Buck Chapter 5 Meta Notes PhL 5.14.15

This meta doc at 11 pages is 1/3 the size of the 31 page raw notes as of 5/15/15.
The book Chapter 5 is 75 pages long and is brimming chock full of theorems (31) and examples.

The Chapter title "Differentiation" seems odd. Section 5.4 with all the chain rule stuff is definitely about Differentiation. But 5.1 and 5.2 seem to have little to do with it. Sec 5.3 regarding the differential of a function (one-row R matrix) does contain derivatives of things. But then the rest of the chapter seems not on title. Well, the elements of the R matrix are after all derivatives, and the Jacobian is det(R). So I guess that really is a central theme for this chapter.

5.1 Transformations 1
5.2 Linear Transformations [229] 1
5.3 Differential of a function [238] 3
5.4 Differentiation of Composite Functions [250] 4
5.5 Differentials of transformations [263] 6
5.6 Inverses of a function of one variable [270] 7
5.7 Inverses of transformations [274] 7
5.8 Implicit Function Theorems [283] 8
5.9 Functional Dependence [287] 10

In general, the topics of this Chapter are "unusual" and I think are not often "treated" in books. It definitely brushes up against tensor doc Chapters 1 and 2, in that both linear and general transformations are addressed, but the overlap is only partial.

5.1 Transformations [221]

Three minor theorems about transformations:

Continuity of f is defined as usual with | | being the E^m norm, though they fail to name that norm.

Theorem 1: [226] If $f : E^n \to E^m$ is continuous, then the **inverse image of an open set S is an open set**. If domain is restricted to D, then that inverse set is open relative to D. [seems a repeat of earlier]

Theorem 2: [226] If $f : E^n \to E^m$ is continuous, then D connected $\Rightarrow$ R is also connected, and same for any subset of D. That is to say S = **connected** $\Rightarrow$ f(S) = **connected**.

Theorem 3: [227]If $f : E^n \to E^m$ is continuous, then D **compact** $\Rightarrow$ f(D) = **compact**.

5.2 Linear Transformations **[229]**

Recall from tensor doc that this is a very limited class of transformation, I wrote a section on this.

Theorem 4: [229] (**Form of Linear Xform**) Just states the *form* for a general *linear* transformation f : $E^n \rightarrow E^m$. Coefficients are real. Transformation is manifestly continuous since all component functions are. The have $\mathbf{y} = \mathbf{y}(\mathbf{x}) = A\mathbf{x}$.

Theorem 5 [233] has n = 3 and shows how dim(range) = rank of the matrix A for r = 3,2,1,0. In each case, the range is a linear space of dimension r, a hyperplane through the origin.

Theorem 6 [233] (**Rank Theorem**) restates the above for general f: $E^n \rightarrow E^m$, general rank r. Here r = m is full rank and the range of linear f is all of E^m which one can think of as a plane of dimension m. For r < m, the range r is a plane of dimension r. These planes all contain the origin since 0→0 no matter what r is. [see notes above, so this theorem cannot have r = m unless n ≥ m. But OK.]

Comment: Bucks are getting the reader familiar with the idea that rank less than full rank results in a range that is less than the full range; reducing rank reduces dimensionality. Here this is all in the context only of *linear* transformations, and all the reduced dimension surfaces pass through the origin. Later this same idea will be applied to general transformations, and the matrix of interest is then R instead of F.

Theorem 7 [236] shows that there exists a constant B such that if Y = AX, then |Y| < B|X| for all X. (seems a bit lonely and out of place, but I guess used in some proof).

The main point of this section is how the *rank* of a *linear* transformation *matrix* (think R of tensor doc) affects the dimensionality of the range. My version of theorems 4,5,6 is this:

Case 1: n = m.

dim(image) = m - N	N = 0,1,2..m	N = dimension of nullspace = nullity
rank ≤ m	suspect: dim(image) = rank = m or less	

Case 2: n < m. If nullity N = 0, then seems to me range is surface of dimension n within E^m . So the max dimensionality of the image is that n of the domain:

dim(image) = n - N	N = 0,1,2..m	N = dimension of nullspace = nullity
rank ≤ n	suspect: dim(image) = rank = n or less	

Once you have N = n, everything maps into the origin so no point increasing N further.

Case 3: n > m. In this case you are already "projecting" from the get-go. You have a nullity of n-m right off the bat. In this case I suspect

dim(image) = m - N_{ex}	N_{extra} = 0,1,2..m	N = n-m+N_{extra}
rank ≤ m	suspect: dim(image) = rank = m or less	

The point is that as the nullspace grows in dimension, the range shrinks in dimension by the same amount. In tensor doc I only dealt with m = n and perhaps Stakgold only did that as well. I quote him in n dimensions:

$R_A = (N_{A^*})^{\perp}$ $\Leftrightarrow$ rank = r = dim(R_A) = n - dim(N_A)

As the nullspace of A* grows, the range of A shrinks.

5.3 Differential of a function **[238]**

This "differential" is the R matrix with one row since mapping into E^1.

Notation conversion:

Bucks	Tensor Doc
$\mathbf{D}_\beta(p_0)$	$\hat{\beta} \bullet \nabla f(\mathbf{p}_0) \equiv \nabla_\beta f$
$\mathbf{D}_\alpha\mathbf{D}_\beta(p_0)$	$\hat{\alpha} \bullet \nabla [\ \hat{\beta} \bullet \nabla f(\mathbf{p}_0)] \equiv \nabla_\alpha\nabla_\beta f$
Δp small but finite	$\Delta\mathbf{x}$ small but finite
dT	R (but might not be square in Buck)
df	R but with only one row
$\Delta f(\Delta p)$	$f(\mathbf{p} + \Delta\mathbf{p}) - f(\mathbf{p})$, $\Delta\mathbf{p}$ = finite
$df(\Delta p)$	$\nabla f \bullet \Delta\mathbf{p} = R\Delta\mathbf{p}$ when R is one row

What they call the differential dT is what I call the R matrix. Various example given.

Theorem 8: [243] **(Approximation Theorem).** In the following, d**r** is differential, Δ**r** is just small. Then

$\Delta f(\Delta\mathbf{r}) = f(\mathbf{r} + \Delta\mathbf{r}) - f(\mathbf{r}) = df(\Delta\mathbf{r}) + R(\Delta\mathbf{r}) = \nabla f \bullet \Delta\mathbf{r} + R(\Delta\mathbf{r})$

So normally one might refer to the last term as order$(\Delta r)^2$ but they do this in more detail.

Some mean value theorems"

$f(p+\Delta p) = f(p) + f'(p')\ \Delta p$	1D MVT	
$f(\mathbf{p}+\Delta\mathbf{p}) = f(\mathbf{p}) + f_1(\mathbf{p}')\ \Delta x + f_2(\mathbf{p}'')\ \Delta y$	**2D MVT**	// p 245 A

The theorems here all seem uninteresting, I just quote them all:

Theorem 9 [246] then says $(\mathbf{D}_\beta f) = df(\beta)$ by which they mean $(\mathbf{D}_\beta f) = \nabla_\beta f = \nabla f \bullet \boldsymbol{\beta} = df(\beta)$ so it is df going in the β direction a unit distance.

Theorem 10 [246] (**Local Maximum Theorem**) says that at a local interior **maximum**, $\partial_i f = 0$ for all variables. A proof is given. Seems pretty reasonable to me!

Theorem 11 [247] says that if df = 0 for all points in domain D, then **f = constant.** D must be open and connected. For disconnected you could have different constant in each piece. For not-open, funny things could happen at a boundary point.

Theorem 12 [247] says that **if $\partial_z \mathbf{f} = 0$ in a convex open domain D**, **then f is not a function of z** and is the same for all values of z. Nice picture on page 248 shows why convex is needed, case is f(x,y).

Theorem 13 [248] **Four Corners Theorem**: is about a double integral over a finite rectangle. It shows that the double integral of $\partial_{xy} f$ gives a simple combination of the values of f at the four corners. Sort of a generalization of the 1D perfect differential idea. But now f must be C" to get that $\partial_{xy} f$ to exist.

Corollary: From the above fact, it follows that $\partial_{xy}\mathbf{f} = \partial_{yx}\mathbf{f}$ on D, which must be an open set. [Later in these Ch 5 notes I prove this fact again because I forgot it was right here!]

Comment: I think they are just building a little stash of theorems they can call upon later. Not much really happens for me in this section. In the previous section 5.2 they talked only about *linear* transformations, but now I guess they are allowing for *general* transformations, and they the "differential" is the R matrix.

5.4 Differentiation of Composite Functions [250]

Now comes more interesting stuff. By "**composite function**" they really mean a variable dependency tree such as that shown page 251 top. Yes, this tree applies to a set of "composite" functions:

example: $w = f(u,v) = f(g(x,y),h(x))$
$u = g(x,y)$
$v = h(x)$

Any **variable tree** implies a certain set of **chain rules**. The non-tiered tree case page 252 brings up ambiguity of notation such that for example $\partial w/\partial z$ has multiple meanings, but using numbered derivatives removes all this ambiguity. I never realized that whole issue and resolution before.

A first derivative with chain rule is usually pretty simple such as

$w = F(x,y,t)$ $\qquad \partial_t w = F_1 \partial_t x + F_2 \partial_t y + F_3$

where you see the numbered derivative idea. Computing then $\partial_t^2 w$ is a bit of a mess and in this case there are 11 terms!

I have two proofs that $\partial_{ij} = \partial_{ji}$ in the notes.

The chain rule business only gets one theorem:

Theorem 14 [250] **(Chain Rule)** applies to Fig 5-8 only! This figure is like a neural network in clean layers with nothing jumping ahead 2 or more layers. Each path segment spans one layer only. Each variable is a function only of variables one layer to the right. The two chain rules are (5-11).

Now we come to the idea of **implicit dependence**, this is important to me. You are given a set of equations like this,

$F(x,y,u,v) = 0$
$G(x,y,u,v) = 0$

where x,y are independent variables and u and v are dependent. You don't have any known u = u(x,y), or for v, but these are implicit in these equations. Bucks show exactly how you can compute things like $\partial u/\partial x$ just from the two equations above. You get a **Cramer's rule thing**. For example

$$\partial u/\partial x = [F_4G_1 - F_1G_4] / [F_3G_4 - F_4G_3]$$

where again appear those "numbered" derivatives. This idea might apply to some situation where I have an initial equation and then a set of constraint equations, the Lagrange multiplier thing, etc etc. Coming later I think.

Bucks then give two examples (Ex 6 and Ex7, pages 256 and 257) where they change variables in a 1D wave equation and in a 2D Laplace equation. Both examples are interesting, and require the messy computing of second derivatives. I noted earlier how these tend to me messy. These examples however do NOT represent examples of the F=0 and G=0 idea discussed just above because these example all have derivatives in the equations of interest (wave and Laplace).

Example 8 [258] is extremely interesting to me. When I did **thermo** in Zemansky and elsewhere, I never understood how you switch around from say P and T being constant to V and T being constant, and this example really explains it all! When I next deal with thermo, I will come back here where the Bucks have put things in a completely general framework! Ah, so many years ago! Thermo has a new relevance I think in a green world.

Now a new theorem:

Theorem 15 [259] . (**Mean Value Theorem in N dimensions**)

The claim is this: $f(\mathbf{p}_2) - f(\mathbf{p}_1) = \nabla f(\mathbf{p}^*) \bullet (\mathbf{p}_2 - \mathbf{p}_1)$ $\quad f: E^n \rightarrow E^1$

where $\mathbf{p}^*$ lies somewhere on the line segment joining $\mathbf{p}_1$ and $\mathbf{p}_2$.

First, compare this to the 1D "MVT of Diff Calc" from Section 2.7

$$y(b) - y(a) = y'(c)(b-a)$$

so you see the obvious way it is generalized to N dimensions. You see this is like the remainder term in a Taylor series expansion in N dimensions where you don't have any terms other than this remainder. I don't have a graphical interpretation as I do for the 1D MVT. The proof is based on the 1D MVT using a clever function $F(t) = f(\mathbf{p}_1 + t\mathbf{\Delta})$ with $\mathbf{p}_2 - \mathbf{p}_1 = \mathbf{\Delta}$

The tour de force conclusion of this Section is the **Taylor Expansion** in N dimensions, though they only do it in 2 dimensions. They provide a Remainder Term which has something evaluated at an unknown point in some range. Here is the expansion:

$$\underset{\text{first}}{f(\mathbf{r}+\Delta\mathbf{r})} = \underset{\text{first}}{f(\mathbf{r})} + \underset{\text{correction terms}}{[\,\Sigma_{m=1}^{n}\,(1/m!)\;(\Delta\mathbf{r}\bullet\nabla)^m f(\mathbf{r})]} + \underset{\text{remainder term}}{(\Delta\mathbf{r}\bullet\nabla)^{n+1} f(\mathbf{r}^*)\;/\,(n+1)!}$$

where **r*** lies between **r** and r+Δ**r**. In 1D this would read

$$f(x+\Delta x) = f(x) + [\,\Sigma_{m=1}^{n}\,(1/m!)\;(\Delta x)^m\,\partial_x^m f(x)] + (\Delta x)^{n+1}\,\partial_x^{n+1} f(x^*)\;/\,(n+1)!$$

and this agrees with the page 127 result Bucks derived earlier. These results are exact even when the distance Δ**r** is not small, that is a new fly in the ointment for me. But if Δr is large, then there is much uncertainty in the location of the point **r*** which lies between **r** and **r** + Δ**r**. Bucks comment on rate of convergence back on page 127, but it really depends entirely on the function f. The MVT is critical in deriving the above expansion(s). This is the best presentation I have on this subject. I comment on my own private derivation in the 2D Taylor theorem. Their theorem does zig-zag but mine does columns first and then rows, and mine has no remainder term! Theirs is much better.

5.5 Differentials of transformations **[263]**

Bucks are now doing their version of my tensor doc. Here are the connections

Bucks	Me in tensor doc
$(dT)_{ij}$	R_{ij}
dT = matrix = "differential"	R = matrix = "Jacobi matrix"
x_i	x_i
y_i	x'_i
T(**p**) // but Bucks have no bolding	**F**(**x**)
	$F_i(\mathbf{x}+d\mathbf{x}) \approx F_i(\mathbf{x}) + \Sigma_k(\partial F_i(\mathbf{x})/\partial x_k)\,dx_k$
	$= F_i(\mathbf{x}) + \Sigma_k R_{ik}\,dx_k$

Buck Theorem 17 [264]:

$\mathbf{T}(\mathbf{p}_0+\Delta\mathbf{p}) = \mathbf{T}(p_0) + dT\,(\Delta\mathbf{p}) + \mathbf{R}(\Delta p)$	
$\mathbf{T}(\mathbf{p}_0+d\mathbf{p}) = \mathbf{T}(p_0) + dT\,(d\mathbf{p})$	$\mathbf{F}(\mathbf{x}+d\mathbf{x}) = \mathbf{F}(\mathbf{x}) + R\,d\mathbf{x}$ // *linearization idea*

So Buck Theorem 17 just says the above line which is just the definition of the linearization. All fine.
[Makes me want to add EQ nums to tensor doc] [it is done!! took 41 days...]

Theorem 18 [265] just says you can concatenate two R matrices in the obvious way. I talk about this in tensor doc in Section 8.10 where I write $R = R_1R_2$ and $J = J_1J_2$.

Theorem 19: [268] (**Mean Value Theorem for Transformation**). In tensor doc notation:

$$\mathbf{F}(\mathbf{x}+\Delta\mathbf{x}) = \mathbf{F}(\mathbf{x}) + R(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*)\Delta\mathbf{x}$$

What this means is that for finite $\Delta\mathbf{x}$ the above is valid, *but* in the R matrix you evaluate each row of elements at a different unknown point. All these points lie in the range $\mathbf{x}$ to $\mathbf{x}+\Delta\mathbf{x}$. Obviously as $\Delta\mathbf{x}\rightarrow 0$, the three points all converge at $\mathbf{x}$. I think this theorem works for $F : E^n \rightarrow E^m$ where there are m points $\mathbf{x}_i^*$. Back in Theorem 15 p 259 we had this same theorem but for $f : E^n \rightarrow E^1$ so there was only one point there called $\mathbf{p}^*$. And of course we know the MVT for $f : E^1 \rightarrow E^1$.

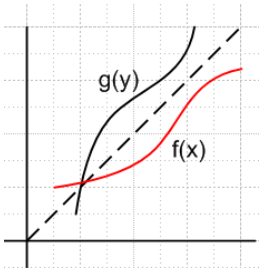
5.6 Inverses of a function of one variable **[270]**

Something a little different here. Page 270 shows plot of $y = x^2$ in solid line. To get the inverse of any function, they show that you must **reflect in the y=x axis, you do not rotate 90 degrees** as I have done. The dotted line here then shows $y = +\sqrt{x}$ and $y = -\sqrt{x}$ as two candidate inverse functions. Wherever you have slope = 0 in the original function, you will have infinite slope in the inverse, and that will be a point where two of the branch solutions meet! Other nice examples: one on page 271 shows that rotation is not the same as reflection, and shows three branch solutions (since a cubic). Then p 272 shows reflection of the sine function and its infinite number of branches.

If there are no locations of zero slope, then this method gives a viable inverse:

Theorem 20 [272] If you have f(x) on a domain [a,b] where f(x) on which $f'(x) \neq 0$, and if range is $[\alpha,\beta]$, and if f(x) is C^1, then the inverse is unique *and* C^1 and can be found from the reflection method and you have then defined here f: D→R with f^{-1}: R→D which is one-to-one.

Here is my picture where the red f(x) has no points of f'(x) = 0. There is a unique inverse, and it is C^1.



5.7 Inverses of transformations **[274]**

So they have talked about inverses of a function, but how do you invert a transformation $E^n \rightarrow E^n$? They show that if $J \neq 0$, then the transformation is "locally invertible", an interesting hedge. This then avoids the issue of multiple sheets and all that stuff -- avoids it for the moment. There are several related theorems. They always require that the transformation of interest be C^1 meaning continuous and differentiable, just as I require in tensor doc.

Theorem 21. [276] (**locally one-to-one theorem**) Let $T:E^n \rightarrow E^n$ and functions of T are C^1 on open set D. Then if $J \neq 0$ over D, the mapping is *locally one to one*.

Theorem 22. [277] If *continuous* T is *one-to-one* on *closed bounded domain* S, then the inverse transformation T^{-1} is also continuous, and of course maps set T(S) back into S.

Theorem 23 [278] If T is class C^1 on D *and* if $J \neq 0$ on all of D *and* if D→T(D) is one-to-one, *then*

(1) T^{-1} is also C^1

(2) $d(T^{-1}) = (dT)^{-1}$ // $S = R^{-1}$ in my language.

Theorem 24 [279] If $T:E^n \to E^n$ is C^1 on open set D and if $J \neq 0$ on D, then T(D) is also an open set.

This is a new twist. We always knew that f cont ⇒ back map takes open to open. But now we know that the back map f^{-1} is continuous as well, so we just apply our usual fact in reverse to conclude that f maps opens to opens!

Theorem 25 [281] (an existence theorem). If $T:E^n \to E^n$ is C^1 on open set D and if $J \neq 0$ on D, *then* :

For any p_0 in D you can find a ball N such that N → T(N) is one-to-one and T(N) is an open set, and you can find a workable T^{-1} inverse transformation, and $(dT^{-1})(dT) = 1$, for me this is SR = 1.

Final Claim [283] If T:D→D* is **locally one-to-one**, there are certain conditions which would make it be **globally one-to-one**. Two of those conditions are that D is compact and that D* is **simply connected**. But this is a Chapter 7 concept, so we have to wait. Vallee Poussin did whatever theorem this is in 1923 (Belgian).

Comment: For many things, the Bucks provide the math rigor that I explicitly omit in tensor doc. It is nice to have a source where I can track down a real proof of something if I need it.

5.8 Implicit Function Theorems **[283]**

A very interesting section. I will state the most general version of the theorem here. You are basically given a set of I+J variables of which I are independent called x_i and of which J are dependent called u_j. You are also given a set of J equations of the form F(all variables) = 0. Your task is to solve these J equations for the J variables u_i. The J equations need not be linear!!! You can use any transformation you want. It's just that the solution can only exist when Jacobian ≠ 0. The theorem does not tell you what the solution is, however.

What is not clear in the Buck presentation is the smallness of things. I think the idea is this: if you zoom in on $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ near some point $\mathbf{x}_0$, you have a linearized world $d\mathbf{x}' = Rd\mathbf{x}$ which you know you can invert to get $d\mathbf{x} = Sd\mathbf{x}'$. Then you can globalize that linear solution to a ball around your point of interest. But the Bucks talk about x,y,z and u,w and the t_i as if they were not small.

So maybe they are not thinking linearized forms as all. Maybe they are just saying $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ and if the Jacobian at some point $\mathbf{x}_0$ does not vanish, then you have $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ near that point and nothing need be linear or small, it just has to be local enough to stay out of trouble. In their $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ they have $\mathbf{x}' = \{t_i\}$ which is fine. and then some of the $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is linear and trivial, but the last J dimensions are non-linear in general which is where this F_j come into play, and their inverses f_j.

In any event, you have to compute R at some point from **F** so you can get J ≠0.

Here then is their stuff in most general form.

To generalize, we could have I independent variables x_i, J dependent variables u_j, and J equations of the form $F_j = 0$. Write

$$F_j(x_1, x_2.....x_I, u_1, u_2....u_J) = 0 \qquad j = 1,2...J$$

We want to solve for these "implicit functions",

$$u_j = \varphi_j(x_1, x_2.....x_I) \qquad j = 1,2...J$$

We treat this as a (J+I)x(J+I) problem and find that Jacobian J = Jacobian(F_i, u_j) which is JxJ. If we avoid points where J = 0, then we first have this initial problem

$t_i = x_i$	i = 1,2..I	
$t_{I+j} = F_j(x_i,u_j)$	j = 1.2..J	the transform
$x_i = t_i$	i = 1,2...I	
$u_j = f_j(t_1..t_I, t_{I+1}..t_{I+J})$	j = 1,2...J	the inverse transform

We then set the higher t's to zero to make the J equations of the form $F_j = 0$ be valid, We then get

$$u_j = f_j(t_1..t_I, 0..0) = f_j(x_1..x_I, 0..0)$$

and there are the solution "implicit functions".

In the local sense, one can regard our transformation as linear if J ≠ 0 and then we have in that local region a set of J equations $F_j = 0$ which we want to solve for the J unknowns u_i in terms of the x_i.

So, there is no restriction that the F_j functions be linear. It's just that we know near a local solution, they will be linear and we know there is always an inverse. The method here does not GIVE you the inverse, it just says it exists. You have to manually solve the problem itself to get a specific solution.

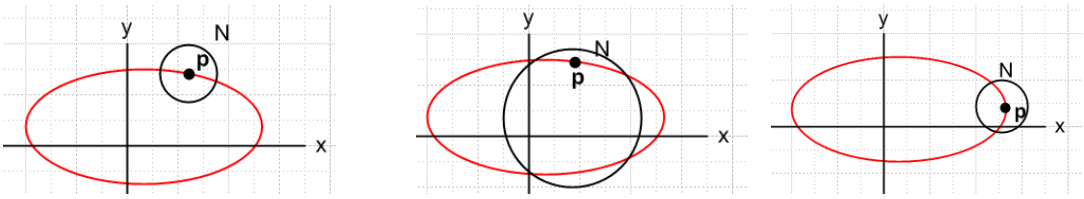
The text does two examples of the above theorem. I will fit them in right here

Example 1 page 284. I = 2 and indep variables are x_1= x and x_2 = y. J = 1 and $F_1(x,y,u_1) = 0$. The solution is then $u_1(x,y) = f_1(x,y,0)$ which is written as $\varphi(x,y) = f(x,y,0)$. They use $z = u_1$.

Example 2 page 285. I = 3 and indeps are x_1= x and x_2 = y and x_3= z. We have J = 2 and the F_1 = F and the F_2 = G. The dependent variables are u_1 = u and u_2 = v. Tiny Example 3 is an example of Example 2.

In the raw notes I show how tensor doc would do this problem. I think the Bucks have put the "point of interest" at $\mathbf{x}_0 = 0$ and perhaps also $\mathbf{x}'_0 = 0$. But OK, let's leave it at that.

The **locality-of-inverse idea** is stated in these pictures which you can interpret in 2D or 3D.



The point p shown on the right is where in 2D you don't have an inverse no matter how much you shrink down the neighborhood ball, and in general this because a point where Jacobian = 0. In the middle you have two inverses. The pictures really talk about a different thing. The red curve is where some F(x,y,z) = 0 and the local piece of the red curve is where you have in implicit function solution z = φ(x,y).

5.9 Functional Dependence **[287]**

Another interesting section.

(a) The first item is a generalization of the notion of **linear dependence** of a set of functions to the **functional dependence** of a set of functions. Suppose for example that u,v,w are each functions of x,y,z. In general if x,y,z wander around in some 3D region in the domain, u,v,w wander around some 3D region in the range. But, if it happened that F(u,v,w) = 0 were true, then somehow the three functions u,v,w are not independent -- they are functionally dependent! In this case, if F is a smooth function, then there is a local implicit function like w = f(u,v) and then that would define a piece of surface in UVW space. So functional dependence is associated with a **reduction in dimension** of a space! In a sense, E^3 got reduced to an E^2 surface in this example. General functional dependence is shown in p 288 A.

(b) Resume this x,y,z→u,v,w transformation example. We know independently that if the R matrix does not have full rank, we also get a reduction in dimension. If rank = 3, no reduction and 3D region in domain maps to 3D region in range. If rank = 2, 3D range maps into a curved 2D surface in the range, and if rank = 1, then 3D range maps into a 1D curve in the range. This is the claim of **Theorem 28 and 29** stated on p 289. These theorems are **proved** in long proofs on pages 289-291.

Note that if there is any reduction below full rank, they you have Jacobian J = 0. So whereas the previous Section 5.8 was concerned with the situation when J ≠ 0, this entire Section 5.9 is concerned with J = 0 and resulting reduction of dimensionality.

(c) It is in the proofs of Thm 28 and 29 that we see the *connection* between dimension reduction due to functional dependence and dimension reduction due to non-maximal rank! The proofs show that if you assume a non-maximal rank, then you end up with functional dependence.

In the $E^3 \rightarrow E^3$ context we have been talking about here, rank 2 means 2D surface in range and that is described by something like w = f(u,v) or two other cyclic ways, such as u = g(w,v). Any of these three relationships defining surfaces are of the form F(u,v,w) = 0 and thus you are stating functional dependence of these three functions! There is for rank 2 only a single relation of dependence.

For rank 1, you instead get something like w = f(u) and v = g(u) or some variation which means for a given u there is only a single w and v, hence you have a curve. So here there are two relations of functional dependence. F(u,w) = 0 and G(u,v) = 0 and you could add the third arg to each to be general.

It is stressed that the particular form of the functional dependence might not work throughout the entire domain and its traced out range.

[I suppose Stakgold must have dealt with this stuff in nullspace language. As rank goes down, size of nullspace goes up, dimensionality of range goes down. But I think Stak was only $E^n \rightarrow E^m$ but not sure. The Bucks have not mentioned nullspace even once! Not in the index!]

Theorems 30 and 31 restate the ideas of Theorems 28 and 29 in $E^3 \rightarrow E^2$ and $E^2 \rightarrow E^3$ contexts. In either of these cases, max rank is 2, and if that is what you have, the range is a 2D thing. But if rank = 1, then range drops to a 1D thing (curve).

For me and my **tensor doc** work, this whole topic **comes to life** when I look at my mapping pictures for polar and spherical coordinates. The subject here are those "special points that have to be fixed up". For each special situation, I have shown have you have rank reduction and corresponding dimensionality reduction of the range. For example, in sphericals J = 0 when r = 0, so the entire main floor of the office building maps into a point, there is a reduction of 2, and so rank = 1.