LECTURE 20

Date of Lecture: April 6, 2020

As always, $\mathbf{K} \in \{\mathbf{R}, \mathbf{C}\}$.

The symbol ☡ is for flagging a cautionary comment or a tricky argument. It occurs in the margins and is Knuth's version of Bourbaki's "dangerous bend symbol".

An n-tuple $(x_1, \dots, x_n)$ of symbols (x_i not necessarily real or complex numbers) will also be written as a column vector when convenient. Thus

$$(x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A map $\boldsymbol{f}$ from a set S to a product set $T_1 \times \cdots \times T_n$ will often be written as an n-tuple $\boldsymbol{f} = (f_1, \dots, f_n)$, with f_i a map from S to T_i, and hence, by the above convention, as a column vector

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

(See Remark 2.2.2 of Lecture 5.)

The default norm on Euclidean spaces of the form $\mathbf{R}^n$ is the Euclidean norm $\|\ \|_2$ and we will simply denote it as $\|\ \|$.

☡ Note that $(x_1, \dots, x_n) \neq [x_1 \dots x_n]$. Each side is the transpose of the other.

1. Maximà and minima continued

Sometimes we need to use both global methods (involving the positive or negative definiteness of the Hessian) as well as Lagrange's multipliers.

1.1. **Example.** Let $K = \{(x, y) \in \mathbf{R}^2 \mid 2x^2 + y^2 \leq 8\}$ and $f\colon \mathbf{R}^2 \to \mathbf{R}$ the map $f(x, y) = x^2 + y^2$. Since K is compact and f continuous f attains a maximum and a minimum on K. Let us find the points. We divide K into two disjoint sets, namely, (a) the open set $V = \{(x, y) \in K \mid 2x^2 + y^2 < 8\}$ and (b) the ellipse $E = \{(x, y) \in K \mid 2x^2 + y^2 = 8\}$. On V we find all the critical points using global methods (i.e. the methods involving the Hessian) and on E we use Lagrange multipliers to find the critical points of $f|_E$.

Now $\boldsymbol{\nabla} f(x, y) = (2x, 2y)$. So the only critical point of $f|_V$ is $\mathbf{0}$. As for the critical points of $f|_E$, if λ is a Lagrange multiplier for the problem of finding these critical points, then at such a point (x, y) of $f|_E$ we have

$$(*) \qquad \begin{cases} 2\lambda x = x \\ \lambda y = y \\ 2x^2 + y^2 = 8 \end{cases}$$

Since $\mathbf{0} \notin E$, one of x or y is non zero, whence either $\lambda = \frac{1}{2}$ or $\lambda = 1$. In the former case $y = 0$ and $x = \pm 2$. In the latter case $x = 0$ and $y = \pm 2\sqrt{2}$. Thus the critical points of $f|_E$ are $(-2, 0)$, $(2, 0)$, $(0, 2\sqrt{2})$, and $(0, -2\sqrt{2})$. The maximum and minimum of $f|_K$ therefore occur amongst the five points $\mathbf{0}$, $(-2, 0)$, $(2, 0)$, $(0, 2\sqrt{2})$, and $(0, -2\sqrt{2})$. We have

$$f(0,0) = 0$$
$$f(\pm 2, 0) = 4$$
$$f(0, \pm 2\sqrt{2}) = 8.$$

It follows that $\mathbf{0}$ is a point of minimum for $f|_K$ and $(0, \pm 2\sqrt{2})$ are two points of maximum for $f|_K$. The minimum value is 0 and the maximum is 8.

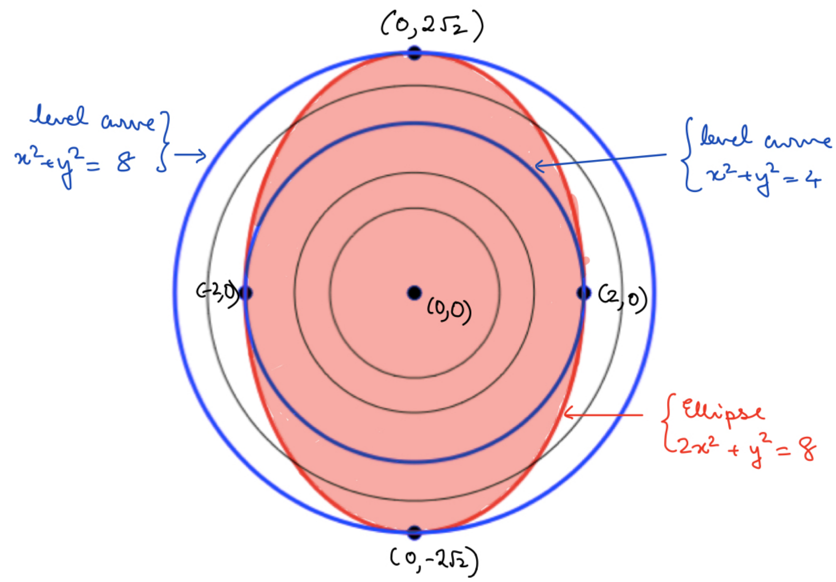

FIGURE 1. The red region together with the bounding ellipse is K. The concentric circles are the level curves of f. The blue level curves are the ones at which the constrained extrema occur.

2. **Complex differentiation**

The complex plane $\mathbf{C}$ is, of course, the same as $\mathbf{R}^2$, at least as a metric space (with the usual Euclidean norm on $\mathbf{R}^2$). However, $\mathbf{C}$ has a richer richer than $\mathbf{R}^2$, the same way that a group has a richer structure than its underlying set. You will learn about the very important effect this richer structure has on the theory of analytic/holomorphic functions on $\mathbf{C}$ in your complex analysis course. Here we will content ourselves with giving a few (very) preliminary results on the structure of such functions.

By an open set in $\mathbf{C}$, we mean an open set in the underlying metric space $\mathbf{R}^2$.

2.1. If $f\colon U \to \mathbf{C}$ is a function on an open set U in $\mathbf{C}$, it makes sense to take about f being differentiable as a map from an open subset of $\mathbf{R}^2$ to $\mathbf{R}^2$. If f is differentiable, its derivative is matrix valued function on U, as we saw in much of the course. But there is another notion of differentiability of f which takes into account the fact that $\mathbf{C}$ is a field and one can divide elements of $\mathbf{C}$ by non zero elements of $\mathbf{C}$ and the two notions are *not* equivalent.

Definition 2.1.1. Let U be an open subset of $\mathbf{C}$ and z_0 a point on U. A function $f\colon U \to \mathbf{C}$ is said to be *(complex) differentiable at* z_0 if the limit

$$\lim_{h\to 0} \frac{f(z_0+h)-f(z_0)}{h}$$

exists. In this case we write $f'(z_0)$ or $\frac{\mathrm{d}f(z)}{\mathrm{d}z}|_{z=z_0}$ for this limit and call it the *derivative of f at a*. We say f is *holomorphic* or *analytic* if it is complex differentiable at all points of U. In this case the map $z \mapsto f'(z)$ on U is called the *derivative of* f and denote it f' or as $\frac{\mathrm{d}f}{\mathrm{d}z}$.

2.1.2. Notations. Let f be a complex valued function on an open subset U of $\mathbf{R}^2$, say $f = u + iv$ where u is the real part of f and v the imaginary part of f. The following notations for f will be used interchangeably.

$$f = u + iv = (u, v) = \begin{bmatrix} u \\ v \end{bmatrix}.$$

Also note that

$$|\ | = \|\ \|$$

on $\mathbf{C} = \mathbf{R}^2$. We use $|\ |$ when we think of elements as complex numbers, and $\|\ \|$ when we think of the same elements as vectors in $\mathbf{R}^2$. The symbols $f_x(a,b)$, $f_y(a,b)$ will denote $(D_1f)(a,b)$ and $(D_2)f(a,b)$ provided these partial derivatives exist. If D_1f exists on all of U, we will often denote it as f_x. Similarly if D_2f exists on all of U, it will be often denoted f_y. Finally, to avoid confusion, when dealing with a complex valued function f on an open subset of $\mathbf{C}$, we prefer writing Df for the real derivative of f, rather then f'. We reserve the latter for the complex derivative of f.

Suppose $f\colon U \to \mathbf{C}$ is differentiable at $z_0 = a + ib$, where U is an open neighbourhood of z_0 in $\mathbf{C}$. Write $f = u + iv$ where $u\colon U \to \mathbf{R}$ and $v\colon U \to \mathbf{R}$ are the real and imaginary parts of f.

$$\begin{aligned} f'(z_0) &= \lim_{t\to 0} \frac{f(a+t+ib)-f(a+ib)}{t} \\ &= \lim_{t\to 0} \frac{u(a+t,b)-u(a,b)}{t} + i\lim_{t\to 0} \frac{v(a+t,b)-v(a,b)}{t} \\ &= u_x(a,b) + iv_x(a,b) \\ &= f_x(z_0). \end{aligned}$$

In other words u_x, v_x, and f_x exist at (a, b) and the above equality holds. Similarly

$$\begin{aligned} f'(z_0) &= \lim_{t\to 0} \frac{f(a+ib+it)-f(a+ib)}{it} \\ &= \lim_{t\to 0} \frac{u(a,b+t)-u(a,b)}{it} + i \lim_{t\to 0} \frac{v(a,b+t)-v(a,b)}{it} \\ &= \frac{u_y(a,b)}{i} + i\frac{v_y(a,b)}{i} \\ &= v_y(a,b) - iu_y(a,b) \\ &= -if_y(z_0). \end{aligned}$$

Once again, this means u_y, v_y, and f_y exist at (a, b) and the above equality holds. We therefore have

$$u_x(z_0) = v_y(z_0) \quad \text{and} \quad u_y(z_0) = -v_x(z_0) \tag{2.1.3}$$

Equivalently, in a more compact form

$$f'(z_0) = f_x(z_0) = -if_y(z_0). \tag{2.1.4}$$

In fact we have a seemingly stronger result.

Proposition 2.1.5. *Let U be an open subset of $\mathbf{C}$, z_0 a point in U, and $f : U \to \mathbf{C}$ a map which is complex differentiable at z_0. Then f is real differentiable at z_0 and*

$$(Jf)(z_0) = \begin{bmatrix} u_x(z_0) & u_y(z_0) \\ -u_y(z_0) & u_x(z_0) \end{bmatrix} = \begin{bmatrix} v_y(z_0) & -v_x(z_0) \\ v_x(z_0) & v_y(z_0) \end{bmatrix}.$$

Proof. If f is real differentiable then clearly, from (2.1.3), the expressions given for $(Jf)(z_0)$ hold. So it remains to show that $Df(z_0)$ exists.

Let $h = h_1 + ih_2$. From (2.1.3) (or (2.1.4)) we get

$$\begin{aligned} hf'(z_0) &= (h_1 + ih_2)(u_x(z_0) - iu_y(z_0)) \\ &= h_1u_x(z_0) + h_2u_y(z_0) + i(-h_1u_y(z_0) + h_2u_x(z_0)) \\ &= h_1u_x(z_0) + h_2u_y(z_0) + i(h_1v_x(z_0) + h_2v_y(z_0)) \\ &= h_1f_x(z_0) + h_2f_y(z_0). \end{aligned} \tag{$*$}$$

Let $J = \begin{bmatrix} u_x(z_0) & u_y(z_0) \\ v_x(z_0) & v_y(z_0) \end{bmatrix}$. To show f is real differentiable, we have to show that

$$\lim_{h\to 0} \frac{1}{\|(h_1,h_2)\|}\left(f(z_0+h) - f(z_0) - J\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}\right) = 0. \tag{$**$}$$

Now

$$\begin{aligned} J\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} &= (h_1u_x(z_0) + h_2u_y(z_0), h_1v_x(z_0) + h_2v_y(z_0)) \\ &= h_1(u_x,(z_0), v_x(z_0)) + h_2(u_y(z_0), v_y(z_0)) \\ &= h_1f_x(z_0) + h_2f_y(z_0) \\ &= hf'(z_0). \end{aligned}$$

The last equality follows from $(*)$. We therefore have

$$\frac{1}{\|(h_1,h_2)\|}\left\|f(z_0+h) - f(z_0) - J\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}\right\| = \left|\frac{f(z_0+h) - f(z) - hf'(z_0)}{h}\right|$$

Since f is complex differentiable at z_0, the right side approaches 0 as $h \to 0$, and hence $(**)$ holds. □

The equations (2.1.3) and its equivalent form (2.1.4) are called the *Cauchy-Riemann* equations. The term is more commonly used when f is holomorphic on U. The following important theorem is an immediate corollary.

Theorem 2.1.6. *Let U be an open subset of $\mathbf{C}$ and $f\colon U \to \mathbf{C}$ a holomorphic function. Let u and v be the real and imaginary parts respectively of f. Then u_x, u_y, v_x, and v_y exist on U and satisfy the* Cauchy-Riemann equations

$$u_x = v_y \quad \text{and} \quad u_y = -v_x.$$

The above can expressed in an equivalent formulation as

$$f' = f_x = -if_y.$$

2.1.7. Important facts about holomorphic functions. We state the following important (and very surprising) facts about holomorphic functions. You will learn the proofs in a course in complex analysis. In what follows U is an open set in $\mathbf{C}$ and $f\colon U \to \mathbf{C}$ a map.

1. If f is holomorphic, u_x and u_y are continuous where u and v are the real and imaginary parts of f. This is a famous result of Édouard Goursat. It is often stated in a slightly different form.

2. If f is holomorphic, then f' is also holomorphic on U. This result, in the generality we have stated it, is also due to Goursat. It is one of the most important theorems in complex analysis. In particular, thought of as a map from U to $\mathbf{R}^2$, f is $\mathscr{C}^\infty$.

3. If f is holomorphic then at each point of $a \in U$, there is a neighbourhood of V of a in U such that f can be written as

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(z-a)^n \qquad (z \in V).$$

This is the main reason a holomorphic function is also called an *analytic* function. Traditionally, an analytic function on an open subset of $\mathbf{C}$ is one which has local power series expansions as above. Such functions can be shown to be holomorphic. In other words, holomorphic functions on U are the same as analytic functions on U.

4. If f is holomorphic on U and non-constant, then f is an *open map*.

5. If f is $\mathscr{C}^1$ and satisfies the Cauchy-Riemann conditions (2.1.3), then it is holomorphic on U. This essentially is Green's Theorem on open sets in $\mathbf{R}^2$ which you will learn about next semester. Green's theorem is part of a larger set of ideas in differential topology and differential geometry which come under the catchall phrase "Stokes' Theorem". You will learn versions of that too next semester.

We point out that **1.** follows from **2.** And it is not hard to show that **2.** follows from **3.**

Conformal mappings. In this subsection we will be slightly informal. You will learn more about these things in later courses.

Recall that the angle between two non-zero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbf{R}^2$ is

$$\theta(\boldsymbol{u},\boldsymbol{v}) = \arccos\left\{\frac{\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right\}.$$

Recall that the range of arccos is $[0,\pi]$. Clearly $\theta(\boldsymbol{u},\boldsymbol{v}) = \theta(\boldsymbol{v},\boldsymbol{u})$, i.e. the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is the angle between $\boldsymbol{v}$ and $\boldsymbol{u}$. Suppose we rotate $\boldsymbol{u}$ in anticlockwise direction (keeping its tail at $\mathbf{0}$) until it is aligned in the same direction as $\boldsymbol{v}$. Then we have covered an angle $\varphi(\boldsymbol{u},\boldsymbol{v})$ equal to either $\theta(\boldsymbol{u},\boldsymbol{v})$ or $2\pi - \theta(\boldsymbol{u},\boldsymbol{v})$. In this case we say $\varphi(\boldsymbol{u},\boldsymbol{v})$ is the *directed angle* from $\boldsymbol{u}$ to $\boldsymbol{v}$. The angle $\varphi(\boldsymbol{u},\boldsymbol{v})$ is not symmetric in $\boldsymbol{u}$ and $\boldsymbol{v}$.

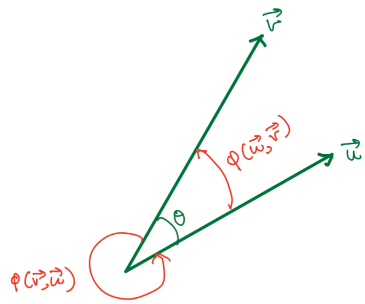

FIGURE 2. The angle $\theta(\boldsymbol{u},\boldsymbol{v})$ between $\boldsymbol{u}$ and $\boldsymbol{v}$, as well as the directed angles $\varphi(\boldsymbol{u},\boldsymbol{v})$ and $\varphi(\boldsymbol{v},\boldsymbol{u})$.

Let $A\colon \mathbf{R}^2 \to \mathbf{R}^2$ be a linear map. It is said to *preserve angles* if it is non singular, and for any pair of non-zero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbf{R}^2$ we have $\theta(\boldsymbol{u},\boldsymbol{v}) = \theta(A\boldsymbol{u},A\boldsymbol{v})$. It is said to *preserve orientation* if it is non-singular, and if for any pair of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ neither of which is a scalar multiples of the other, $\det[\boldsymbol{u}\ \boldsymbol{v}]$ has the same sign as $\det[A\boldsymbol{u}\ A\boldsymbol{v}]$. Since the latter is $(\det A)\det[\boldsymbol{u}\ \boldsymbol{v}]$, this is equivalent to saying $\det A$ is positive. Another way of putting this is as follows. A preserves orientation if and only if the sign of $\varphi(\boldsymbol{u},\boldsymbol{v})$ is the same as the sign of $\varphi(A\boldsymbol{v},A\boldsymbol{u})$ for every pair of non zero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ with distinct lines of support. For those who know cross products, this is equivalent to saying $\boldsymbol{u}\times\boldsymbol{v}$ and $A\boldsymbol{u}\times A\boldsymbol{v}$ point in the same direction (and not in opposite directions).

Suppose Γ is an orthogonal 2×2 matrix such that $\det\Gamma = 1$. Since the first column is of unit length it is equal to $(\cos\theta,\sin\theta)$ for a unique $\theta\in[0,2\pi)$. Since the second column is orthogonal to the first and $\det\Gamma = 1$, the second column must be $(-\sin\theta,\cos\theta)$. Thus $\Gamma = \Gamma_\theta := \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$. Pick any non-zero number α and let $A = \alpha\Gamma$. Then for any pair of non-zero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbf{R}^2$ we have

$$\frac{\langle A\boldsymbol{u},A\boldsymbol{v}\rangle}{\|A\boldsymbol{u}\|\|A\boldsymbol{v}\|} = \frac{\alpha^2\langle\Gamma\boldsymbol{u},\Gamma\boldsymbol{v}\rangle}{\alpha^2\|\Gamma\boldsymbol{u}\|\|\Gamma\boldsymbol{v}\|} = \frac{\langle\boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}.$$

Thus A preserves angles. Moreover, since $\det A = \alpha^2 > 0$, A preserves orientation. Conversely, if A preserves orientations and angles, then it is a scalar multiple of Γ_θ for some $\theta\in[0,2\pi)$. In fact the scalar multiple can be taken to be positive. Indeed, let $\alpha = (\det A)^{1/2}$. Then $\alpha^{-1}A$ has determinant equal to one, and preserves lengths (check!) and hence is orthogonal. Since its determinant is one, $\alpha^{-1}A = \Gamma_\theta$ for a unique $\theta \in [0, 2\pi)$ and $A = \alpha\Gamma_\theta$. It is worth pointing out that $A^t A = AA^t = \left[\begin{smallmatrix} \Delta & 0 \\ 0 & \Delta \end{smallmatrix}\right]$ where $\Delta = \det A = \alpha^2$.

An angle and orientation preserving linear map on $\mathbf{R}^2$ clearly preserves shapes since it preserves angles. In fact, being orientation preserving, it preserves directed angles. It turns out that $\mathscr{C}^1$ maps whose derivatives have this property are closely related to holomorphic functions as we will see below.

Definition 2.1.8. Let U be an open set in $\mathbf{R}^2$ and $f\colon U \to \mathbf{R}^2$ a map. Then f is said to be *conformal* if it is $\mathscr{C}^1$ and at each point z of U, $(Jf)(z)$ is angle and orientation preserving.

Suppose U is open in $\mathbf{C}$ and $f\colon U \to \mathbf{C}$ is complex differentiable at z_0. From Proposition 2.1.5 and (2.1.3) we have

$$\det (Jf)(z_0) = u_x(z_0)^2 + u_y(z_0)^2 = |f'(z_0)|^2. \tag{2.1.9}$$

Thus $Df(z_0)$ is invertible if and only if $f'(z_0) \neq 0$.

To lighten notations, let us use the shorthand $J = Jf(z_0)$ and $\Delta = \det J$. Then using Proposition 2.1.5 and (2.1.9), we have:

$$J^t J = \begin{bmatrix} u_x(z_0) & -u_y(z_0) \\ u_y(z_0) & u_x(z_0) \end{bmatrix} \begin{bmatrix} u_x(z_0) & u_y(z_0) \\ -u_y(z_0) & u_x(z_0) \end{bmatrix} = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta \end{bmatrix} \tag{2.1.10}$$

This looks quite a bit like an orthogonal matrix, except for a scalar factor. We elaborate upon that now. *For the rest of this discussion assume $f'(z_0) \neq 0$.* As we observed J is then invertible and $\Delta \neq 0$. Set

$$\Gamma := |f'(z_0)|^{-1} J.$$

Then (2.1.10) shows that Γ is orthogonal and (2.1.9) shows that

$$\det \Gamma = |f'(z_0)|^{-2} \det J = 1.$$

Complex differentiability at a point is not as interesting as complex differentiability on an open set in $\mathbf{C}$. Moreover, when we only assume differentiability at a point, statements have to be made in an awkward way while chasing an artificial generality. So the results of this discussion are best summarised in the following way (the statement has a little more information towards the end than what can be gathered from the above discussion).

Proposition 2.1.11. *Let $f\colon U \to \mathbf{C}$ be a map on an open subset U of $\mathbf{C}$. If f is holomorphic such that f' is nowhere vanishing on U, then f is conformal. Conversely, if f is conformal then it is a holomorphic map whose derivative is nowhere vanishing.*

Proof. The converse is the only thing left to prove. That follows from item **5.** of §§2.1.7. ☐

Here are some pictures illustrating conformality. The map is $f(z) = 2(z + \frac{1}{4})^2$ and the open set is a small disc around the origin where $f'(z) = 4(z + \frac{1}{4})$ does not vanish. The vector $\boldsymbol{u}$ is the velocity vector at $t = 0$ of the path $\boldsymbol{\gamma}$ given by $\boldsymbol{\gamma}(t) := (t, -(t+1)^2+1)$ and $\boldsymbol{v}$ is the velocity vector at $t = 0$ of $\boldsymbol{\sigma}(t) = (t, (t+1)^2-1)$. It is easy to see that the velocity vectors are $(1, -2)$ and $(1, 2)$ and that $\varphi(\boldsymbol{u}, \boldsymbol{v}) = 2\arctan 2$. Note that $\boldsymbol{\gamma}(0) = \boldsymbol{\sigma}(0) = \mathbf{0}$.

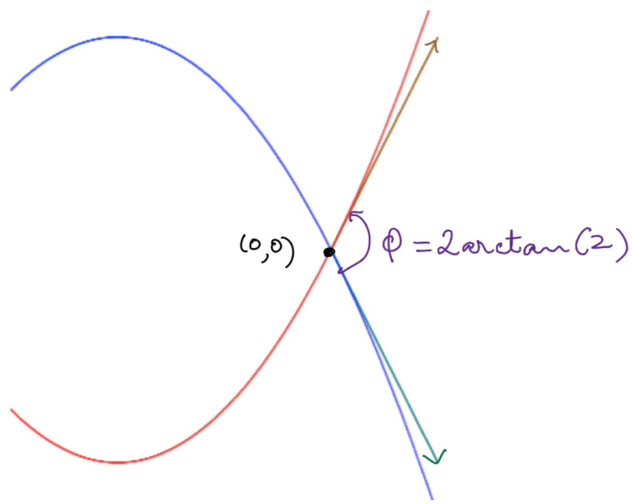

FIGURE 3. The blue curve is $\boldsymbol{\gamma}$ and the red $\boldsymbol{\sigma}$. The arrows give the direction of the velocity vectors.

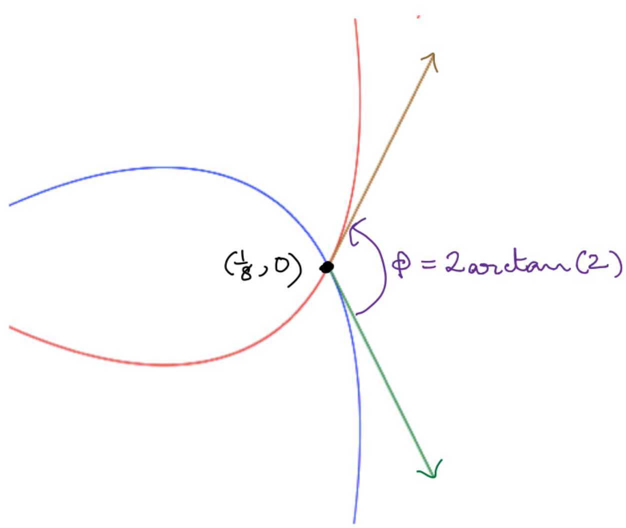

FIGURE 4. The blue curve is $f \circ \boldsymbol{\gamma}$ and the red $f \circ \boldsymbol{\sigma}$. The arrows give the direction of the velocity vectors. The displayed point of intersection is $f(0,0) = (\frac{1}{8}, 0)$.

3. **Sequential compactness**

3.1. A metric space X is said to be *sequentially compact* if every sequence in X has a convergent subsequence. In a later course you will see that this is equivalent to compactness. Let us prove this for X a subset of $\mathbf{R}^n$.

Theorem 3.1.1. *Let S be a subset of $\mathbf{R}^n$. Then S is compact if and only if it is sequentially compact.*

Proof. If $n = 0$ there is nothing to prove. If $n = 1$ this is a theorem from your first analysis course (the so called Bolzano-Weierstrass theorem).

Suppose S is compact. We have to show S is sequentially compact. From above comments, we assume $n \geq 2$. By induction (whose first step is the Bolzano-Weierstrass theorem) we assume that every compact subset of $\mathbf{R}^{n-1}$ is sequentially compact.

For J a non empty subset of $\{1, \ldots, n\}$, say $J = \{j_1, \ldots, j_d\}$, with j_k strictly increasing with respect to k with $1 \leq k \leq d$, we have the projection map $\pi_J \colon \mathbf{R}^n \to \mathbf{R}^d$ given by $\pi_J(x_1, \ldots, x_n) = (x_{j_1}, \ldots, x_{j_d})$. If J is a singleton set, say $J = \{k\}$, it is typographically more convenient to write π_k for $\pi_{\{k\}}$ and we will do so. Let $J^* = \{1, \ldots, n-1\}$. Since π_J is continuous, $S_J := \pi_J(S)$ is compact for every non empty subset J of $\{1, \ldots, n\}$. Let $\{\boldsymbol{x}_m\}$ be a sequence in S. Since S_{J^*} is compact, and since by way of induction we have assumed that every compact subset of $\mathbf{R}^{n-1}$ is sequentially compact, the sequence $\{\pi_{J^*}(\boldsymbol{x}_m)\}_m$ has a convergent subsequence, say $\{\boldsymbol{y}_\nu\}_\nu = \{\pi_{J^*}(\boldsymbol{x}_{m_\nu})\}_\nu$. Let $z_\nu = \pi_k(\boldsymbol{x}_{m_\nu})$. Since the theorem is true for $n = 1$, $\{z_\nu\}$ has a convergent subsequence, say $\{z_{\nu_j}\}$. Then $\{(\boldsymbol{y}_{\nu_j}, z_{\nu_j})\}$ is a convergent subsequence of $\{\boldsymbol{x}_m\}$.

Conversely suppose S is sequentially compact. We have to show is it compact, i.e. show that it is closed and bounded. If $\boldsymbol{\zeta}$ is a limit point of S, there is a sequence $\{\boldsymbol{x}_\nu\}$ in S converging to $\boldsymbol{\zeta}$. Since S is sequentially compact, there is a subsequence $\{\boldsymbol{x}_{\nu_m}\}$ which converges to a point $\boldsymbol{y}$ in S. On the other hand, since $\{\boldsymbol{x}_\nu\}$ converges to $\boldsymbol{\zeta}$, every subsequence converges to $\boldsymbol{\zeta}$. Hence $\boldsymbol{\zeta} = \boldsymbol{y}$, whence $\boldsymbol{\zeta} \in S$. Thus S is closed. It remains to show that S is bounded. Suppose it is not. Then for each $m \in \mathbf{N}$ there exists $\boldsymbol{x}_m \in S$ with $\|\boldsymbol{x}_m\| \geq m$. It is clear that no subsequence of $\{\boldsymbol{x}_m\}$ can be convergent. Indeed, if $\{\boldsymbol{x}_{m_\nu}\}$ is a subsequence of $\{\boldsymbol{x}_m\}$ then $m_\nu \to \infty$ as $\nu \to \infty$ by definition of a subsequence. Since $\|\boldsymbol{x}_{m_\nu}\| \geq m_\nu$, clearly $\{\boldsymbol{x}_{m_\nu}\}$ is not convergent. This contradicts the sequential compactness of S. Hence S is bounded. □

Remark 3.1.2. This proof does not work on an arbitrary metric space. However, it is true that a space is compact if and only if it is sequentially compact. Our proof used the Heine-Borel theorem which is not available for a general metric space. You will see the proof of the theorem for arbitrary metric spaces next semester in Analysis III.

About these notes. This lecture was supposed to be given on April 6, 2020. Classes got suspended on March 17 because of the coronavirus COVID 19 pandemic, and all teaching moved online. These course notes are a reasonably faithful record of the lectures given (before the shutdown) at the Chennai Mathematical Institute (CMI) in the January-April semester of 2019-20. The course is Analysis II, a core course for first year undergraduates at CMI. For more material related to this course, visit `https://www.cmi.ac.in/~pramath/teaching.html#ANA2`. If you have comments on these notes, or on related course material, please send an email to pramath@cmi.ac.in.

References

[A] L. V. Ahlfors, *Complex Analysis*, McGraw-Hill, New-York, 1979.

[MT] J.E. Marsden and A.J. Tromba, *Vector Calculus*, Fifth edition, W.H. Freeman and Company, New York, 2003.

[St] J. Stewart, *Calculus*, Sixth edition, Thomson Brooks/Cole, 2008.