Analog Sorting

Theory and evaluation

Lusa Zhan
Columbia University
Department of Computer Science
Department of Mathematics
lz2371@columbia.edu

Yipeng Huang
Columbia University
Department of Computer Science
yipeng@cs.columbia.edu

Table 1: Frequency of Special Characters

Non-English or Math	Frequency	Comments
Ø	1 in 1,000	For Swedish names
π	1 in 5	Common in math
\$	4 in 5	Used in business
Ψ_1^2	1 in 40,000	Unexplained usage

ABSTRACT

Keywords

linear algebra; ordinary differential equations; algorithms

1. INTRODUCTION

2. BACKGROUND

2.1 Classical sorting algorithms

Table for comparing important efficient sorting algorithms, for example quick & merge sort.

How does analog sorting compare to other sorting algorithms in terms of complexity?

2.2 The QR algorithm

The analog sorting method we show is related to the classical QR algorithm. While the QR algorithm is a discrete algorith operating step-by-step, the analog sorting algorithm does so in continuous time [5, 3, 4].

The QR algorithm finds the eigenvalues and eigenvectors of a square matrix. The eigenvalue problem is as follows:

$$A_0 x = \lambda x$$

The QR algorithm operates as follows: Given the QR decomposition, each step proceeds as:

$$A_k = Q_k R_k$$

$$A_{k+1} = R_k Q_k$$

© 2016 ACM. ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

A side effect of the QR algorithm is that the eigenvalues of the original matrix end up in sorted order along the diagonal. This is a useful property in eigenvalue problems, where users of the algorithm are interested in finding the largest few eigenvalues. But few researchers have pointed out that this may be useful in itself, for sorting. Other algorithms for finding the eigenvalues and eigenvectors include the Jacobi eigenvalue algorithm, and the divide-and-conquer algorithm. It seems that the QR algorithm is the algorithm among these that sorts the eigenvalues.

2.3 Hamiltonian systems

The finite Toda lattice system of ODEs belongs to special class of ODEs called Hamiltonian systems. Hamiltonian systems are an important and efficient way to describe classical mechanics. Because of their importance, special ODE solvers called symplectic solvers have been developed specifically to solve Hamiltonian systems.

A Hamiltonian system is characterized by a total energy scalar $\hat{\mathcal{H}}$. The components of the system are described by vectors p for momenta and q for positions. The system obeys the laws of motion:

$$dp/dt = -dH/dq = f(q)dq/dt \qquad = dH/dp = g(p) \qquad (1)$$

Conceptually, we encode the keys we like to sort as the momenta of particles. The momenta of the particles carries them to a final state that represents the sorted system.

2.4 Toda lattice, double bracket

The finite Toda lattice is an ODE system that is related to the QR algorithm. If you plot the evolution of the Toda lattice ODE with respect to time, the values of the ODE at integer time steps is the intermediate states of the QR algorithm. The common property of the finite Toda lattice and the QR algorithm is that both preserve the eigenvalues of the matrix they operate on [1].

The finite Toda lattice is a Hamiltonian system with the form:

$$\hat{\mathcal{H}}(p,q) = \frac{1}{2}\sum_{1}^{n} p_k^2 + \sum_{1}^{n-1} exp(q_k - q_{k+1})$$

This basic form can changed in two ways, using change of variables, into the equations for analog sorting described in [2].

The central idea of analog sorting as described in the papers is using the Toda Lattice, which is described using the Hamiltonian

$$H(p,q) = \frac{1}{2}\sum_{1}^{n} p_k^2 + \sum_{1}^{n-1} exp(q_k - q_{k+1})$$

A system of equations are associated with this, defined as the Hamiltonian equations (can be derived using total differential)

$$\dot{p_k} = -\frac{\partial H}{\partial q_k}$$

$$\dot{q_k} = \frac{\partial H}{\partial p_k}.$$

Calculating this, we get

$$\dot{p_k} = exp(q_{k-1} - q_k) - exp(q_k - q_{k+1})$$

$$\dot{q_k} = p_k$$

where $e^{x_0 - x_1} = e^{x_n - x_{n+1}} = 0$

2.5 Toda Lattice - ODE form

This system of ODEs can be changed to the system described in (http://hrl.harvard.edu/analog/) using a change of variables.

$$x_k = -\frac{1}{2}p_k$$

$$y_k = \frac{1}{2}e^{(q_k - q_{k+1})/2}$$

In that case,

$$\begin{aligned}\dot{x_k} &= -\frac{1}{2}\dot{p_k} \\ &= -\frac{1}{2}(exp(q_{k-1} - q_k) - exp(q_k - q_{k+1})) \\ &= -\frac{1}{2}(4y_{k-1}^2 + 4y_k^2) \\ &= 2y_k^2 - 2y_{k-1}^2\end{aligned}$$

and

$$\begin{aligned}\dot{y_k} &= \frac{1}{2}e^{(q_k - q_{k+1})/2}(\frac{\dot{q_k} - \dot{q}_{k+1}}{2}) \\ &= y_k\frac{p_k - \dot{p}_{k+1}}{2} \\ &= y_k(x_{k+1} - x_k)\end{aligned}$$

Taking into account the boundary conditions $y_0 = y_n = 0$ and we get the desired system of ODEs.

2.6 Toda Lattice - Jacobi Matrix

The connection to the Double Bracket notation can be made through the Jacobi Matrix form of the Toda Lattice. The Jacobi Matrix is given by

$$H = \begin{bmatrix} x_1 & y_1 & 0 & \dots & 0 \\ y_1 & x_2 & y_2 & \dots & 0 \\ & & \ddots & & \\ & & y_{n-2} & x_{n-1} & y_{n-1} \\ 0 & \dots & & y_{n-1} & x_n \end{bmatrix}$$

This is also the form that H is in for analog sorting.

In order to get the double bracket form, we need a diagonal matrix $N = diag(n, n-1, \dots, 1)$, so we need

$$N = \begin{bmatrix} n & 0 & \dots & 0 \\ 0 & n-1 & & \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & 1 \end{bmatrix}$$

We then get

$$HN = \begin{bmatrix} nx_1 & (n-1)y_1 & \dots & 0 \\ ny_1 & (n-1)x_2 & \dots & 0 \\ \vdots & & \ddots & 0 \\ & & 2x_{n-1} & y_{n-1} \\ 0 & \dots & 2y_{n-1} & x_n \end{bmatrix}$$

$$NH = \begin{bmatrix} nx_1 & ny_1 & \dots & 0 \\ (n-1)y_1 & (n-1)x_2 & \dots & 0 \\ \vdots & & \ddots & 0 \\ & & 2x_{n-1} & 2y_{n-1} \\ 0 & \dots & y_{n-1} & x_n \end{bmatrix}$$

$$HN - NH = \begin{bmatrix} 0 & -y1 & \dots & & 0 \\ y_1 & 0 & -y_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & y_{n-2} & 0 & -y_{n-1} \\ 0 & & \dots & y_{n-1} & 0 \end{bmatrix}$$

$$H[H,N] = \begin{bmatrix} y_1^2 & -x_1y_1 & -y_1y_2 & 0 & \dots & 0 \\ x_2y_1 & -y_1^2 + y_2^2 & -x_2y_2 & 0 & \dots & \\ y_1y_2 & \ddots & & \ddots & & 0 \\ \vdots & & & & y_{n-1}^2 - y_{n-2}^2 & -x_{n-1}y_{n-1} \\ 0 & \dots & & & x_ny_{n-1} & -y_{n-1}^2 \end{bmatrix}$$

$$[H,N]H = \begin{bmatrix} y_1^2 & -x_2y_1 & -y_1y_2 & 0 & \dots & 0 \\ x_1y_1 & y_1^2 - y_2^2 & -x_3y_2 & 0 & \dots & \\ y_1y_2 & \ddots & & \ddots & & 0 \\ \vdots & & & & -y_{n-1}^2 + y_{n-2}^2 & -x_ny_{n-1} \\ 0 & \dots & & & x_{n-1}y_{n-1} & y_{n-1}^2 \end{bmatrix}$$

Therefore, we get

$$[H,[H,N]] = \begin{bmatrix} 2y_1^2 & y_1(x_2 - x_1) & 0 & & \dots \\ y_1(x_2 - x_1) & -2y_1^2 + 2y_2^2 & y_2(x_3 - x_2) & 0 & \dots \\ 0 & \ddots & & \ddots & \\ \vdots & & & & 2y_{n-1}^2 - 2y_n^2 \\ 0 & \dots & & & y_{n-1}(x_n - x \end{bmatrix}$$

which is exactly $\dot{H}$

3. RELATED WORK

Below is a list of articles that are related to this topic:

- Bloch, Anthony M., and Alberto G. Rojo. "Sorting: The Gauss Thermostat, the Toda Lattice and Double Bracket Equations,." *Three Decades of Progress in Control Sciences* (2010): 35-48.

- Guseinov, Gusein Sh. "A Class of Complex Solutions to the Finite Toda Lattice." *Mathematical and Computer Modelling* 57.5-6 (2013): 1190-202.

- Tomei, Carlos. "The Toda Lattice, Old and New." *JGM Journal of Geometric Mechanics* 5.4 (2013): 511-30. Web.

3.1 Role of N

In order to understand the role of N fully, we need to consider *Theorem 1.5* in Helmke and Moore. To summarize for our purposes, the theorem states the following:

Given $H = [H, [H, N]]$ for each N, converges to an equilibrium as $t \to \infty$. If $N = \text{diag}(\mu_1, \ldots, \mu_n)$ where $\mu_1 > \cdots > \mu_n$, then the Hessian of $f_N(H) = \frac{1}{2}\|N - H\|^2$ is nonsingular and negative definite.

That means that the linearization at a critical point is

$$\dot{\xi}_{ij} = -(\lambda_{\pi(i)} - \lambda_{\pi(j)})(\mu_i - \mu_j)\xi_{ij}$$

where $\xi = [H_\infty, N]$. Since the Hessian is negative definite, ξ must be at the max, so H_∞ must be sorted.

3.2 Role of y_k

This section aims to explain the choice of initial values for the y_k's to start the algorithm.

First of all, as mentioned previously, $y_0 = y_n = 0$. To analyze the initial values of y_k, we consider the relation to $\dot{x}_k$. To recall,

$$\dot{x_k} = 2y_k^2 - 2y_{k-1}^2$$
$$\dot{y_k} = y_k(x_{k+1} - x_k)$$

If $y_k = 0$, then $\dot{x}_k$ would be 0, meaning $x_k(t) = x_k(0), \forall t$. This means that the values along the diagonal of H will remain constant and never get sorted.

In addition, the y_k values should be as small as possible (but still positive). Recall that the eigenvalues will be the values along the diagonal of H as $t \to \infty$. Consider the determinant $f_n = \det(H - \lambda I)$, where n is the size of the matrix. Since H is a tridiagonal matrix, this determinant can be formulated using a recurrence relation on the size of the matrix n.

$$f_n = (x_n - \lambda)f_{n-1} - y_{n-1}^2 f_{n-2}$$
$$f_1 = x_1 - \lambda$$

We can use this to calculate the first few f_n and set them to 0 (what we would do to calculate the eigenvalues):

$$f_1 = x_1 - \lambda = 0$$
$$f_2 = (x_2 - \lambda)(x_1 - \lambda) - y_1^2 = 0$$
$$f_3 = (x_3 - \lambda)(x_2 - \lambda)(x_1 - \lambda) - (x_3 - \lambda)y_1^2 - (x_1 - \lambda)y_2^2 = 0$$

Intuitively, the eigenvalues will be closest to x_k if the y_k values are smaller.

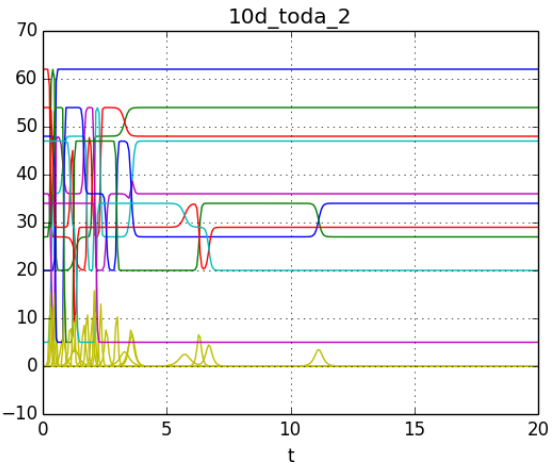

Figure 1: Simulated sorting a 10 element vector.

3.3 Realizing the analog sorter

After reviewing the mathematics of analog sorting, we now discuss how the analog sorter is put in practice.

The analog sorter works as follows: We construct matrix H(t=0) with the real numbers we want to sort on the diagonal. Naturally H(t=0) has eigenvalues consisting of the same real numbers. We set up a special ODE that involves the vector x, and another vector consisting of the natural numbers. This vector of the natural numbers provides the "discreteness" for the algorithm. Specifically, this ODE is the finite Toda lattice ODE, which preserves the eigenvalues of H(t), while reordering the elements on the diagonal to the sorted sequence. This solve this ODE on an analog computer. The final steady state of the analog output would have the original elements of the vector x, but now in sorted order. For example, the first integrator would have the lowest magnitude element of x.

There are some variations to this idea. One is we can tackle the finite Toda lattice system directly as a Hamiltonian system. Another is we can reverse the roles of the sort keys and the indices. This is described by Brockett briefly in page 802 of [2].

4. METHODOLOGY

We validate the functionality of analog sorting using a prototype analog computer chip. In order to explore how analog sorting works in larger systems, we use an ODE solver built on the odeint solver in the Python SciPy library.

5. EVALUTION

How does analog sorting compare to other sorting algorithms for example quick, merge etc in terms of complexity?

5.1 Functionality validation of analog sorting

We see in this picture that the sorting system has periods of swapping, interspersed in quiet periods of little change.

5.2 Time cost of the discrete QR algorithm

Time cost of overall QR loop: how many iterations of qr til convergence? Since we know the ODE is analogous to QR algorithm, they should take the same amount of time. Time

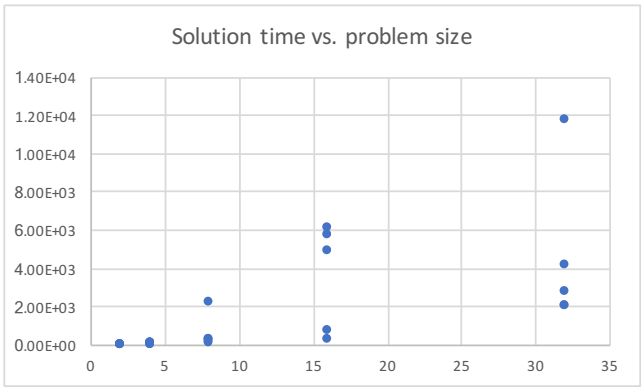

Figure 2: Preliminary time to convergence vs. problem size.

cost of QR step: Numerical Recipes 3rd Edition p585 says the QR algorithm takes O(N) time for symmetric tridiagonal matrices.

5.3 Time cost of analog sorting

In terms of time, the sorter takes at least O(N) time because of the time it takes just for signals to propagate across the circuit. Another issue is the time it takes for the ODE to settle to its final value. This is preliminary data showing the time to convergence of analog sorting, plotted against problem size.

The axes are the time to solution, plotted against N, the number of real numbers we are sorting. The data points come from multiple random trials at each N. The set of real numbers is randomly generated from a chosen dynamic range. It appears that as the N size increases, the average time grows linearly with respect to N. Furthermore, the variance of the solution time, measured as the standard deviation, also grows.

Other variables to consider include:

1. the dynamic range of real numbers to sort
2. the initial values of the off-diagonals

5.4 Hardware cost of analog sorting

The analog sorter takes up O(N) amount of circuit components to sort N elements.

6. APPLICATIONS & CONCLUSIONS

In this work, we validate that analog sorting works in practice. While this idea has been presented several times in mathematical literature, this is the first where it is tested in practice, using analog computer hardware. Using real measurements and simulated experiments, we evaluate the performance and efficiency of analog sorting. We compare this technique against classical sorting algorithms.

In terms of concrete impact of our findings, analog sorting can be a new way to sort in future computer architectures. Sorting is the underpinning of big data. Increasingly, approximate sorts are useful.

In terms of the impact of this work on the future of analog computer techniques, we show in this paper two key ideas. One, using the soliton behavior of the finite Toda lattice, we are able to swap two values in an analog computer without having to use a third capacitor as a buffer. Two, it is possible to perform a typically "discrete" task using continuous-time hardware. From this discrete primitive operation, we can build other discrete algorithms.

7. ACKNOWLEDGMENTS

This section is optional; it is a location for you to acknowledge grants, funding, editing assistance and what have you. In the present case, for example, the authors would like to thank Gerald Murray of ACM for his help in codifying this *Author's Guide* and the **.cls** and **.tex** files that it describes.

8. REFERENCES

[1] A. M. Bloch and A. G. Rojo. *Sorting: The Gauss Thermostat, the Toda Lattice and Double Bracket Equations,*, pages 35–48. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[2] R. W. Brockett. Dynamical systems that sort lists, diagonalize matrices and solve linear programming problems. In *Proceedings of the 27th IEEE Conference on Decision and Control*, pages 799–803 vol.1, Dec 1988.

[3] M. T. Chu. On the continuous realization of iterative processes. *SIAM Review*, 30(3):375–387, 1988.

[4] M. T. Chu. A list of matrix flows with applications. In *in Hamiltonian and Gradients Flows, Algorithms and Control*, pages 87–97, 1994.

[5] P. Deift, T. Nanda, and C. Tomei. Ordinary differential equations and the symmetric eigenvalue problem. *SIAM Journal on Numerical Analysis*, 20(1):1–22, 1983.

APPENDIX