MATH 435 SUPPLEMENT: MULTI-VARIABLE CHAIN RULE NOTATION

JOSEPH HELFER

Fix an open set $U \subset \mathbb{R}^2$.

Recall that *diffeomorphism* means by definition "smooth bijection with smooth inverse".

Now fix a second open set $\widetilde{U} \subset \mathbb{R}^2$ and a diffeomorphism $\Phi\colon \widetilde{U} \to U$.

Given a smooth function $\sigma\colon U \to \mathbb{R}^n$, we have its *reparametrization* $\tilde{\sigma}\colon \widetilde{U} \to \mathbb{R}^n$ with respect to Φ, defined by $\tilde{\sigma} = \sigma \circ \Phi$.

Write u, v for the coordinates on U and $\tilde{u}, \tilde{v}$ for the coordinates on $\widetilde{U}$, so that we write σ_u, σ_v and $\tilde{\sigma}_{\tilde{u}}, \sigma_{\tilde{v}}$ for the partial derivatives of σ and $\tilde{\sigma}$ respectively.

On page 79 of the textbook, in the proof of Proposition 4.2.7, we find the multivariable chain rule given as

$$\tilde{\sigma}_{\tilde{u}} = \frac{\partial u}{\partial \tilde{u}}\sigma_u + \frac{\partial v}{\partial \tilde{u}}\sigma_v \qquad \tilde{\sigma}_{\tilde{v}} = \frac{\partial u}{\partial \tilde{v}}\sigma_u + \frac{\partial v}{\partial \tilde{v}}\sigma_v. \tag{1}$$

Note that this does not quite make sense since $\tilde{\sigma}_{\tilde{u}}, \tilde{\sigma}_{\tilde{v}}$ are functions on $\widetilde{U}$, whereas σ_u, σ_v are functions on U. Also, what does it mean to differentiate the coordinates u, v on U with respect to the coordinates $\tilde{u}, \tilde{v}$ on $\widetilde{U}$?

The answer to the second question is given in the book, where it is written "let $(u, v) = \Phi(\tilde{u}, \tilde{v})$". That is, since we have the map $\Phi\colon \widetilde{U} \to U$, we have for each point $(\tilde{u}, \tilde{v}) \in \widetilde{U}$ a corresponding point $(u, v) = \Phi(\tilde{u}, \tilde{v})$, so we can consider u, v as functions on $\widetilde{U}$.

(Of course, we could also consider $\tilde{u}, \tilde{v}$ as functions on U, by setting $(\tilde{u}, \tilde{v}) = \Phi^{-1}(u, v)$. This is a general phenomenon: whenever we have a map between two subsets of $\mathbb{R}^n$, we can consider the coordinates on the codomain as being functions on the domain.)

This also gives an answer to the first question: σ_u and σ_v are each functions of u, v, but in turn we are considering u and v as functions of $\tilde{u}, \tilde{v}$, hence σ_u, σ_v become functions of $\tilde{u}, \tilde{v}$.

So we might write: $\sigma_u = \sigma_u(u, v) = \sigma_u(u(\tilde{u}, \tilde{v}), v(\tilde{u}, \tilde{v}))$.

In short, this is simply a convenient abuse of notation, whereby, having fixed a diffeomorphism $\Phi\colon \widetilde{U} \to U$ we use the same notation for a function on U and its reparametrization on $\widetilde{U}$. Hence, if we wrote out (1) more precisely, it would be

$$\tilde{\sigma}_{\tilde{u}} = \frac{\partial u}{\partial \tilde{u}}(\sigma_u \circ \Phi) + \frac{\partial u}{\partial \tilde{v}}(\sigma_v \circ \Phi) \qquad \tilde{\sigma}_{\tilde{v}} = \frac{\partial v}{\partial \tilde{u}}(\sigma_u \circ \Phi) + \frac{\partial v}{\partial \tilde{v}}(\sigma_v \circ \Phi).$$

This is a bit messier, so to keep things tidy, we can write (1) instead, at the price of having to be attentive and remember to mentally replace σ_u and σ_v by their reparametrizations.

There are many instances of this kind of abuse of notation in the textbook. A significant one appears on page 86, in the proof of Proposition 4.4.2, since a comment is made there explaining what's going on with the notation.

There, the context is that we have an open set $U \subset \mathbb{R}^2$, a curve $\Gamma\colon (\alpha, \beta) \to U$, a smooth function $\sigma\colon U \to \mathbb{R}^3$, and we are considering the composite $\gamma = \sigma \circ \Gamma\colon (\alpha, \beta) \to \mathbb{R}^3$. The chain rule is then

Date: 2023-01-20.

given as

$$\dot{\gamma} = \sigma_u \dot{u} + \sigma_v \dot{v}. \tag{2}$$

Here, both abuses of notation from before are appearing again.

First of all, writing $\dot{u}$ and $\dot{v}$ suggests that the coordinates u and v are functions of $t \in (\alpha, \beta)$. As before, what is happening is that the function $\Gamma\colon (\alpha, \beta) \to U$ associates to each $t \in (\alpha, \beta)$ a point $(u, v) = \gamma(t)$ in U, hence we can consider u and v to be functions of t.

Secondly, the left-hand side of the equation is a function on (α, β), whereas σ_u, σ_v are both functions on U. As before, what is happening is that σ_u, σ_v are both functions of u, v, but we are considering u, v as functions of t, so that σ_u, σ_v become functions of t.

So we might write: $\sigma_u = \sigma_u(u, v) = \sigma_u(u(t), v(t))$.

Hence, if we wrote out (2) more precisely, it would be

$$\dot{\gamma} = (\sigma_u \circ \Gamma)\dot{u} + (\sigma_v \circ \Gamma)\dot{v},$$

but we might prefer to write the more concise (2), at the cost of having to remember to interpret the notation correctly.