INSPIRING EXERCISES FOR UNDERGRADUATES

TZEE-CHAR KUO

ABSTRACT. There are two ways to learn mathematics: to discover (research) and to rediscover. Inspiring exercises are used to guide students at all levels to rediscover the essential meaning of various individual pieces of mathematics. Their rôle has become all the more important at a time when Bourbakisation is the dominating fashion. While mathematics is decorated with a thick layer of cosmetics, such as utmost generalities, excessive formalism and rigour, subtle proofs, etc, its natural beauty (intuitive ways of thinking, simplicity of ideas, etc) is obscured. From a pedagogical point of view, to say the least, it would be most desirable if a theorem, or an idea, can be fully explained by, hence rediscovered from, one or two simple examples. These kinds of illustrative examples actually exist everywhere, and at all levels. We can use them as inspiring exercises. In this paper, we give five sets of examples, beginning with a simple one on Abel's Identity, followed by examples on Hensel's Lemma, Finitely Generated Abelian Groups, Baire's Category Theorem and the Weierstrass Preparation Theorem. Solutions, hints and discussions are provided in the last section.

Dedicated to Professor Saunders Mac Lane for his 90th birthday

1. Abel's Identity

Look at the following stairway and rediscover the Abel identity:

$$a_1h_1 + a_2h_2 + \cdots + a_nh_n$$
$$= a_1(h_1 - h_2) + (a_1 + a_2)(h_2 - h_3) + \cdots + (a_1 + \cdots + a_n)h_n.$$

This is true in any number system (ring); the stairway suggests a proof.

EXERCISE. Consider the power series

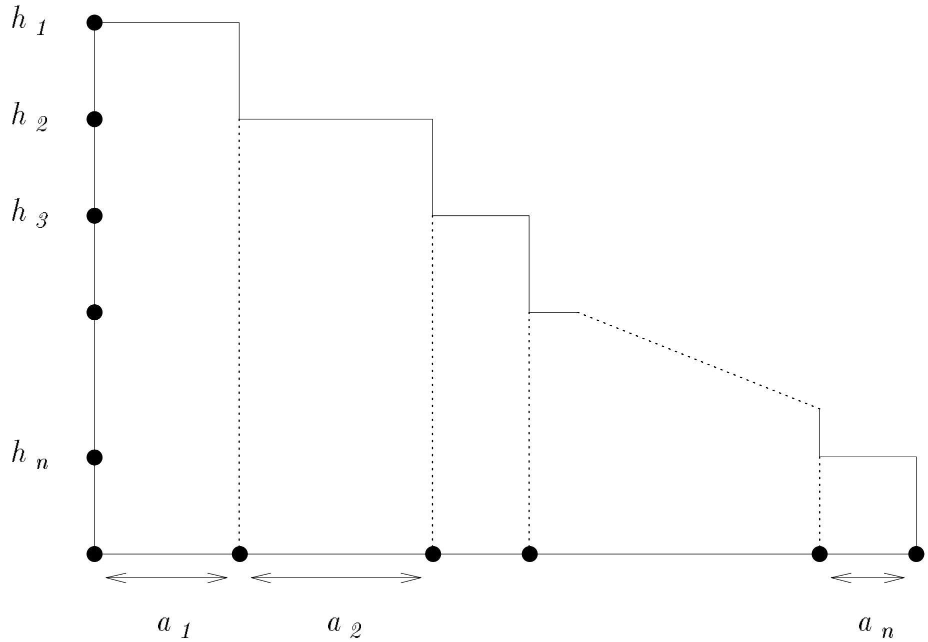
$$x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n+1}\frac{x^n}{n} + \cdots$$

which clearly converges uniformly on $[-1+\delta, 1-\delta], \delta > 0$, and converges at $x = 1$. Show that it converges uniformly on $[-1+\delta, 1]$. (Therefore, $1 - \frac{1}{2} + \frac{1}{3} - \cdots = \log 2$.)

EXERCISE. Show that $\sum_{n=1}^{\infty} \frac{\cos nx}{n}$ converges uniformly on $[\delta, 2\pi - \delta]$.

1991 *Mathematics Subject Classification.* 00A05, 00A07, 00A35, 00A30.

Several standard theorems on uniform convergence are proved in this way. See, for example, [2], Chapter XII.



2. HENSEL'S LEMMA

Consider the following formal power series

$$\sigma(x,y) = xy + H_3(x,y) + \cdots + H_n(x,y) + \cdots ,$$

where $H_n(x,y)$ denotes a homogeneous form of degree n, also called an n-form.

An important observation is that this series always factors:

$$\sigma(x,y) = [x + K_2(x,y) + K_3 + \cdots][y + L_2 + L_3 + \cdots],$$

where K_2, L_2, K_3, L_3, ..., can be calculated recursively (but not uniquely) as follows. Every monomial term of H_n is divisible either by x or by y (many by both). Hence we can write

$$H_3(x,y) = xL_2(x,y) + yK_2(x,y),$$

where K_2, L_2 are 2-forms (not unique). Similarly, 3-forms K_3, L_3 can be found such that

$$xL_3 + yK_3 = H_4 - K_2L_2.$$

By continuing in this way we can find all the terms in a factorisation of $\sigma(x,y)$.

Similarly, a series of the form

$$x^a y^b + H_{a+b+1}(x,y) + \cdots$$

also always factors.

EXERCISE (Hensel's Lemma). (Compare[1],[5].) Given forms $K_a(x,y)$ and $L_b(x,y)$ of the indicated degrees. Suppose they are relatively prime. Show that any formal power series of the form

$$\sigma(x,y) = K_a(x,y)L_b(x,y) + H_{a+b+1} + \cdots,$$

admits a factorisation of the form

$$\sigma(x,y) = [K_a + K_{a+1} + \cdots][L_b + L_{b+1} + \cdots],$$

where K_{a+1}, L_{b+1}, ..., can be calculated recursively. (A proof is given is §6.)

There is also a "weighted" version of the lemma. Let positive integers $w(x)$, $w(y)$ be assigned as the "weights" of x and y. For example, if we choose $w(x) = 3$ and $w(y) = 2$, then $x^2 + y^3$ is a weighted 6-form, while

$$(x^2 + y^3) + xy^2 + (x^2y + y^4) + (x^3 + xy^3) + \cdots$$

is a weighted Taylor series.

Let $K_a(x,y)$, $L_b(x,y)$ be relatively prime weighted forms of the indicated degrees. A weighted Taylor series of the form

$$\sigma(x,y) = K_a(x,y)L_b(x,y) + \cdots$$

admits a factorisation

$$\sigma(x,y) = [K_a + K_{a+1} + \cdots][L_b + L_{b+1} + \cdots]$$

where K_{a+1}, L_{b+1}, ..., can be calculated recursively.

The proof is only slightly more complicated than the homogeneous case. The reader is referred to [4] for details.

3. Finitely generated abelian groups

Consider the abelian group

$$\mathbb{Z} \oplus \mathbb{Z} = \{(m,n) | m,n \in \mathbb{Z}\}$$

of integral lattice points in the plane, with addition

$$(m,n) + (p,q) = (m+p, n+q),$$

and "scalar multiplication"

$$k(m,n) = (km, kn), \qquad k \in \mathbb{Z}.$$

A given set of elements, or "vectors", say

$$\vec{\mathbf{v}}_1 := (m_1, n_1), \quad \ldots, \quad \vec{\mathbf{v}}_N := (m_N, n_N),$$

generates a subgroup consisting of elements of the form

$$k_1\vec{\mathbf{v}}_1 + \cdots + k_N\vec{\mathbf{v}}_N, \qquad k_i \in \mathbb{Z}.$$

EXERCISE (Important). Is it true that every subgroup of $\mathbb{Z} \oplus \mathbb{Z}$ is generated by at most two elements? (Answer given in §6.)

Let $\mathbb{S}$ be a subgroup generated by two elements, say

$$\vec{\mathbf{u}} = (m, n), \qquad \vec{\mathbf{v}} = (p, q).$$

Using Cramer's Rule, it is easy to show that $\mathbb{S} = \mathbb{Z} \oplus \mathbb{Z}$ if and only if

$$\begin{vmatrix} m & p \\ n & q \end{vmatrix} = \pm 1.$$

(This condition means that the parallelogram spanned by $\vec{\mathbf{u}}, \vec{\mathbf{v}}$ has area 1.)

A subgroup, $\mathbb{S}$, generated by two elements $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, is the image of a homomorphism

$$h : \quad \mathbb{Z} \oplus \mathbb{Z} \to \mathbb{Z} \oplus \mathbb{Z}$$

which sends $\vec{\mathbf{e}}_1 := (1, 0)$ to $\vec{\mathbf{u}}$ and $\vec{\mathbf{e}}_2 := (0, 1)$ to $\vec{\mathbf{v}}$.

As in Vector Space Theory, h is represented by the matrix

$$\begin{pmatrix} m & p \\ n & q \end{pmatrix}.$$

Observe that a column operation on the matrix corresponds to a change of basis in the source space. For example, using basis $\{\vec{\mathbf{e}}_1 + \vec{\mathbf{e}}_2, \vec{\mathbf{e}}_2\}$ in the source space, h is represented by

$$\begin{pmatrix} m+p & p \\ n+q & q \end{pmatrix}.$$

A row operation corresponds to a change of basis in the target space. For example, using basis $\{\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2 - k\vec{\mathbf{e}}_1\}$ in the target space, h is represented by

$$\begin{pmatrix} m+kn & p+kq \\ n & q \end{pmatrix}.$$

It is well-known that, by a repeated application of the Euclidean Division Algorithm, $\left(\begin{smallmatrix} m & p \\ n & q \end{smallmatrix}\right)$ can be brought to a diagonal matrix $\left(\begin{smallmatrix} d & 0 \\ 0 & d' \end{smallmatrix}\right)$ by a sequence of row and column operations, where d divides d'. This means that there exist basis $\{\vec{\mathbf{b}}_1, \vec{\mathbf{b}}_2\}$ in the source space, and $\{\vec{\mathbf{c}}_1, \vec{\mathbf{c}}_2\}$ in the target space such that

$$h(\vec{\mathbf{b}}_1) = d\vec{\mathbf{c}}_1, \quad h(\vec{\mathbf{b}}_2) = d'\vec{\mathbf{c}}_2.$$

It follows that $\mathbb{S}$ is generated by $d\vec{\mathbf{c}}_1$ and $d'\vec{\mathbf{c}}_2$, and hence

$$(\mathbb{Z} \oplus \mathbb{Z})/\mathbb{S} \cong \mathbb{Z}_d \oplus \mathbb{Z}_{d'}.$$

Note the following special cases: when $d \neq 0 = d'$, the right hand side is $\mathbb{Z}_d \oplus \mathbb{Z}$, $\mathbb{S}$ is generated by a non zero element; when $d = d' = 0$, $\mathbb{S}$ is the zero subgroup; and in case $d \neq 0 \neq d'$, the above is the direct sum of two finite cyclic groups.

Now, an abelian group $\mathbb{A}$ generated by two elements is, of course, the image of a homomorphism

$$\sigma : \quad \mathbb{Z} \oplus \mathbb{Z} \longrightarrow \mathbb{A}$$

which sends $\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2$ to the generators of $\mathbb{A}$. Hence $\mathbb{A}$ is isomorphic to $(\mathbb{Z} \oplus \mathbb{Z})/\mathbb{S}$, $\mathbb{S}$ the kernel of σ.

EXERCISE. Show that every subgroup of

$$\mathbb{Z} \oplus \mathbb{Z} \oplus \cdots \oplus \mathbb{Z} \qquad (n \text{ copies})$$

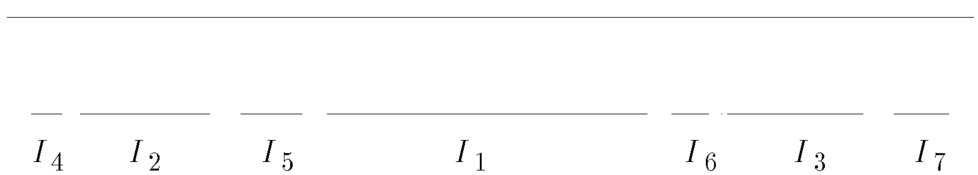
is generated by at most n elements, whence derive that every abelian group generated by n elements is the direct sum of at most n cyclic groups.

The structure theorem of finitely generated modules over a principal ideal domain is proved in *exactly* the same way. (See any text book.)

4. Cantor set and Baire Category Theorem

The real number system, $\mathbb{R}$, is complete. That is, every Cauchy sequence converges. Using the completeness property, one can answer all the questions in this section.

First, take an infinite sequence of closed, mutually disjoint, intervals $I_n = [a_n, b_n]$ in the open interval $(0, 1)$, as indicated:



EXERCISE. Suppose each I_n is sufficiently long. Can they exhaust $(0, 1)$:

$$(0, 1) = [a_1, b_1] \cup \cdots \cup [a_n, b_n] \cup \cdots ?$$

A similar construction gives rise to the celebrated Cantor set. Consider the following sequence of closed sets, F_n, obtained recursively by "purging the middle $\frac{1}{3}$":

$$F_0 := [0, 1]; \quad F_1 := [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]; \quad F_2 := [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1], etc.$$

The Cantor set is, by definition, $\mathfrak{C} := F_0 \cap \cdots \cap F_n \cap \cdots$.

Let us collect the vertices:

$$\mathbb{V} := \{0, 1, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \cdots\}.$$

Of course, $\mathfrak{C}$ is closed and contains $\mathbb{V}$.

EXERCISE. Determine whether $\mathfrak{C}$ contains any point other than the vertices.

Next, a harder question. The rationals, $\mathbb{Q}$, is countable:

$$\mathbb{Q} = \{r_1, r_2, \cdots, r_n, \cdots\}.$$

For every integer $k \geq 1$, let

$$\mathbb{J}_k := \bigcup_{n=1}^{\infty}(r_n - \frac{\epsilon}{2^{k+1+n}}, r_n + \frac{\epsilon}{2^{k+1+n}}), \qquad \epsilon > 0, \quad \text{a constant.}$$

Each $\mathbb{J}_k$ is an open set containing $\mathbb{Q}$, having small measure:

$$m(\mathbb{J}_k) \leq \sum_{n=1}^{\infty} \frac{2\epsilon}{2^{k+1+n}} = \frac{\epsilon}{2^k}.$$

QUESTION: Is it true that $\bigcap_{k=1}^{\infty} \mathbb{J}_k = \mathbb{Q}$?

Consider $\mathbb{N}_k := \mathbb{R} - \mathbb{J}_k$. This is closed and excludes $\mathbb{Q}$, hence does not contain any open interval.

A subset, $\mathbb{N}$, of $\mathbb{R}$ is called *nowhere dense* if the closure, $\overline{\mathbb{N}}$, does not contain any open interval. Each $\mathbb{N}_k$ is a typical nowhere dense set. Baire's Category Theorem asserts that $\mathbb{R}$ cannot be a countable union of nowhere dense subsets.

In particular, the answer to the above question must be "No". For otherwise, there would be a contradiction:

$$\mathbb{R} = \mathbb{N}_1 \cup \{r_1\} \cup \mathbb{N}_2 \cup \{r_2\} \cup \cdots .$$

EXERCISE. Prove Baire's Category Theorem. (This is no harder than showing that the answer to the above question is "No".)

5. Weierstrass Preparation Theorem

Consider a polynomial equation

$$a_n(w)z^n + \cdots + a_k(w)z^k + \cdots + a_0(w) = 0$$

where $a_i(w)$ are continuous (or analytic) functions of w, $|w|$ small.

EXERCISE. Suppose $a_0(0) = \cdots = a_{k-1}(0) = 0 \neq a_k(0)$. Show that as $w \to 0$, exactly k roots $\to 0$.

QUESTION. Are there k continuous (or analytic) functions $z_1(w), \cdots, z_k(w)$, each defined for all w near 0, such that for every given value of w, $\{z_1(w), \cdots, z_k(w)\}$are *the* k roots near 0?

More generally, consider a convergent power series in z, w:

$$p(z, w) = \sum c_{ij} z^i w^j.$$

Let us suppose

$$p(z, 0) = a_k z^k + a_{k+1} z^{k+1} + \cdots, \qquad a_k \neq 0.$$

Thus, as above, when $w = 0$, $z = 0$ is a root of multiplicity k. We say, in this case, that $p(z, w)$ is *regular* in z of order k.

EXERCISE. Show that there exist $\epsilon > 0, \delta > 0$ such that for each given value of w, with $|w| < \delta$, there are exactly k roots $z_1(w), \cdots, z_k(w)$ of $p(z, w) = 0$, with $|z_i(w)| < \epsilon$.

Now, let us state the celebrated Weierstrass Preparation Theorem:

Theorem 5.1. *Suppose $p(z,w)$ is regular in z, say of order k. There exist $\epsilon > 0$, $\delta > 0$, such that*

$$p(z,w) = u(z,w)[z^k + a_1(w)z^{k-1} + \cdots + a_k(w)], \quad |z| < \epsilon, \quad |w| < \delta,$$

where $u(z,w)$ is analytic, $u(0,0) \neq 0$, and each $a_i(w)$ is analytic, $a_1(0) = \cdots = a_k(0) = 0$.

We call $u(z,w)$ an analytic unit, and

$$W(z,w) := z^k + a_1(w)z^{k-1} + \cdots + a_k(w)$$

the Weierstrass polynomial associated to $p(z,w)$.

This theorem reduces the study of $p(z,w)$, near 0, to that of a polynomial-like function, which is easier to handle, whence is called the "Preparation Theorem".

Now we describe the motivation and the proof.

Having been assured of exactly k roots, $z_1(w), \cdots z_k(w)$ (possibly repeated), in the disk $|z| < \epsilon$, an answer to the above question is now important.

Suppose each $z_i(w)$ is analytic, then

$$W(z,w) := \prod_{i=1}^{k}[z - z_i(w)]$$

would of course have analytic coefficients. The rest of the proof then runs as follows.

The quotient function $p(z,w)/W(z,w)$, as a function of z (for each fixed value of w), has only removable singularities, and is non-vanishing. By a standard theorem in Several Complex Variables Theory, called the Riemann Removable Singularities Theorem ([3]),

$$u(z,w) := \frac{p(z,w)}{W(z,w)}$$

is analytic in z,w, and $u(0,0) \neq 0$. This would complete the proof.

But, unfortunately, the answer to the question is "No". Take, as a simple example, the equation $z^2 - w = 0$. For each w, fixed, there are two roots $z_1(w), z_2(w)$. Yet, for a simple topological reason, it is clear that no continuous function $z(w)$, which is defined for *all* w in a neighbourhood of 0 and satisfies this equation, can exist. (This phenomenon is the starting point of Riemann Surface Theory.)

Thus, each $z_i(w)$ is merely a function, not even continuous. The most interesting observation, vital for Weierstrass' Theorem, is that their symmetric functions

$$\mathbb{S}_1 := z_1(w) + z_2(w), \qquad \mathbb{S}_2 := z_1(w)z_2(w),$$

being 0 and $-w$ respectively, are actually analytic functions of w.

This is true for polynomials in general. Let the roots of

$$z^k + a_1(w)z^{k-1} + \cdots + a_k(w) = 0, \qquad a_i(w) \quad analytic,$$

be denoted by $z_1(w), \cdots, z_k(w)$, for each given w. Their symmetric functions:

$$\mathbb{S}_1 = \sum z_i(w), \qquad \mathbb{S}_2 = \sum z_i(w)z_j(w), \quad etc,$$

being the coefficients $a_i(w)$, each up to a sign, are analytic.

A vital assertion for the Preparation Theorem is that this is also true for a power series $p(z, w)$, regular in z, say of order k, not just for polynomials.

EXERCISE. Let $z_1(w), \cdots, z_k(w)$ denote the roots of $p(z, w) = 0$, in the disk $|z| < \epsilon$. Show that the following symmetric functions

$$\sigma_\ell := \sum z_i(w)^\ell, \qquad \ell = 0, 1, \ldots,$$

are analytic functions in a neighbourhood of $w = 0$.

To prove this, one need a generalised version of Rouché's Theorem. (See §6.)

As is well known, the two systems of symmetric functions $\{\mathbb{S}_1, \mathbb{S}_2, \cdots\}$ and $\{\sigma_1, \sigma_2, \cdots\}$, each can be expressed as polynomials of the other. For example, $\sigma_2 = \mathbb{S}_1^2 - 2\mathbb{S}_2$, etc. Hence the coefficients of $W(z, w)$ are indeed analytic functions of w. We have therefore proved the Preparation Theorem.

6. Discussions and Solutions

Abel's Identity. Since $\sum(-1)^n \frac{1}{n}$ converges,

$$|\frac{1}{n} + \cdots + (-1)^k \frac{1}{n+k}| < \epsilon, \text{ for } n \geq N.$$

For $0 \leq x \leq 1$, we have $x^n \geq x^{n+1}$, and

$$\begin{aligned} &|\frac{x^n}{n} - \cdots + (-1)^k \frac{x^{n+k}}{n+k}| \\ &= |\frac{1}{n}(x^n - x^{n+1}) + \cdots + [\frac{1}{n} + \cdots + (-1)^k \frac{1}{n+k}](x^{n+k} - 0)| \qquad \text{(Abel)} \\ &< \epsilon|[(x^n - x^{n+1}) + \cdots + (x^{n+k} - 0)]| \\ &= \epsilon x^n \leq \epsilon. \end{aligned}$$

For the second series, note that $\cos nx + \cdots + \cos mx$ is the real part of

$$e^{inx} + \cdots + e^{imx} = e^{inx} \frac{1 - e^{i(m-n+1)x}}{1 - e^{ix}}$$

which is bounded for $x \in [\delta, 2\pi - \delta]$. A simple application of Abel's identity completes the proof.

Hensel's lemma. The proof is elementary, using only the Euclidean Division Algorithm. First, we can assume, without loss of generality, that

$$K_a(1,0) \neq 0 \neq L_b(1,0).$$

This can be achieved, say, by a generic linear change of variables

$$x = X, \qquad y = Y + cX.$$

For example, $x^2y + xy^2$ does not have this property, but

$$x^2y + xy^2 = c(c+1)X^3 + \cdots ,$$

whence taking any $c \neq 0, -1$, would do.

Now, take any $k \geq a + b + 1$. We claim that every given monomial $x^i y^j$, $i + j = k$, can be expressed in the form

$$R(x,y)K_a(x,y) + S(x,y)L_b(x,y) = x^i y^j \tag{6.1}$$

where R, S are $(k-a)$ and $(k-b)-$forms respectively.

Once this is proved, the terms in the factorisation can be recursively calculated.

Let us consider $K_a(z,1)$ and $L_b(z,1)$. They are relatively prime polynomials in a single variable z of degree a and b respectively. (This is called dehomogenisation.) Hence there exist polynomials $A(z)$, $B(z)$, such that

$$A(z)K_a(z,1) + B(z)L_b(z,1) = z^i. \tag{6.2}$$

We can assume $degA < b$. This can be achieved by dividing A by L_b:

$$A(z) = Q(z)L_b(z,1) + R(z), \qquad degR < b.$$

Thus (6.2) can be replaced by

$$R(z)K_a(z,1) + S(z)L_b(z,1) = z^i \tag{6.3}$$

where $S(z) := Q(z)K_a(z,1) + B(z)$.

We must have $degS \leq k - b$, since no term of degree $> k$ can appear in this equation.

Now replace z by x/y in (6.3) and multiply throughout by y^{i+j} to yield (6.1).

Abelian Groups. Yes, every subgroup of $\mathbb{Z} \oplus \mathbb{Z}$ is generated by at most two elements. Given $\mathbb{S}$, to find generators $\vec{\mathbf{g}}_1$ and $\vec{\mathbf{g}}_2$, let us consider

$$I = \{n \in \mathbb{Z} | \exists m, (n,m) \in \mathbb{S}\}$$

which is clearly an ideal.

Hence I consists of the integral multiples of an integer d: $I = \{kd | k \in \mathbb{Z}\}$.

Take d' such that $\vec{\mathbf{g}}_1 := (d, d') \in \mathbb{S}$. Then for any $(n,m) \in \mathbb{S}$, there exists k such that $(n,m) - k\vec{\mathbf{g}}_1$ has the form $(0,e)$.

Now, collect elements of $\mathbb{S}$ of the form $(0,e)$. This is a subgroup generated by a single element, say $\vec{\mathbf{g}}_2$. And then $\mathbb{S}$ is generated by $\vec{\mathbf{g}}_1, \vec{\mathbf{g}}_2$. (In case $d = 0$, $\vec{\mathbf{g}}_1 = \vec{\mathbf{g}}_2$.)

Cantor Set. We first show that the I_k's, as indicated, can not exhaust $(0,1)$. Let us define a sequence $\{c_n\}$ as follows. Take $c_1 = b_1$, the end point of I_1. Then, as I_3 is the first interval which appears to the right of b_1, we take $c_2 = a_3$, the initial point of I_3. Now, I_6 is the first interval which appears to the immediate left of I_3, we take $c_3 = b_6$, etc. Thus, c_{2n-1} is increasing, and c_{2n} is decreasing, $c_{2n-1} < c_{2n}$.

Let $\alpha = \lim c_{2n-1}, \beta = \lim c_{2n}$. (If the $I_k's$ are sufficiently long, $\alpha = \beta$.)

Take any I_N, N fixed. The sequence $\{c_n\}$, by construction, will eventually "dance away" from I_N. Hence α, β can not be in I_N, for any N; $(0,1)$ is not exhausted.

This idea can also be used to show that the Cantor Set is not countable.

Suppose $\mathfrak{C}$ is countable: $\mathfrak{C} = \{x_1, \cdots, x_n\}$. We can then choose a sequence $\{y_n\}$ with the following properties:

1. y_n is a vertex of F_n,
2. y_n and x_n are not in the same component interval of F_n,
3. y_{n+i}, for all $i \geq 0$, belong to a same component interval of F_n.

This is a Cauchy sequence whose limit cannot be x_k, for any given k. Hence $\mathfrak{C}$ cannot be countable, and must contain non vertex points.

Finally, let us describe a proof of Baire's Category Theorem.

Let $\{N_n\}$ be a sequence of nowhere dense sets in $\mathbb{R}$. We can choose a sequence of closed intervals $[a_n, b_n]$ which "dances away" from every given N_k. More precisely, we have: $[a_n, b_n] \cap \mathbb{N}_n = \emptyset$, and $[a_n, b_n] \supseteq [a_{n+1}, b_{n+1}]$, for all n. Then, clearly, $\lim a_n$, $\lim b_n \notin \bigcup_{k=1}^{\infty} N_k$.

We have actually proved more. Let (a, b) be any given interval. We can then choose the first interval $[a_1, b_1]$ in (a, b). It follows that the union $\bigcup N_k$ does not contain all the points of (a, b). Hence the complement $\mathbb{R} - \bigcup N_k$ is dense in $\mathbb{R}$ and is uncountable.

The full statement of Baire's Category Theorem asserts that in a complete metric space, a countable union of nowhere dense subsets is a residue set in the sense that its complement is dense and uncountable. The proof is the same.

Preparation Theorem. Let $p(z)$ be an analytic function. Take a contour, $\mathfrak{C}$, in the z-plane. Suppose $p(z) = 0$ has no root lying on $\mathfrak{C}$.

There is a formula expressing the total number of roots, counting multiplicities, of $p(z) = 0$ in the interior of $\mathfrak{C}$. This is Roché's Theorem, as follows.

Let a be a root of $p(z) = 0$ of multiplicity m:

$$p(z) = (z-a)^m g(z), \qquad g(a) \neq 0,$$

where $g(z)$ is analytic.

Take $\epsilon > 0$, sufficiently small, then, clearly,

$$m = \frac{1}{2\pi i} \oint_{|z-a|=\epsilon} \frac{p'(z)}{p(z)} dz.$$

Therefore the total number of roots within $\mathfrak{C}$ is

$$\frac{1}{2\pi i}\oint_{\mathfrak{C}} \frac{p'(z)}{p(z)}dz \qquad \text{(Rouché)}.$$

This formula enables us to calculate the number of roots without knowing exactly where they are. Its usefulness to us is obvious. Adding a sufficiently small perturbation to $p(z)$ will not change the values of the above two integrals, because they are integers.

Hence, if $\mathfrak{C}$ is a contour containing n roots of $p(z)$, the perturbed equation also has exactly n roots within $\mathfrak{C}$. If a is a root of multiplicity m, the perturbed equation also has exactly m roots near a.

For the first two exercises, we can simply consider $p(z, w)$, $|w|$ small, as a perturbation of $p(z, 0)$.

For the last exercise, take an analytic function in one variable, $p(z)$, with roots $z_1, \cdots, z_k$, within a given contour $\mathfrak{C}$. Then

$$\sigma_\ell := z_1^\ell + \cdots + z_k^\ell = \frac{1}{2\pi i}\oint_{\mathfrak{C}} \frac{z^\ell p'(z)}{p(z)}dz.$$

This can be proved by an easy calculation of the residues. When $\ell = 0$, this is Rouché's Theorem.

Now, we can treat w as a parameter and then

$$\sum_{j=1}^{k} z_j(w)^\ell = \frac{1}{2\pi i}\oint_{|z|=\epsilon} \frac{z^\ell \frac{\partial p}{\partial z}}{p(z, w)}dz.$$

Since the denominator $p(z, w)$ never vanishes on $|z| = \epsilon$, this integral is an analytic function of w.

Concluding Remark. The best way to understand a theorem is to discover, by oneself, a proof for the first non-trivial special case. In this way, one would be led to encounter illustrative examples, and be able to catch the essential meaning of the theorem. The general case may become quite straightforward.

As an experiment, a student who knows some elementary topology may try this method to learn the theory of covering spaces, which is normally the content of a chapter in a standard text book on Riemann Surfaces or Algebraic Topology.

Two typical examples of covering spaces are:

$$p_1 : \mathbb{R}^1 \to \mathbb{S}^1 \quad \text{and} \quad p_2 : \mathbb{R}^2 \to \frac{\mathbb{R}^2}{\mathbb{Z} \oplus \mathbb{Z}},$$

where $p_1(x) = e^{ix}$, and p_2 is the quotient map which sends (x, y) to its equivalence class. Here (x, y) is equivalent to $(x + n, y + m)$ for all $n, m \in \mathbb{Z}$; the target space is a torus.

Now, carefully read the definition of a path and the statement of the Path Lifting Theorem. Then prove the theorem for p_2. (Yes, you can do it.)

It is then easy to read the rest of the chapter, to see how a loop downstairs acts on the space upstairs, generating a group action, etc.

Thus, the whole theory is fully exposed by one single example.

This phenomenon, as our examples have shown, is not at all isolated in mathematics.

A theorem, and its proof, are just one entity; illustrative examples are the soul. For a student, "soul-searching" is a key to understanding and research; Bourbakisation is not.

When I was a student at Chicago (a very long time ago), I took an elementary reading course with Professor Saunders Mac Lane. I saw him once a week. But there was very little reading, only inspiring exercises, one after another.

That was, and still is, the best course I have ever had.

Acknowledgement. This paper is based on my University Special Lecture for Academic Year 1996 at Hyogo University of Teacher Education, and written up during my visit at l'Université d'Angers in 1998. I wish to express my deep appreciation to the two institutions and PRIA for their supports, and especially to Professors Satoshi Koike, Jean-Michel Granger and Adam Parusiński, along with their colleagues, for their hospitality and help. Thanks are due to Professor Shinya Sakitani and again to Satoshi for valuable suggestions. I also wish to thank Professor Michael Polyak for drawing the two figures in this paper, and to Adam again, for teaching me patiently the LaTex skill. Dr Laurentiu Paunescu's help in the final preparation of the text has been most helpful, to him, special thanks.

This work is partially supported by an ARC Institutional Grant.

References

[1] M.F. Atiyah, I.G. MacDonald, *Introduction to Commutative Algebra,* (1969) Addison-Wesley, Reading, MA.
[2] E. Fisher, *Intermediate Real Analysis*, Undergraduate Text in Mathematics, (1980) Springer-Verlag, New York, Heidelberg, Berlin.
[3] R.C. Gunning, H. Rossi, *Functions of Several Complex Variables,* (1965) Prentice-Hall, Englewood Cliff, N. J.
[4] T.C. Kuo, *A Simple Algorithm for Deciding Primes in* $\mathbb{K}[[x,y]]$, Can. J. Math. Vol **47** (4), 1995 pp.80 1-816.
[5] A. Seidenberg, *Elements of the Theory of Algebraic Curves,* (1968) Addison-Wesley, Reading, MA.

School of Mathematics and Statistics, University of Sydney, Sydney, NSW, 2006, Australia
E-mail address: tck@maths.usyd.edu.au