
 Shiftoperator

$$\Delta x' = -\frac{f_i}{\epsilon_i - \lambda}$$

$\leftarrow$  wenn es der kleinste Eigenwert ...

4) Es ist sehr aufwendig, die Hessian Matrix  
 zu bestimmen.

In dem Fall gibt es iterative Methoden, die eine  
 Näherung der Hesse'schen Matrix ermöglichen.

z. B. BFGS-Methode

$$\underline{H}_0 = Id \Rightarrow \text{erster Schritt} \Rightarrow \text{Steepest Descent} \\ \text{Line Search.}$$

$$\underline{H}_{k+1} = \underline{H}_k + \underline{U}_k + \underline{V}_k$$

wobei  $y = g_{k+1} - g_{k+1}$  und  $s_k = x_{k+1} - x_k$

mit 
$$\underline{U}_{k+1} = \frac{y_k \cdot y_k^T}{y_k \cdot s_k} \quad \underline{V}_{k+1} = - \frac{H_k \cdot s_k \cdot s_k^T \cdot H_k}{s_k^T \cdot H_k \cdot s_k}$$

Die Entwicklung kann ohne die Anwendung einer Matrix ausgesetzt werden. (Formel ohne Herleitung)

$$\underline{H}_{k+1}^{-1} = \underline{H}_k^{-1} + \frac{(s_k^T \cdot y_k + y_k^T \cdot H_k^{-1} \cdot y_k)(s_k - s_{k+1})}{(s_k^T \cdot y_k)^2} - \frac{s_k \cdot H_k^{-1} \cdot y_k + y_k \cdot H_k^{-1} \cdot s_k}{s_k^T \cdot y_k}$$

Wichtig: Hier muss keine Matrix gespeichert werden!

Meist ein guter Wahl, die Besetzung der H Matrix zu vermeiden.

Beispiel für Gradientenfreie Methode:  
Nelder-Mead

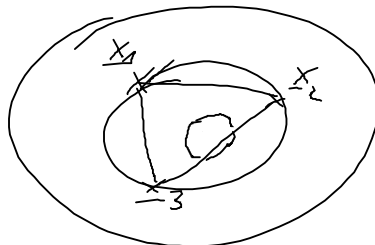
Manchmal ist die Besetzung eines Gradienten zu aufwendig, dann z.B. Nelder-Mead verwenden.

Idee: Wenn wir ein N-Dimensionales Problem haben, dann wird es ein Simplex von N+1 Dimensionen geben:  
z.B. N=2 => Dreieck

Die Vektoren sind sortiert, d.h.

$$\underline{x}_1, \underline{x}_2, \underline{x}_3, \dots, \underline{x}_{N+1} \text{ gilt}$$

$$f(\underline{x}_1) \leq f(\underline{x}_2) \leq \dots \leq f(\underline{x}_{N+1})$$



Der schlechteste Wert bei jeder Iteration ersetzt:

$$x_{N+1}^{(k+1)} = (1+\mu) \frac{1}{N} \sum_{i=1}^N x_i - \mu x_{N+1}$$

Das wird für verschiedene Werte von  $\mu$  versucht.

$\mu = \mu_0, \mu = 2\mu_0, \frac{1}{2}\mu_0, -\frac{1}{2}\mu_0$  dies ergibt eine Verbesserung  $\Rightarrow$  stop!

Entspricht Reflexion, Expansion, Kontraktion ...

(Achtung, kann auch zu nicht minimalem Konvergenz  $\Rightarrow$  Problem)

### Allgemeine Bemerkung zu Optimierung

1) Das Koordinatensystem spielt eine wichtige Rolle. Die richtige Koordinate kann die Konvergenz beschleunigen.

Beispiel: 
$$F_{\text{Morse}} = \left[ 1 - e^{-x} \right]^2$$

Iteration in  $x$ ,  $y = e^{-x}$  und  $z = e^x$

mittel schnell (5. Iteration)	schnell (2. Iteration)	langsam (7. Iteration)
----------------------------------	---------------------------	---------------------------

Genauso spielt auch die richtige Barriere eine Rolle.

2.) Oft gibt es auch Nebenbedingungen!

1) Man addiert zu zu optimierender Funktion  $C[x]$ , ein Kostenfunktion  $C_{NB}[x]$  hinzu, die Abweichen der Nebenbedingung mit hohem Funktionswert bestraft und optimiert  $C[x] + C_{NB}[x]$

Nachteil: Nebenbedingung werden wir  
Approximativ eingehalten.

2) Lagrange Multiplikatoren Nebenbed.

$$L(x) = f(x) - \sum_i \lambda_i (g_i(x) - c_i)$$

$\lambda_i$  wird zusätzlich optimiert und  
 $L$  minimiert. Optimal; physikalisch relevant.

(Tatsächlich ist das System unterbestimmt  
und  $i$  Dimensionen müsste entfernt werden,  
ist aber zu kompliziert.)

3.) Projektionsmethode.....

IV.11 Zeitliche Dynamik: Schrödungsgl.,

Blockgleichung, Dichtematrix Theorie

Bislang haben wir uns vorrangig Eigenwertprobleme  
und stationäre Probleme angesehen.

Jetzt schauen wir uns zeitabhängige Probleme  
an.

1. Beispiel: zeitabhängige Schrödungsgl. in  
Raum diskretisiert.

$$i\hbar \partial_t |\psi\rangle = H |\psi\rangle$$

z.B. für unser Exziton Problem.

$$H = H_0 + H_{el-Licht}$$

bisherige Exzitonproblem  
ohne Bewegung + (potential)      Wechselwirkung mit Licht  
(lineare Optik)

$$H_{el-L} = \delta(r_e - r_n) \underline{d} \cdot \underline{E}(r_n) + \dots$$

$$= \delta(r) \underline{d} \cdot \underline{E}$$

↑  
räumlich homogen Anregung

### Bewegungsgl.

$$i\hbar \partial_t \psi(r) = \left( -\frac{\hbar^2}{2m} \Delta + V_{ext}(r) \right) \psi(r, t) + \delta(r) \underline{d} \cdot \underline{E}(t) + \dots \text{Rephasing}$$

Delta Funktion, kann durch Zerlegen in Intervalle und Integrieren gelöst werden.

Damit ist es z.B. möglich, die zeitliche Dynamik bei optischer Anregung zu berechnen.

Vorteil: im Vergleich zu Eigenwertproblem nur  $\psi$  (unabhängig jener Zustände) muss in Spalten erhalten werden. Zeitlösung löst Eigenwertproblem für relevanten Zustände implizit.

Nachteil: Eigenwerte sind meist nicht bekannt.

Kann direkt in Nichtlinearer Optik erweitert werden (z.B. Biexzitonen, Rephasierung z.B. Quantenprojektion und Jump-Operatoren).

2. Beispiel: Blochgleichungen (z.B. in der Dichtematrixtheorie)

$$H = H_0 + H_{el-L}$$

$$H_0 = \sum_n \epsilon_n |n\rangle\langle n|$$

↑  
Energie

↑  
Vielteilchenzustand

$$H_{el-L} = \sum_{nm} d_{nm} \underline{E}(t) |n\rangle\langle m|$$

↑  
Dipolmoment.

z. B. Zweiwirbelssystem

$$H_0 = \frac{1}{2} \epsilon_1 |1\rangle \langle 1| + \frac{1}{2} \epsilon_2 |2\rangle \langle 2|$$

$$H_{\text{ext}} = \frac{1}{2} d_{12} \cdot \underline{E}(r, t) |1\rangle \langle 1| + \frac{1}{2} d_{21} \cdot \underline{E}(r, t) |2\rangle \langle 1|$$

Verwende Liouville von Nambu gl.:

$$\partial_t \rho = -\frac{i}{\hbar} [H, \rho] \quad P = d_{12} \frac{\langle 1| \rho |2\rangle}{\rho_{12}}$$

Blodgleichung:

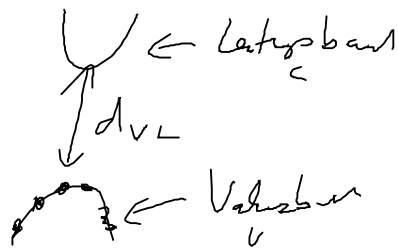
$$\partial_t \rho_{12} = -i(\epsilon_1 - \epsilon_2) \rho_{12} - i E \cdot d_{21} (\rho_{12} - \rho_{21})$$

$$\begin{aligned} \partial_t \rho_{11} &= -i E \cdot d_{21} \rho_{21} + i E \cdot d_{12} \rho_{12} \\ &= 2 \operatorname{Im}(E \cdot d_{21} \rho_{21}) \end{aligned}$$

Beispiel für Nichtlinear Gleichung in  $\underline{E}$ , aber auch das ist in Prinzip linear Problem.

### 3. Beispiel: Halbleiter Blod Gleichung

2. Bandstruktur



Freie Bewegung

$$H_0 = \sum_k \epsilon_{vk} a_{vk}^\dagger a_{vk} + \sum_{lk} \epsilon_{lk} a_{lk}^\dagger a_{lk}$$

$$H_{\text{ext}} = \sum_k d_{ll} \cdot \underline{E} a_{vk}^\dagger a_{lk} + \text{h. a.} + \text{Coulomb}$$

$$H_{\text{Coulomb}} = \sum_{\substack{l, l' \\ l \neq l'}} V_q a_{l, l+q}^\dagger a_{l', l-l}^\dagger a_{l', l} a_{l, l+q}$$

(Interband wie Interband Prozess)

$$P(t) = \sum_n P_n d_n \quad P_n = \langle a_{c, l}^\dagger a_{v, l} \rangle$$

$\Rightarrow$  Halbleiterblockglieder (im kohärenten Limit, Heitler-Fold)

$$\partial_t P_k = -\frac{i}{\hbar} (\epsilon_{v, k} - \epsilon_{c, k} + i\gamma) P_k + \frac{i}{\hbar} \left( \sum_{q \neq 0} V_q (f_{c, k+q} - f_{v, k+q}) P_k \right) + \frac{i}{\hbar} \left( d_{v, k} E(t) + \sum_{q \neq 0} V_q P_{k+q} \right) (n_{c, k} - n_{v, k})$$

Stoerterm z. B.

$$n_1 n_2 (1 - n_3) (1 - n_4)$$

$\Rightarrow$  Mittelwert.