BE 159: Signal Transduction and Mechanics in Morphogenesis

Justin Bois

Caltech

Winter, 2019

This document was prepared at Caltech with financial support from the Donna and Benjamin M. Rosen Bioengineering Center.

9 Continuum mechanics I: conservation of mass

As we move into the mechanics of morphogenesis, we need to develop a mathematical framework, similar to our use of mass action kinetics in our studies of signaling. We have already seen some of the results of this analysis in our discussion of Turing patterns and the Ben-Zvi, et al. paper, where we already used some results pertaining to diffusive transport that we will soon derive. We will discuss this more formally now.

We will be working up to a hydrodynamic theory for active viscous nematic fluids, which we will use to model the actomyosin cortex in a developing *C. elegans* embryo. For a recent review of the theory of these types of complex materials, see this paper by Jülicher, Grill, and Salbreux.

9.1 Assumptions about continua

We will be treating cells and tissues as continua, meaning that we do not consider discrete molecules, or in some cases discrete cells. When is this a reasonable thing to do? When can we neglect molecular details?

We can think of an obvious example where it is ok to treat objects as continua. Let's say we are engineering a submarine. We want to design its shape and propeller such that it moves efficiently through water. Do we need to take into account the molecular details of the water? Definitely not! We only need to think about *bulk properties* of the water; its density and viscosity (both of which are temperature dependent). We can also define a velocity of water as a continuum as opposed to thinking about the trajectories of every molecule. So, clearly there are situations where the continuum treatment of a fluid is valid and in fact preferred.

Similarly, we do not need to know all of the details of the metal of the submarine. We would again need to know only bulk properties, such as its stiffness and thermal expansivity. We can also treat solids as continua.

There are also cases where we cannot use a continuum approximation. For example, if we are studying an aquaporin, we might want to analyze the electrostatic interactions as a single water molecule passes through. Clearly here we need to have a molecular/atomistic description of the system, at least of the water molecule itself.

So, when can we use a continuum description instead of a discrete one? We will have a more precise answer for this as we develop the theory in a moment, but for now, we'll just say that we need plenty of particles so that we can average their effects.

9.2 A preliminary: indicial notation

In order to more easily work our way through our treatment of continuum mechanics, we will introduce **indicial notation**, which is a convenient way to write down vectors, matrices and differential operators. This technique was invented by Albert Einstein.

Before plunging in, I note that we shouldn't trivialize or fear new notation. To quote Feynman, "We could, of course, use any notation we want; do not laugh at notations; invent them, they are powerful. In fact, mathematics is, to a large extent, invention of better notations."

The main concept behind indicial notation is a **tensor**. A tensor is a system of components organized by one or more indices that transform according to specific rules under a set of transformations. Tensors are parametrization-independent objects. The **rank** of a tensor is the number of indices it has. To help keep things straight in your mind, you can think of a tensor as a generalization of a scalar (rank 0 tensor), vector (rank 1 tensor) and a matrix (rank 2 tensor).[7] That's a mouthful, and quite abstract, so it's better to see how they behave with certain operations.

9.2.1 Contraction

The **contraction** of a tensor involves summing over like indices. For example, say we have two rank 1 tensors, a_i and b_i. Then, their contraction is

$$a_i b_i = a_1 b_1 + a_2 b_2 + \cdots . \tag{9.1}$$

It is convention in indicial notation to **always sum over like indices**. So, if a_i and b_i represent vectors in Cartesian three-space, which they usually will in our studies, they have components like (a_x, a_y, a_z). Then $a_i b_i = a_x b_x + a_y b_y + a_z b_z$ is the vector dot product. This relates to vector notation you might already be used to seeing.

$$a_i b_i = \mathbf{a} \cdot \mathbf{b}. \tag{9.2}$$

Note that in indicial notation, a_i represents the perhaps more familiar **a**. Note also that the tensor operations we are defining are parametrization independent, and they need not, in general, represent Cartesian coordinates.

We have just seen that contraction of two rank one tensors gives a rank zero tensor. Similarly, we can contract a rank two tensor with a rank one tensor, which is equivalent to a matrix-vector dot product.

$$A_{ij} b_j = c_i. \tag{9.3}$$

---

[7] We will not talk about covariant and contravariant tensors in this class, since they are not necessary for what we are studying.

Since we summed (or contracted) over the index j, the index i remains. It is helpful to write it out for the case of $i, j \in \{x, y, z\}$.

$$A_{ij} = \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix}, \tag{9.4}$$

and $b_j = (b_x, b_y, b_z)$. Then, we have

$$c_i = A_{ij} b_j = \begin{pmatrix} A_{xx}b_x + A_{xy}b_y + A_{xz}b_z \\ A_{yx}b_x + A_{yy}b_y + A_{yz}b_z \\ A_{zx}b_x + A_{zy}b_y + A_{zz}b_z \end{pmatrix}. \tag{9.5}$$

This is equivalent to $\mathsf{A} \cdot \mathbf{b}$ in notation you may be more accustomed to. Note that $A_{ij} b_i$ is equivalent to $\mathsf{A}^\mathsf{T} \cdot \mathbf{b}$. Stated explicitly, with indicial notation on the left hand side and vector notation on the right hand side,

$$\text{entry } i \text{ of } A_{ij} b_j = (\mathsf{A} \cdot \mathbf{b})_i, \tag{9.6}$$

$$\text{entry } j \text{ of } A_{ij} b_i = (\mathsf{A}^\mathsf{T} \cdot \mathbf{b})_j. \tag{9.7}$$

9.2.2 Direct product

We can also make higher order tensors from lower order ones. For example, $a_i b_j$ gives a second order tensor.

$$a_i b_j = \begin{pmatrix} a_x b_x & a_x b_y & a_x b_z \\ a_y b_x & a_y b_y & a_y b_z \\ a_z b_x & a_z b_y & a_z b_z \end{pmatrix}. \tag{9.8}$$

Comparing to vector notation, entry i, j of $a_i b_j = (\mathbf{a} \otimes \mathbf{b})_{ij}$.

9.2.3 Differential operations

You have probably seen the gradient operator before. In Cartesian coordinates, it is

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right). \tag{9.9}$$

In indicial notation, this is ∂_i. So, the gradient of a scalar function f is $\partial_i f$, which is a rank 1 tensor, as we would expect. In perhaps more familiar notation, we would write this as ∇f. The divergence of a vector v_i, commonly written as $\nabla \cdot \mathbf{v}$, is $\partial_i v_i$. This is a contraction of the differential operator with the vector. The Laplacian of a scalar, commonly written as $\nabla^2 f$ or Δf, is $\partial_i \partial_i f$.

9.2.4 Trace and matrix multiplication

We can define the trace of a rank 2 tensor as the sum of the diagonal elements.

$$A_{ii} = A_{xx} + A_{yy} + A_{zz}. \tag{9.10}$$

Note that we could multiply matrices and then take the trace. Comparing to familiar notation,

$$A_{ij}B_{ij} = \mathrm{tr}(\mathsf{A}^{\mathsf{T}} \cdot \mathsf{B}). \tag{9.11}$$

In other words, the contracted indices tell us what to sum. Simple matrix multiplication is

$$A_{ij}B_{jk} = (\mathsf{A} \cdot \mathsf{B})_{ik}. \tag{9.12}$$

9.2.5 The Levi-Civita symbol

We represent cross products with the **Levi-Civita symbol**. This is defined as

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if } ijk = xyz, yzx, zxy \\ -1 & \text{if } ijk = zyx, yxz, xzy \\ 0 & \text{otherwise.} \end{cases} \tag{9.13}$$

Thus, we can represent the vector cross product as

$$\text{entry } i \text{ in } \varepsilon_{ijk}u_jv_k = (\mathbf{u} \times \mathbf{v})_i. \tag{9.14}$$

The curl of a vector field is

$$\text{entry } k \text{ in } \varepsilon_{ijk}\partial_i v_j = (\nabla \times \mathbf{v})_k = (\mathrm{curl}\,\mathbf{v})_k. \tag{9.15}$$

9.3 Conservation of mass

Now that we have the mathematical notation in place, we will proceed to derive conservation laws for a continuous material. To do this, consider a piece of space within a material, which we will call a **volume element**. The volume element has an outward normal vector n_i, as shown in Fig. 12.

Now, let's say that this volume element has material in it with a density ρ. Then, the total mass of material inside the volume element is

$$m = \int \mathrm{d}V \rho, \tag{9.16}$$

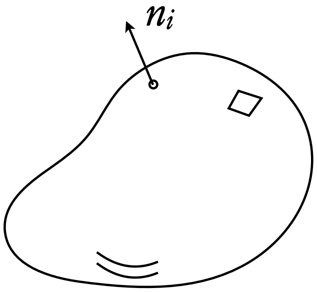

Figure 12: Drawing of a three-dimensional volume element with outward normal n_i.

where the integral is over the volume. Now, the time rate of change of mass in the volume must be equal to the net flow of mass into the control volume. The mass flow rate *out* of control volume per unit area is $n_i(\rho v_i)$, where v_i is the velocity of material. So, the rate of change of mass is

$$\partial_t \int \mathrm{d}V \rho = -\int \mathrm{d}S\, n_i(\rho v_i), \tag{9.17}$$

where the second integral is over the surface of the control volume.

We will now make use of the **divergence theorem**, also known as Gauss's theorem or the Gauss divergence theorem, which states that for any closed surface, any continuously differentiable tensor field F_i satisfies

$$\int \mathrm{d}V\, \partial_i F_i = \int \mathrm{d}S\, n_i F_i. \tag{9.18}$$

This generalizes for higher rank tensor fields. E.g.,

$$\int \mathrm{d}V\, \partial_j T_{ij} = \int \mathrm{d}S\, n_j T_{ij}, \tag{9.19}$$

for a rank 2 tensor. Taking our tensor field as ρv_i, we apply the divergence theorem to get

$$\partial_t \int \mathrm{d}V \rho = -\int \mathrm{d}V\, \partial_i(\rho v_i). \tag{9.20}$$

We can take the time derivative inside the integral sign and rearrange to get

$$\int \mathrm{d}V\, (\partial_t \rho + \partial_i(\rho v_i)) = 0. \tag{9.21}$$

This must be true for all arbitrary control volumes, which means that the integrand must be zero, or

$$\partial_t \rho + \partial_i(\rho v_i) = 0. \tag{9.22}$$

We define the operator

$$\frac{\mathrm{d}}{\mathrm{d}t} \equiv \partial_t + v_i \partial_i \tag{9.23}$$

as the **material derivative** (also known as the substantial derivative), which is the time derivative in the **co-moving frame**. The second term in its definition in effect puts the observer moving along with this control volume in the material. Using the definition of the material derivative and the chain rule, we have

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho\, \partial_i v_i. \tag{9.24}$$

If ρ does not change, i.e., if the material is **incompressible**, the result is that the velocity field is divergenceless (also called solenoidal), or

$$\partial_i v_i = 0. \tag{9.25}$$

This result is called the **continuity equation**.

9.4 Conservation of mass for each species

The same analysis applies for the conservation of mass for a given species k. We will write k as a superscript with the understanding that repeated superscript indices are *not* implicitly summed over (though we may explicitly sum over them if we like). We start with the analog of equation (9.17). We define ρ^k as the density of species k and v_i^k as the velocity of particles of type k, and we have

$$\int \mathrm{d}V\, \partial_t \rho^k = -\int \mathrm{d}S\, n_i(\rho^k v_i^k) + \text{net production of } k \text{ by chemical reaction}. \tag{9.26}$$

I have added the production of k by chemical reaction (in words) to this equation, since we need to consider this as well. We can write this using the stoichiometric coefficients for chemical reaction l, ν_l^k, and their respective rates, r_l.

$$\int \mathrm{d}V\, \partial_t \rho^k = -\int \mathrm{d}S\, n_i(\rho^k v_i^k) + \int \mathrm{d}V M^k \nu_l^k\, r_l, \tag{9.27}$$

where M^k is the molar mass of species k. The expressions for r_l are typically given by mass action expressions. Now, we can apply the divergence theorem and rearrange, giving

$$\partial_t \rho^k = -\partial_i(\rho^k v_i^k) + M^k \nu_l^k\, r_l. \tag{9.28}$$

To both sides of this equation, we add $\partial_i(\rho^k v_i)$. The result is

$$\partial_t \rho^k + \partial_i(\rho^k v_i) = \frac{\mathrm{d}\rho^k}{\mathrm{d}t} + \rho^k \partial_i v_i = -\partial_i(\rho^k(v_i^k - v_i)) + M^k \nu_l^k r_l. \tag{9.29}$$

We define the **diffusive mass flux** by $j_i^k = \rho^k(v_i^k - v_i)$. This is the relative movement of species k compared to the center of mass, or **barycentric** movement. So we have

$$\frac{\mathrm{d}\rho^k}{\mathrm{d}t} = -\rho^k \partial_i v_i - \partial_i j_i^k + M^k \nu_l^k r_l. \tag{9.30}$$

We can re-write this equation in terms of the number density (the concentration) of species k instead of the mass density. It is as simple as dividing the entire equation by the molar mass.

$$\frac{\mathrm{d}c^k}{\mathrm{d}t} = -c^k \partial_i v_i - \frac{1}{M^k}\partial_i j_i^k + \nu_l^k r_l. \tag{9.31}$$

It is common to also use the symbol j_l^k for the **diffusive particle flux**, which is the diffusive mass flux divided by the molar mass. This double notation can be confusing, and we will avoid using it here.

Deriving an expression for the diffusive particle flux is nontrivial[8], and we will not do it here. We will take as given **Fick's first law**, which states that

$$\frac{j_i^k}{M^k} = -D^k \partial_i c^k, \tag{9.32}$$

where D^k is the (strictly positive) diffusion coefficient of species k. Using this expression, we arrive at the **reaction-diffusion-advection** equation,

$$\partial_t c^k = -\partial_i(c^k v_i) + \partial_i(D^k \partial_i c^k) + \nu_l^k r_l. \tag{9.33}$$

The diffusion coefficient is often constant (though not always, especially in developmental systems when phosphorylation states can alter the effective diffusion coefficient), so we get

$$\partial_t c^k = -\partial_i(c^k v_i) + D^k \partial_i \partial_i c^k + \nu_l^k r_l. \tag{9.34}$$

The first term on the right hand side describes the change in concentration as a result of being embedded in a moving material (advection). The second term describes diffusion, and the last describes chemical reaction. These are the same equations that we encountered in studying Turing patterns, sans the advective term. We see now that the equation is derived simply by accounting for all of the mass in an arbitrary volume element.

---

[8] It involves invoking the Second Law of Thermodynamics and some symmetry arguments. Importantly, it requires little else, and arrives naturally without substantial assumptions.

9.5 Shoring up when we can use continua

From the above, we can see what the criteria are for using continuum mechanics. We have to be able to define volume elements large enough to contain enough particles such that each volume element has a well defined average and does not experience large fluctuations. The volume elements must be small enough that we can define spatial derivatives of these average quantities. So, we need to have a system big enough and full enough to contain many sufficiently big volume elements. This seems restrictive, but in practice, continuum mechanics has been very successful even in describing phenomena on very small length scales.

9.6 General conservation law

Instead of counting mass, let's count any other conserved quantity that is a property of the material; let's call it ξ. If j_i is the flux of ξ out of the volume element Then, we have

$$\partial_t \int \mathrm{d}V\, \xi = -\int \mathrm{d}S\, n_i j_i, \tag{9.35}$$

or, upon applying the divergence theorem and considering that the volume element is arbitrary,

$$\partial_t\, \xi = -\partial_i j_i. \tag{9.36}$$

This tells us that the local time rate of change of a conserved quantity is given by the divergence of a flux, an important general result.