

eDNA

Richard Arnold

2024-07-16

Particle dispersal

If there is a mass released at time t_0 at location \mathbf{x}_0 then it is transported by advection (water flow) and dispersion (turbulence).

Let $p(\mathbf{x}, t | \mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)$ be the proportion of the mass released that is found at location \mathbf{x} at time t , given water flow $\mathbf{v}()$ and dispersion $\mathbf{a}()$. Ψ is a set of additional parameters controlling the flow/dispersion model.

A simple model is a simple Gaussian spread in 2D:

$$\begin{aligned} p(\mathbf{x}, t | \mathbf{x}_0, t_0, \mathbf{v}, \mathbf{a}, \Psi) &= \frac{I(t - t_0 \geq 0)}{2\pi\sqrt{a_x a_y}(t - t_0 + t_\varepsilon)} \exp\left(-\frac{(x - x_0 - v_x(t - t_0))^2}{2a_x(t - t_0 + t_\varepsilon)} - \frac{(y - y_0 - v_y(t - t_0))^2}{2a_y(t - t_0 + t_\varepsilon)}\right) \quad (1) \\ &= g(\mathbf{x}, t - t_0 | \mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) \quad (2) \end{aligned}$$

where a small positive temporal offset $\Psi = \{t_\varepsilon\}$ is added to avoid a singularity at $t = t_0$.

We have

$$\iint p(\mathbf{x}, t | \mathbf{x}_0, t_0, \mathbf{v}, \mathbf{a}, \Psi) d\mathbf{x} = 1$$

for all times $t > t_0$.

Note that $g()$ has a symmetry in which we can interchange \mathbf{x} and \mathbf{x}_0 and at the same time reverse the velocity field, so for any time interval $u \geq 0$:

$$g(\mathbf{x}, u | \mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = g(\mathbf{x}_0, u | \mathbf{x}, -\mathbf{v}, \mathbf{a}, \Psi) \quad (3)$$

```
source("funcs.R")
```

```
nx <- 101
ny <- 101
xmin <- 0; xmax <- 100
ymin <- 0; ymax <- 100
xvec <- seq(from=xmin, to=xmax, length=nx)
yvec <- seq(from=ymin, to=ymax, length=ny)

x0 <- mean(c(xmin, xmax))
y0 <- mean(c(ymin, ymax))
t0 <- 0
vx <- 5; vy <- 1
ax <- 2^2; ay <- 4^2
teps <- 0.01
```

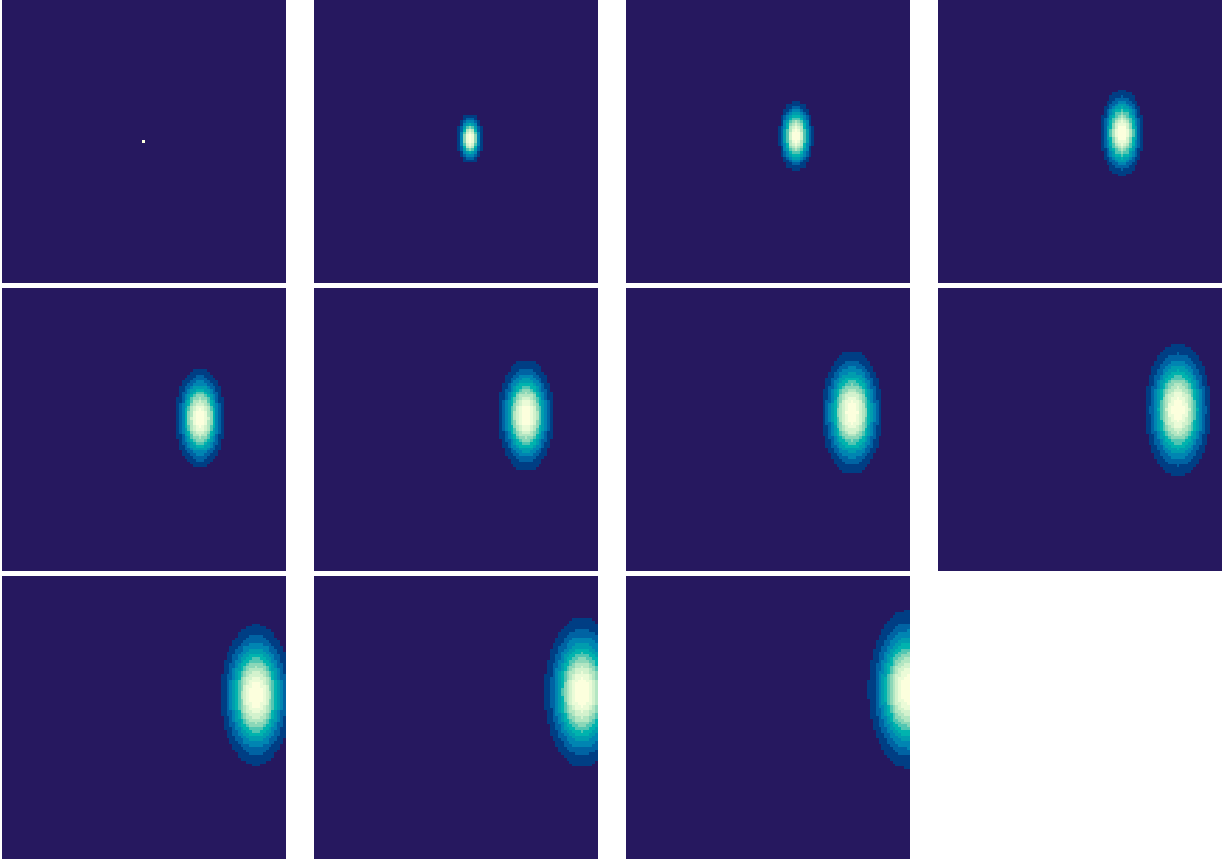
For example here is the spread over time of mass released at $\mathbf{x}_0 = (50, 50)^T$ at time $t = 0$ with constant flow field $\mathbf{v} = (5, 1)^T$ and dispersion $\mathbf{a} = (4, 16)^T$ and with $t_\varepsilon = 0.01$.

Forward tracking: $p(\mathbf{x}, t | \mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)$: $t_0, \dots, t_0 + 10$ (no decay)

```

par(mfrow=c(3,4))
par(mar=0.1*c(1,1,1,1))
colvec <- hcl.colors(12, "YlOrRd", rev = TRUE)
colvec <- hcl.colors(12, "YlGnBu")#, rev = TRUE)
for(t in 0:10) {
  amat <- outer(xvec,yvec, dispfunc, t=t,
               x0=x0,y0=y0,t0=0, vx=vx,vy=vy, ax=ax,ay=ay, teps=teps)
  #if(t==0) amax <- max(amat)
  amax <- max(amat)
  breaks <- seq(from=0, to=amax, length=length(colvec)+1)
  image(xvec, yvec, amat, asp=1, col=colvec, breaks=breaks, axes=FALSE)
  #image(xvec, yvec, amat,asp=1, main=bquote("Forwards:" ~ t==.(t)), xlab="x", ylab="y",
  #      col=colvec, breaks=breaks)
}

```



A simple backtracking model simply reverses the flow field. If we are interested in the likely origin of mass observed at location \mathbf{x}_1 at time t_1 we can backtrack by reversing $\mathbf{v}()$.

We might assume that the likelihood that the observed mass was released at location \mathbf{x} at time $t < t_1$ is given by

$$\tilde{p}(\mathbf{x}, t | \mathbf{x}_1, t_1, \mathbf{v}(), \mathbf{a}(), \Psi) = g(\mathbf{x}, t_1 - t | \mathbf{x}_1, -\mathbf{v}(), \mathbf{a}(), \Psi) \quad (4)$$

$$= p(\mathbf{x}, -t | \mathbf{x}_1, -t_1, -\mathbf{v}(), \mathbf{a}(), \Psi) \quad (5)$$

In the simple Gaussian model above this is

$$\tilde{p}(\mathbf{x}, t | \mathbf{x}_1, t_1, \mathbf{v}(), \mathbf{a}(), \Psi) \quad (6)$$

$$= p(\mathbf{x}, t_1 | \mathbf{x}_1, t, -\mathbf{v}(), \mathbf{a}(), \Psi) \quad (7)$$

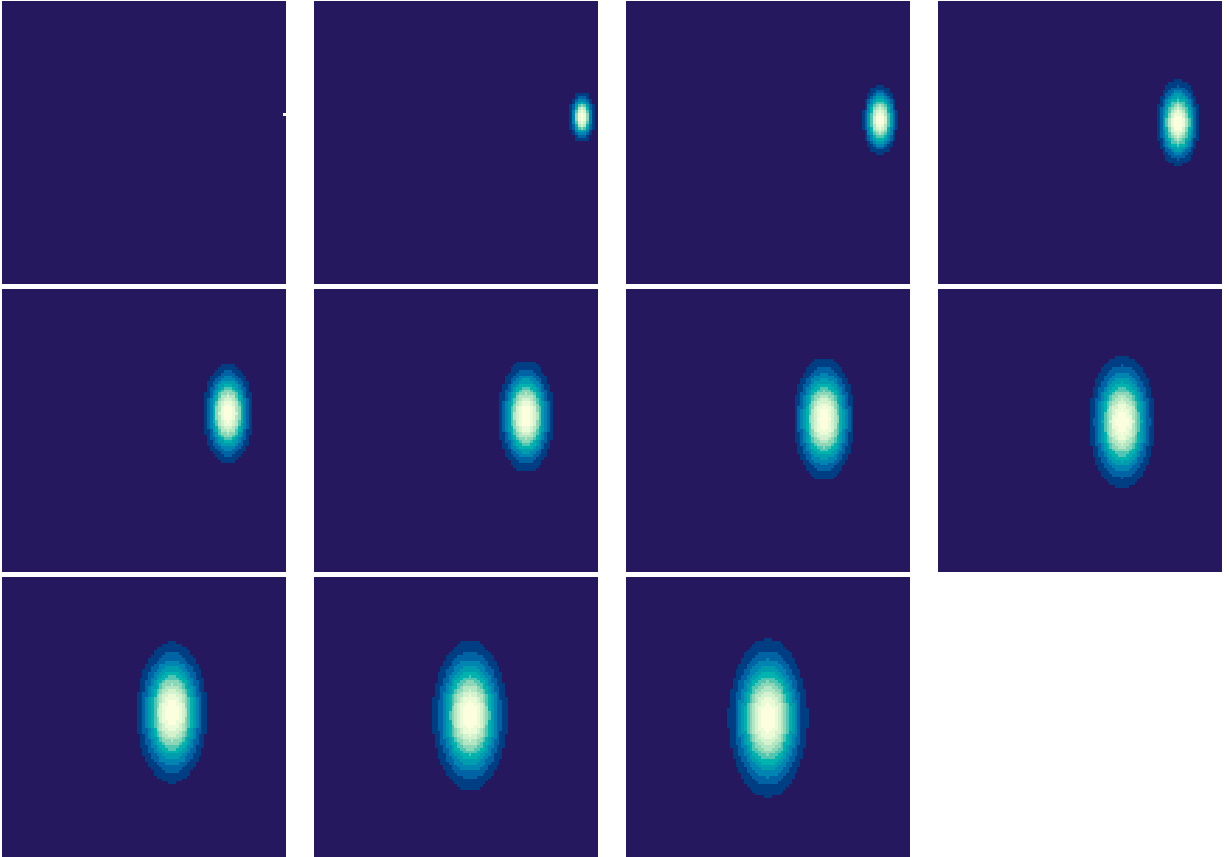
$$= p(\mathbf{x}, -t | \mathbf{x}_1, -t_1, -\mathbf{v}(), \mathbf{a}(), \Psi) \quad (8)$$

$$= g(\mathbf{x}, t_1 - t | \mathbf{x}_1, -\mathbf{v}(), \mathbf{a}(), \Psi) \quad (9)$$

$$= \frac{I(t_1 - t \geq 0)}{2\pi\sqrt{a_x a_y}(t_1 - t + t_\varepsilon)} \exp\left(-\frac{(x - x_1 + v_x(t_1 - t))^2}{2a_x(t_1 - t + t_\varepsilon)} - \frac{(y - y_1 + v_y(t_1 - t))^2}{2a_y(t_1 - t + t_\varepsilon)}\right) \quad (10)$$

Back tracking: $\tilde{p}(\mathbf{x}, t | \mathbf{x}_1, t_1, \mathbf{v}(), \mathbf{a}(), \Psi) = p(\mathbf{x}, -t | \mathbf{x}_1, -t_1, -\mathbf{v}(), \mathbf{a}(), \Psi)$: $t_1, \dots, t_1 - 10$ (no decay)

```
par(mfrow=c(3,4))
par(mar=0.1*c(1,1,1,1))
x1 <- 100; y1 <- 60
t1 <- 10
for(t in 10:0) {
  amat <- outer(xvec,yvec, dispfunc, t=-t,
                x0=x1,y0=y1,t0=-t1, vx=-vx,vy=-vy, ax=ax,ay=ay, teps=teps)
  #if(t==10) amax <- max(amat)
  amax <- max(amat)
  breaks <- seq(from=0, to=amax, length=length(colvec)+1)
  image(xvec, yvec, amat,asp=1, col=colvec, breaks=breaks, axes=FALSE)
  #image(xvec, yvec, amat,asp=1, main=bquote("Backwards: " ~ t==.(t)), xlab="x", ylab="y",
  #      col=colvec, breaks=breaks)
}
```



Continuous release with decay

Assume that at location \mathbf{x}_0 there is a continuous release of mass at a rate N_0 kg/s. Further assume that the mass decays at a rate

$$d(t|T_h, T_m) = 2^{-t/T_h} I(0 \leq t \leq T_m)$$

where T_h is the decay half life and T_m is the maximum age at which all particles have disintegrated. If we set $\lambda = (1/T_h) \log 2$ then

$$d(t|T_h, T_m) = e^{-\lambda t} I(0 \leq t \leq T_m)$$

The steady state mass in the system is then

$$M_0 = \int_{-\infty}^t N_0 d(t - t_0|T_h, T_m) dt_0 \quad (11)$$

$$= N_0 \int_{-\infty}^t e^{-\lambda(t-t_0)} I(t_0 > t - T_m) dt_0 \quad (12)$$

$$= N_0 \int_{t-T_m}^t e^{-\lambda(t-t_0)} dt_0 \quad (13)$$

$$= N_0 \int_0^{T_m} e^{-\lambda u} du \quad (14)$$

$$= \frac{N_0}{\lambda} [1 - e^{-\lambda T_m}] \quad (15)$$

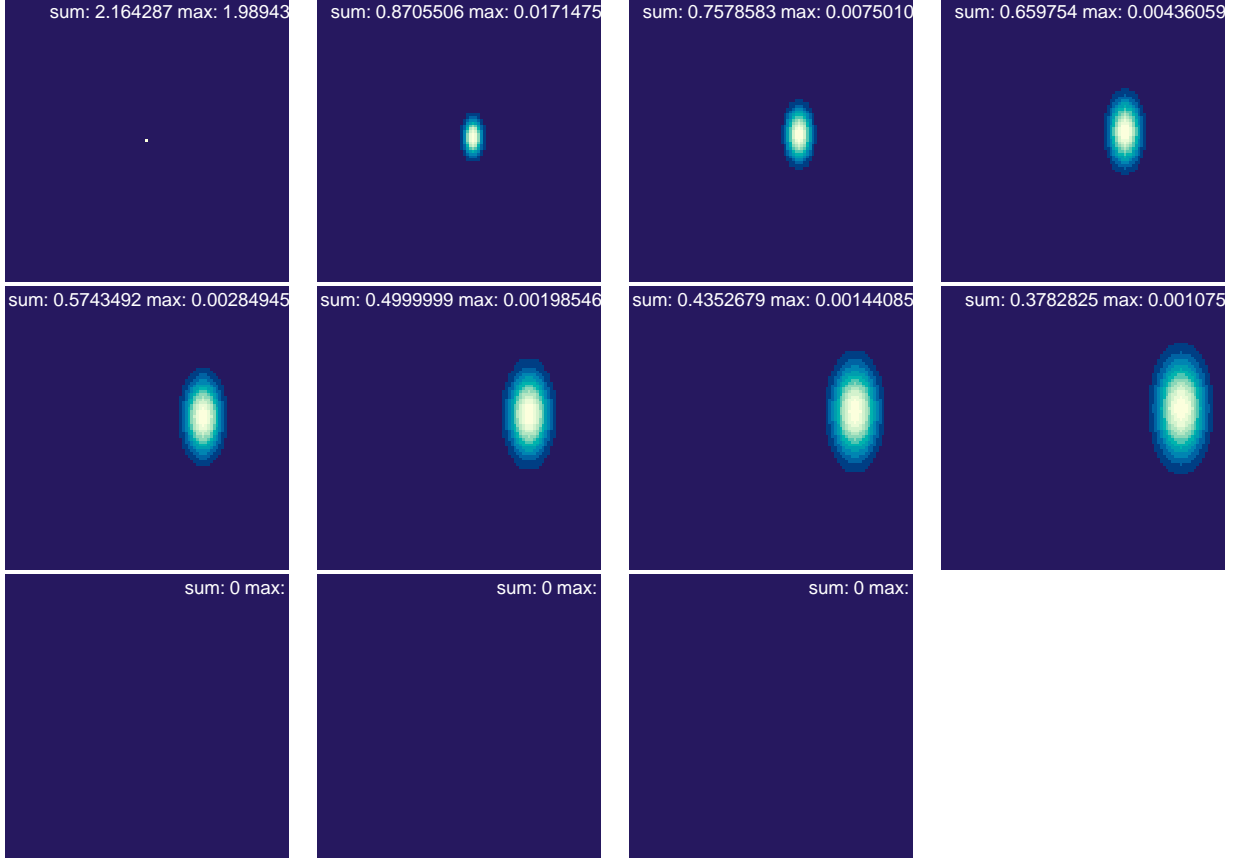
$$= \frac{N_0 T_h}{\log 2} [1 - e^{-\lambda T_m}] \quad (16)$$

The mass density transiting per unit time at location \mathbf{x} at time t emitted from a point source at \mathbf{x}_0 at time t_0 is

$$\frac{d\rho(\mathbf{x}, t|N_0, \mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)}{dt} = N_0 d(t - t_0|T_h, T_m) p(\mathbf{x}, t|\mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)$$

Density rate $d\rho(\mathbf{x}, t|N_0, \mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)/dt$, $t = t_0, \dots, t_0 + 10$ (with decay). (Scales differ between panels: note max and sum values.)

```
par(mfrow=c(3,4))
par(mar=0.1*c(1,1,1,1))
n0 <- 1
th <- 5
tm <- 7
for(t in 0:10) {
  amat <- outer(xvec,yvec, denratefunc, t=t,
               x0=x0,y0=y0,t0=0, vx=vx,vy=vy, ax=ax,ay=ay, teps=teps, n0=n0,th=th,tm=tm)
  #amat <- log(1+amat)
  #if(t==0) {
  #  amin <- min(amat)
  #  amax <- max(amat)
  #}
  amin <- min(amat)
  amax <- max(amat)
  breaks <- seq(from=amin, to=amax, length=length(colvec)+1)
  image(xvec, yvec, amat, asp=1, col=colvec, breaks=breaks, axes=FALSE)
  #image(xvec, yvec, amat,asp=1, main=bquote("Forwards with decay:" ~ t==(t)), xlab="x", ylab="y",
  #      col=colvec, breaks=breaks)
  mtext(bquote("sum:" ~.(sum(amat))~"max:" ~.(max(amat))), side=3, adj=1, cex=0.6, line=-1, col="white")
}
```



The total (equilibrium) mass density present is

$$\rho(\mathbf{x}, t | N_0, \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) \quad (17)$$

$$= \rho(\mathbf{x} | N_0, \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) \quad (18)$$

$$= \int_{-\infty}^t N_0 d(t - t_0 | T_h, T_m) p(\mathbf{x}, t | \mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi) dt_0 \quad (19)$$

$$= N_0 \int_{t-T_m}^t e^{-\lambda(t-t_0)} g(\mathbf{x}, t - t_0 | \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) dt_0 \quad (20)$$

$$= N_0 \int_0^{T_m} e^{-\lambda u} g(\mathbf{x}, u | \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) du \quad (21)$$

$$= N_0 h(\mathbf{x} | \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) \quad (22)$$

where

$$h(\mathbf{x} | \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) = \int_0^{T_m} e^{-\lambda u} g(\mathbf{x}, u | \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) du$$

The function $h()$ retains the same symmetry as $g()$, so since

$$g(\mathbf{x}, u | \mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = g(\mathbf{x}_0, u | \mathbf{x}, -\mathbf{v}, \mathbf{a}, \Psi)$$

it follows that

$$h(\mathbf{x}, u | \mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = h(\mathbf{x}_0, u | \mathbf{x}, -\mathbf{v}, \mathbf{a}, \Psi)$$

In the Gaussian dispersal example

$$h(\mathbf{x} | \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) = \frac{1}{2\pi\sqrt{a_x a_y}} \int_0^{T_m} \frac{e^{-\lambda u}}{u + t_\varepsilon} \exp\left(-\frac{(x - x_0 - v_x u)^2}{2a_x(u + t_\varepsilon)} - \frac{(y - y_0 - v_y u)^2}{2a_y(u + t_\varepsilon)}\right) du$$

```

#th <- 5
#tm <- 7
#ax <- 2^2; ay <- 5^2
#amat <- array(0,dim=c(nx,ny))
#nc <- 100
#for(i in 1:nc) {
#  t <- i/nc * 10
#  amat <- amat + outer(xvec,yvec, denratefunc, t=t,
#                        x0=x0,y0=y0,t0=0, vx=vx,vy=vy, ax=ax,ay=ay, teps=teps, n0=n0,th=th,tm=tm)
#}
#image(xvec, yvec, amat,asp=1,
#      main=bquote(t[h]==.(th) ~ ", " ~ t[m]==.(tm)),
#      xlab="x", ylab="y",col=colvec)

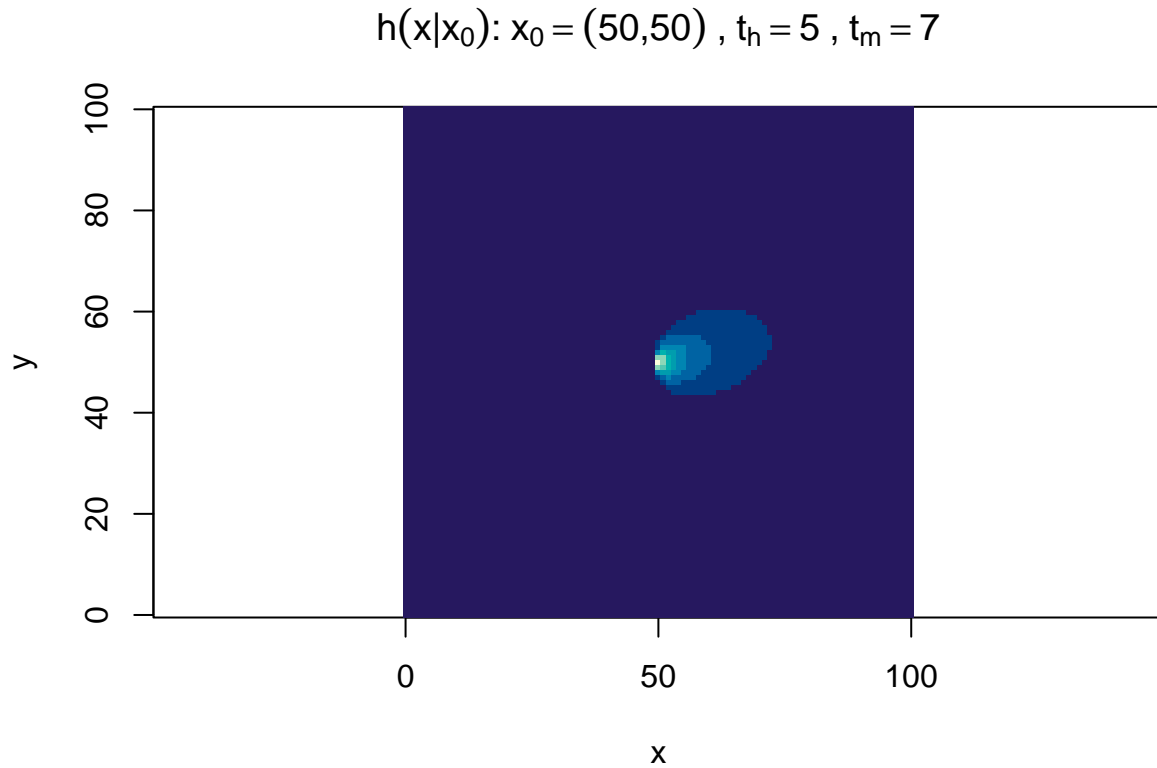
```

Equilibrium Density: $\rho(\mathbf{x}|N_0, \mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = N_0 h(\mathbf{x}|\mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi)$

```

th <- 5
tm <- 7
ax <- 2^2; ay <- 5^2
amat <- n0*outer(xvec,yvec, hfunc,
                 x0=x0,y0=y0, vx=vx,vy=vy, ax=ax,ay=ay, teps=teps, th=th,tm=tm)
image(xvec, yvec, amat, asp=1,
      main=bquote(h(x*"|"*x[0])*": " ~ x[0]==(. (x0)*", "*.(y0)) ~ ", " ~ t[h]==.(th) ~ ", " ~ t[m]==.(tm)),

```



The above figure plots, for a continuous fixed source of rate N_0 at location \mathbf{x}_0 , the equilibrium density at all locations \mathbf{x} : $N_0 h(\mathbf{x}|\mathbf{x}_0)$.

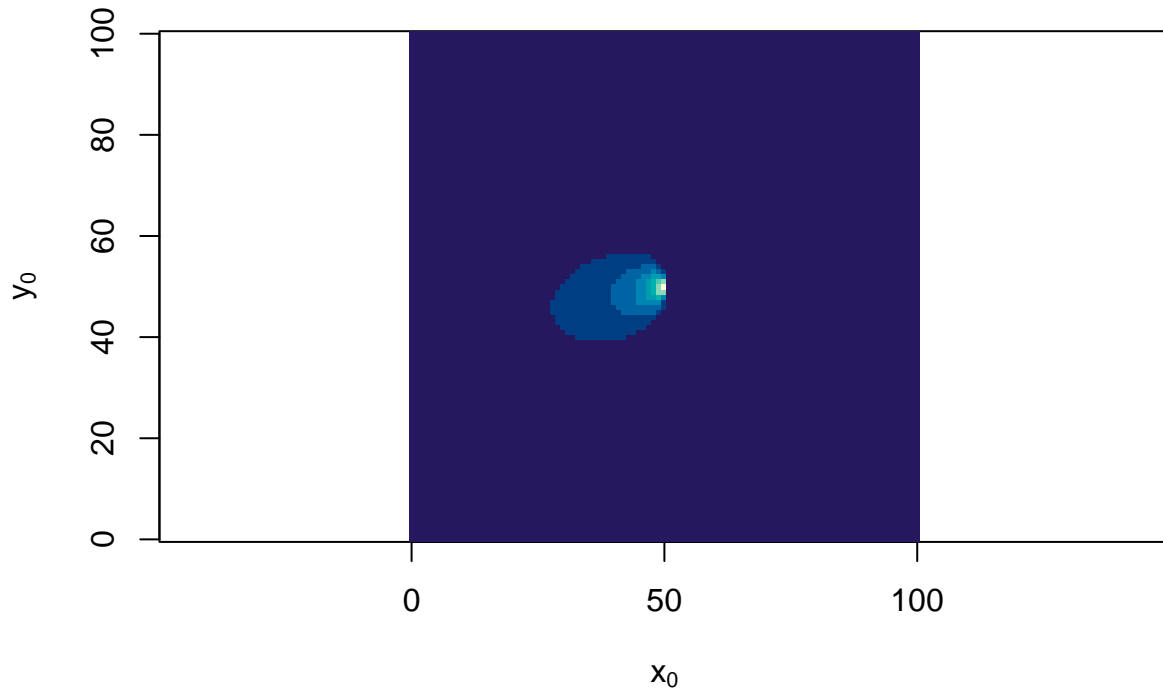
Now plot at each \mathbf{x}_0 the equilibrium density at \mathbf{x}_1 (in the centre of the diagram) that would result from a continuous source at \mathbf{x}_0 . i.e. $N_0 h(\mathbf{x}_1|\mathbf{x}_0)$

$$\begin{aligned} N_0 \tilde{h}(\mathbf{x}_0|\mathbf{x}_1, \mathbf{v}, \mathbf{a}, \Psi) &= N_0 h(\mathbf{x}_1|\mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) \\ &= N_0 h(\mathbf{x}_0|\mathbf{x}_1, -\mathbf{v}, \mathbf{a}, \Psi) \end{aligned}$$

$$N_0 \tilde{h}(\mathbf{x}_0|\mathbf{x}_1, \mathbf{v}, \mathbf{a}, \Psi) = N_0 h(\mathbf{x}_1|\mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = N_0 h(\mathbf{x}_0|\mathbf{x}_1, -\mathbf{v}, \mathbf{a}, \Psi)$$

```
th <- 5
tm <- 7
ax <- 2^2; ay <- 5^2
x1 <- 50; y1 <- 50
rho1 <- 1; mu <- 1
amat <- outer(xvec,yvec, function(x0vec,y0vec) {
  apply(cbind(x0vec,y0vec),1,function(x0y0) {
    hfunc(x=x1,y=y1,x0=x0y0[1],y0=x0y0[2],
    vx=vx,vy=vy, ax=ax,ay=ay, teps=teps, th=th,tm=tm)
  })
})
dim(amat) <- c(length(xvec),length(yvec))
image(xvec, yvec, amat, asp=1,
  main=bquote(h(x[1]*"|"*x[0])*": " ~ x[1]==(. (x0)*", "*.(y0)) ~ ", " ~ t[h]==.(th) ~ ", " ~ t[m]==.(tm)),
  xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
```

$h(\mathbf{x}_1|\mathbf{x}_0): \mathbf{x}_1 = (50,50) , t_h = 5 , t_m = 7$

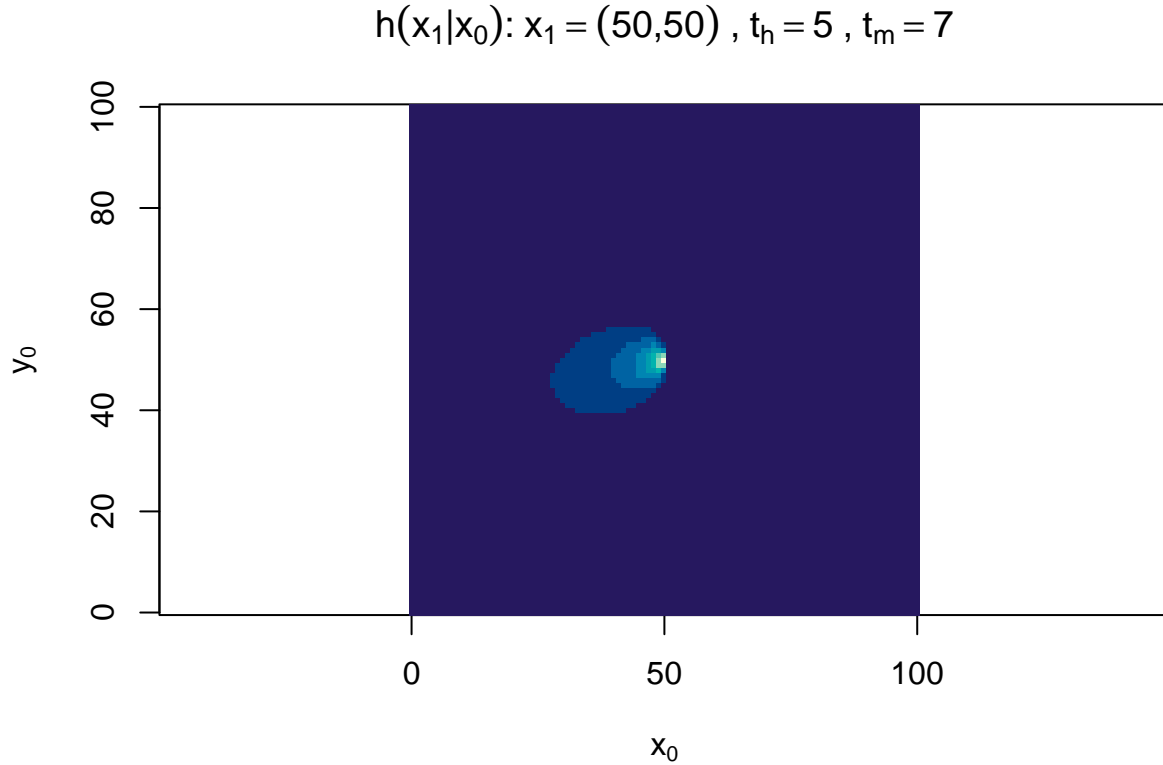


(Computing this the other way using the symmetry of $h()$)

```

th <- 5
tm <- 7
ax <- 2^2; ay <- 5^2
x1 <- 50; y1 <- 50
rho1 <- 1; mu <- 1
amat <- n0*outer(xvec,yvec, hfunc,
                 x0=x1,y0=y1, vx=-vx,vy=-vy, ax=ax,ay=ay, teps=teps, th=th,tm=tm)
image(xvec, yvec, amat, asp=1,
      main=bquote(h(x[1]*"|"*x[0])*":" ~ x[1]==(. (x0)*", "*.(y0)) ~ "," ~ t[h]==.(th) ~ "," ~ t[m]==.(tm)),
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)

```



Probabilistic Inversion

Now assume that at some location \mathbf{x}_1 we observe a density ρ_1 , and that this is an observation made with error:

$$\rho_1 | \rho_1^*, \sigma_e \sim N(\rho_1^*, \sigma_e^2)$$

where the true density is ρ_1^* and the error variance is σ_e^2 .

Assume that there is a single continuous source of unknown strength N_0 and unknown location \mathbf{x}_0 . Priors for the strength and location are $\pi(N_0)$ and $\pi(\mathbf{x}_0)$ respectively.

Assume that the velocity field $\mathbf{v}()$ and the diffusion field $\mathbf{a}()$ are known.

Using the results from the previous section, conditional on N_0 and \mathbf{x}_0 the observed density at \mathbf{x}_1 is

$$\rho_1^* = N_0 h(\mathbf{x}_1 | \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi)$$

which we write for brevity as

$$\rho_1^* = h_1^*(\mathbf{x}_0) = h_1^*$$

but noting its dependence on \mathbf{x}_0 .

It follows that the joint distribution of the unknown source strength N_0 and location \mathbf{x}_0 is

$$\begin{aligned} p(N_0, \mathbf{x}_0 | \rho_1, \mathbf{x}_1) &\propto \pi(N_0) \pi(\mathbf{x}_0) \exp\left(-\frac{1}{2\sigma_e^2} [\rho_1 - N_0 h_1^*(\mathbf{x}_0)]^2\right) \\ &\propto \pi(N_0) \pi(\mathbf{x}_0) \exp\left(-\frac{(h_1^*)^2}{2\sigma_e^2} \left[N_0 - \frac{\rho_1}{h_1^*}\right]^2\right) \end{aligned}$$

A suitable prior for N_0 is the Exponential(μ) distribution

$$\begin{aligned} N_0 &\sim \text{Exp}(\mu) \\ \pi(N_0) &= \mu e^{-\mu N_0} \end{aligned}$$

This can be marginalised out of the posterior distribution as follows. Firstly the joint distribution is

$$\begin{aligned} p(N_0, \mathbf{x}_0 | \rho_1, \mathbf{x}_1) &\propto \pi(\mathbf{x}_0) e^{-\mu N_0} \exp\left(-\frac{(h_1^*)^2}{2\sigma_e^2} \left[N_0 - \frac{\rho_1}{h_1^*}\right]^2\right) \\ &\propto \pi(\mathbf{x}_0) \exp\left(-\frac{(h_1^*)^2}{2\sigma_e^2} \left[N_0 - \left(\frac{\rho_1}{h_1^*} - \frac{\mu\sigma_e^2}{(h_1^*)^2}\right)\right]^2 - \frac{\rho_1\mu}{h_1^*} + \frac{\mu^2\sigma_e^2}{2(h_1^*)^2}\right) \end{aligned}$$

Integrating out N_0 leads to

$$p(\mathbf{x}_0 | \rho_1, \mathbf{x}_1) \propto \pi(\mathbf{x}_0) \frac{\sigma_e}{h_1^*} \exp\left(-\frac{\rho_1\mu}{h_1^*} + \frac{\mu^2\sigma_e^2}{2(h_1^*)^2}\right)$$

If we let $\sigma_e \rightarrow 0$ (indicating very accurate observations) we are left with a very simple form for the posterior:

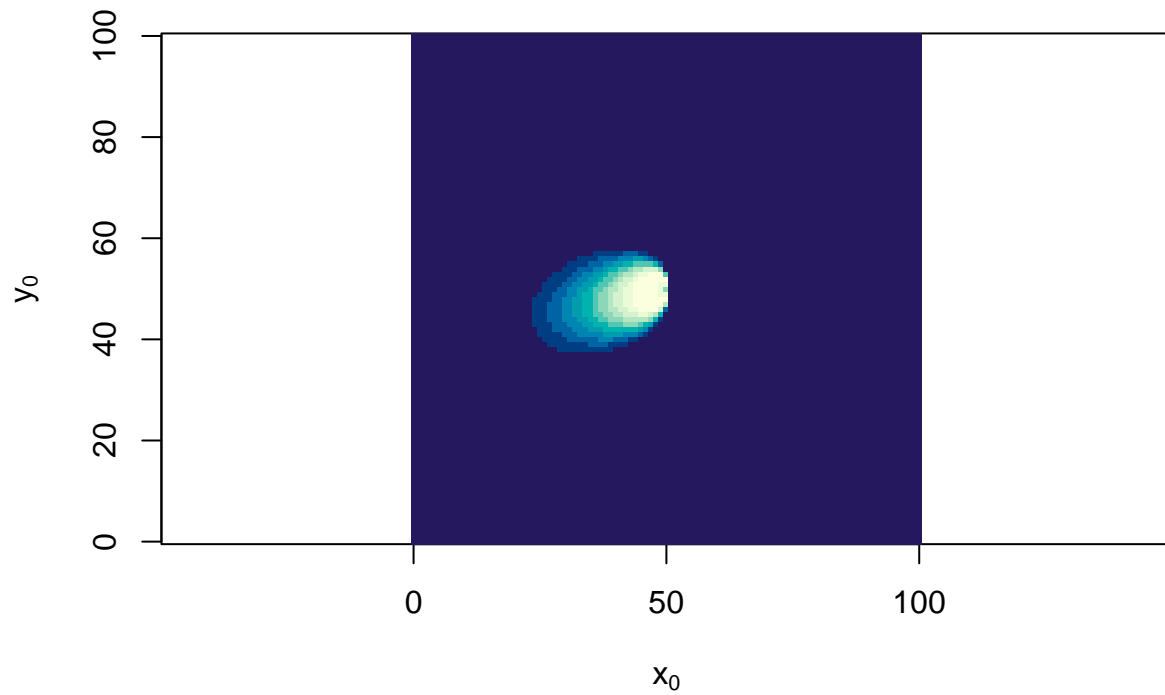
$$p(\mathbf{x}_0 | \rho_1, \mathbf{x}_1) \propto \pi(\mathbf{x}_0) \frac{1}{h_1^*} e^{\frac{\rho_1\mu}{h_1^*}}$$

which we plot below for a uniform prior on the source location $\pi(\mathbf{x}_0) \propto 1$.

```
th <- 5
tm <- 7
ax <- 2^2; ay <- 5^2
x1 <- 50; y1 <- 50
rho1 <- 1; mu <- 1
amat <- n0*outer(xvec,yvec, hfunc,
                 x0=x1,y0=y1, vx=-vx,vy=-vy, ax=ax,ay=ay, teps=teps, th=th,tm=tm)

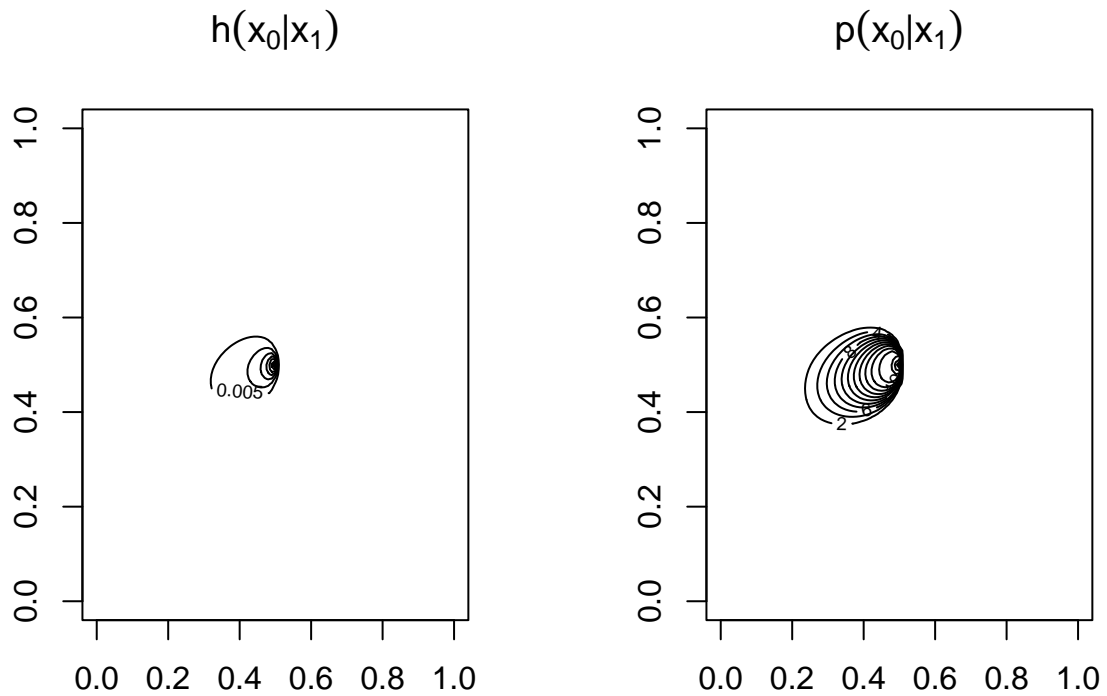
qmat <- exp(-max(amat)/3.0/amat) * (1/amat)
image(xvec, yvec, qmat, asp=1,
      main=bquote(p(x[0]*"|"*x[1])*":" ~ x[1]==(. (x0)*","*(.y0)) ~ "," ~ t[h]==.(th) ~ "," ~ t[m]==.(tm)
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
```

$p(x_0|x_1): x_1 = (50,50) , t_h = 5 , t_m = 7$



Plotting again, but as a contour map

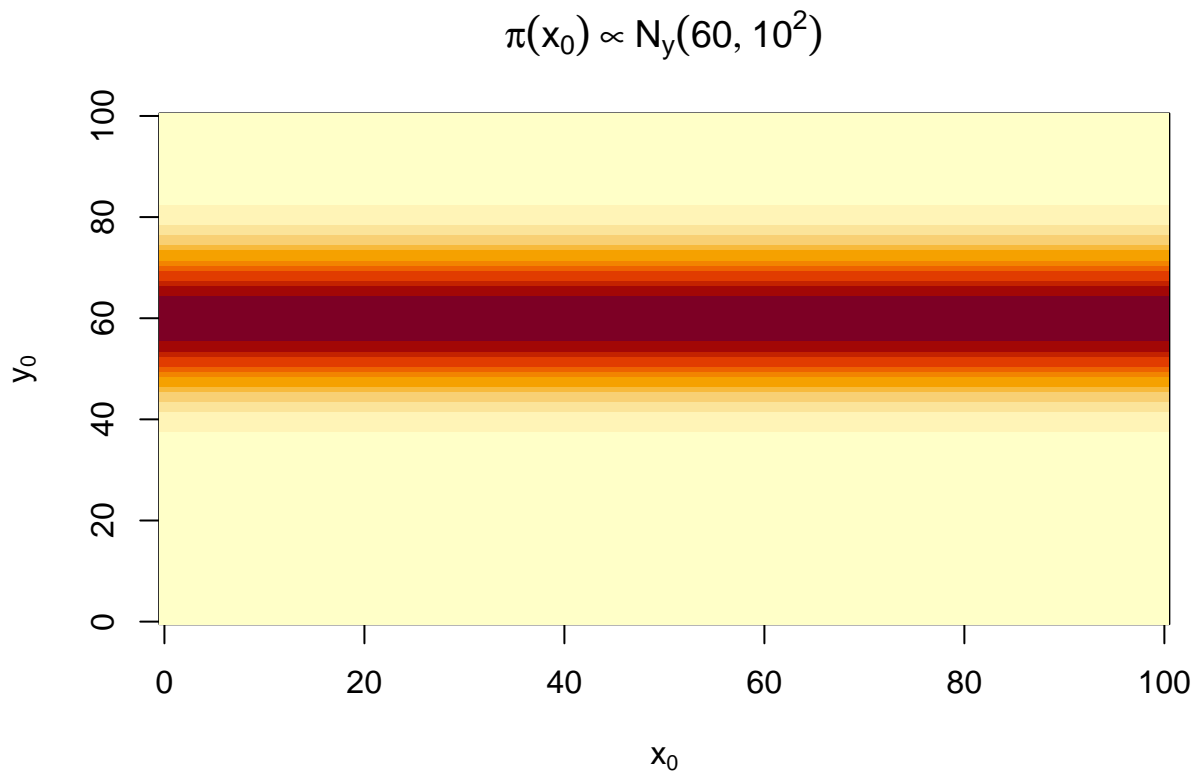
```
par(mfrow=c(1,2))
contour(amat, main=expression(h(x[0]*"|"*x[1])))
contour(qmat, main=expression(p(x[0]*"|"*x[1])))
```



A different prior: e.g. normally distributed in y (centered on y_a with standard deviation σ_a) and uniform in x :

$$\pi(\mathbf{x}_0) \propto \exp\left(-\frac{1}{2\sigma_a^2} [x_{0y} - y_a]^2\right)$$

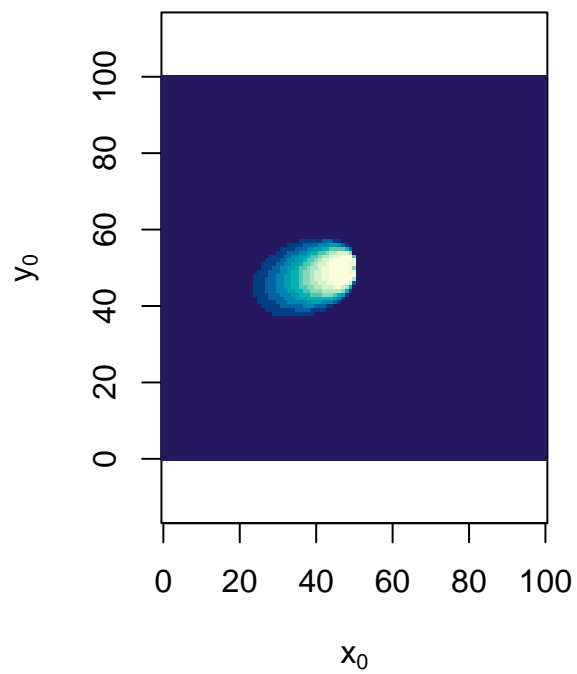
```
ya <- 60
sigmaa <- 10
p0mat <- t(array(exp(-0.5*(yvec-ya)^2/sigmaa^2), dim=dim(amat)))
image(xvec, yvec, p0mat, main=bquote(pi(x[0]) %prop% N[y](.(ya),.(sigmaa)^2)),
      xlab=expression(x[0]), ylab=expression(y[0]))
```



The effect of the prior is shown at right below

```
par(mfrow=c(1,2))
image(xvec, yvec, qmat, asp=1,
      main=bquote(p(x[0]*"|"*x[1])*": " ~ pi(x[0]) %prop% 1),
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
image(xvec, yvec, p0mat*qmat, asp=1,
      main=bquote(p(x[0]*"|"*x[1])*": " ~ pi(x[0]) %prop% N[y](.(ya),.(sigmaa)^2)),
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
```

$$p(x_0|x_1): \pi(x_0) \propto 1$$



$$p(x_0|x_1): \pi(x_0) \propto N_y(60, 10^2)$$

