# Estimation of growth parameters for the Drosophila wing disc from the sequence of its micrographs using the Growth as Random Iterated Diffeomorphisms Model

Nataliya Portman
Ph.D. Candidate

Department of Applied Mathematics
University of Waterloo

September 11, 2008

# Acknowledgement

Dr. Ulf Grenander, a founder of Pattern Theory
Division of Applied Mathematics, Brown University
Providence, Rhode Island, USA.

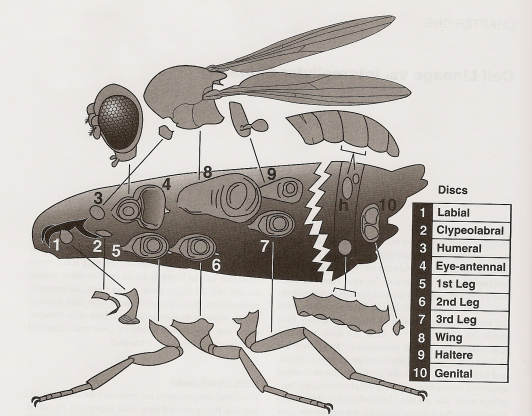
http://www.dam.brown.edu/ptg/

# Overview of the talk

- Larval development of normal Drosophila wing disc
- A pattern theoretic model for biological growth called GRID
- Inference of growth magnitude of the wing disc using GRID Model:
  - Formulation of unconstrained optimal control optimization problem
  - Solution of the inference problem

# Larval development of normal Drosophila wing disc

The fly exterior assembly from separate parts

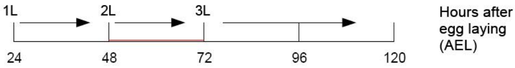


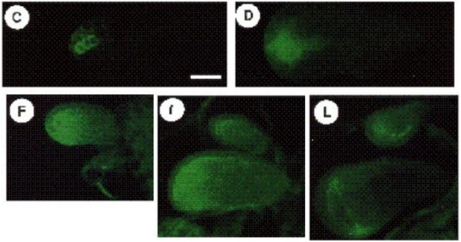

The schematic drawing of the Drosophila larva and adult structures made by imaginal discs.

# Larval development of normal Drosophila wing disc

A glimpse into the inner working of the wing disc

The Wingless protein plays a major role in patterning of the different elements of the adult fly wing

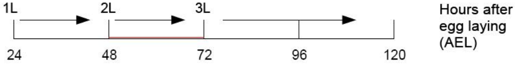


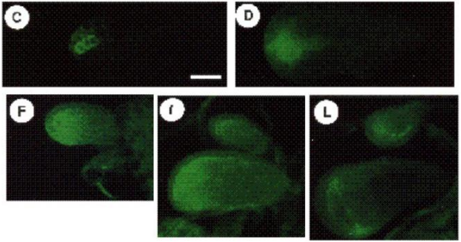




Confocal micrographs showing the dynamics of Wingless(Wg) expression during larval disc growth.

# Larval development of normal Drosophila wing disc

A glimpse into the inner working of the wing disc
The Wingless protein plays a major role in patterning of the different elements of the adult fly wing




(C)Middle second instar
(D)Late second instar, scale bar=25$\mu$m
(F,I)Early third instar
(L)Middle third instar,scale bar=50$\mu$m

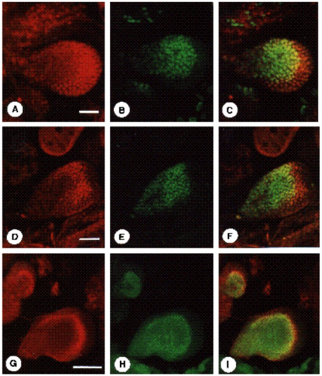
Confocal micrographs showing the dynamics of Wingless(Wg) expression during larval disc growth.

# Larval development of normal Drosophila wing disc

The Wingless protein functions in relation to the other key regulatory genes:

**1.**Regulation of the boundary of the Apterous(Ap) gene expression
**2.**Promotion of the Vestigial(Vg) gene expression

Ap is required for the formation of the entire wing
Vg is required for identifying the wing subfield

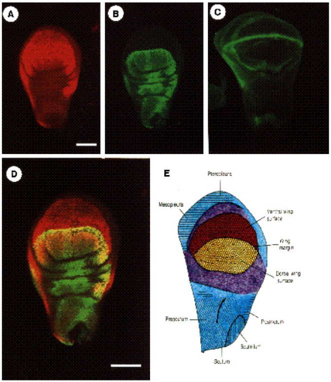



Development of the Vg and Ap expression patterns during the wing disc growth

(A,B,C)Middle second instar,
(D,E,F)Late second instar,
(G,H,I)Early third instar
discs expressing Vg, Ap and both proteins correspondingly

# Larval development of normal Drosophila wing disc

The mature wing disc as a geographic map of the adult fly wing and body elements




The late third instar (120 hours AEL) expression patterns of the key regulatory genes subdividing the wing disc into discrete subregions.

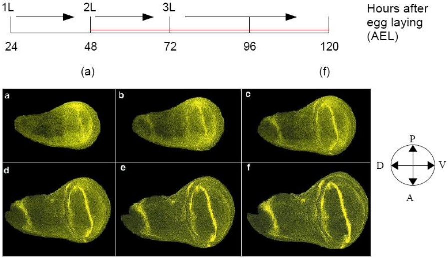
The expression patterns of Vg protein (A), Ap protein (B) and Wg (C).

Overlapping of the Vg(red) and Ap(green) expression patterns (D).

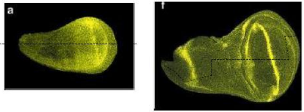
Wing fate map showing the dorsal/ventral boundary as a future wing margin, the ventral and dorsal wing surfaces and notal(body wall) region (E).
scale bar in (A-D)is 100 $\mu$m.

# Larval development of normal Drosophila wing disc

Biological data taken into a GRID
(Growth as Random Iterated
Diffeomorphisms)
mathematical model:
Brightness intensities of the image pixels corresponding to the levels of expression (or densities)of Wingless gene in the cells, $\mathrm{I}(\mathrm{x,t})$




Growth of the wing disc is subject
to other restrictions
that are not taken into account:
**1.** Spatial-temporal patterns of Ap and Vg gene products
**2.** Evolution of the boundary between Notum and Wing regions
**3.** Evolution of the Anterior/Posterior boundary in the wing disc




We ignore the effects of "secondary" regulatory genes to avoid additional complexity of the complex inference problem.

# Larval development of normal Drosophila wing disc

Biological facts that make 2D GRID modeling of the wing disc growth possible:

**1.** Cells divide randomly and uniformly throughout the wing disc at larval stage of development.
**2.** Cell number doubles on average every 9 hours during the second and early third instar.
**3.** Most cell movements are due to passive displacements (newborn cells pushing extant ones).
**4.** The disc epithelium is one cell thick like the ectoderm.

# A pattern theoretic model for biological growth called GRID

During growth, the wing disc locally expands and contracts changing its shape and internal structure.
$\Downarrow$
Its growth pattern can be represented by subsequent high-dimensional diffeomorphic transformations of the coordinates of the initial cellular field into the coordinates of the grown cellular field.

In the population of wing discs the growth patterns exhibit variability from one disc to another.
$\Downarrow$
It is natural to model the growth pattern as a composition of random transformations.

$$\Omega(t) = \phi^{({\xi^n}_{seed}, t_n)} \circ \phi^{({\xi^{n-1}}_{seed}, t_{n-1})} \circ \dots \circ \phi^{({\xi^1}_{seed}, t_1)} \Omega(t_0)$$

These random transformations $\phi^{({\xi^i}_{seed}, t_i)}$ are biologically motivated with the underlying probability laws characterizing the intensity of elementary biological events (cell divisions, cell death, cell movements).

# A pattern theoretic model for biological growth called GRID

The growth pattern is defined as a cumulative growth deformation composed of elementary deformations $\phi^{(\xi^{i}{}_{seed}, t_i)}$

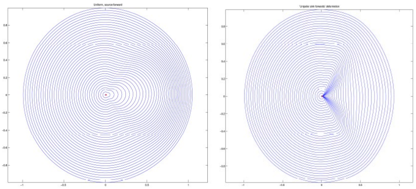
Construction of the elementary $\phi$-map:

**1.** Place a focal point of local growth, called a *seed*, according to a spatial Poisson process on a time-varying coordinate system called *Darcyan*;

**2.** Deform the neighborhood around the seed using a radial deformation function $k(\tau)$

$$\phi(x(\xi, t)) = x(\xi, t) + (x(\xi, t) - x(\xi_{seed}, t)) \cdot k(\tau) \cdot exp^{-(\|x(\xi,t)-x(\xi_{seed},t)\|)^2/step^2}$$

*step* is the range of influence of the current seed $\xi$ at time $t$



Elementary deformations "uni-source-forward" and "uni-sink-forward"

# A pattern theoretic model for biological growth called GRID

Illustrations of GRID model

We build a growth pattern out of several elementary "uni-source-forward" maps. The result is a pure growth in one preferential direction.

*(Click on the empty screen below to see a movie)*

# A pattern theoretic model for biological growth called GRID

## Illustrations of GRID model

Similarly, we build a growth pattern out of several elementary "uni-sink-forward" maps.
The result is a pure decay in one preferential direction.

*(Click on the empty screen below to see a movie)*

# A pattern theoretic model for biological growth called GRID

Continuous time approximation of the GRID model:
The growth pattern $\{x(\xi, t), t > t_0\}$ is a Poisson-driven Markov process described by the stochastic differential equation

$$dx = \int_{\xi_{t_j} \in \Xi} d\phi^{\xi_{t_j}}(x)\mu(d\xi, dt) \tag{1}$$

subject to the initial condition

$$x(\xi, t_0) = x_0(\xi) \tag{2}$$

$dx$ is the growth increment occurring in time interval $[t, t + dt]$
$\xi_{t_j}$ is a random seed placement distributed in $[t_0, \infty] \times \Xi$ with Poisson intensity $\Lambda = \lambda_t \cdot \lambda_x$
$d\phi^{\xi_{t_j}}(x) = (x(\xi, t) - x(\xi_{t_j})) \cdot k(\tau) \cdot exp^{-\|(x(\xi,t)-x(\xi_{t_j}))\|^2/step^2}$ is the jump size for the points $x(\xi)$ at time $t_j > t$
$\{\mu(T, B), T \subset [t_0, \infty), B \subset \Xi\}$ is a time-space Poisson process such that $E(\mu(T, B)) = \int_T \lambda_t dt \int_B \lambda_x dx$, where $\lambda_x$ is the intensity of events per unit volume.

# A pattern theoretic model for biological growth called GRID

Continuous space-time approximation of the GRID model:

The displacements $d\phi^{\xi_{t_j}}(x)$ are small and random and spaced so close in time that the resultant change in the position of the organism appears as a continuous motion.

Letting $\Delta t \to 0$ we obtain the instantaneous growth equation for a fixed seed $\xi_{seed}$

$$\frac{\partial x(\xi,t)}{\partial t} = (x(\xi,t) - x(\xi_{seed},t)) \cdot k(\tau) \cdot exp^{-\|(x(\xi,t)-x(\xi_{seed},t))\|^2/step^2} \tag{3}$$

The solution set consists of variable trajectories $x(\xi,t)$ in the absolute space.

At each time step $t_n$ the position of $\Omega(t)$ changes due to the single discrete decision of a cell.

A visible growth pattern observed in images is a result of a large number of cell decisions.

Letting a number of events $N \to \infty$ on the RHS of (3) we obtain the macroscopic growth equation

$$\frac{\partial x(\xi,t)}{\partial t} = \int_{\xi_{seed}\in\Xi} (x(\xi,t) - x(\xi_{seed},t)) \cdot k(\xi_s) \cdot exp^{-\|(x(\xi,t)-x(\xi_{seed},t))\|^2/step^2} F(d\xi_s) \tag{4}$$

# A pattern theoretic model for biological growth called GRID

Macroscopic growth equation

$$\frac{\partial x(\xi,t)}{\partial t} = \int_{\xi_{seed}\in\Xi} (x(\xi,t) - x(\xi_{seed},t)) \cdot k(\xi_{seed}) \cdot exp^{-\|(x(\xi,t)-x(\xi_{seed},t))\|^2/step^2} F(d\xi_{seed}) \tag{5}$$

$F$ is the Poisson intensity measure underlying the placement of seeds.
On the RHS of (5) we have the average value of the growth increment taken over all seeds.
Equation (5) predicts the internal structure and shape of the grown organism in the average sense.
Introducing a bounded measure $a(\xi) = k(\xi) \cdot f(\xi)$ called **the growth magnitude** we obtain

$$\frac{\partial x(\xi,t)}{\partial t} = \int_{\xi_{seed}\in\Xi} (x(\xi,t) - x(\xi_{seed},t)) \cdot exp^{-\|(x(\xi,t)-x(\xi_{seed},t))\|^2/step^2} a(\xi_{seed}) d(\xi_{seed}) \tag{6}$$

$a(\xi) > 0$ means expansion
$a(\xi) < 0$ means contraction

# Inference of growth magnitude of the wing disc using GRID Model

Inference problem formulation: Representation of the given biological data

Figure: Source and target micrographs $I_1(x)$ and $I_2(x)$ ($\sim$ 11 hours later)

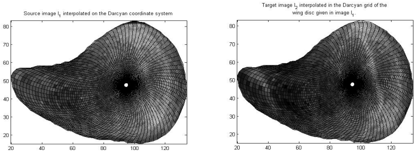



Figure: Source and target images $I_1(x(\xi), 0)$ and $I_2(x(\xi), T)$ interpolated on the Darcyan grid of the wing disc in $I_1$

# Inference of growth magnitude of the wing disc using GRID Model

Inference problem formulation:

Given two observed images $I_1(x(\xi), 0)$ and $I_2(x(\xi), T)$ of Wingless gene expression pattern estimate the growth magnitude $a(\xi)$ and the diffeomorphic transformation $\phi(x(\xi), t)$, $0 \leqslant t \leqslant T$,underlying the dynamics of the observed expression pattern.

We are led to the unconstrained continuous optimal control problem of the form

$$\hat{a} = argmin \int_{\Xi} (I_2(x(\xi), T) - I_1(\phi_T(x(\xi), 0))^2 \left| \frac{\partial(x_1, x_2)}{\partial(\xi_1, \xi_2)} \right| d\xi \qquad (7)$$

subject to

$$\frac{\partial \phi(x(\xi), t)}{\partial t} = \int_{\xi_{seed} \in \Xi} (\phi(x(\xi), t) - x(\xi_{seed}, t)) \cdot exp^{-\|(\phi(x(\xi),t) - x(\xi_{seed},t))\|^2 / step^2} \cdot \qquad (8)$$

$\cdot a(\xi_{seed}, t) d\xi_{seed}$ subject to the initial conditions $\phi(x(\xi), 0) = x(\xi)$.

$\hat{a} = \lim_{t \to T} a(\xi, t)$ and $\phi_T(x(\xi), 0)$ is such that $I_1(\phi_T(x(\xi), 0)) \backsim I_2(x(\xi), T)$.

# Inference of growth magnitude of the wing disc using GRID Model

Inference problem formulation: Space-time discretization

$$\hat{a} = argmin \sum_{\xi_j \in \Xi} (I_2(x(\xi_j), T) - I_1(\phi_T(x(\xi_j), 0)))^2 \left| \frac{\partial(x_1, x_2)}{\partial(\xi_{1_j}, \xi_{2_j})} \right| \tag{9}$$

subject to

$\phi_T(x(\xi_i), 0) = x(\xi_i) + \sum_{t=1}^{T} \sum_{\xi_{seed_j} \in \Xi} (\phi(x(\xi_i), t) - x(\xi_{seed_j}, t)) \cdot exp^{-\|(\phi(x(\xi_i), t) - x(\xi_{seed_j}, t))\|^2 / step^2} \cdot a(\xi_{seed_j}, t)$

for all seeds $\xi_i$, $1 \leq i \leq M$

# Solution of the inference problem

**Step 1** Construction of the Darcyan coordinate system of the wing disc seen in $I_1$ and $I_2$ using the Level Set Method

**Step 2** Image preprocessing:

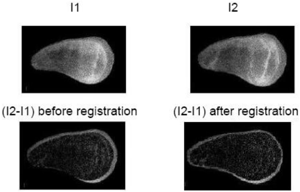
**(i)**Image registration about the centre of mass using Principal Component Analysis(PCA)

Given the Darcyan grid nodes $x_i = x(\xi_{1_i}, \xi_{2_i}),\ ,1 \leq i \leq N$, we define the covariance matrix $K = \frac{1}{N}\sum_{n=1}^{N}\{(x_i - x_c)(x_i - x_c)^T\}$ and its eigen vectors for $I_1$ and $I_2$.

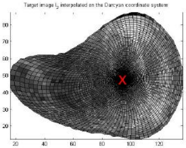
$x_c$ is the origin of the Darcyan coordinate system.

The eigen vectors of $K$ give the principal axes of both Darcyans. Then we register $I_1$ with $I_2$ by aligning the principal axis of their Darcyans. That is,

$x_{i\,I_1(reg)} = R\ x_{i\,I_1} + t$, where $R$ is the rotation matrix and $t$ is the translation vector.




X is the origin of the Darcyan coordinate system of the wing disc in I2



# Solution of the inference problem

Image preprocessing:

**(ii)**(proposed) Salt and pepper noise removal

**Step 3** Implementation of the conjugate gradient method to find the optimal value of the growth magnitude $\hat{a(\xi)}$.

The choice of this method is caused by factors:

(1) The steepest descent method is inefficient:It takes many small steps to reach the minimum. In a general multidimensional case, it does not take to the minimum.

(2) In $R^N$ for convex functions, the conjugate gradient method has a nice property: It constructs either a finite sequence $\{a_i\}$, whose last element $a_K$ minimizes the energy functional or an infinite sequence $\{a_i\}$ such that $lim_{i\to\infty} a_i = \hat{a}$ minimizes the energy functional.

(3) Its Polak-Ribiere version is useful in finding the minimum of a nonconvex function in $R^N$: The repeated cycles of $N$ iterations will eventually converge to the minimum.

# Solution of the inference problem

Polak-Ribiere conjugate gradient algorithm :

**0.** Select the initial growth magnitude value $a(\xi)^0 = 0$.

**1.** Compute the gradient of the energy functional

$$F(a(\xi)^0) = \sum_{\xi_j \in \Xi} (I_2(x(\xi_j), T) - I_1(\phi_0(x(\xi_j), 0)))^2 \left|\frac{\partial(x_1, x_2)}{\partial(\xi_{1_j}, \xi_{2_j})}\right| \qquad (10)$$

*At* $t = 0$, $\phi_t(x(\xi_j), 0) = x(\xi_j)$

**2.** Set $g^0(\xi) = h^0(\xi, t) = -gradF(a(\xi)^0$ *initialization part*

**3.** If $g^0(\cdot) = 0$ stop; else, set i=0 and go to step 4

**4.** Compute a scalar $\lambda_i$ such that

$F(a(\cdot)^i + \lambda_i \cdot h_i(\cdot)) = min_\lambda\{F(a(\cdot)^i + \lambda \cdot h_i(\cdot))|\lambda > 0\}$

**5.**Set $a(\cdot)^{i+1} = a(\cdot)^i + \lambda_i h(\cdot)^i$

**6.** Compute $\phi_{i+1}(x(\cdot), 0) = x^{i+1}(\cdot)$ by solving the initial value problem

$\frac{\partial x(\xi,t)}{\partial dt} = \int_{\xi_{seed} \in \Xi} (x(\xi, t) - x(\xi_{seed}, t)) \cdot exp^{-\|(x(\xi),t)-x(\xi_{seed},t))\|^2/step^2} a(\xi, t) d\xi_{seed}$

subject to $x(\xi, t = i) = x(\xi)^i$

The discrete solution is $\mathbf{x}(\cdot)^{\mathbf{i+1}} = \mathbf{x}(\cdot)^{\mathbf{i}} + \sum_{\xi_{\mathbf{seed_j}}} \theta(\mathbf{x}(\cdot)^{\mathbf{i}} - \mathbf{x}(\xi_{\mathbf{seed_j}})^{\mathbf{i}})\mathbf{a}(\xi_{\mathbf{seed_j}})^{\mathbf{i+1}}$

# Solution of the inference problem

Polak-Ribiere conjugate gradient algorithm (continued) :

After we have updated the Darcyan grid according to
$\phi_{i+1}(x(\cdot),0) = x(\cdot)^{i+1} = x(\cdot)^i + \sum_{\xi_{seed_j}} \theta(x(\cdot)^i - x(\xi_{seed_j})^i) a(\xi_{seed_j})^{i+1}$

**7.** Compute gradient $F(a(\cdot)^{i+1})$

**8.** If $grad(F(a(\cdot)^{i+1})) = 0$ then stop,
else set $g_{i+1}(\cdot) = -grad(F(a(\cdot)^{i+1})),\ \ h_{i+1}(\cdot) = g_{i+1}(\cdot) + \gamma_i\ h_i(\cdot)$

(Fletcher-Reeves algorithm)$\gamma_i = \frac{\|g_{i+1}(\cdot)\|^2}{\|g_i(\cdot)\|^2}$

(Polak-Ribiere algorithm)$\gamma_i = \frac{\|(g_{i+1}(\cdot)-g_i(\cdot))g_{i+1}(\cdot)\|^2}{\|g_i(\cdot)\|^2}$

go to line minimization step 4.

Ideally, we would like to stop (the stopping time $t = T$) when the distance measure between images $I_2(x(\xi), T)$ and $I_1(\phi_T(x(\xi)), 0)$ becomes reasonably close to 0.

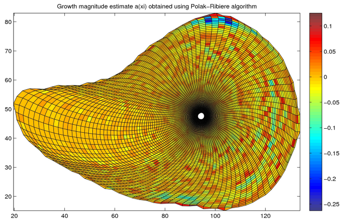
However, as experimental results show, we are not able to iterate until full convergence since further calculations destroy the Darcyan coordinate system of the growing disc.

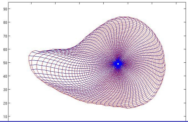
# Solution of the inference problem

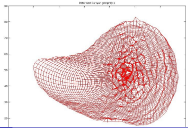
Experiment(1)

We apply Polak-Ribiere minimization algorithm to $100 \times 100$ Darcyan grid representing the wing disc given in $I_1$.

This means that we have 10,000 unknown components of the growth magnitude vector $a(\xi)$. Let *step* parameter be 1 and the number of iterations be 30. In 30 iterations the distance between the images has reduced from 51 to 30.





# Solution of the inference problem

## Experiment(1)

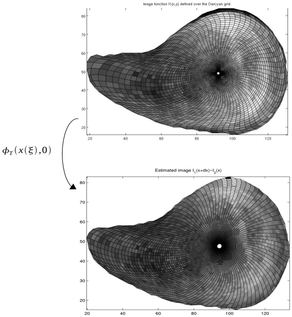
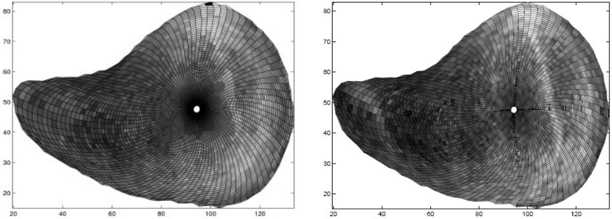



Figure: Source image $I_1(x(\xi))$ and Target image estimate $I_1(\phi_T^{-1}(x(\xi),0))$, $T = 30$

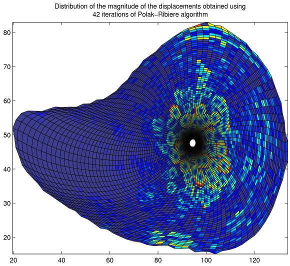
# Solution of the inference problem

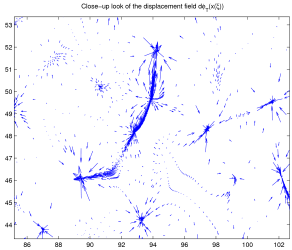
Experiment(1): Comparison of the estimated target image $I_1(\phi_T^{-1}(x(\xi),0))$, $T = 30$, and the target image $I_2(x(\xi))$



# Solution of the inference problem

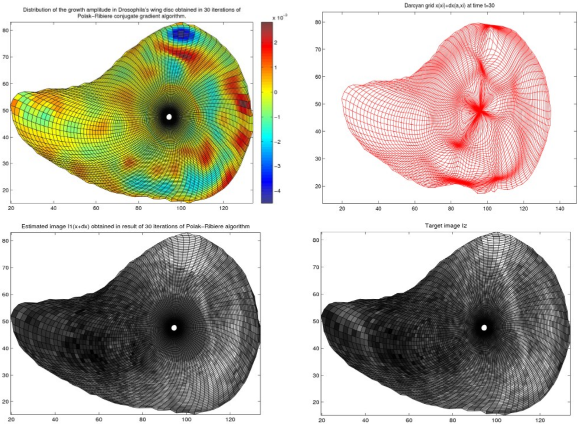
Experiment (1): Color-coded plot of the magnitude of the displacements $\phi_T(x(\xi), 0) - x(\xi, 0)$ appearing in a pattern of concentric circles accumulated near the origin.



# Solution of the inference problem

Experiment(2). Set the *step* parameter equal to 5 and run 30 iterations of the Polak-Ribiere algorithm.
The distance between the images has been reduced from 51 to 33.



# Solution of the inference problem

We have obtained the estimate of the growth magnitude $a(\xi)$ over the cellular field of the wing disc for different values of the *step* parameter.

For a smaller range of influence the regions of contraction and expansion appear more localized.

In both cases these regions are combined with one another.

This distribution can describe growth with regions in red designating cell division locations.

It is possible that the dividing cells push out the neighboring cells causing local contractions of the cellular field.

For a smaller *step* parameter it looks like cell divisions have been taking place in the uniform manner in the ventral part of the wing disc.

# Future analysis of the macroscopic growth integro-differential equation

Conjecture of the macroscopic growth equation in discrete form
$\Delta x = \sum_{xi_{seed} \in \Xi} \theta(x(\xi) - x(\xi_{seed}))a_{seed}$ , $\xi = (\xi_1, \xi_2)$ where $\xi, \xi_{seed}$span a finite range of integers.

We can write this equation as $\Delta x = \Theta a$, where $\Theta(z) = zexp(-\frac{\|z\|^2}{step^2})$.
We would like to study the properties of the operator $\Theta$ and the numerical stability of the inverse problem $a = \Theta^{-1}\Delta x$.
If the eigen-values of the operator $\Theta$ are purely imaginary and close to 0 then the solution to the inverse problem is not well-defined.

# Summary

- We have introduced the GRID Model based on fundamental biological principles of growth.
  - Elementary deformations result from discrete biological events (mitosis, cell death, cell movement)
  - Growth patterns are variable due to random occurrences of active gene sites or 'seeds' modeled as a Poisson point process
  - The equation of motion is formulated in terms of Darcyan coordinates independent of the absolute space and related to the organism shape changes
- We have demonstrated estimation of the growth magnitude directly from the micrographs of the Wingless expression pattern using the macroscopic growth equation.
  - Image preprocessing step removes rigid transformation and noise
  - Polak-Ribiere conjugate gradient algorithm constructs a finite sequence of optimal controls that converges to the optimal value of the growth magnitude. Then the optimal diffeomorphic flow matching the source and the target images is generated automatically.
- We have performed computations of the growth magnitude field and the corresponding diffeomorphic map with different values of the *step* parameter. Increasing *step* leads to a different growth mode.

# Literature

**[1]** U.Grenander, A.Srivastava, S.Saini, *A Pattern-Theoretic characterization of Biological Growth*,IEEE Transactions on Medical Imaging, Vol.26, No.2, May 2007.

**[2]** U.Grenander,*On the Mathematics of Growth*, Quart. Appl. Math. 65/2007, pp.205-257.

**[3]** N.Portman, U.Grenander, E.Vrscay, *New Computational Methods for the Construction of "Darcyan" Biological Coordinate Systems*, ICIAR Proceedings, Vol.4633/2007, pp.143-156.

**[4]** S.B.Carroll, *Endless Forms Most Beautiful*, W.W.Norton Company,Inc., New York, 2005.

**[5]** J.A.Williams, S.W.Paddock, S.B.Carroll,*Pattern Formation in a Secondary Field:a Hierarchy of Regulatory Genes Subdivides the Developing Drosophila Wing Disc into Discrete subregions*, Development 117/1993, pp.571-584.

**[6]** E.Polak,*Computational Methods in Optimization;a Unified Approach*, New York, Academic Press, 1971.