Biomathematics:  
A course on some applications of dynamical systems

Spatial dynamical systems and diffusion

Anders Källén  
MatematikCentrum  
LTH  
anderskallen@gmail.com
Introduction

In this chapter we introduce dynamical models that are described by differential equations which depends on time plus another variable. This latter is often, but not always, a space variable, and in such cases what we consider are models describing motion in this space.

About non-stationary flows in one dimension

Consider a large number of identical particles that moves in one dimension. The space location is described by a coordinate $x$. The particles could be cars on a street or pollution in a river. All movement in a different direction is neglected. In order to describe the movement of the particles we will use the notions density and flow.

Density, or concentration, of particles is described by a smooth function $\rho(x, t)$.

What else can an observer measure? He can measure the number of particles that pass by per time unit, which is the flow $q(x, t)$, which also is assumed smooth. By convention $q > 0$ means that the net flow is from left to right.

There is a relation between density and flow, and in order to discuss it we first assume that no particles are created or destroyed. Choose an interval $[a, b]$ on the $x$-axis. The total number of particles in this interval at time $t$ is given by

$$N(t) = \int_a^b \rho(x, t)dx.$$ 

It means that the derivative $N'(t)$ is the change in the number of particles per time unit in the interval. But such a change is the result of flow over the endpoints of the interval. More precisely, there is a net inflow $q(a, t)$ over the left endpoint and a net outflow $-q(b, t)$ over the right endpoint. Thus we have

$$N'(t) = q(a, t) - q(b, t).$$

If we allow particles to be created or disappear in the interval we need to describe the creation rate with a function $k(x, t)$. In the interval $[a, b]$ the net creation is then

$$\int_a^b k(x, t)dx$$

particles per unit time. This leads to the balance equation

$$N'(t) = q(a, t) - q(b, t) + \int_a^b k(x, t)dx.$$ 

But we have that

$$q(a, t) - q(b, t) = -\int_a^b \partial_x q(x, t)dx,$$ 

so the right hand side in the balance equation can be written

$$\int_a^b (-\partial_x q(x, t) + k(x, t))dx.$$
In the left hand side we can take the derivative inside the integral sign:

\[ N'(t) = \frac{d}{dt} \int_{a}^{b} \rho(x, t) dx = \int_{a}^{b} \partial_{t} \rho(x, t) dx, \]

assuming that the derivative \( \partial_{t} \rho \) is continuous.

This gives us the relation

\[ \int_{a}^{b} \partial_{t} \rho(x, t) dx = \int_{a}^{b} (-\partial_{x} q(x, t) + k(x, t)) dx, \]

and since it is true for all intervals \([a, b]\) this requires that

\[ \partial_{t} \rho(x, t) = -\partial_{x} q(x, t) + k(x, t). \]

This is our basic partial differential equation.

The flow can be written

\[ q(x, t) = \rho(x, t) u(x, t), \]

where \( u(x, t) \) is the average velocity of the particles in the point \( x \) by time \( t \). Inserting that in the PDE we get

\[ \partial_{t} \rho + \partial_{x} (\rho u) = k. \]

This equation is called the \textit{continuity equation} for the situation.

**Example 1** Consider an organic pollution which is emitted from a factory into a small river and follow it along. The water velocity of the river is constant, \( c \), and we assume that in the river bacteria break down the pollution at a rate proportional to the concentration \( \rho(x, t) \) of the pollution. Here \( x \) is the distance along the river to the factory downstream. These assumptions lead to the equation

\[ \partial_{t} \rho + c \partial_{x} \rho = -\mu \rho, \]

A flow, like in the previous example, in which we know the velocity \( u(x, t) \) is called a \textit{convective flow}.

**Example 2** A female insect sit in a particual place and releases pheromones, for which the concentration is given by a function \( c(x, t) \). Male insects moves towards here guided by these pheromones so that they move in the direction in wich the concentration gradient is maximal, and with a speed which is proportional to this:

\[ u(x, t) = \alpha \partial_{x} c(x, t), \]

Here the proportionality factor \( \alpha \) can depend on both \( x \) and \( t \). If the concentration of male insect is given by \( \rho(x, t) \) the assumptions lead to the following equation

\[ \partial_{t} \rho = -\partial_{x} (\alpha n \partial_{x} c). \]

A movement which is much the opposite of a convective flow is diffusion.
Example 3 Consider a unicellular cell in a still water which moves by random movement. A simple law to describe this is named Fick’s law which says that the flow is in the opposite direction to the concentration gradient

\[ q = -D \partial_x \rho. \]

This gives us the equation

\[ \partial_t \rho = -\partial_x (-D \partial_x \rho) = D \partial_x^2 \rho. \]

The equation

\[ \partial_t \rho = D \partial_x^2 \rho \]

is called the diffusion equation and the constant \( D \) is called the diffusion coefficient. Its unit is \((\text{length})^2/\text{time}\) and measure how efficiently the particles disperse from high to low density. For example, in blood, haemoglobin molecules have a diffusion coefficient of the order \(10^{-7} \text{ cm}^2\text{s}^{-1}\), while that for oxygen in blood is of the order of \(10^{-5} \text{ cm}^2\text{s}^{-1}\).

Remark \( D \) can be dependent on both \( x, t \) (e.g. as a function of \( \rho(x,t) \)), and then diffusion equation looks slightly different:

\[ \partial_t \rho = \partial_x(D \partial_x \rho). \]

Example 4 To model the motion of bacteria as they consume a diffusible substrate to which they are chemotactically attracted may look like

\[
\begin{align*}
\frac{\partial b}{\partial t} &= \frac{\partial}{\partial x}\left( \mu \frac{\partial b}{\partial x} - \chi b \frac{\partial s}{\partial x} \right) + (f(s) - k_e)b, \\
\frac{\partial s}{\partial t} &= D \frac{\partial^2 s}{\partial x^2} - Y^{-1} f(s)b,
\end{align*}
\]

where \( b \) is bacterial density, \( s \) substrate density, \( Y \) the yield, \( f(s) \) the substrate dependent growth rate and \( k_e \) the constant death rate.

It is generally believed the leucocyte motion is a combination of chemotaxis and random motion like this.

A final example is a variation of the discussion above.

Example 5 Propagation of the Action Potential along an Axon. To derive a balance equation for the charge that incorporates the effect of transport in the axial direction let \( x \) be the distance along the axon and define

\[ q(x,t) = \text{charge density} \]

\[ J(x,t) = \text{current, i.e. flux of charged particles along the axon.} \]

\[ \sigma(x,t) = \text{rate at which charge enters or leaves axon through its membrane} \]
In these entities we have the balance equation
\[
\frac{\partial q}{\partial t} = -\frac{\partial J}{\partial x} + \sigma.
\]

In addition to this we have that

a) If \( v(x, t) \) is the voltage across the membrane
\[
q(x, t) = 2\pi r C v(x, t),
\]
where \( r \) is the radius of the axon and \( C \) the capacitance of the axonal membrane.

b) The local source of charge \( \sigma(x, t) = -2\pi r I_i(x, t) \), where \( I_i \) is the net ionic current into the axon.

c) Ohm’s law states that
\[
J = -\frac{\pi r^2}{R} \frac{\partial v}{\partial x},
\]
where \( R \) is intracellular resistivity.

From this it follows that the action potential equation is
\[
\frac{\partial v}{\partial t} = \frac{a}{2RC} \frac{\partial^2 v}{\partial x^2} - \frac{I_i}{C}.
\]

In addition we need a model for \( I_i \).

A second look at diffusion

Diffusion is an important concept, so there is reason to have a closer look at it from other perspectives.

Consider \( N + 1 \) cells attached to each other lying in a row. Along the parts where they are attached there are small holes in the cell wall through which a certain substance can flow. Within a cell this substance is always in equilibrium, but the rate with which it flows between cells is proportional to the difference in the concentrations in the two cells.

Let \( c_i(t) \) be the concentration of the substance in cell number \( i \) at time \( t \). The assumption above means that the flow into this cell from the cell with number \( i + 1 \) occurs with a rate \( \mu(c_{i+1}(t) - c_i(t)) \), so if the concentration is higher in cell \( i + 1 \) the substance flows into cell \( i \) and vice versa. Similarly, the exchange rate of substance with cell \( i - 1 \) is \( \mu(c_{i-1}(t) - c_i(t)) \). Thus the rate with which the concentration changes in cell \( i \) is given by
\[
c'_i(t) = \mu(c_{i+1}(t) - c_i(t) + c_{i-1}(t) - c_i(t)) = \mu(c_{i+1} + c_{i-1} - 2c_i(t)).
\]

Note that this assumes the cell has two neighbour cells.

If the cells are attached in a ring – which is common – this equation is valid for all cells in the ring. The only problem is our notation, since now cell 0 is to the right of cell \( N \), whereas the cell to the left of cell 0 is cell \( N \). Mathematically we express this to say that
we have *periodic boundary conditions*:

\[ c_{-1}(t) = c_N(t), \quad c_{N+1}(t) = c_0(t) \]

and then we use the same system of equations as above.

We can alternatively assume that cells 0 and \( N \) at the end of the row is special. We can e.g. assume that in those cells there is a fast reaction that completely eliminates the substance. In those cells we can therefore assume that there never was any substance:

\[ c_0(t) = c_N(t) = 0. \]

In this situation we talk about absorbing boundary conditions, and those are the ones we are primarily interested in, but in slightly different context.

The situation we are interested is more like the following. Replace the row of cells with a row of trees in which there is a pest which can move to adjacent trees in a way similar to above. We then spray the boundary trees with some pesticide which kills the pest. The question then is what happens with the pest in the remaining trees.

The system above is a large linear system of differential equations and can therefore in principle be solved with methods from linear algebra. But this becomes difficult when \( N \) is large, and that is precisely the situation we are interested in.

Instead we assume the cells are evenly spread over an interval, which we can take as \([0, 1]\).

It means that cell number \( i \) is centered around the point \( x_i = i/N \). Assume now there is a smooth function \( c(x,t) \) such that

\[ c(x_i, t) = c_i(t). \]

When \( h = 1/N \) we can now write the system of equations above as

\[
\partial_t c(x_i, t) = \mu (c(x_i+h, t) - 2c(x_i, t) + c(x_i-h, t)) \frac{c(x_i + h, t) - 2c(x_i, t) + c(x_i - h, t)}{h^2}.
\]

But \( \mu \) depends \( h \) and if we assume that

\[ \mu(h)h^2 \rightarrow D \quad \text{as} \quad h \rightarrow 0, \]

we see that in the limit we get the diffusion equation

\[ \partial_t c = D \partial_{xx}^2 c. \]

The periodic boundary conditions now means that \( c \) should be a 1-periodic function, whereas the absorbing boundary conditions means that

\[ c(0, t) = c(1, t) = 0. \]

We close this section with an example where \( x \) is not a space variable. Another example is given in the appendix. It is important to note that PDEs do not need to address time and space.
Example 6 We have previously discussed Leslie’s model, which was a generalisation of ordinary geometric growth of an animal population in which we also let different age classes have different fertility. In doing so we divided age into a finite number of age classes. In this example will look at how the model changes when we instead consider age a continuous variable.

Let \( x(t, a) \) be the number of females of age \( a \) at time \( t \). If we sample this population at time points \( t_i = i\Delta t \), the Leslie model means that, since \( t_i + 1 = t_i + \Delta t \),

\[
x(t_i + \Delta t, a + \Delta t) = p(a)x(t_i, a) \iff x(t_i + \Delta t, a + \Delta t) - x(t_i) = -(1 - p(a))x(t_i, a_i).
\]

But \( p(a) \) depends on how often we sample, i.e. on \( \Delta t \), and a reasonable hypothesis is that

\[
1 - p(a) = \text{fraction that dies in the time interval} \ (t_i, t_{i+1}) = \mu(a)\Delta t
\]

(at least to first order). If we let \( \Delta t \to 0 \), we see that the rows (2 : \( m \)) in the reproduction matrix of the Leslie model corresponds to the PDE

\[
\frac{\partial x}{\partial t} + \frac{\partial x}{\partial a} = -\mu(a)x.
\]

In order to get something that corresponds to the first row in this matrix we let \( f(a) \) denote the fertility (number of offsprings per year) of a female of age \( a \). The equation that corresponds to the first row in the reproduction matrix is then

\[
x(t, 0) = \int_{-\infty}^{\infty} f(a)x(t, a) da.
\]

This equation is called Von Foerster’s equation.

About solutions to the diffusion equation

Now consider the diffusion equation

\[
\partial_t u = D \partial^2_{xx} u
\]

and assume the equation applies for all \( x \) and \( t > 0 \). For uniqueness of the solution we need to specify what happens at the beginning, i.e. specify a function \( u_0(x) \) such that

\[
u(x, 0) = u_0(x).
\]

Assuming this we want to see what can be said about the solution \( u(x, t) \) for \( t > 0 \).

We start with the following observation. If we let

\[
K(x, t) = \frac{1}{2\sqrt{\piDt}} e^{-x^2/(4Dt)},
\]

we have a function that fullfills the equation when \( t > 0 \) and for which we have that

\[
\int_{-\infty}^{\infty} K(x, t) dx = 1.
\]
The latter fact is known from probability function, since function is the density function for a gaussian distribution with mean zero and variance $2Dt$.

Letting $t \to 0$, the function $K(x,t)$ will be more and more concentration around $x=0$. In the limit it means that we have a start population of one in the single point $x=0$ – the so-called Dirac measure which is denoted $\delta_0$.

If we change the origin to another point $a$, i.e. have the problem

$$\partial_t u = D\partial^2_{xx}u, \quad u(x,0) = \delta_0(x-a)$$

we get the solution $u(x,t) = K(x-a,t)$, and if we also change mass one to mass $m$ in this point, the solution will be $mK(x-a,t)$.

Based on this we can derive an expression for the solution to the problem

$$\partial_t u = D\partial^2_{xx}u, \quad u(x,0) = u_0(x).$$

The solution consists of the sum of all solutions that different points contribute with from the start. More precisely: if in the point $y$ start with $u_0(y)dy$ individuals, then this point contributes with $K(x-y,t)u_0(y)dy$ individuals to the point $x$ at time $t$. The total number of particles in the point $x$ at time $t$ is therefore given by the sum of these:

$$u(x,t) = \int_{-\infty}^{\infty} K(x-y,t)u_0(y)dy.$$ 

With this we have derived a formula for what the solution to the diffusion equation on the real line looks like.

**Example 7** The equation

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} + r \rho$$

has been applied to model the dispersal of bacteria along a tube. Då kommer funktionen $n(x,t) = e^{-rt} \rho(x,t)$ att uppfylla diffusionsekvationen, vilket man använder vid lösandet av den.

The discussion above applies to the situation when we consider the diffusion equation on the whole line. Another type of problem is to find the solution of the problem

$$\partial_t u = D\partial^2_{xx}u, \quad 0 \leq x \leq 1, \quad u(0,t) = u(1,t) = 0.$$ 

Here we can make a change of variable in the space variable and get an equation in the alternative form

$$\partial_t u = \partial^2_{xx}u, \quad 0 < x < L, \quad u(0,t) = u(L,t) = 0.$$ 

We choose to analyse this, since this form we need the result in the next chapter.

We start by looking for solutions of a very special form:

$$u(x,t) = a(t)b(x),$$

where we assume that $a(0) = 1$ in order to get uniqueness. If we insert this into the differential equation we get

$$a'(t)b(x) = a(t)b''(x), \quad b(0) = b(L) = 0.$$
If we divide with \(a(t)b(x)\) this implies that we should have that
\[
\frac{a'(t)}{a(t)} = \frac{b''(x)}{b(x)},
\]
where the left hand side depends only on \(t\) and the right hand side only on \(x\). Consequently they must both be equal to a constant \(\lambda\), from which we then have the two equations
\[
a'(t) = \lambda a(t), \quad b''(x) = \lambda b(x).
\]
But here \(b(x)\) should vanish in both \(x = 0\) and \(x = L\), which is only possible if \(\lambda\) is negative, \(\lambda = -\mu^2 < 0\). And then
\[
b''(x) + \mu^2 b(x) = 0, \quad b(0) = 0 \implies b(x) = A \sin(\mu x)
\]
for some constant \(A\). The additional condition \(b(L) = 0\) means that \(\mu L\) must be a zero of the sinus function, so
\[
\mu L = n\pi
\]
for some integer \(n\). This means that \(\mu\) cannot be any number, but only of the form
\[
\mu = \frac{n\pi}{L}.
\]
Possible choices for \(b\) are therefore
\[
b(x) = A \sin\left(\frac{n\pi x}{L}\right)
\]
for some constant \(A\) and some integer \(n\), which we can assume to be positive (figure out why for yourself).

From this it is easy to determine \(a(t)\) as
\[
a(t) = e^{-\frac{\mu^2 x^2}{L^2} t}
\]
In summary, if
\[
u_0(x) = A \sin\left(\frac{n\pi x}{L}\right)
\]
for some integer \(n\), the solution to the problem is
\[
u(x, t) = A e^{-\frac{\mu^2 x^2}{L^2} t} \sin\left(\frac{n\pi x}{L}\right).
\]
But this is a very special initial condition, so it will not solve all problems. But it actually does. Every function such that
\[
u_0(0) = u_0(L) = 0
\]
can be written
\[
u_0(x) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right)
\]
for some constants \(A_n\). The corresponding solution to the diffusion equation is then given by
\[
u(x, t) = \sum_{n=1}^{\infty} A_n e^{-\frac{\mu^2 x^2}{L^2} t} \sin\left(\frac{n\pi x}{L}\right).
\]
We do not concern ourselves with why this is true (Fourierseries), the important point is that if we can solve the equation for functions of the form \(a(t)b(x)\), we an also solve it for all (reasonable) initial conditions.
**Exercises**

**Exercise 1** This exercise discusses the concept *genetic cline*.

a) Consider a long line of sites, and at each site is located a population that carries a one-locus, two-allele genetic trait. Let $g_{m,n}$ denote the frequency of the gene pool at site $m$ that is type a. If the sites are isolated, the reproduction of the population would be described by the recursion

$$g_{m,n+1} = f(g_{m,n}), \quad f(g) = \frac{\rho g^2 + \sigma g(1-g)}{\rho g^2 + 2\sigma g(1-g) + \tau(1-g)^2}.$$ 

We suppose that at each sampling time the population reproduces and then does a random walk to neighboring sites. Thus

$$g_{m,n+1} = \lambda(f(g_{m+1,n}) + f(g_{m-1,n})) + (1 - 2\lambda)f(g_{m,n}).$$ 

Solve this problem numerically for 100 sites using periodic boundary conditions. Suppose that at some sites $\rho > \sigma > \tau$ and at others $\rho < \sigma < \tau$. Using computer simulations, describe the population’s final distribution of genotypes. An irregular distribution of gene pool proportions is called a *genetic cline*.

b) Show that if we replace $g_{m,n}$ by a smooth function $g(x,y)$ for which $g(m\delta x, n\delta t) = g_{m,n}$ for small sampling steps $\delta x$ and times $\delta t$, then the recursion scheme is approximated by the diffusion equation

$$\frac{\partial g}{\partial t} = D\frac{\partial^2 g}{\partial x^2} + g(1-g)((\alpha - 2\beta + \gamma)g - \alpha + \beta).$$ 

c) If $D = 1$ and cubic polynomial is $g(1-g)(4-8g)$, then show that

$$g(x,t) = \frac{1}{2}(\tanh x + 1)$$

is a solution of the diffusion approximation. Therefore, we have a convenient mathematical example of a cline. Explain how this cline is formed and maintained by selective pressures.

**Exercise 2** The larvae of the parasitic worm (*Trichostrongylus retortaeformis*) hatch from eggs in sheep and rabbit excreta. The larvae disperse randomly on the grass and are consequently eaten by sheep and rabbits. In the intestines the cycle starts again. Consider the one-dimensional problem in which the larvae disperse with constant diffusion and have a mortality proportional to the population. Show that $n$ satisfies

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} - \mu n$$

(positive constants) where $n$ is the larvae population. Find the population distribution at any $x$ and $t$ arising form $N_0$ larvae being released at $x = 0$ at $t = 0$. Show that as $t \to \infty$ the population dies out.
If the larvae lay eggs at a rate proportional to the population of the larvae:

\[
\frac{\partial E}{\partial t} = \lambda n
\]

where \( E(x,t) \) is the egg population density, show that in the limit as \( t \to \infty \) a nonzero spatial distribution of eggs persists.

**Exercise 3** Consider the density-dependent diffusion model for insect dispersal which includes a linear death process:

\[
\frac{\partial n}{\partial t} = D_0 \frac{\partial}{\partial x} \left( \left( \frac{n}{n_0} \right)^m \frac{\partial n}{\partial x} \right) - \mu n.
\]

If \( Q \) insects are released at \( x = 0 \) at \( t = 0 \), show that the population wavefront reaches a finite distance \( x_{\text{max}} \) from \( x = 0 \) as \( t \to \infty \), where

\[
x_{\text{max}} = \frac{r_0}{(\mu n t_0)^{1/(m+2)}}.
\]

**Appendix: A diffusion approximation to the Fisher-Wright model**

In order to illustrate a situation when the \( x \)-variable is not a space variable in a diffusion equation we will not reconsider the Fisher-Wright model and adding selection and mutations to it. Doing that transforms the relatively simple model into one that is hard to analyse. In order to get something we can analyse we will look at an approximation to the model when \( N \) is large.

We start more general and consider a Markov chain with \( M+1 \) different states \( E_0, E_1, \ldots, E_M \). We assume that jumps between states occur at times \( t_n = n\Delta t \). As usual \( p_{ij} = P(E_i|E_j) \) are the transition probabilities and we also introduce the stochastic variable \( \xi_n \) which notes the state we are in after \( n \) transitions. More precisely \( \xi_n = k \) if the system is in the state \( E_k \) at that time. We also set \( x_k = k/M \).

Now assume that

a) There is a smooth function \( p(x,t) \), defined for \( 0 \leq x \leq 1 \) and \( t \geq 0 \), such that

\[
P(\xi_n = k) = p(x_k, t_n) \frac{1}{M}.
\]

b) There is a smooth function \( p(x, y, t) \) such that

\[
P(\xi_n = i|\xi_0 = j) = p(x_i, x_j, t_n) \frac{1}{M}.
\]

In other words, we make smooth interpolation between the absolute probabilities and the transition probabilities in the model.

The obvious statement

\[
P(\xi_n = k) = \sum_{j=0}^{M} P(\xi_n = k|\xi_0 = j)P(\xi_0 = j)
\]
is now expressed as

\[ p(x_k, t_n) = \sum_{j=0}^{M} p(x_k, x_j, t_n) p(x_j, 0) \frac{1}{M} \approx \int_{0}^{1} p(x_k, y, t_n) p(y, 0) dy, \]

where the approximation refers to the sum as an Riemann sum for the integral. Also introduce the start distribution

\[ f(x) = p(x, 0). \]

When \( M \to \infty \) this turns into

\[ p(x, t) = \int_{0}^{1} p(x, y) f(y) dy, \quad t = t_n. \]

Furthermore

\[ P(\xi_{n+m} = i|\xi_0 = j) = \sum_{k=0}^{M} P(\xi_{n+m} = i|\xi_m = k) P(\xi_m = k|\xi_0 = j), \]

which means that

\[ p(x_i, x_j, t_n + t_m) = \sum_{k=0}^{M} p(x_i, x_k, t_n) p(x_k, x_j, t_m) \frac{1}{M} \approx \int_{0}^{1} p(x_i, z, t_n) p(z, x_j, t_m) dz \]

in the same sense as above. Letting \( N \to \infty \) we get

\[ p(x, y, t + s) = \int_{0}^{1} p(x, z, t)p(z, y, s) dz. \]

From this it follows that

\[ p(x, y, t + \Delta t) = \int_{0}^{1} p(x, z, \Delta t)p(z, y, t)dz. \]

Now let \( \phi(x) \) be an arbitrary function such that \( \phi(0) = \phi(1) = \phi'(0) = \phi'(1) = 0 \) and consider

\[ \int_{0}^{1} \phi(x)(p(x, y, t+\Delta) - p(x, y, t))dx = \int_{0}^{1} \int_{0}^{1} (\phi(x)p(x, z, \Delta t)p(z, y, t)dz)dx - \int_{0}^{1} \phi(x)p(x, y, t)dx. \]

If we change the integration order in the double integral and change the variable names we get

\[ \int_{0}^{1} p(x, y, t) \int_{0}^{1} (\phi(z)p(z, x, \Delta t)dz)dx \]

which means that the expression is

\[ \int_{0}^{1} p(x, y, t) (\int_{0}^{1} p(z, x, \Delta t)(\phi(z) - \phi(x))dz)dx \]

where we have used that

\[ \int_{0}^{1} p(z, x, \Delta t)dz = 1. \]
We now add a very natural assumption: during a short time interval the process can only make one small jump. Mathematically we express this to be that during a short time interval $\Delta t$ the probability that the total jump is $> \epsilon$ from a point $x$ is given by the integral
\[
\int_{|z-x|>\epsilon} p(z, x, \Delta t)dz
\]
and we assume this is small as $\Delta t \to 0$. More precisely, it is assumed to be $o(\Delta t)$. And this for every given $\epsilon > 0$ (we need smaller $\Delta t$ for smaller $\epsilon$). It is close to the assumption we did for the diffusion process above.

All this means that it is only the integral
\[
\int_0^1 p(x, y, t)\left(\int_{|z-x|\leq\epsilon} p(z, x, \Delta t)(\phi(z) - \phi(x))dz\right)dx
\]
that is of interest, and if we expand $\phi$ in its Taylor series around $x$:
\[
\phi(z) = \phi(x) + \phi'(x)(z-x) + \frac{1}{2}\phi''(x)(z-x)^2 + \ldots
\]
and plug that into the integral we get
\[
\int_0^1 p(x, y, t)\left(\phi'(x)m(x, \Delta t) + \frac{1}{2}\phi''(x)v(x, \Delta t)\right)dx,
\]
where
\[
m(x, \Delta t) = \int_0^1 (z-x)p(z, x, \Delta t)dz, \quad v(x, \Delta t) = \int_0^1 (z-x)^2p(z, x, \Delta t)dz.
\]

Note that we go rather quickly here: the integral above should really be the an integral over the the interval $|z-x| \leq \epsilon$, but points outside it does not give negligible contributions when $\Delta t$ is small, so the approximation is valid.

Here
\[
m(x, 0) = \int_0^1 (z-x)p(z, x, 0)dz = 0,
\]
since $p(z, x, 0) = 0$ when $z \neq x$, and similarly for $v(x, \Delta t)$. If we assume we can throw away higher order terms and also assume that $m(x, \Delta t) = m(x)\Delta t(1+o(1)$ and similarly for $v(x, \Delta t)$, we get
\[
\int_0^1 \phi(x)p(x, y, t + \Delta t) - p(x, y, t)\frac{\Delta t}{\Delta t} dx = \int_0^1 p(x, y, t)(\phi'(x)m(x)(1+o(1) + \frac{1}{2}\phi''(x)v(x)(1+o(1)))dx.
\]
Let $\Delta t \to 0$ we see that
\[
\int_0^1 \phi(x)\partial_t p(x, y, t)dx = \int_0^1 (\phi'(x)(m(x)p(x, y, t)) + \frac{1}{2}\phi''(x)(v(x)p(x, y, t)))dx.
\]
In the right hand side we now make a partial integration and get
\[
\int_0^1 \phi(x) \partial_t p(x, y, t) dx = \int_0^1 \phi(x)(-\partial_x(m(x)p(x, y, t)) + \frac{1}{2}\partial_{xx}^2(v(x)p(x, y, t))) dx.
\]
Since \( \phi \) is an arbitrary function it follow that \( p(x, y, t) \) must solve Fokker-Planks equation
\[
\partial_t p(x, y, t) = -\partial_x(m(x)p(x, y, t)) + \frac{1}{2}\partial_{xx}^2(v(x)p(x, y, t))
\]
where
\[
m(x) = \lim_{\Delta t \to 0} ((\Delta t)^{-1} \int_0^1 (z-x)p(z, x, \Delta t) dz, \quad v(x) = \lim_{\Delta t \to 0} ((\Delta t)^{-1} \int_0^1 (z-x)^2 p(z, x, \Delta t) dz
\]
Not ignoring higher order terms in the Taylor expansion would give higher order momenst, but in biological applications we often have that \( m(x) \), called the drift, and \( v(x) \), called the variance, are of order \( 1/M \) whereas higher moments are of at least order \( 1/M^2 \) and can therefore be ignored for large \( M \).

For the Fisher-Wright model we have
\[
m(x) = sx(1-x) - ux + v(1-x), \quad v(x) = \frac{x(1-x)}{N}.
\]
That the factor \( x(1-x) \) appears in both functions is natural: for \( x \) to change we need to replace an individual of one type by an individual of the other type and the rate with which this can happen depends on the product of their frequencies.