Lecture 10

Domenico Valloni

1 Integral elements and integral extensions

The idea of "integral elements" begins in algebraic number theory (Kummer and Dedekind) so I am going to explain its origin first as a motivation.

The starting point is the ring $\mathbb{Z}$ together with its fraction field $\mathbb{Q} = \mathrm{Frac}(\mathbb{Z})$. In algebraic number theory, one studies number fields, which are field extensions $\mathbb{Q} \subset L$ of finite degree, e.g., $L = \mathbb{Q}(\sqrt{d})$ for d a square-free integer. The question is then: is there a way to construct a subring $\mathbb{Z}_L \subset L$ which - in some sense - plays the same role as $\mathbb{Z}$ for $\mathbb{Q}$? The answer is not obvious, and basically requires to define when an element $x \in L$ is integral.

Definition 1. Let $R \subset S$ be an extension of rings (i.e., an injective ring morphism). An element $s \in S$ is said to be integral over R if it satisfies a monic polynomial equation with coefficients in R, that is, if

$$s^n + r_{n-1}s^{n-1} + \cdots + r_1 s + r_0 = 0$$

for some $n \geq 1$ and $r_i \in R$ (monic means that $r_n = 1$).

Example. Let R be a UFD, so in particular a domain, and let F be its fraction field. We consider the ring extension $R \subset F$, and we now compute the elements of F which are integral over R. Take $f \in F$ and write $f = a/b$ with $a, b \in R$, $b \neq 0$ and $\gcd(a, b) = 1$ (note that since R is a UFD, we have a well-defined notion of gcd). Assume that f is integral over R, so there are $r_i \in R$ such that

$$f^n + r_{n-1}f^{n-1} + \cdots + r_1 f + r_0 = 0, \ \ i.e.$$

$$(a/b)^n + r_{n-1}(a/b)^{n-1} + \cdots + r_1(a/b) + r_0 = 0,$$

by multiplying both sides by b^n we then get

$$a^n + r_{n-1}a^{n-1}b + \cdots + r_1 a b^{n-1} + r_0 b^n = 0.$$

Note that this last equation takes place in R, and it implies that b divides a. Since a and b are coprime by assumption, then b is necessarily a unit in R, hance $f = a/b \in R$. So the only elements of F integral over R are precisely the elements of R.

In particular, the only elements of $\mathbb{Q}$ integral over $\mathbb{Z}$ are the integers, and the only elements of $K(x)$ integral over $K[x]$ are the polynomials (where K denotes a field).

Back to our original motivation, given a number field $L/\mathbb{Q}$, we would like to define the integers of L as $\mathbb{Z}_L \coloneqq \{x \in L \colon x \text{ is integral over } \mathbb{Z}\}$. The problem with this is that it is hard to show that it is a ring: in general, given a ring extension $R \subset S$ and $s_1, s_2 \in S$, it is very hard to show by hand that $s_1 + s_2$ or $s_1 \cdot s_2$ are integral over R knowing that s_1 and s_2 are, because there is not a general formula to cook up a monic equation satisfied by the sum or the product. In the end, this problem is solved using a clever applycation of Cayley-Hamilton (which we already encountered):

Theorem 2 (Cayley-Hamilton)**.** *Let R be a ring and let M be a finitely generated module. Let $\phi \colon M \to M$ be a R-module morphism. Then, there is a monic polynomial $P(x) \in R[x]$ such that $P(\phi) = 0$.*

Proof. One begins by proving the theorem for M free, i.e., $M = R^n$. In this case, any $\phi \colon M \to M$ can be written as a $n \times n$ matrix $A = (a_{ij}) \in M_{n\times n}(R)$ with coefficients in R. Then one defines $P(x) = \det(x \operatorname{Id}_n - A)$. This is clearly a monic polynomial of degree n. The proof that $P(\phi) = 0$ is the same as in linear algebra.

Now, if M is a finitely generated module, we choose a surjection $\pi \colon R^n \twoheadrightarrow M$. Using the universal property of free modules, it is easy to find a dotted arrow which makes the following diagram commute:

$$\begin{array}{ccc} R^n & \xrightarrow{\pi} & M \\ \vdots & & \downarrow{\scriptstyle \phi} \\ R^n & \xrightarrow[\pi]{} & M \end{array}$$

let us call it $\psi \colon R^n \to R^n$. Now we have a monic polynomial $P(x) \in R[x]$ such that $P(\psi) = 0$, i.e.,

$$\psi^n + r_{n-1}\psi^{n-1} + \cdots + r_1\psi + r_0 = 0$$

(recall that ψ^k is the k-fold composition $\psi \circ \psi \circ \cdots \psi$). But $\pi \circ \psi^k = \phi^k \circ \pi$ and therefore

$$(\phi^n + r_{n-1}\phi^{n-1} + \cdots + r_1\phi + r_0) \circ \pi = 0$$

Finally, since π is surjective, this implies that $P(\phi) = 0$. □

So, the solution of our problem is contained in the following theorem, which gives another characterization for integral elements:

Theorem 3. *Let $R \subset S$ be a ring extension and let $s \in S$. The following are equivalent:*

1. *s is integral over R;*
2. *The subring $R[s] \subset S$ generated by s is a finitely generated R-module;*
3. *There is a subring $T \subset S$ which is a finitely generated R-module and which contains s.*

Proof. Let us show that $(1) \implies (2)$. Recall that $R[s]$ is nothing but the image of the map $R[x] \to S$ which sends x to s. Since $R[x]$ is generated as a R-module by $1, x, x^2, \cdots$, it follows that also $R[s]$ is generated by $1, s, s^2, \cdots$ as a R-module. Since s is integral by assumption, it satisfies a relation of the form $s^n + r_{n-1}s^{n-1} + \cdots + r_1 s + r_0 = 0$, i.e., $s^n = -r_{n-1}s^{n-1} - \cdots - r_1 s - r_0$. This shows that in fact $R[s]$ is generated only by $1, s, \cdots, s^{n-1}$ as a R-module, because any power s^k for $k \geq n$ can be written as a R-linear combination of $1, s, \cdots, s^{n-1}$. So $R[s]$ is a finitely generated R-module.

The fact that $(2) \implies (3)$ is obvious, as we can take $T = R[s]$.

Finally, we prove that $(3) \implies (1)$. Since T is a ring and $s \in T$, we can consider the map $\phi\colon T \to T$ given by multiplication by s. This is clearly R-linear. Since T is finitely generated by assumption as a R-module, we can apply Cayley-Hamilton, and deduce that there is a monic polynomial $P(x) \in R[x]$ such that $P(\phi) = 0$, i.e.,

$$\phi^n + r_{n-1}\phi^{n-1} + \cdots + r_1\phi + r_0.$$

Note that this equation takes place in $\mathrm{End}_R(T)$. Since $\phi(1) = 1 \cdot s$ by evaluating the equation above at 1 we get

$$s^n + r_{n-1}s^{n-1} + \cdots + r_1 s + r_0 = 0,$$

which shows that s is integral over R. □

Definition 4. A ring extension $R \subset S$ is said to be

1. Integral, if every element of S is integral over R;
2. Finite, if S is a finitely generated R-module (with respect to its natural R-module structure).

Note that every finite extension is integral by the theorem above. On the other hand, an integral extension is finite if and only if it is generated by finitely many elements as a ring, i.e., if there are $s_1, \cdots, s_n \in S$ such that $R[s_1, \cdots, s_n] = S$.

Proposition 5. *Both properties are transitive, i.e., if $R \subset S$ and $S \subset T$ are integral/finite ring extensions, then also $R \subset T$ is integral/finite.*

Proof for finite case. We begin with $R \subset S$ finite. Since S is a finitely generated R-module, there is a surjection $R^n \twoheadrightarrow S$ of R-modules. Since T is a finitely generated S-module, there is a surjection $S^m \twoheadrightarrow T$ of $S-$modules. From this it follows that there is a surjection $R^{nm} \cong (R^n)^m \to T$ of R-modules, hence T is finite. □

Before proving the other part of the statement, we need a lemma:

Lemma 6. *Let $R \subset S$ be a ring extension, and let $s_1, \cdots, s_n \in S$ be integral over R. Then, the subring $R[s_1, \cdots, s_n] \subset S$ is a finitely generated R-module.*

Proof. We prove it by induction on n. If $n = 1$, this is Theorem 3. Now, we assume that the statement is true for values smaller than n. Since s_n is integral over R, it is in particular integral over $R[s_1, \cdots, s_{n-1}]$. Hence, $R[s_1, \cdots, s_n]$ is a finitely generated $R[s_1, \cdots, s_{n-1}]$-module by Theorem 3. Since also $R[s_1, \cdots, s_{n-1}]$ is a finitely generated R-module by induction, the proof follows by the previous result. □

Proof of Proposition 5 for integral extensions. To prove the statement in case both extensionsa are integral, we need to show that every $t \in T$ is integral over R. Since t is integral over S by assumption, we can find a monic polynomial $P(x) \in S[x]$ such that $P(t) = 0$. Write $P(x) = x^n + s_{n-1}x^{n-1} + \cdots + s_1 x + s_0$. Since every element of S is integral over R by assumption, we can use the previous lemma to deduce that $S' \coloneqq R[s_0, s_1, \cdots, s_{n-1}] \subset S$ is a finitely generated R-module. But then t is integral over S' by construction, hence $S'[t]$ is a finitely generated S'-module. This shows that $S'[t]$ is a finitely generated R-module, and we can use Theorem 3 to deduce that t is integral over R. □

We finally prove that integral elements form a ring:

Corollary 7. *Let $R \subset S$ be a ring extension. Then the set*

$$S' = \{s \in S \colon s \text{ is integral over } R\}$$

is a subring of S. We call it the integral closure of R in S.

Proof. Clearly $0, 1 \in S'$. If $s \in S'$ then clearly also $-s \in S'$. Now, take $s_1, s_2 \in S$. To show that also $s_1 + s_2$ and $s_1 s_2$ belong to S', we simply use Lemma 6: since s_1 and s_2 are integral over R, we deduce that $R[s_1, s_2] \subset S$ is a finitely generated R-module. But then $s_1 + s_2$ and $s_1 s_2$ belong to $R[s_1, s_2]$, and therefore are integral over R, because both contained in a subring of S which is a finitely generated R-module. □

Finally, let us give an example of an integral extension which is not finite. Consider the ring $\mathbb{Z}[\sqrt{2}, \sqrt[3]{2}, \sqrt[4]{2}, \sqrt[5]{2}, \cdots]$. This is clearly integral over $\mathbb{Z}$, but it is not finite.

Definition 8. Assume that R is a domain and let F be its fraction field. The integral closure of R in F is called the normalization of R. A domain R is said to be normal if it is equal to its normalization.

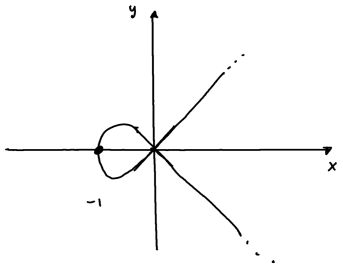
For instance, we showed at the beginning of the lecture that every UFD is normal. To conclude the lecture, we make some examples on the meaning of normality. Whoever will take the next semester course on algebraic curves, will see how the notion of normality is related to the notion of singularities. We give some informal insights (this is not exam material). So take K an algebraically closed field, and let $R = K[x, y]/(f)$ be a domain (so f is irreducile). Then by the Nullstellensatz, we know that $m - \mathrm{Spec}(R)$ corresponds one-to-one to $V(f) = \{(x, y) \in K^2 \colon f(x, y) = 0\}$, which is an irreducible curve (because R is a domain). We have also explained how one should interpret R as the ring of polynomial functions on $V(f)$. Now, $F \coloneqq \mathrm{Frac}(R)$ should be interpreted as the field of meromorphic functions on $V(I)$, i.e., functions which are not well-defined at some point of the curve (where they have poles).

For instance, $m - \mathrm{Spec}(K[x])$ corresponds to the line K and $\mathrm{Frac}(K[x]) = K(x)$. Any element of $f \in K[x]$ yields a function $K \to K$ given by $k \mapsto f(k)$. Consider now an element $f/g \in K(x)$ and let $V(g) \subset K$ be the finitely many points $k \in K$ such that $g(k) = 0$. Let $U = K \setminus V(g)$. This is a Zariski open of K, by definition, and f/g yields a well-defined function $U \to K$ given by $k \mapsto f(k)/g(k)$. Note that if $K = \mathbb{C}$ this is a meromorphic function, with poles contained in $V(g)$.

Going back to our original situation and notation, R being normal then means that the only meromorphic functions on $V(f)$ which are integral over R are the polynomial functions themselves.

Example.

1. The node. Consider $R = K[x,y]/(y^2 - x^2 - x^3)$. We know from previous lectures that R is a domain. To visualize $V(y^2 - x^2 - x^3)$ we can assume $K = \mathbb{R}$. Note that when $x \to 0$ the term x^3 is negligible, and the equation looks like $(y-x)(y+x) = 0$. Drawing the zero locus in $\mathbb{R}^2$ will result in something like this:



Although we do not know what smooth or singular means, we can guess that $(0,0)$ is a singular point of the curve. Let us now show that R is not normal. By abuse of notation, we denote by $x, y \in R$ the images of $x, y \in K[x,y]$ under the quotient $K[x,y] \to R$. Clearly R is generated as a K-algebra by x and y. Let $t = y/x \in F = \mathrm{Frac}(R)$. Our first claim is that the natural inclusion $K(t) \subset F$ is an equality. Surely also F is generated by x and y; now $t^2 = y^2/x^2 = 1 + x$, so $x \in K(t)$. But then $y = tx$ also belongs to $K(t)$, from which the claim follows.

Our second claim is that $t \notin R$. By absurd, assume that $t \in R$. Now, the equation $tx = y$ holds in R. We lift this equation to $K[x,y]$: since by assumption $t \in R$ we can find $T \in K[x,y]$ whose image in R is equal to t, then we have $T \cdot x - y \in (y^2 - x^2 - x^3)$, i.e., $T \cdot x - y = (y^2 - x^2 - x^3) \cdot P$ for some $P \in K[x,y]$. But plugging $x = 0$ yields the equation $-y = y^2 P(0,y)$ which is clearly impossible.

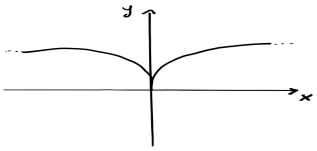
Finally, note that $t \in F$ is integral over R, since $t^2 = 1 + x$. This shows that R is not normal. Interpreting t as a meromorphic function on the curve, we see that t is well-defined at every point but $(0,0)$, but not because it has a pole in it! The fact is that t is indetermined at $(0,0)$ because it can take both values ± 1 (as the equation $t^2 = 1 + x$ also shows) i.e., it is not continuous. To explain this better, still working in $\mathbb{R}^2$, let U be our curve without the point $(0,0)$. Then, the map $U \to \mathbb{R}$ given by $(x,y) \to y/x$ is a well-defined continuous map. Now, we said already that the curve looks like $(x-y)(x+y) = 0$ when x is very close

to 0. Thus, the curve looks like $\{(x,x)\} \cup \{(x,-x)\}$ when x is very small. But the function t has constant value 1 on the first set, and constant values -1 on the second, so it cannot be extended continuously to $(0,0)$.

Let us now compute the integral closure of R in F. We already know that this is a ring which contains t. Note that $R[t] = K[t]$ as subrings of F (why?). Finally, $F = K(t)$ and since $K[t]$ is integrally closed in $K(t)$ because $K[t]$ is a UFD, we conclude that $K[t]$ must be the integral closure of R in F (why?).

Finally, the (finite) ring extension $R \subset K[t]$ yields a map of maximal-spectrum (this is not true in general, but always true for finite ring extensions, as we shall see in later lectures). Thus we get a map $f\colon K \to V(y^2 - x^2 - x^3)$. Since $x = t^2 - 1$ and $y = xt = t(t^2 - 1)$, we see that the map corresponds to $f(t) = (t^2 - 1, t(t^2 - 1))$. Note what is happening here: we started with the irreducible curve $V(y^2 - x^2 - x^3)$ which has a singular point at $(0,0)$. Taking the normalization of R gave us a smooth curve (namely, the line K) and a map $K \to V(y^2 - x^2 - x^3)$ which "resolved the singularities" of the curve. This is a general phenomena for curves: the normalization will always yield a resolution of singularities.

2. The cusp: Another singularity is the cusp $V(x^2 - y^3) \subset K^2$. Again working over $\mathbb{R}$ to get some geometric insights, we see that if $(x,y) \in V(x^2 - y^3)$ then $y \geq 0$ and that therefore the curve is just the graph of the function $y = (x^{1/3})^2$. Again,



we notice that the point $(0,0)$ looks bad. We check that $R = K[x,y]/(x^2 - y^3)$ is not normal (we know that it is a domain). Once again, we let $t = x/y \in \mathrm{Frac}(R)$. Then $t^2 = y$, from which we deduce that t is integral over R. We show that $t \notin R$ as we did before: we assume by absurd that $t \in R$ and we let $T \in K[x,y]$ be a polynomial which is mapped to t under the quotient map. Then we must have that $T(x,y)y - x \in (x^2 - y^3)$ i.e., there is polynomial $P \in K[x,y]$ such that $T(x,y)y - x = P(x,y)(x^2 - y^3)$. By putting $y = 0$ we obtain a contradiction. But note the difference from the previous case: the function t is actually well-defined and continuous on the curve $V(x^2 - y^3)$, sending $(0,0)$ to 0. The reason why we should not consider this as a regular function is more subtle: although t yields a continuos function, this is not differentiable at $(0,0)$!

Again as before, we can check that $\mathrm{Frac}(R) = K(t)$ and that $K[t]$ is the integral closure of R in $\mathrm{Frac}(R)$. The induced map $K \to V(x^2 - y^3)$ is given by

$t \mapsto (t^3, t^2)$ and it is again a resolution of singularities.

3. Finally, let us go back to the original question of this lecture, and let us give a hint on how to compute the integral closure of $\mathbb{Z}$ in $L = \mathbb{Q}(\sqrt{d})$ where d is a square-free integer. We know now that this is a ring, and we denote it by $\mathbb{Z}_L$. The Galois group of $L/\mathbb{Q}$ is isomorphic to $\mathbb{Z}/2$. Now, we can write any $x \in L$ as $a+b\sqrt{d}$ for $a, b \in \mathbb{Q}$ and the non-trival element of $\mathrm{Gal}(L/\mathbb{Q})$ acts as $a + b\sqrt{d} \mapsto a - b\sqrt{d}$. To determine $\mathbb{Z}_L$ one proceeds in the following way: first, note that $\sqrt{d} \in \mathbb{Z}_L$ since it satisfies the monic equation $x^2 - d$. Therefore $\mathbb{Z}[\sqrt{d}] \subset \mathbb{Z}_L$ always. Next, note that if $x \in \mathbb{Z}_L$ then also $\bar{x} \in \mathbb{Z}_L$ (why?). From this it follows that if $x = a + b\sqrt{d} \in \mathbb{Z}_L$ then $2a = x + \bar{x} \in \mathbb{Z}_L$ and $a^2 - db^2 = x\bar{x} \in \mathbb{Z}_L$. Next, note that $2a, a^2 - db^2 \in \mathbb{Q}$, and therefore $2a, a^2 - db^2 \in \mathbb{Z}$. This gives us that if $x \in \mathbb{Z}_L$ then $2a \in \mathbb{Z}$. So, we check two cases. First, if $a \in \mathbb{Z}$ then $a^2 \in \mathbb{Z}$ too so that $db^2 \in \mathbb{Z}$. Since d is square-free, this forces $b \in \mathbb{Z}$ too. This means that if $a \in \mathbb{Z}$ then $x \in \mathbb{Z}[\sqrt{d}]$. Now, assume that $a = \alpha/2$ with $\alpha \in \mathbb{Z}$ odd, and let us look at the condition $\alpha^2/4 - db^2 \in \mathbb{Z}$, which means that $\alpha^2 - 4db^2 \in 4\mathbb{Z}$. If d is even, this cannot be solved for any $b \in \mathbb{Q}$ (why?). So if d is even we get that $\mathbb{Z}_L = \mathbb{Z}[\sqrt{d}]$. Now, assume that d is odd. In this case, we necessarily have $b = \beta/2$ with $\beta \in \mathbb{Z}$ an odd number (why?) and the condition becomes $\alpha^2 - d\beta^2 \in 4\mathbb{Z}$. The only squares in $\mathbb{Z}/4$ are $[0]$ and $[1]$. Since both α and β are odd, we have that $[\alpha^2] = [\beta^2] \mod 4$. Again, since d is odd, $[d] = [1]$ or $[d] = [3]$ in $\mathbb{Z}/4$, from which it follows that if there is a solution then necessarily $d \equiv 1 \mod 4$. So we have showed that if $d \equiv 3 \mod 4$ we have $\mathbb{Z}_L = \mathbb{Z}[\sqrt{d}]$ necessarily. So the last case to study is $d \equiv 1 \mod 4$. Check that in this case $\delta = (1 + \sqrt{d})/2 \in \mathbb{Z}_L$. Finally, if $x = \alpha/2 + \beta/2\sqrt{d} \in \mathbb{Z}_L$ (with α, β odd) then also $x - \delta \in \mathbb{Z}_L$. But then $x - \delta \in \mathbb{Z}[\sqrt{d}]$ (why?). This proves that if $d \equiv 1 \mod 4$ then the ring of integers of L is $\mathbb{Z}[\delta]$.