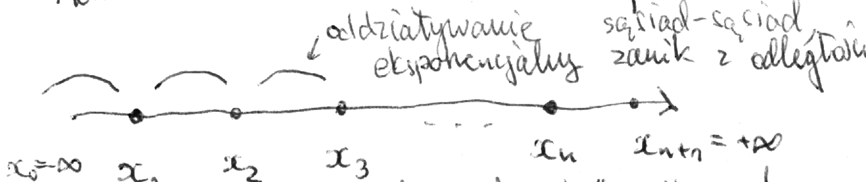
„O potokach diagonalizujących macierze" 31.08.2011
(Tomasz Tkocz, „Wisełka 2011")

Celem tego referatu jest przedstawienie metody diagonalizacji macierzy symetrycznych za pomocą potoku. Zrobimy to na dwóch przykładach.

Pierwszy z nich ma interesującą motywację fizyczną. Rozważmy n cząstek umieszczonych na prostej w punktach $x_1, \dots, x_n$. Hamiltonian układu ma postać



$$H = \underbrace{\sum_{k=1}^{n} \frac{1}{2}(\dot{x}_k)^2}_{T} + \underbrace{\sum_{k=1}^{n-1} e^{x_k - x_{k+1}}}_{V}.$$

Rys. "the truncated Toda lattice" – natura oddziaływania jest bilardowa.

Ustalamy warunki początkowe ($x_k(0) = \dots$, $\dot{x}_k(0) = \dots$) i oczywiście pytamy jaka będzie ewolucja układu. W szczególności interesuje nas zachowanie się układu w nieskończoności, $t \to \infty$. Może istnieje jakiś stan stacjonarny?

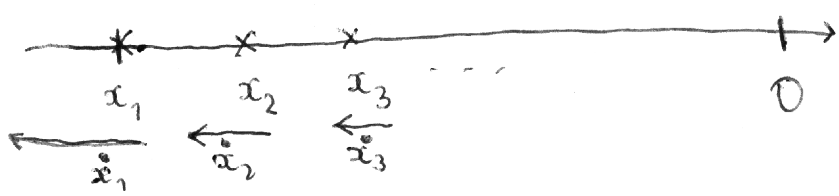
Intuicja fizyczna jest taka, że nasz układ dąży przy $t \to \infty$ do minimalizacji energii oddziaływania ($H = \text{const}$).

$$V \longrightarrow 0$$

Ma to miejsce, gdy

$$x_k - x_{k+1} \longrightarrow -\infty,$$

czyli np. gdy $\dot{x}_k \to u_k$, $|u_1| > |u_2| > \dots > |u_n|$, $u_k < 0$



(jest dryf – potok w lewo z oddalaniem się cząstek)

Spróbujmy teraz uściślić tę heurezę. Mając hamiltonian wypisujemy równania ruchu

$$(*) \quad \begin{cases} \ddot{x}_1 = -e^{x_1 - x_2} \\ \ddot{x}_k = -\frac{\partial}{\partial x_k} V = -e^{x_k - x_{k+1}} + e^{x_{k-1} - x_k} \\ \ddot{x}_n = e^{x_{n-1} - x_n} \end{cases}$$

Przepiszemy je teraz trikowo. Określamy

$$a_k := -\tfrac{1}{2}\dot{x}_k, \qquad k \leq n$$

$$b_k := \tfrac{1}{2} e^{(x_k - x_{k+1})/2}, \qquad k < n.$$

i stwierdzamy, że (*) jest równoważna

$$(**) \begin{cases} \dot{a}_k = -\tfrac{1}{2}\ddot{x}_k = 2(b_k^2 - b_{k-1}^2), & b_0 \equiv b_n \equiv 0 \\ \dot{b}_k = b_k(a_{k+1} - a_k) & (\leftarrow \text{def. } a_k, b_k). \end{cases}$$

Teraz dzieje się cud, bo powyższe można zapisać macierzowo

$$L := \begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & b_3 & \\ & & \ddots & \ddots & b_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix}, \quad B := \begin{bmatrix} 0 & b_1 & & & \\ -b_1 & 0 & b_2 & & \\ & -b_2 & 0 & b_3 & \\ & & \ddots & \ddots & b_{n-1} \\ & & & -b_{n-1} & 0 \end{bmatrix}$$

$$(***) \quad \dot{L} = [B, L].$$

[Pierwszy zauważył to Flaschka, [Fl 74]. Rozwiązanie które zaraz zobaczymy pochodzi od Mosera, [Mo 75]. Artykuł numeryków Deifta, Nandy, Tomei, [DNT 81], dokładnie omawia równanie (***) podając dokładne oszacowania tempa zbieżności.]

Naszym celem jest udowodnić, że

$$b_k \approx e^{(x_k - x_{k+1})/2} \xrightarrow[t\to\infty]{} 0,$$

czyli, że

$$L(t) \xrightarrow[t\to\infty]{} \text{macierz diagonalna}.$$

Dokonamy tego w kilku prostych uwagach.

⚠ (1) (energia $\mathcal{H}$ jest stała)

$$\mathcal{H} = \sum \tfrac{1}{2}(\dot{x}_k)^2 + \sum e^{(x_k - x_{k+1})/2}$$
$$= 2\sum a_k^2 + 4\sum b_k^2 = 2\left(\sum a_k^2 + 2\sum b_k^2\right)$$
$$= 2\|L\|^2 = 2\,\mathrm{tr}\, L^T L.$$

Ale

$$(\mathrm{tr}\, L^T L)^{\bullet} = 2\,\mathrm{tr}\, L^T \dot{L} = 2\,\mathrm{tr}\, \underbrace{L^T}_{L}(BL - LB)$$
$$= 2\,\mathrm{tr}(LBL - LLB) = 0. \quad \square$$

(2) (istnieje n całek pierwszych ruchu $\frac{d}{dt}\lambda_1 = \dots = \frac{d}{dt}\lambda_n = 0$)

Istotnie, pokażemy że potok $\dot{L} = [B, L]$ jest izospektralny, tzn. zachowuje spektrum macierzy L_0. ~~Aby~~ to zobaczyć rozważmy rozwiązanie zagadnienia Cauchy'ego

$$\begin{cases} \dot{V} = BV \\ V(0) = I \end{cases}$$

Ewoluuje ono w macierzach ortogonalnych $O(n)$, bo

$$(V^T V)^{\bullet} = \dot{V}^T V + V^T \dot{V} = V^T \underbrace{B^T}_{-B} V + V^T B V$$
$$= 0.$$

Niech $L_0 := L(0)$. Zauważmy, że

$$(V L_0 V^T)^{\bullet} = \dot{V} L_0 V^T + V L_0 \dot{V}^T = B V L_0 V^T + V L_0 V^T B^T$$
$$= B(V L_0 V^T) - (V L_0 V^T) B$$
$$= [B, V L_0 V^T],$$

więc z jednoznaczności potoku

$$L(t) = V L_0 V^T;$$

w szczególności

$$\mathrm{Spec}(L(t)) = \mathrm{Spec}(L(0)). \quad \square$$
$$= \{\lambda_1, \dots, \lambda_n\}$$

całki pierwsze

(3) Jest zbieżność do stanu stacjonarnego – macierzy diagonalnej

$$L(t) \xrightarrow[t\to\infty]{\|\cdot\|} \mathrm{diag}(\{\lambda_1,\dots,\lambda_n\}).$$

Jak już wiemy, wystarczy uzasadnić, że

☺ $\forall_k \quad b_k(t) \xrightarrow[t\to\infty]{} 0$.

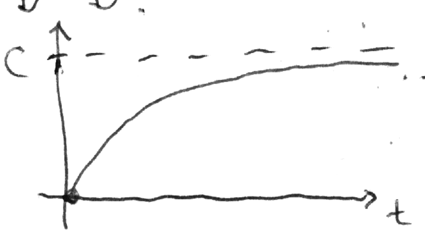
Zauważmy, że

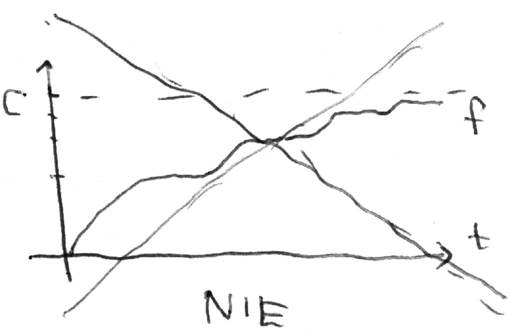
$$(\underbrace{a_1+\dots+a_k}_{S_k})^\circ = 2b_k(t)^2$$

Wobec (1), S_k, $\dot S_k$, $\ddot S_k = 4b_k\dot b_k = 4b_k^2(a_{k+1}-a_k)$ są ograniczone. Ponadto $\dot S_k \geq 0$. Teraz ☺ wynika z lematu.

LEMAT. Niech $f\colon [0,\infty) \xrightarrow{C^2} [0,C]$ ma ograniczone pochodne f', f''. Jeśli $f' \geq 0$, to $f' \xrightarrow[t\to\infty]{} 0$.

D-D.





Przypuśćmy przeciwnie, że istnieje $\varepsilon>0$ i $t_n \nearrow \infty$, że

$$f'(t_n) > \varepsilon.$$

Z jednostajnej ciągłości f' (wszak $|f''| \leq M$) istnieje δ, że

$$f'(t) > \varepsilon/2, \qquad |t-t_n|<\delta.$$

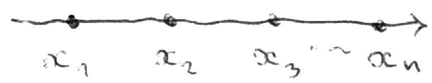
Ustalmy N i wycałkujmy tę nierówność po zbiorze $\bigcup_{n=1}^{N}(t_n-\delta, t_n+\delta)$

$$N\delta\varepsilon < f(t_N+\delta) \overbrace{- f(t_N-\delta) + f(t_{N-1}+\delta)}^{\leq 0} - f(t_{N-1}-\delta) + \dots - f(t_1-\delta) < f(t_N+\delta) - f(t_1-\delta)$$

f niemalejąca!

i gdy $N\to\infty$ mamy sprzeczność z $f \leq C$. $\square$

Podsumujmy to, co tu się stało.



analiza układu n cząstek „kraty Tody" $\Longleftrightarrow$ potok w trójdiagonalnych macierzach symetrycznych

$$\dot{L} = [B(L), L]$$

para Laxa

$$\mathcal{H} = 2\|L\|^2 \approx \sum a_k^2 + \sum b_k^2$$

prędkości oddziaływanie

Pokazaliśmy, że

$$\begin{bmatrix} a & b & 0 \\ b & \ddots & \ddots \\ 0 & \ddots & \end{bmatrix} = L(t) \longrightarrow \text{diag},$$

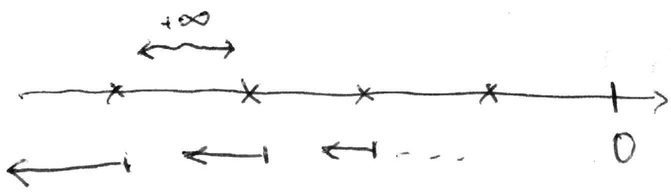
czyli

$$\sum b_k^2 \longrightarrow 0 \qquad \text{(oddziaływanie znika)}$$

$$-\tfrac{1}{2}\dot{x}_k = a_k \longrightarrow \lambda_k \qquad \text{(prędkości się ustalają)}$$

Jeśli więc ustalimy wartości początkowe tak, aby

$\lambda_1 > \dots > \lambda_n > 0$, będziemy mieli dryf układu w lewo dzięki naszemu eksponencjalnemu oddziaływaniu bilardowemu (sąsiad - sąsiad).



Na koniec można zapytać, czy podobnie nie udałoby się zdiagonalizować dowolnej macierzy symetrycznej, niż koniecznie trójdiagonalnej. Numeryk odpowiedziałby, że łatwo sprowadzić dowolną macierz symetryczną do trójdiagonalnej, potem odpaliłby omówiony potok i po kłopocie. My postąpimy inaczej.

Rozważmy potok

$$\begin{cases} \dot{L} = [L, [L, D]] \\ L(0) = L_0 \in MSym_{n\times n}(\mathbb{R}), \end{cases}$$

gdzie D jest ustaloną, niezależną od czasu macierzą diagonalną o wyrazach parami różnych. (Ten potok jest „stopnia" 3 a poprzednio mieliśmy zależność kwadratową. Pochodzi on z pracy Brocketta, [Br 91], gdzie omówiono szereg jego zastosowań – oprócz diagonalizacji może on sortować listy albo optymalizować pewne zagadnienia programowania liniowego.).

Dlaczego on działa? Rozumujemy identycznie jak poprzednio.

⚠ (1) $\|L\|^2 = \text{const}$,

bo się liczy $(\operatorname{tr} LL^T)^\circ = 0$ (przydaje się, że potok płynie w macierzach symetrycznych, co wynika z (2))

(2) spektrum jest zachowane,

bo się określa macierz diagonalizującą

$$\begin{cases} \dot{V} = V[V^T L_0 V, D] \\ V(0) = I \end{cases}$$

i sprawdza

- ortogonalność $(VV^T)^\circ = 0$
- $L := V^T L_0 V$ spełnia $\dot{L} = [L, [L, D]]$.

(3) $L \xrightarrow[t\to\infty]{\|\cdot\|}$ diag,

bo sztuczką tutaj jest zauważenie, że

$$(\operatorname{tr} LD)^{\cdot} = \|[L,D]\|^2 = \sum_{i\neq j} L_{ij}^2 (\lambda_i - \lambda_j)^2 .$$

$\lambda_i - \lambda_j \neq 0$

Zawężając więc temat do funkcji: $f(t) = \operatorname{tr} LD$ mamy, że jej pochodna, tzn. $\|[L,D]\|^2$ dąży do 0, co oznacza, że wyrazy niediagonalne L zbiegają do 0. $\square$

Dalej można by się zastanawiać nad tempem zbieżności L do macierzy diagonalnej. W pracy [DNT83] dla poprzedniego potoku

$$\dot{L} = [B, L]$$

pokazano oszacowania

$$|a_k(t) - \lambda_k| \leq c e^{-2\mu t},$$

$$b_k(t)^2 \leq c e^{-2\mu t}, \qquad c = c(L(0)),$$

$$\mu = \min_{k=1,\dots,n-1} \{\lambda_k - \lambda_{k+1}\}.$$

Te nierówności, jak i dla potoku $\dot{L} = [L,[L,D]]$ to jednak temat na inną historię.

LITERATURA

[Br91] R.W. Brockett, "Dynamical Systems That Sort Lists, Diagonalize Matrices, and Solve Linear Programming Problems", Linear Algebra and Its Applications 146 (1991), 79-91.

[DNT83] P. Deift, T. Nanda, C. Tomei, "Ordinary Differential Equations and the Symmetric Eigenvalue Problem", SIAM J. Numer. Anal., 20 (1983), 1-22.

[Fl74] H. Flaschka, "The Toda lattice, I", Phys. Rev. B, 9 (1974), 1924-1925.

[Mo75] J. Moser, "Finitely Many Mass Points on the Line Under the Influence of an Exponential Potential - An Integrable System", in Dynamical Systems Theory and Applications, New York - Berlin - Heidelberg, 1975, 467-497.

PODZIĘKOWANIA dla Arkadiusza Trawińskiego za wprowadzenie mnie w tą tematykę.