Application of differential elimination to biological modelling using BLAD

François Boulier
Équipe calcul formel,
Laboratoire d'Informatique Fondamentale de Lille
Université Lille 1

Linz, May 11th, 2006

Differential equations : simplification of systems

The following DAE (Hairer, Wanner II) has derivation index 2.
The unknowns are three functions $x(t)$, $y(t)$ and $z(t)$.

$$\left\{\begin{array}{rcl} \dot{x}(t) & = & 0.7 \cdot y(t) + \sin(2.5 \cdot z(t)) \\ \dot{y}(t) & = & 1.4 \cdot x(t) + \cos(2.5 \cdot z(t)) \\ 1 & = & x^2(t) + y^2(t). \end{array}\right.$$

There is a missing ODE for numerical integration : $\dot{z}(t) = ???$.

Differential elimination software permit to extract it automatically

Differential elimination methods require polynomial systems

$$\left\{\begin{array}{rclrcl} \dot{x}(t) & = & 0.7 \cdot y(t) + s(t) & \dot{s}(t) & = & 2.5 \cdot \dot{z}(t) \cdot c(t) \\ \dot{y}(t) & = & 1.4 \cdot x(t) + c(t) & \dot{c}(t) & = & -2.5 \cdot \dot{z}(t) \cdot s(t) \\ 1 & = & x^2(t) + y^2(t) & 1 & = & s^2(t) + c^2(t). \end{array}\right.$$

A demo of Rosenfeld–Gröbner using the MAPLE diffalg package.

Differential equations : simplification of systems

The following DAE (Hairer, Wanner II) has derivation index 2. The unknowns are three functions $x(t)$, $y(t)$ and $z(t)$.

$$\begin{cases} \dot{x}(t) &= 0.7 \cdot y(t) + \sin(2.5 \cdot z(t)) \\ \dot{y}(t) &= 1.4 \cdot x(t) + \cos(2.5 \cdot z(t)) \\ 1 &= x^2(t) + y^2(t). \end{cases}$$

There is a missing ODE for numerical integration : $\dot{z}(t) = ???$.

Differential elimination software permit to extract it automatically

Differential elimination methods require polynomial systems

$$\begin{cases} \dot{x}(t) &= 0.7 \cdot y(t) + s(t) & \dot{s}(t) &= 2.5 \cdot \dot{z}(t) \cdot c(t) \\ \dot{y}(t) &= 1.4 \cdot x(t) + c(t) & \dot{c}(t) &= -2.5 \cdot \dot{z}(t) \cdot s(t) \\ 1 &= x^2(t) + y^2(t) & 1 &= s^2(t) + c^2(t). \end{cases}$$

A demo of Rosenfeld–Gröbner using the MAPLE diffalg package.

diffalg versus BLAD

diffalg

I wrote the first version for MAPLE in 1995 while a postdoc at the university of Waterloo.
Much improved by E. Hubert, A. Wittkopf, F. Lemaire.
It is designed for interactive use. Not so easy to manipulate :

- how to choose rankings ?
- terrific worst case complexity.

The *Bibliothèques Lilloises d'Algèbre Différentielle*

First version in 2004 http ://www.lifl.fr/~boulier/BLAD.
Started in 2000. Four versions thrown away.
Open source C libraries (LGPL). 36000 lines of C code.
Designed to be software components.
Provide functionnalities to overcome the above problems.
A less fancy demo of Rosenfeld–Gröbner using BLAD.

diffalg versus BLAD

diffalg

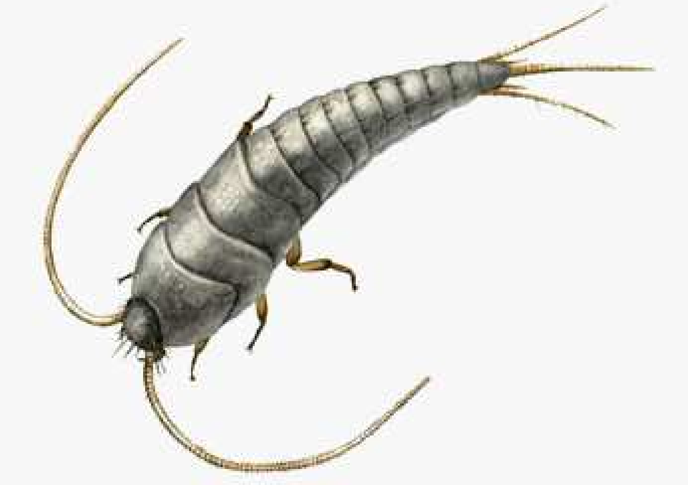
I wrote the first version for MAPLE in 1995 while a postdoc at the university of Waterloo.
Much improved by E. Hubert, A. Wittkopf, F. Lemaire.
It is designed for interactive use. Not so easy to manipulate :

- how to choose rankings ?
- terrific worst case complexity.

The *Bibliothèques Lilloises d'Algèbre Différentielle*

First version in 2004 http ://www.lifl.fr/~boulier/BLAD.
Started in 2000. Four versions thrown away.
Open source C libraries (LGPL). 36000 lines of C code.
Designed to be software components.
Provide functionnalities to overcome the above problems.
A less fancy demo of Rosenfeld–Gröbner using BLAD.

The lepism (or silver fish)



Logiciels pour l'Estimation de Paramètres et l'Identification Systématique de Modèles

Statement of the problem

Given

- a parametric ODE system (four parameters k_e, V_e, k_{12}, k_{21}) :

$$\begin{aligned}\dot{x}_1(t) &= -k_{12}\,x_1(t) + k_{21}\,x_2(t) - \frac{V_e\,x_1(t)}{k_e + x_1(t)},\\ \dot{x}_2(t) &= k_{12}\,x_1(t) - k_{21}\,x_2(t).\end{aligned}$$

- some measures :
  $x_1(t)$ is observed ;
  $x_2(t)$ is not observed

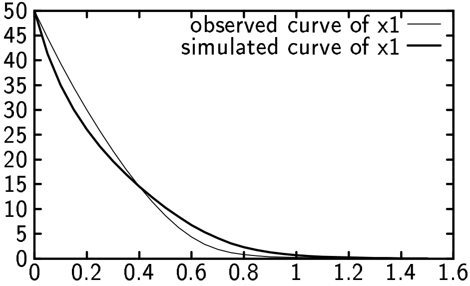
t	$x_1(t)$
$0.00000e-01$	$5.00000e+01$
$0.50000e-01$	$4.45078e+01$
...	
$1.50000e+00$	$4.95270e-02$

- possibly some extra information : $x_2(0) = 0$; $k_e = 7$.

Estimate the values of the unknown parameters V_e, k_{12}, k_{21}.

There exists a purely numerical method

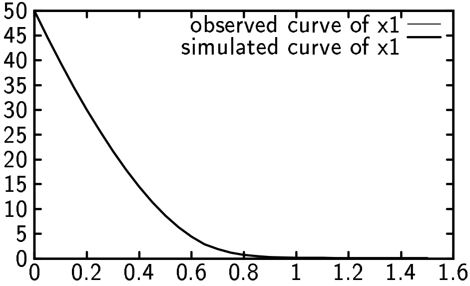
1. Give random values to k_{12}, k_{21}, V_e.
2. Integrate numerically the ODE and get a graph for $x_1(t)$.



3. If the error is too large, update k_{12}, k_{21}, V_e by the Levenberg–Marquardt method and goto step 2.

There exists a purely numerical method

1. Give random values to k_{12}, k_{21}, V_e.
2. Integrate numerically the ODE and get a graph for $x_1(t)$.



The Levenberg–Marquardt method ends in a wrong local minimum

$$k_{21} = .16, \quad k_{12} = .76, \quad V_e = 82.8.$$

Should be

$$k_{21} = .5, \quad k_{12} = 3, \quad V_e = 101.$$

Differential elimination for guessing good initial values

$$\dot{x}_1(t) = -k_{12}\,x_1(t) + k_{21}\,x_2(t) - \frac{V_e\,x_1(t)}{k_e + x_1(t)},$$
$$\dot{x}_2(t) = k_{12}\,x_1(t) - k_{21}\,x_2(t).$$

1. Eliminate the non observed variable $x_2(t)$ using Rosenfeld–Gröbner or, better, PARDI.

$$\ddot{x}_1\,(x_1+k_e)^2 + [k_{12}+k_{21}]\,\dot{x}_1\,(x_1+k_e)^2 + [V_e]\,\dot{x}_1\,k_e + [k_{21}\;V_e]\,x_1\,(x_1+k_e).$$

2. Evaluate the ODE for many different values of t.
   By linear least squares, estimate the [parameters blocks].
3. Solve the parameters blocks w.r.t. parameters :

$$k_{12} = 0.45, \quad k_{21} = 1.65, \quad V_e = 87.29.$$

4. Run the optimization method starting from these values.

Some numerical difficulties

Evaluation of the ODE

Difficult to numerically evaluate the derivatives.
Try to transform most of the ODE as an integral equation.

Getting the parameters from the [blocks]

Some numerical difficulties

Evaluation of the ODE

Difficult to numerically evaluate the derivatives.
Try to transform most of the ODE as an integral equation.

Getting the parameters from the [blocks]

Theoretically simple. Name the blocks

$$k_{12} + k_{21} = b_1, \quad V_e = b_2, \quad k_{21}\, V_e = b_3.$$

Then perform an algebraic elimination

$$(k_{12},\, k_{21},\, V_e) \gg (b_1,\, b_2,\, b_3).$$

The complexity of computing an ideal of relations !

Some numerical difficulties

Evaluation of the ODE

Difficult to numerically evaluate the derivatives.
Try to transform most of the ODE as an integral equation.

Getting the parameters from the [blocks]

More efficient : directly solving the system

$$k_{12} + k_{21} = value_1, \quad V_e = value_2, \quad k_{21}\, V_e = value_3.$$

- Algebraic solving : Solves the identifiability question. Beware algebraic relations amongs blocks !
- Numerical solving : Again a Newton method.

Summary

A very interesting collaboration

- We work with researchers from applied maths who carry out real examples
- We develop a software implementing their methods.
- Article at ICPSS04 (w. L. Denis–Vidal, T. Henin, F. Lemaire).

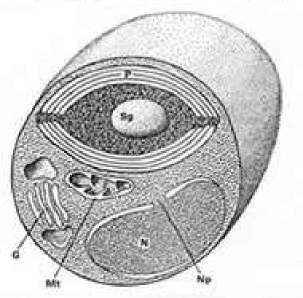
Strong connection with non commutative algebra

Identification Systématique de Modèles (black vs white box)

LÉPISME is pursued with the ANR project proposal IMAGES

Investigation du Métabolisme végétal des Acides Gras par Étude Systémique
Project led by a biological lab. *Génie Enzymatique et Cellulaire*.
Modelling the biosynthesis of fatty acids and oil in oilseed embryos.

Ostreococcus tauri



A minimal organism
Discovered in 1994
11 millions of pairs of bases

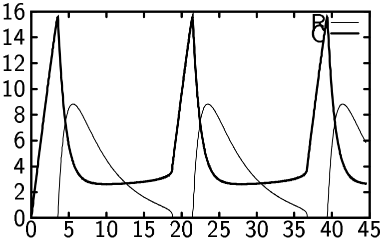
Institut de Recherches Interdisciplinaire
A pluridisciplinary group
F.-Y. Bouget (observatoire océanologique de Banyuls)
M. Lefranc (physics, Lille)

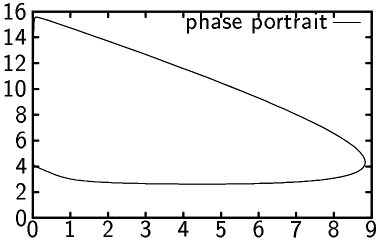
Goal : modelling the circadian clock
What makes a dynamical system oscillate ?

The example of Vilar, Kueh, Barkai, Leibler

2 genes, 7 variables, 15 parameters.

It oscillates only on restricted ranges of parameters.





The VKBL idea : searching oscillations using the Poincaré–Bendixson theorem

- Poincaré–Bendixson. It only holds in the two variables setting. It gives sufficient conditions so that oscillations occur.
- Model reduction. If $g(x,y) \gg f(x,y)$ then a system

$$\dot{x}(t) = f(x(t),\, y(t)), \quad \dot{y}(t) = g(x(t),\, y(t))$$

may possibly be approximated by

$$\dot{x}(t) = f(x(t),\, y(t)), \quad \not{\dot{y}}(\not{t})\ 0 = g(x(t),\, y(t)).$$

- Get rid of 5 of the 7 variables.
  Get parameters values w.r.t. which Poincaré–Bendixson holds.
  Check these values produce oscillations for the initial system.

Differential elimination for realizing VKBL idea

$$
\begin{aligned}
\dot{D}_A(t) &= \theta_A(1 - D_A(t)) - \gamma_A\, D_A(t)\, A(t),\\
\dot{D}_R(t) &= \theta_R\,(1 - D_R(t)) - \gamma_R\, D_R(t)\, A(t),\\
\dot{M}_A(t) &= \alpha'_A\,(1 - D_A(t)) + \alpha_A\, D_A(t) - \delta_{M_A}\, M_A(t),\\
\dot{M}_R(t) &= \alpha'_R\,(1 - D_R(t)) + \alpha_R\, D_R(t) - \delta_{M_R}\, M_R(t),\\
\dot{A}(t) &= \beta_A\, M_A(t) + \theta_A\,(1 - D_A(t)) + \theta_R\,(1 - D_R(t)) - A(t)\,(\cdots),\\
\dot{R}(t) &= \beta_R\, M_R(t) - \gamma_C\, A(t)\, R(t) + \delta_A\, C(t) - \delta_R\, R(t),\\
\dot{C}(t) &= \gamma_C\, A(t)\, R(t) - \delta_A\, C(t).
\end{aligned}
$$

Differential elimination for realizing VKBL idea

$$
\begin{aligned}
\cancel{\dot{D}_A(t)}\;{\color{red}0} &= \theta_A(1 - D_A(t)) - \gamma_A\, D_A(t)\, A(t),\\
\cancel{\dot{D}_R(t)}\;{\color{red}0} &= \theta_R\,(1 - D_R(t)) - \gamma_R\, D_R(t)\, A(t),\\
\cancel{\dot{M}_A(t)}\;{\color{red}0} &= \alpha'_A\,(1 - D_A(t)) + \alpha_A\, D_A(t) - \delta_{M_A(t)}\, M_A(t),\\
\cancel{\dot{M}_R(t)}\;{\color{red}0} &= \alpha'_R\,(1 - D_R(t)) + \alpha_R\, D_R(t) - \delta_{M_R}\, M_R(t),\\
\cancel{\dot{A}(t)}\;{\color{red}0} &= \beta_A\, M_A(t) + \theta_A\,(1 - D_A(t)) + \theta_R\,(1 - D_R(t)) - A(t)\,(\cdots),\\
\dot{R}(t) &= \beta_R\, M_R(t) - \gamma_C\, A(t)\, R(t) + \delta_A\, C(t) - \delta_R\, R(t),\\
\dot{C}(t) &= \gamma_C\, A(t)\, R(t) - \delta_A\, C(t).
\end{aligned}
$$

Differential elimination for realizing VKBL idea

$$
\begin{aligned}
\cancel{\dot{D}_A(t)}\ \color{red}{0} &= \theta_A(1 - D_A(t)) - \gamma_A\, D_A(t)\, A(t),\\
\cancel{\dot{D}_R(t)}\ \color{red}{0} &= \theta_R\,(1 - D_R(t)) - \gamma_R\, D_R(t)\, A(t),\\
\cancel{\dot{M}_A(t)}\ \color{red}{0} &= \alpha'_A\,(1 - D_A(t)) + \alpha_A\, D_A(t) - \delta_{M_A(t)}\, M_A(t),\\
\cancel{\dot{M}_R(t)}\ \color{red}{0} &= \alpha'_R\,(1 - D_R(t)) + \alpha_R\, D_R(t) - \delta_{M_R}\, M_R(t),\\
\cancel{\dot{A}(t)}\ \color{red}{0} &= \beta_A\, M_A(t) + \theta_A\,(1 - D_A(t)) + \theta_R\,(1 - D_R(t)) - A(t)\,(\cdots),\\
\dot{R}(t) &= \beta_R\, M_R(t) - \gamma_C\, A(t)\, R(t) + \delta_A\, C(t) - \delta_R\, R(t),\\
\dot{C}(t) &= \gamma_C\, A(t)\, R(t) - \delta_A\, C(t).
\end{aligned}
$$

For some ranking, Rosenfeld–Gröbner gives the VKBL system :

$$
\begin{aligned}
\dot{R}(t) &= f(R(t),\ C(t),\ A(t))\\
\dot{C}(t) &= g(R(t),\ C(t),\ A(t))\\
\color{red}{A^2(t)} &= h(R(t), C(t))
\end{aligned}
$$

Differential elimination for realizing VKBL idea

$$
\begin{aligned}
\cancel{\dot{D}_A(t)}\ 0 &= \theta_A(1 - D_A(t)) - \gamma_A\, D_A(t)\, A(t),\\
\cancel{\dot{D}_R(t)}\ 0 &= \theta_R\,(1 - D_R(t)) - \gamma_R\, D_R(t)\, A(t),\\
\cancel{\dot{M}_A(t)}\ 0 &= \alpha'_A\,(1 - D_A(t)) + \alpha_A\, D_A(t) - \delta_{M_A(t)}\, M_A(t),\\
\cancel{\dot{M}_R(t)}\ 0 &= \alpha'_R\,(1 - D_R(t)) + \alpha_R\, D_R(t) - \delta_{M_R}\, M_R(t),\\
\cancel{\dot{A}(t)}\ 0 &= \beta_A\, M_A(t) + \theta_A\,(1 - D_A(t)) + \theta_R\,(1 - D_R(t)) - A(t)\,(\cdots),\\
\dot{R}(t) &= \beta_R\, M_R(t) - \gamma_C\, A(t)\, R(t) + \delta_A\, C(t) - \delta_R\, R(t),\\
\dot{C}(t) &= \gamma_C\, A(t)\, R(t) - \delta_A\, C(t).
\end{aligned}
$$

For some less intuitive ranking, Rosenfeld–Gröbner gives even better :

$$
\begin{aligned}
\dot{A}(t) &= f(R(t),\, C(t),\, A(t))\\
\dot{C}(t) &= g(R(t),\, C(t),\, A(t))\\
R(t) &= h(C(t), A(t))
\end{aligned}
$$

Yakoubsohn idea. Differential elimination may provide different sets of coordinates to describe constraint varieties.

A difficulty which turns out to be an advantage

- VKBL model does not apply to *ostreococcus tauri* for biological reasons.
- Due to lack of information, we look for the simplest models.
- Our model : one gene + something.
- Design helped by a software mixing
  - computer algebra (BLAD) : algebraic elimination (to compute equilibria), complex roots isolation (to compute eigenvalues)
  - numerical methods : ODE integration (DOP853), FFT (to compute periods).
- A recent result of F. Lemaire, A. Sedoglavic and Asli Ürgüplü : the study of some symmetries of the model indicate how to act on parameters to obtain a 24 hours period.

Prospects

Modelling in biology

- is very challenging,
- tells us which parts of the theory/software to develop.
- the situation regularly improves (e.g. measuring one cell)

Clarify the applications of differential elimination

To eventually find out how this theory applies . . .

- Index reduction, parameters estimation, model reduction
- A very local role, complementary to numerical methods

Develop BLAD

The previous applications were known but not implemented . . .

- Keep BLAD small and specialized.