Implementing Life Phenomena: Examples of Pattern Formation and Cell Differentiation (Ryo Doi)

“Development” as a Supernatural Phenomenon

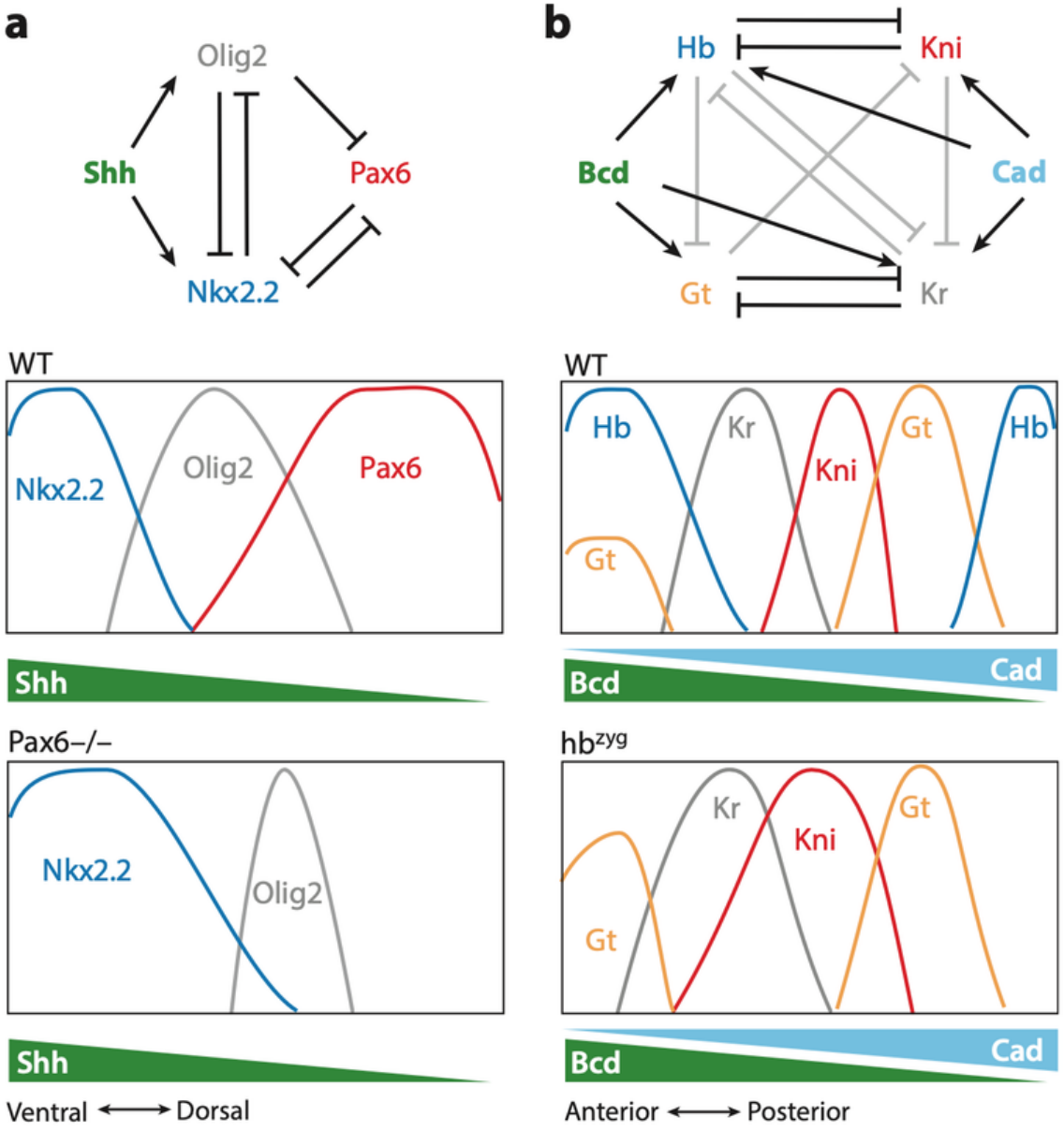
The phenomenon of biological development is extremely mystical, and I even think (perhaps with a bit of exaggeration) that there is no other physical phenomenon as fascinating as this.

Delving a bit into history, there was a time in the field of developmental biology when two opposing hypotheses existed: preformationism and epigenesis. The former proposed that each organ was preformed in a small shape within the reproductive cells (known as a homunculus), and that an adult was formed through the enlargement of these during development. On the other hand, epigenesis suggested that each organ was newly created from scratch in each generation. Today, it is common knowledge, even among middle and high school students, that preformationism is incorrect and epigenesis is correct. However, it is not surprising that preformationism was somewhat supported at the time. The fact that our bodies originally were clumps of proteins wrapped in a lipid membrane about 0.1mm in size remains hard to believe even today.

The formation of tissues and organs in living organisms is essentially realized through two main elements. One is the concentration gradient of morphogens. Morphogens are a collective term for chemicals in the body that control organ morphogenesis through their concentration gradients during development. It is (at least within the biological community) well-known that in fruit flies, the anterior-posterior axis is determined by morphogen proteins such as bicoid, nanos, and hunchback, which induce body axis formation.

The other element is the regulation of gene expression. Unlike morphogens, which operate at the scale of cell populations (tissues), gene expression regulation occurs within cells. This involves cells detecting stimuli from morphogen concentration gradients or neighboring cells via receptors, then through information received, leading to epigenomic modifications such as

chromatin modification and DNA methylation, as well as regulation of expression by transcription factors. This mechanism cleverly adjusts the type, amount, and timing of proteins expressed.



An example of position-dependent gene expression through morphogen concentration gradients and gene regulatory networks. For instance, in the diagram on the right, the gradient of the morphogen Sonic Hedgehog (Shh) exists along both the anterior-posterior and dorsal-ventral axes. It has been discovered that, as shown in the graph above, a gene regulatory network exists which causes the expression levels of Nkx2.2, Olig2, and Pax6 to change in a position-dependent manner when there is a gradient of Shh. It has also been shown that the distribution changes when Pax6 is knocked out. (Cited from Kicheva A, Briscoe J. Control of Tissue Development by Morphogens. Annu Rev Cell Dev Biol. 2023 Oct 16;39:91-121. doi: 10.1146/annurev-cellbio-020823-011522. Epub 2023 Jul 7. PMID: 37418774.)

One of the points that should definitely be focused on in the developmental process is its robustness. Put more simply, it's the "strength" against fluctuations and noise. While living organisms are often compared to machines, in this respect, it can be said that living organisms are far superior systems compared to machines.

Firstly, the signal-to-noise ratio (S/N ratio) between living organisms and machines is significantly different. In essence, biological systems are filled with noise. The nature of this noise can be varied, including unintended mechanical fluctuations, thermodynamic inevitabilities during reactions between biomolecules, non-equilibrium fluctuations, and quantum fluctuations due to the microscopic scale of biomolecules.

I am deeply interested in how, in such an environment, biomolecules and cells manage to transmit information within their interactions, and how they construct and maintain robust and plastic biological systems.

Mathematics of Pattern Formation

Diffusion Equation

Morphogens, which can induce tissue formation through their concentration gradients, might seem like "magic powder," but they are ultimately just chemical substances. Therefore, their movement in a fluid is principally governed by the diffusion equation.

Let's start by considering the behavior of a morphogen in one-dimensional space. Let $c(x,\ t)$ be the concentration of molecules at position x and time t. The change in the amount of substance between x and $x + \Delta x$ during the time Δt is

$$(c(x,\ t + \Delta t) - c(x,\ t)\Delta x$$

Furthermore, if we define the flux density $J(x)$ as the number of molecules passing through a unit area per unit time, then the change in the amount of substance between x and $x + \Delta x$ during Δt is

$$(J(x) - J(x + \Delta x))\Delta t$$

Since the change in concentration over time must balance with the net flux,

$$\frac{c(x,\ t + \Delta t) - f(x,\ t)}{\Delta t} = \frac{J(x) - J(x + \Delta x)}{\Delta x}$$

Taking the limit, we obtain

$$\frac{\partial c(x,\ t)}{\partial t} = \frac{\partial J(x)}{\partial x}$$

This equation illustrates how the concentration gradient of a substance leads to diffusion across space over time.

Here's Fick's first law:

$$J = -D\frac{\partial c(x,\ t)}{\partial x}$$

where D is the diffusion constant, we obtain:

$$\begin{aligned}\frac{\partial c(x,\ t)}{\partial t} &= \frac{\partial}{\partial x}\left(D\frac{\partial c(x,\ t)}{\partial x}\right) \\ &= D\frac{\partial^2 c(x,\ t)}{\partial x^2} \\ &= D\nabla^2 c(x,\ t)\end{aligned}$$

This is the one-dimensional diffusion equation. Similarly, by the same argument, we can obtain the three-dimensional diffusion equation:

$$\frac{\partial c(x,\ y,\ z,\ t)}{\partial t} = D\nabla^2 c(x,\ y,\ z,\ t)$$

Next, let's find the (analytical) solution to the one-dimensional diffusion equation. Since this equation is first-order in time and second-order in space with respect to partial derivatives, one condition for $c(x,\ t)$ at a certain time and two conditions for $c(x,\ t)$ at a certain position are required. Therefore, as initial conditions we set:

$$\begin{aligned}C(x,\ 0) &= C_0 x \\ C(\pm\infty,\ t) &= 0\end{aligned}$$

First, since the Fourier transform of $c(x,\ t)$ is:

$$c(x,\ t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} c(\hat{k},\ t)e^{ikx}dk$$

By substituting this into the diffusion equation, we get:

$$\int_{-\infty}^{\infty} \left(\frac{d\hat{c}}{dt} + Dk^2 \hat{c} \right) e^{ikx} dk = 0$$

For the above equation to hold for any x, the integrand must be zero, hence:

$$\frac{d\hat{c}}{dt} + Dk^2 \hat{c} = 0$$

This ordinary differential equation can be easily solved using the method of separation of variables, giving:

$$c(\hat{k},\ t) = A(k) e^{-Dk^2 t}$$

Substituting this back into the original Fourier transform equation, we have:

$$c(x,\ t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k) e^{ikx - Dk^2 t} dk$$

Considering the initial condition $C(x,\ 0) = C_0(x)$, we have:

$$c_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c_0(x) e^{ikx} dx$$

By inverse Fourier transform, we find:

$$A(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c_0(x) e^{ikx} dx$$

Therefore:

$$\begin{aligned} c(x,\ t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dx' c_0(x') e^{ik(x-x') - Dk^2 t} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx' c_0(x') e^{-\frac{(x-x')^2}{4Dt}} \left[\int_{-\infty}^{\infty} dk e^{-Dtk - \frac{i(x-x')^2}{2Dt}} \right] \end{aligned}$$

Using Gaussian integral, we can rewrite this as:

$$c(x,\ t) = \frac{1}{2\sqrt{\pi Dt}} \int_{-\infty}^{\infty} dx' c_0(x') e^{\frac{(x-x')^2}{4Dt}}$$

This provides the solution to the one-dimensional diffusion equation for any initial condition $C_0(x)$.

For example, let's set the initial condition as:

$$c_0(x) = \delta(x)$$

With this, due to the properties of the delta function, we obtain:

$$c(x,\ t) = \frac{1}{2\sqrt{\pi D t}} e^{-\frac{x^2}{4Dt}}$$

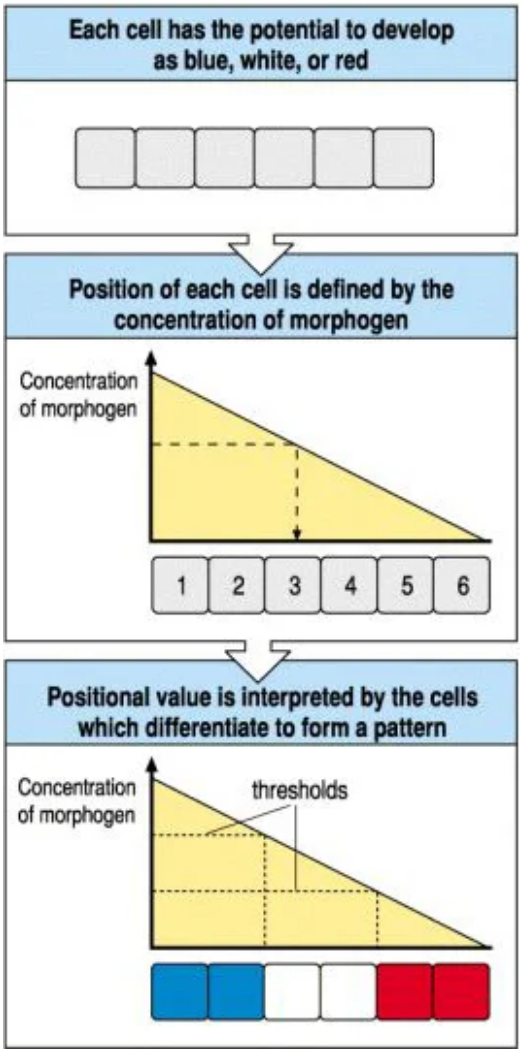
This solution takes the form of a Gaussian distribution, which might be easier to remember. As another condition, let's assume that the concentration of the substance at both ends of a one-dimensional space remains constant over time. Specifically, we impose the conditions:

$$\begin{aligned} c(0,\ t) &= c_{\max} \\ c(L,\ t) &= c_{\min} \end{aligned}$$

Considering the steady state $\frac{\partial c}{\partial t} = 0$, it's easy to see that the solution to the diffusion equation is:

$$c(x) = (c_{\min} - c_{\max})\frac{x}{L} + c_{\max}$$

In other words, in the steady state, it shows a linear gradient. From this result, developmental biologists once thought that each cell senses the surrounding morphogen concentration and simply differentiates into a specific type of cell according to that concentration (this is referred to as the "French flag model" due to its analogy). However, as can be seen from the analytical solution we derived, in this model, the morphogen concentration depends on the scale of the system L. It can easily be imagined that this would be significantly affected by the growth of the organism or fluctuations.



The French Flag Model (Quoted from https://bastiani.biology.utah.edu/courses/3230/DB%20Lecture/Lectures/a8Pattern.html)

Let's consider another situation where decomposition proportional to the concentration occurs across all spaces, namely:

$$\frac{\partial c(x,\ t)}{\partial t} = -kc(x,\ t) + D\frac{\partial^2 c(x,\ t)}{\partial x^2}$$

Let's consider the steady state where $\frac{\partial c}{\partial t} = 0$. Setting the boundary conditions as:

$$c(0,\ t) = c_{\max}$$
$$c(\infty,\ t) = 0$$

The solution can be relatively easily calculated to be:

$$c(x) = c_{\max}\exp\left(-\frac{x}{\lambda}\right)$$

That is, in the steady state, it shows an exponential gradient. Unlike the previous model, this model does not depend on the size of the system but is characterized by λ. Such exponential concentration gradients are well-known, for example, in the bicoid protein of fruit flies, and are known to be robust against perturbations.

In the next subsection, let's numerically solve the development equations under these two conditions to check if they broadly match the analytical solutions.

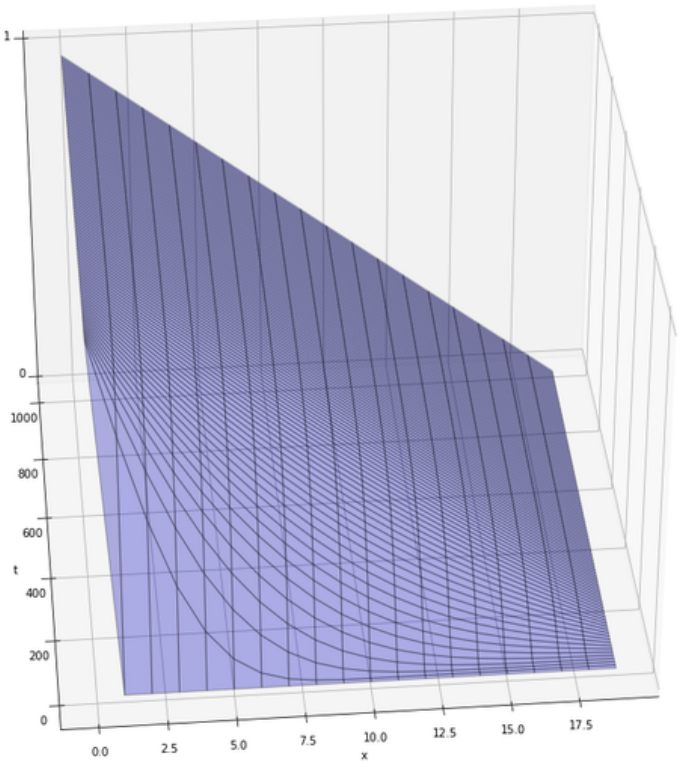
Example of Numerical Analysis for a Simple Evolution Equation

First, note that both simulations aim to calculate and visualize the temporal changes in the concentration of one type of morphogen at 20 points, $x = 0, 1, 2, \ldots, 19$.

First, we'll compute the scenario where the morphogen concentration at $x = 0$ is always kept constant. The code is as follows.

```
# import libraries
import numpy as np
import matplotlib.pyplot as plt

# set parameters
dx=0.05
dt=0.1
Du=0.01

# initial conditions and constant concentration at x=0
u0=np.linspace(0,0,20)
u0[0]=1
up=np.linspace(0,0,20)
up[0]=1

# diffusion term
def left(u):
    x1=np.roll(u,1)
    x1[0]=x1[1]
    return x1
def right(u):
    x2=np.roll(u,-1)
    x2[19]=x2[18]
```

```
    return x2
def diffusion(u):
    ans=Du*(left(u)+right(u)-2*u)/dx/dx
    return ans
def dudt(u):
    ans=u+dt*(up+diffusion(u))
    return ans

# position and time
x0=np.linspace(0,19,20)
x1=np.linspace(0,999,1000)

# insert results to matrix y
y=np.zeros((1000,20))
for s in range(1000):
        u0[0]=1
        u0[19]=0
        y[s,:]=u0
        u0=dudt(u0)

# visualization
xx0,xx1=np.meshgrid(x0,x1)
plt.figure(figsize=(15,15))
ax=plt.subplot(projection="3d")
ax.plot_surface(
    xx0,
    xx1,
    y,
    rstride=10,
    cstride=1,
    alpha=0.3,
    color="blue",
    edgecolor="black",
)
ax.set_zticks((0,1))
ax.view_init(35,-95)
plt.show()
```



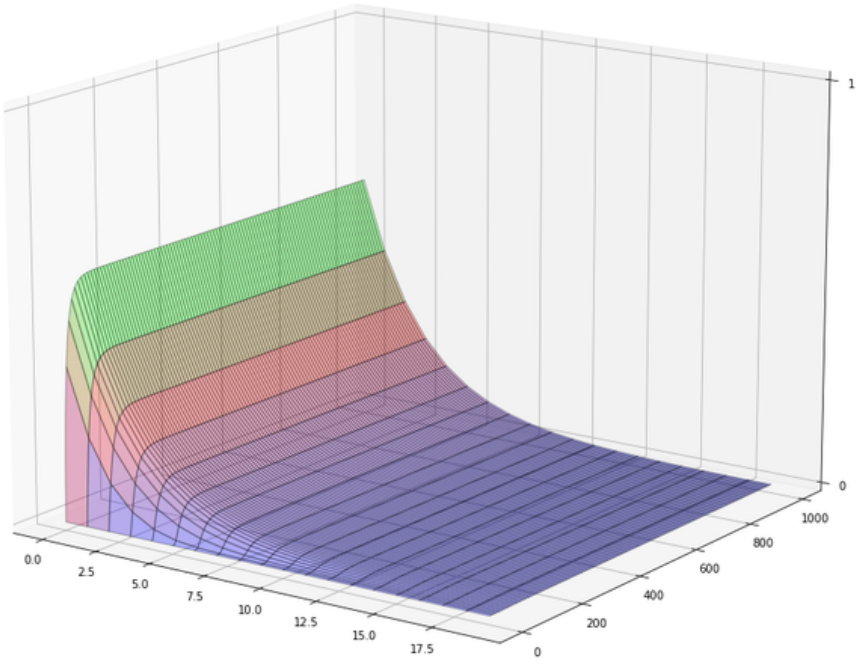
Under this condition, it is observed that a gradient proportional to the position is formed in the steady state.

Next, consider the scenario where the morphogen concentration at $x = 0$ is always kept constant, and the morphogen is decomposed at each point in proportion to its concentration.

```
# import libraries
import numpy as np
import matplotlib.pyplot as plt

# set parameters
dx = 0.05
dt = 0.1
```

```
deg = 0.5
Du = 0.01

# initial conditions and constant concentration at x=0
u0 = np.linspace(0, 0, 20)
up = np.linspace(0, 0, 20)
up[0] = 1

# diffusion term
def left(u):
    x1 = np.roll(u, 1)
    x1[0] = x1[1]
    return x1

def right(u):
    x2 = np.roll(u, -1)
    x2[19] = x2[18]
    return x2

def diffusion(u):
    ans = Du * (left(u) + right(u) - 2 * u) / dx / dx
    return ans

def dudt(u):
    ans = u + dt * (up - deg * u + diffusion(u))
    return ans

# position and time axes
x0 = np.linspace(0, 19, 20)
x1 = np.linspace(0, 999, 1000)

# store calculation results in matrix y
y = np.zeros((1000, 20))
for s in range(1000):
        y[s, :] = u0
        u0 = dudt(u0)

# visualization
xx0, xx1 = np.meshgrid(x0, x1)
plt.figure(figsize=(15, 15))
```

```
ax = plt.subplot(projection="3d")
ax.plot_surface(
    xx0,
    xx1,
    y,
    rstride=10,
    cstride=1,
    alpha=0.3,
    color="blue",
    edgecolor="black",
    cmap='brg',
)
ax.set_zticks((0, 1))
ax.view_init(15, -55)
plt.show()
```



Under this condition, it is observed that an exponential gradient is formed in relation to position in the steady state.

Consequently, (as expected) it has been confirmed that in both scenarios, the numerical solutions align well with the results obtained from analytical solutions.

In the models we've considered so far, we've examined gradients that arise locally when a specific morphogen emerges within limited boundary conditions. These are feedforward systems, and it is well-known that they can form certain patterns (such as stripes) to some extent (for example, the segmentation formation in fruit flies can be considered as an extension of such models). However, for more diverse and robust tissue formation, we hope for the existence of more complex molecular entities. For instance, it is well-known that in systems where synthesis

and decomposition occur everywhere in space, the uniform solution becomes unstable, and periodic pattern formations can emerge.

This reflects the two inductive factors of tissue formation mentioned at the beginning: the concentration gradient of morphogens and the impact of gene regulatory networks. In other words, until now, the discussion has focused solely on the concentration gradients of morphogens, but it is thought that each cell receives inputs of various types of molecules and regulates gene expression internally, either promoting or inhibiting it, thus the cell itself also produces outputs. Such systems are generally called reaction-diffusion systems, and in the context of the models discussed, the reaction term corresponds to gene expression within the cell, while the diffusion term corresponds to the temporal development of concentration gradients of morphogens on the tissue level.

The Reaction-Diffusion System (1-Dimensional)

The reaction-diffusion system was historically proposed by Alan Turing in 1952. Therefore, the periodic patterns characteristically seen in reaction-diffusion systems are called Turing Patterns. Despite the underlying equations for Turing Patterns being relatively simple, they have been very valuable in theoretical biology for explaining self-organization due to their powerful implications.

Let's consider the temporal development of a reaction-diffusion system in one-dimensional space. Introducing two types of morphogens, molecule p and molecule q, and assuming their interactions within each cell are described by the reaction equations:

$$\frac{dp}{dt} = f(p,\ q) = 0.6p - q - p^3$$
$$\frac{dp}{dt} = g(p,\ q) = 1.5p - 2q$$

Therefore, the reaction-diffusion equations can be written as:

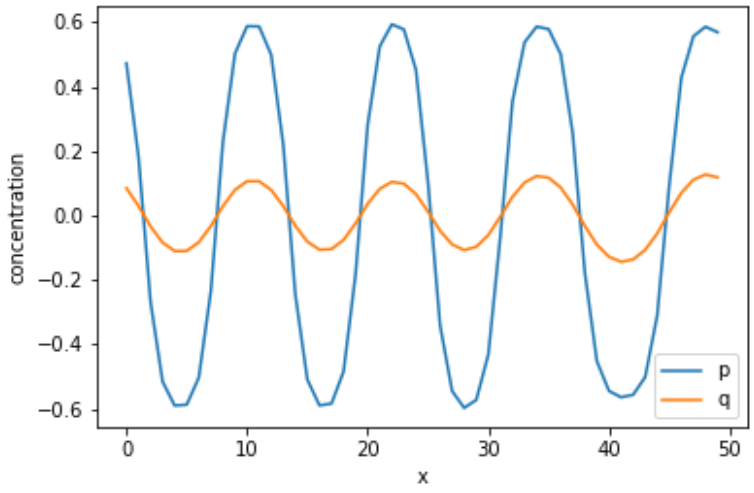
$$\frac{\partial p}{\partial t} = f(p,\ q) = D_p \frac{\partial^2 p}{\partial x^2}$$
$$\frac{\partial q}{\partial t} = q(p,\ q) + D_q \frac{\partial^2 p}{\partial x^2}$$

where D_p and D_q are the diffusion coefficients for each molecule. Let's use Python for numerical analysis of this.

```
# import libraries
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt

# set parameters
dx = 0.02
dt = 0.01

# set initial conditions with random values
np.random.seed(seed=1)
p = 0.01 * np.random.rand(50)
np.random.seed(seed=2)
q = 0.01 * np.random.rand(50)

# reaction terms
def f(p, q):
    return 0.6 * p - q - p**3

def g(p, q):
    return 1.5 * p - 2 * q

# diffusion terms
def diffusion(x):
    return np.roll(x, 1) + np.roll(x, -1) - 2 * x

def diffusionP(x):
    return dp * diffusion(x) / dx**2

def diffusionQ(x):
    return dq * diffusion(x) / dx**2

dp = 0.0002
dq = 0.01

# calculate matrices p and q
def pAfterDt(p, q):
    return p + dt * (f(p, q) + diffusionP(p))
```

```
def qAfterDt(p, q):
    return q + dt * (g(p, q) + diffusionQ(q))

for i in range(5000):
    p = pAfterDt(p, q)
    q = qAfterDt(p, q)

# visualization
plt.plot(range(50), p, label="p")
plt.plot(range(50), q, label="q")
plt.xlabel("x")
plt.ylabel("concentration")
plt.legend()
plt.savefig("reaction_diffusion_one_dimension.png")
plt.show()
```

By executing this, you will obtain the figure below. Please note the periodic concentration gradients that did not appear in the previous models. Let's qualitatively explain this result.



Example of a One-Dimensional Reaction-Diffusion System

Looking closely at the development equations, the synthesis of molecule p is promoted by p itself and inhibited by molecule q, and strongly inhibited when the concentration of p is too high. Meanwhile, the synthesis of molecule q is promoted by molecule p and inhibited by q itself. The difference in the diffusion coefficients of molecules p and q is also important. In this model, the diffusion

coefficient of molecule q is set significantly larger than that of molecule p. This means that molecule p acts locally, and once the concentration at a certain point becomes high, it catalytically increases in concentration. Molecule q is synthesized in response to the increase in p but diffuses rapidly, acting to inhibit the synthesis of p around the peak of p. This interaction forms the observed pattern. Later, we will provide a quantitative interpretation of this.

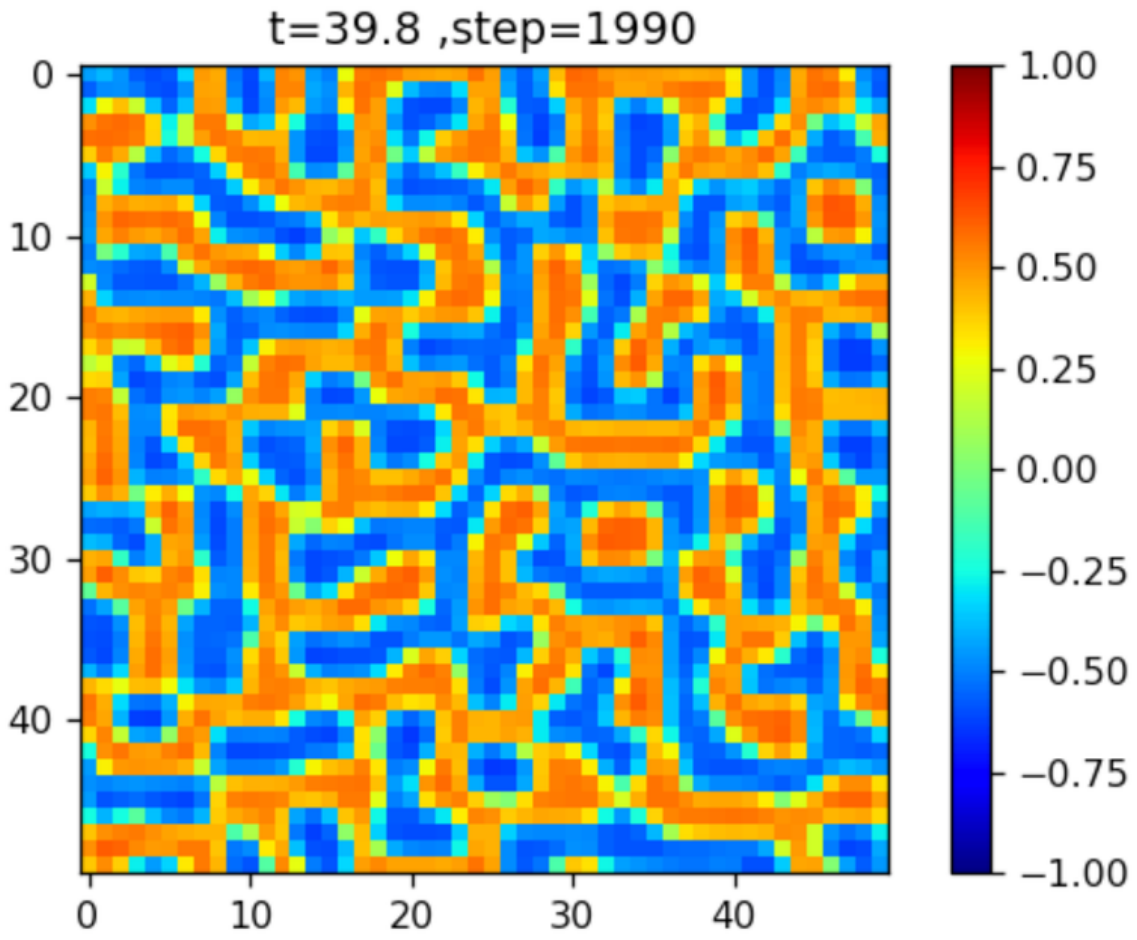
Reaction-Diffusion System (2-Dimensional)

Let's extend the discussion to a two-dimensional space. The reaction equations remain the same as before, but the difference is that a single cell receives inputs from four directions: up, down, left, and right.

```
# import libraries
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt

# set parameters
dx = 0.04
dt = 0.02
dp = 0.0002
dq = 0.01

# set initial conditions with random values
np.random.seed(seed=1)
p = 0.01 * np.random.rand(50, 50, 1)
np.random.seed(seed=2)
q = 0.01 * np.random.rand(50, 50, 1)

# reaction terms
def f(a, b):
    return 0.6 * a - b - a**3

def g(a, b):
    return 1.5 * a - 2.0 * b

# diffusion term
def diffusion(x):
    delta_diffusion = np.roll(x, 1, axis=0) + np.roll(x, -1, ax
```

```
    return delta_diffusion / (dx**2)

# calculation
for i in range(2000):
    next_p = p[:, :, -1] + dt * (f(p[:, :, -1], q[:, :, -1]) +
    next_q = q[:, :, -1] + dt * (g(p[:, :, -1], q[:, :, -1]) +
    p = np.dstack([p, next_p])
    q = np.dstack([q, next_q])

# visualization
%matplotlib nbagg
from matplotlib.animation import PillowWriter, FuncAnimation

def animate(i):
    clock = dt * i * 10
    step = i * 10
    plt.cla()
    plt.imshow(p[:, :, i * 10], interpolation='nearest', vmin=-
    plt.title("t={0:.1f}, step={1:3d}".format(clock, step))

fig = plt.figure()
plt.colorbar(plt.imshow(p[:, :, -1], interpolation='nearest', v
anim = FuncAnimation(fig, animate, repeat=False, frames=200, in
plt.show()
```

I encourage you to run this in Google Colaboratory or a similar environment. You should see patterns gradually emerging as shown in the figure below. Note that since the seed values for the random numbers set as initial conditions are specified, the results will slightly vary with each run.



Example of a One-Dimensional Reaction-Diffusion System

In development, the kind of repetitive patterns we've just examined can be found in various locations, such as body segments, bones in limbs, and patterns on the skin of fish. Although it has not been conclusively proven that Turing Patterns truly appear in biology, there have been various observations and experiments that support the presence of Turing Patterns in living organisms.

Turing Instability

This section requires some knowledge of dynamical systems, so those unfamiliar may skip it.

The general form of the multi-variable reaction-diffusion equations we have discussed so far can be expressed as:

$$\frac{\partial c_i(x,\ t)}{\partial t} = f_i(\mathbf{c}) + D_i \frac{\partial^2 c_i(x,\ t)}{\partial x^2}$$

For simplicity, let's initially ignore the diffusion term. In this case, the fixed points are combinations of $c_i(x,\ t)$ that satisfy:

$$\frac{\partial c_i(x,\ t)}{\partial t} = 0$$

These fixed points are given by a set of $c_i(x,\ t)$ values, $\mathbf{c}^*$. The Jacobian at these fixed points, denoted by J, determines the linear stability of the system through the eigenvalues λ of J (knowledge from the theory of dynamical systems).

Let's examine whether this solution becomes unstable due to the effect of diffusion, under the condition that the real parts of all eigenvalues of J are negative (stable solution).

For simplicity, assume that $c_i(x,\ t)$ near the fixed point can be expressed as:

$$c_i(x,\ t) = c_i^* + C_i(t)\sin(kx)$$

Substituting this into the previous equation gives:

$$\frac{\partial C_i}{\partial t} = -D_i k^2 C_i + \sum_{j=1}^{N} J_{ij} C_j = \sum_{j=1}^{N} (J_{ij} - D_i k^2 \delta_{ij}) \delta C_j$$

Thus, if we denote:

$$M_{ij} = J_{ij} = D_i k^2 \delta_{ij}$$

Then, we can treat M equivalently to J in the absence of diffusion terms. Therefore, denoting the eigenvalues of M as $\lambda_i(k)$, if the real part of λ_i is negative for all i, then $\mathbf{c}^*$ is stable. Conversely, if there is even one positive, it is unstable. When the real part is zero, a steady spatial periodic pattern emerges, which is the essence of Turing instability.

For clarity, let's consider when $N = 2$. The eigenvalues of M_{ij} are the solutions to the quadratic equation:

$$\lambda^2 - (J_{11} - D_1 k^2 + J_{22} - D_2 k^2)\lambda + (J_{11} - D_1 k^2)(J_{22} - D_2 k^2) - J_{12} J_{21} = 0$$

The conditions for both eigenvalues to have negative real parts, by the relations between solutions and coefficients, are:

$$J_{11} + J_{22} < (D_1 + D_2) k^2$$
$$J_{12} J_{22} < (J_{11} - D_1 k^2)(J_{22} - D_2 k^2)$$

Thus, when there's no diffusion ($k = 0$), the system is linearly stable near the fixed point if $J_{22} < 0$, $J_{21} < 0$ for $J_{11} > 0$, $J_{12} > 0$ and $J_{22} < 0$, $J_{12} < 0$ for $J_{11} > 0$, $J_{21} > 0$. The former system is called substrate-depletion, and the latter activator-inhibitor.

Furthermore, from the earlier equations, the condition for stability to easily break down with the addition of diffusion terms is $D_2 \gg D_1$. This matches the assumed relationship between the diffusion coefficients of molecules p, q in the previous sections, hence theoretically explaining the emergence of Turing instability and the formation of unique patterns.

Mathematics of Cell Differentiation

Originally, the essence was supposed to begin from here, but due to time constraints (and uncertainty about realization), I would like to introduce just the introduction as a preview for next time.

Dynamical Systems Model of Cell Differentiation

The topics from here are based on the following three papers. Those interested are encouraged to read them.

The discussion from here is based on three papers, which I highly recommend for those interested in exploring further:

https://www.science.org/doi/full/10.1126/science.1224311

https://www.science.org/doi/10.1126/science.aar4362

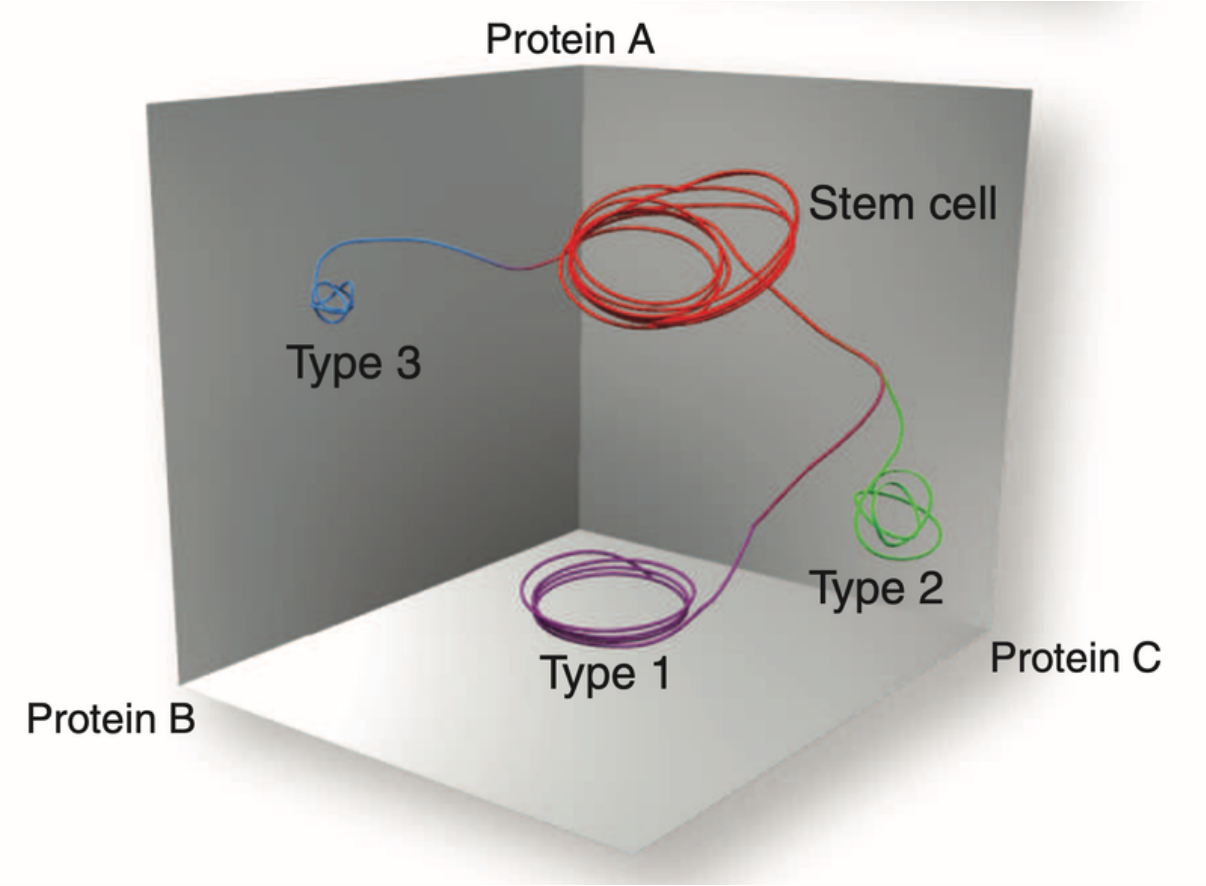
A very famous metaphor for cell differentiation is Waddington's epigenetic landscape. It aptly illustrates how stem cells located upstream differentiate over time into various cell types downstream, and how this flow is irreversible.

The molecular biological foundation of cell differentiation is epigenomic regulation. Naturally, all cells in our body share (almost) the same DNA. The variety of cells arises from differences in which parts of the shared DNA are "read." This clever regulation is primarily managed through epigenomic control, with histone modification and DNA methylation being typical examples.

This epigenomic change generally begins with the sensing of exogenous factors, which then goes through various types of cellular signal cascades, and ultimately, expression regulation by transcription factors is achieved. Here again, gene regulatory networks play a crucial role, and it can be considered that once stem cells differentiate, they stabilize by forming positive feedback loops in gene groups characteristic of their cell type.

Mathematical modeling of cell differentiation can effectively utilize dynamical systems theory. This is because each cell contains a vast variety of genes, mRNA, proteins, various kinases, and other metabolites, which differ from cell to cell. Each cell can be considered as a point in a high-dimensional space woven by the expression levels of these molecules. This concept aligns with the phase space idea in dynamical systems theory. Moreover, cells do not scatter randomly in this high-dimensional space; groups of cells showing similar phenotypes cluster closely together. In other words, in the high-dimensional space woven by numerous biological molecules, there are regions where cells tend to cluster (which can also be described as regions where cells tend to converge at the end of differentiation), corresponding to each post-differentiation cell type. These "clusters" are called attractors in the language of dynamical systems theory.

To summarize briefly, a single cell can be represented as a point in a high-dimensional space based on the expression levels of its internal genes, mRNA, proteins, etc. Cell differentiation can be understood as the irreversible transition from the state of a stem cell to one of several attractors as a result of interactions within and between cells.

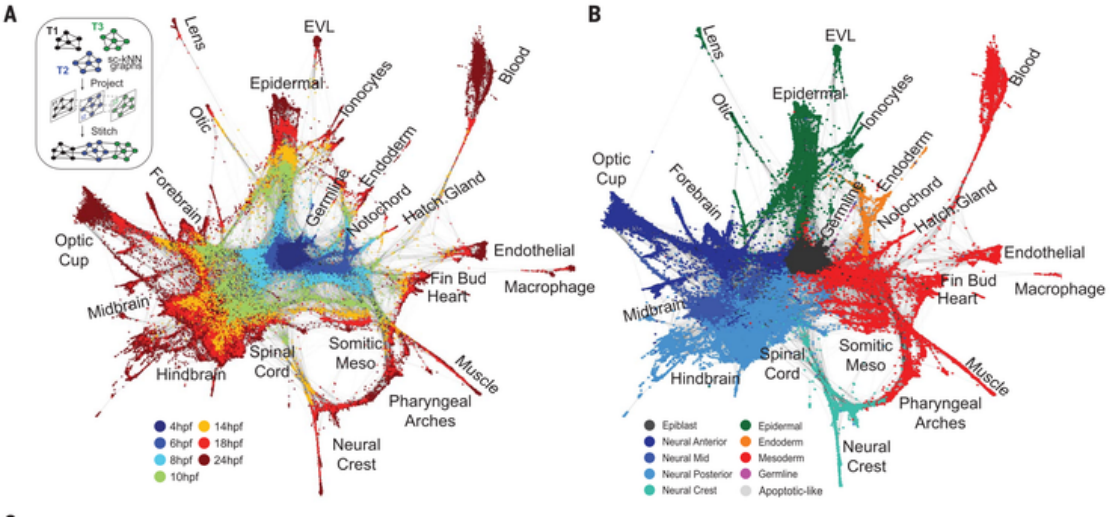


Dynamical Systems Model of Cell Differentiation (Cited from Chikara Furusawa, Kunihiko Kaneko, "A Dynamical-Systems View of Stem Cell Biology." Science 338, 215-217 (2012))

Visualization of Cell Differentiation through Single-Cell Transcriptome Analysis

In 2018, a fascinating paper was published in Science that performed single-cell transcriptome analysis on cells from zebrafish embryos at various times after fertilization. The expression data obtained was visualized using a dimensionality reduction technique called t-SNE, revealing the temporal changes in cell differentiation.

https://www.science.org/doi/10.1126/science.aar4362



Visualization of Cell Differentiation by t-SNE (Cited from Wagner, et al. "Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo." Science. 2018 Jun 1;360(6392):981-987. doi: 10.1126/science.aar4362. Epub 2018 Apr 26. PMID: 29700229; PMCID: PMC6083445.)

Not only for the high-dimensional phase space model of cell types but also for the general field of biological research, it can be said that (putting it bluntly) life is incredibly complex. Even a single cell is composed of a vast number of components.

Therefore, before the 2010s, the foundation of life science research was largely driven by hypothesis-driven molecular biological approaches, focusing on collecting local and fragmentary knowledge, such as "what happens when a specific gene is knocked out" or "molecule A phosphorylates molecule B." However, with recent technological advancements and the proliferation of analytical methods, data-driven approaches have rapidly increased.

This paper is emblematic of such a trend, as it translates the states of individual cells in a high-dimensional space, which are not intuitively recognizable to us humans, into a form that can be somewhat recognized and understood.

Network Structure Enabling Cell Differentiation

In 2011, a paper was published that proposed a model explaining cell differentiation based on changes in protein expression levels within each cell, considering interactions between cells and cell division, assuming five genes and their proteins inside the cell, across all possible gene regulatory networks they could form.

Originally, there was a plan to implement the algorithm from this paper within this article, but due to delays in publication, it has been postponed. There might be a

sequel where the implementation will be attempted.

The paper screened gene regulatory networks that could reproduce the phenomenon of cell differentiation through numerical simulation and then examined common network motifs among them. Further, based on these results, they designed gene regulatory networks that could induce hierarchical cell differentiation and confirmed their realization through numerical simulations.

Works Cited

1. Theoretical Biology of the Cell - A dynamical-systems Perspective -
2. Hassei no Suuri [Mathematics of Development]
3. Developmental Biology 5th Edition
4. Suzuki N, Furusawa C, Kaneko K. Oscillatory protein expression dynamics endows stem cells with robust differentiation potential. PLoS One. 2011;6(11):e27232. doi: 10.1371/journal.pone.0027232. Epub 2011 Nov 3. PMID: 22073296; PMCID: PMC3207845.
5. Kicheva A, Briscoe J. Control of Tissue Development by Morphogens. Annu Rev Cell Dev Biol. 2023 Oct 16;39:91-121. doi: 10.1146/annurev-cellbio-020823-011522. Epub 2023 Jul 7. PMID: 37418774.
6. Chikara Furusawa, Kunihiko Kaneko, A Dynamical-Systems View of Stem Cell Biology.Science**338**,215-217(2012).
7. Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. Science. 2018 Jun 1;360(6392):981-987. doi: 10.1126/science.aar4362. Epub 2018 Apr 26. PMID: 29700229; PMCID: PMC6083445.