Math 55a, Fall 2004

Eleventh Assignment, Solutions
Adapted from Andrew Cotton and George Lee

Problem 1

(a) We start with a few observations:

• Any element $a \in K$ is algebraic and is the root of the monic irreducible $X - a \in K[X]$. From our work in the last assignment, we thus have that $\{p \in K[X] \mid p(a) = 0\}$ is the ideal generated by $X - a$. That is, $p(a) = 0$ if and only if $(X - a) \mid p$.

• Given $f_1, \ldots, f_k \in K[X]$, $(\sum_{i=1}^k f_i)' = \sum_{i=1}^k f_i'$. Also, given polynomials $\alpha = aX^i$ and $\beta = bX^j$, notice that $(\alpha\beta)' = (i + j)abX^{i+j} = \alpha\beta' + \alpha'\beta$. Then suppose we have two general polynomials $q = \sum_{i=0}^m a_i X^i$ and $r = \sum_{j=0}^n b_j X^j$. Write $\alpha_i = a_i X^i$ for $i = 0, 1, \ldots, m$ and $\beta_j = b_i X^j$ for $j = 0, 1, \ldots, n$; then

$$\begin{aligned}(pq)' &= \left(\sum_{i,j} \alpha_i\beta_j\right)' = \sum_{i,j} (\alpha_i\beta_j)' = \sum_{i,j} (\alpha_i\beta_j' + \alpha_i'\beta_j) \\ &= \sum_{i,j} \alpha_i\beta_j' + \sum_{i,j} \alpha_i'\beta_j = pq' + pq',\end{aligned}$$

the "product rule" of differentiation.

• Using the above result and induction, or using the binomial theorem, we find that for $n \geq 1$, $(X-a)^n$ has formal derivative $n(X-a)^{n-1}$.

• Suppose $p \in K[X] - \{0\}$ and $a \in K$. Let N be the largest nonnegative integer less than or equal to $\deg p$ such that $(X - a)^n \mid p$. One must exist because $(X - a)^0 \mid p$. Then for $m > N$, $(X - a)^m \not| \, p$, while $(X - a)^m \mid p$ for $m < N$. Hence p has a root of of order exactly n at a for $n = N$ but not for $n \neq N$ — that is, this n exists and is uniquely determined.

• If $p \in K[X]$ and $n = \deg p > 0$, we claim that p' is nonzero. Indeed, write $p = \sum_{i=0}^n a_i X^i$ where $a_n \neq 0$. Then $p' = \sum_{i=0}^{n-1} (i+1)a_{i+1}X^i$. Because $\mathrm{char}(K) = 0$, $n \neq 0$ so the coefficient na_n of X^{n-1} is nonzero. Therefore, $p' \neq 0$. Observe that this claim is *not* necessarily true if K has some nonzero characteristic κ, because then X^κ has positive degree but formal derivative 0.

We now prove that if $p \in K[X]$ vanishes to order exactly $n > 0$ at $a \in K$, then p' is a nonzero polynomial that vanishes to order exactly $n - 1$. Write $p = (X - a)^n r$ for some nonzero polynomial $r \in K[X]$ that does not vanish at a. Writing $q = (X - a)^n$, we have

$$\begin{aligned} p' &= q'r + qr' \\ &= n(X-a)^{n-1}r + (X-a)^n r' \\ &= (X-a)^{n-1}\underbrace{(nr + (X-a)r')}. \end{aligned}$$

$X - a$ divides $(X - a)r'$. But because $\mathrm{char}(K) = 0$, $n \neq 0$, so $(nr)(a) = nr(a) \neq 0$. It follows that $X - a$ does not divide nr or the underbraced quantity. Furthermore, as argued in the fifth initial observation, $p' \neq 0$. Hence p' is nonzero and vanishes to order exactly $n - 1$ at a, as claimed.

Now, if a polynomial q vanishes to order exactly $n > 0$ at a, then $(X - a)^n$ and thus $X - a$ divides q, so $q(a) = 0$. So suppose that p vanishes to order exactly $n \geq 0$. If $n = 0$ then $(X - a) \nmid p$ so that $p^{(n)}(a) = p(a) \neq 0$, as needed. Otherwise suppose that $n > 0$. Then by induction on k, $p^{(k)}$ vanishes to order exactly $n - k$ at a for $k = 0, 1, \ldots, n$. Thus, $p(a) = p'(a) = p^{(2)}(a) = \cdots = p^{(n-1)}(a) = 0$. However, $p^{(n)}$ is nonzero and vanishes to order exactly 0, implying that $(X - a) \nmid p^{(n)}$ and $p^{(n)}(a) \neq 0$.

To prove the other direction of the claim, suppose that $p(a) = p'(a) = \cdots = p^{(n-1)}(a) = 0$ but $p^{(n)}(a) \neq 0$. Then p must be nonzero because otherwise $p^{(n)}(a) = 0$. From our final initial observation, it vanishes to order exactly m for some nonnegative integer m. But from the previous paragraph, m is then the smallest nonnegative integer such that $p^{(m)}(a) \neq 0$; so it must equal n, as desired.

(b) Suppose we have a field K and an extension field L of K. Let d_1 be the greatest common divisor of $p, q \in K[X]$ when viewed as polynomials in $K[X]$; and let d_2 be their greatest common divisor when viewed as polynomials in $L[X]$. Then there exist $r, s \in K[X] \subset L[X]$ such that $pr + qs = d_1$; so since (in $L[X]$) we know $d_2 \mid pr$, $d_2 \mid qs$ we know that $d_2 \mid d_1$. As a corollary, if $d_1 = 1$ then d_2 must be constant; so if p, q are relatively prime in $K[X]$ then they must be relatively prime in $L[X]$.

If p is a constant polynomial the desired result is trivially true. Otherwise, since p' is nonzero and has degree one less than p, $p \nmid p'$; then since p is irreducible, p and p' are relatively prime in $K[X]$. Then calling the given extension field L, from our first paragraph p and p' are also relatively prime in $L[X]$.

But if p has a root of order at least 2 at a, from part (a) we know that $p'(a) = 0$. Then $X - a$ divides both p and p' in $L[X]$, so they are *not* relatively prime—a contradiction. Therefore p only has simple zeroes.

This proof fails if $\mathrm{char}(K) \neq 0$. Indeed, here is a sketch of several counterexamples to the desired result if we omit this requirement. Suppose that $\mathrm{char}(K) = \kappa$ and that there exists $b \in K$ such that $c^\kappa \neq b$ for all $c \in K$. (Note that κ must be prime (why?). Also, although $\mathbb{Z}/\kappa\mathbb{Z}$ does not have this property (why?), certain extension fields of it would (can you find one?).) We leave it up to the reader to then prove that $X^\kappa - b$ is irreducible in $K[X]$. But then consider any splitting field of $X^\kappa - b$ where this polynomial has root c. Then $(X - c)^\kappa = X^\kappa - c^\kappa = X^\kappa - b$, because all the intermediate terms in the expansion of $(X - c)^\kappa$ have coefficients divisible by κ and are thus 0. Therefore, $X^\kappa - b$ does *not* have only simple roots.

Problem 2

(a) First, $K[\alpha, \beta]$ contains both $K[\alpha]$ and β, so it contains $(K[\alpha])[\beta]$. And $(K[\alpha])[\beta]$ contains K, α, and β, so it contains $K[\alpha, \beta]$. Therefore $K[\alpha, \beta] = (K[\alpha])[\beta]$, and similarly

$$K[\alpha_1, \ldots, \alpha_n] = (((K[\alpha_1])\,[\alpha_2]) \cdots [\alpha_{n-1}])\,[\alpha_n]$$

(which we could also write $K[\alpha_1][\alpha_2] \cdots [\alpha_n]$).

Now suppose the theorem is true when $L = K[\alpha, \beta]$. Then if L is *any* finite extension of K, write $L = K[\alpha_1, \alpha_2, \ldots, \alpha_n]$ for some $\alpha_i \in L$. If $n \leq 2$ we are clearly done; otherwise (using induction) suppose the claim is true for all smaller n. Writing $K' = K[\alpha_1, \ldots, \alpha_{n-2}]$, we have $L = K'[\alpha_{n-1}, \alpha_n] = K'[\beta]$ for some $\beta \in L$. Then $L = K[\alpha_1, \ldots, \alpha_{n-2}, \beta]$; so by the induction hypothesis, we are done.

(b) Simply take the splitting field of p and consider the irreducible factors of q of degree 2 or more in this splitting field. Take the splitting field of one of these, and repeat until q splits.

(c) Because K has characteristic zero, it has infinitely many elements since $\mathbb{N}$ injects into K.

From Problem 1, all the β_j are distinct so that $\beta - \beta_j \neq 0$ for $2 \leq j \leq s$. Then there are finitely many values of the form $(\alpha_i - \alpha)(\beta - \beta_j)^{-1}$ (where $1 \leq i \leq r$, $2 \leq j \leq s$); so since K is infinite, there is some element in K not of this form. Let c be such an element; by construction,

$\alpha_i + c\beta_j \neq \alpha + c\beta$ for $1 \le i \le r$, $2 \le j \le s$.

(d) Write $p = \sum_{i=0}^n k_i X^i$ for $k_i \in K$. Then $\tilde{p} = \sum_{i=0}^n k_i(\zeta - cX)^i$. But

$$k_i(\zeta - cX)^i = \sum_{j=0}^{i} \binom{i}{j} k_i \zeta^{i-j}(-c)^j X^j.$$

(Here, $\binom{i}{j}$ is the unit 1 added to itself $\binom{i}{j}$ times.) But $k_i \in K \subset K[\zeta]$; $\zeta^{i-j} \in K[\zeta]$; and $c \in K[\zeta] \implies (-c)^j \in K[\zeta]$. Thus each $k_i(\zeta - cX)^i$ expands to some polynomial in $K[\zeta][X]$; so adding all such terms, $\tilde{p}$ is in $K[\zeta][X]$ as well.

Next, $\tilde{p}$ and q cannot be relatively prime in $K[\zeta][X]$ because as explained in (b), that would imply they are also relatively prime in $E[X]$. But this is impossible because in $E[X]$ we have $\tilde{p}(\beta) = q(\beta) = 0$ so that $X - \beta$ is a common divisor of both $\tilde{p}$ and q in $E[X]$.

Now suppose that r is the greatest common divisor of $\tilde{p}$ and q in $K[\zeta][X]$; we already know that $(X - \beta) \mid r$ in $E[X]$. Since $\tilde{p}$ and q split into linear factors over $E[X]$, so does r. So if r isn't a constant multiple of $X - \beta$ in $K[\zeta][X]$, then $(X - \beta_j) \mid r$ in $E[X]$ for some other $\beta_j \neq \beta$. But then $0 = \tilde{p}(\beta_j) = p(\zeta - c\beta_j)$, so $\zeta - c\beta_j = \alpha_i$ for some α_i; and by the choice of c, this is impossible.

Therefore r *is* a constant multiple of $X - \beta$, and $X - \beta$ is the greatest common divisor of $\tilde{p}$ and q in $K[\zeta][X]$.

(e) Since $X - \beta$ must be in $K[\zeta][X]$ from part (d), β must be in $K[\zeta]$. Therefore $K[\zeta]$ contains K, β, and also $\zeta - c\beta = \alpha$ — so $K[\zeta] \supset K[\alpha, \beta]$. Conversely, $K[\alpha, \beta]$ contains both K and $\alpha + c\beta = \zeta$ — so $K[\alpha, \beta] \supset K[\zeta]$. Therefore $K[\zeta] = K[\alpha, \beta]$, as desired.

(f) We used that $\mathrm{char}(K) = 0$ in part (c); the proof there fails because K might be finite if it has nonzero characteristic. As a conterexample, let $K = \mathbb{Z}_5$; $L = K[\sqrt[4]{2}]$; $p = q = X^4 - 2$; and $\alpha = \beta = \sqrt[4]{2}$. (The polynomial $X^4 - 2$ is irreducible in $\mathbb{Z}_5$ because it has no roots and thus no linear factors; and some algebra shows we can't split it into a product of two quadratics $(X^2 + aX + b)(X^2 + cX + d)$.) Then

$(\alpha, \alpha_2, \alpha_3, \alpha_4) = (\beta, \beta_2, \beta_3, \beta_4) = (\sqrt[4]{2}, 2\sqrt[4]{2}, 3\sqrt[4]{2}, 4\sqrt[4]{2})$; and

$$\begin{aligned}
\alpha + 0 \cdot \beta &= 1 \cdot \sqrt[4]{2} = \alpha_1 + 0 \cdot \beta_2, \\
\alpha + 1 \cdot \beta &= 2 \cdot \sqrt[4]{2} = \alpha_3 + 1 \cdot \beta_4, \\
\alpha + 2 \cdot \beta &= 3 \cdot \sqrt[4]{2} = \alpha_2 + 2 \cdot \beta_3, \\
\alpha + 3 \cdot \beta &= 4 \cdot \sqrt[4]{2} = \alpha_2 + 3 \cdot \beta_4, \\
\alpha + 4 \cdot \beta &= 0 \cdot \sqrt[4]{2} = \alpha_2 + 4 \cdot \beta_2.
\end{aligned}$$

Also, we used that $\mathrm{char}(K) = 0$ in problem 1, as noted there.

Problem 3

(a) Define $f : V^* \times W \to \mathrm{Hom}(V, W)$ by $f(\phi, w)(v) = \langle \phi, v \rangle w$. $f(\phi, w)$ is indeed a linear map from V to W since it is the composition of the two linear maps $v \mapsto \langle \phi, v \rangle$ (from V to K) and $k \mapsto kw$ (from K to W).

f is clearly canonical. Fix $v \in V$. Then fixing w, $\phi \mapsto f(\phi, w)(v)$ is the composition of the linear maps $\phi \mapsto \langle \phi, v \rangle$ (from V^* to K) and $k \mapsto kw$ (from K to W). Fixing ϕ, $w \mapsto f(\phi, w) = \langle \phi, v \rangle w$ is clearly linear. Therefore, $(\phi, w) \mapsto f(\phi, w)(v)$ is bilinear for all v, which in turn implies that f is bilinear.

Therefore, there exists a unique linear map $\psi : V^* \otimes W \to \mathrm{Hom}(V, W)$ corresponding to f; this is the canonical map we are looking for. It is nonzero because some $\phi \in V^*$ is nonzero, so that $f(\phi, w)$ and hence ψ are nonzero as well. (Also observe that $\psi(\phi \otimes w) = f(\phi, w)$ is the function $v \mapsto \langle \phi, v \rangle w$ for any $(\phi, w) \in V^* \times W$; we use this fact later.)

(b) The canonical map above is always injective; but it is an isomorphism if and only if V and W are *not* both infinite dimensional.

First we prove that ψ as given in part (a) is injective. Suppose that $\psi(a) = 0$ for some $a \in V^* \otimes W$. Pick bases for V^* and W and look at the corresponding basis for $V^* \otimes W$. We can thus write a as the sum $\sum_{i=1}^{m} \phi_i \otimes w_i$ where $\phi_i \in V^*$ and each w_i is in the basis of W; assume without loss of generality that the w_i are distinct (because otherwise we could combine the corresponding ϕ_i).

Then $\psi(a) = \sum_{i=1}^{m} \psi(\phi_i \otimes w_i)$ is the map $v \mapsto \sum_{i=1}^{m} \langle \phi_i, v \rangle w_i$. For $\psi(a)(v)$ to equal 0 we then must have $\langle \phi_i, v \rangle = 0$ for all i. But because this is true for all $v \in V$, we must have $\phi_i = 0$ for all i; and therefore $a = 0$.

Thus, $\mathrm{Ker}(\psi) = \{0\}$ and ψ is indeed injective.

Because ψ is an injective homomorphism, to prove it is an isomorphism it suffices to prove that ψ is surjective — that is, any $g \in \mathrm{Hom}(V, W)$ is in its image. Fix such a g. First suppose that V is finite dimensional with basis $\{v_1, v_2, \ldots, v_k\}$; let $\{v_1^*, v_2^*, \ldots, v_k^*\}$ be the corresponding dual basis. Since $v_j^*(v)$ is the "v_j-projection of v" — the coefficient of v_j when v is written as a linear combination of the v_i — we have $v = \sum_{i=1}^k \langle v_i^*, v\rangle v_i$. Then consider

$$a = \sum_{i=1}^{k} v_i^* \otimes g(v_i)$$

in $V^* \otimes W$. For any $v \in V$ we have

$$\psi(a)(v) = \sum_{i=1}^{k} \langle v_i^*, v\rangle g(v_i) = g\left(\sum_{i=1}^{k} \langle v_i^*, v\rangle v_i\right) = g(v),$$

and $\psi(a) = g$, as desired.

Next suppose that W is finite dimensional with basis $\{w_1, w_2, \ldots, w_k\}$ and corresponding dual basis $\{w_1^*, w_2^*, \ldots, w_k^*\}$. Again suppose $g \in \mathrm{Hom}(V, W)$. For each i, the map $w_i^* \circ g$ is a composition of two homomorphisms and hence is a member of V^*. Then setting

$$a = \sum_{i=1}^{k} (w_i^* \circ g) \otimes w_i$$

in $V^* \otimes W$, for any $v \in V$ we have

$$\psi(a)(v) = \sum_{i=1}^{k} \left(w_i^* \circ g\right)(v) w_i = \sum_{i=1}^{k} \langle w_i^*, g(v)\rangle w_i = g(v).$$

Hence, $\psi(a) = g$, as desired.

Therefore, if either V or W is finite dimensional then ψ is an isomorphism. Now suppose instead that V and W are both infinite dimensional with bases $\{v_\alpha \mid \alpha \in A\}$ and $\{w_\beta \mid \beta \in B\}$ respectively; let $\{v_1, v_2, \ldots\}$ and $\{w_1, w_2, \ldots\}$ be countable subsets of these bases. Then consider the homomorphism $g \in \mathrm{Hom}(V, W)$ that maps v_i to w_i for all $i \in \mathbb{N}$; and that maps all other v_α to, say, 0.

Now given any $a \in V^* \otimes W$, we can write $a = \sum_{i=1}^{k} \kappa_i \otimes \lambda_i$ for $(\kappa_i, \lambda_i) \in V^* \times W$. Regardless of the choice of $v \in V$, the value

$$\psi(a)(v) = \sum_{i=1}^{k} \langle \kappa_i, v \rangle \lambda_i$$

always lies in the span of the finitely many λ_i. But since $\{w_1, w_2, \ldots\}$ is infinite dimensional, at least one w_j is not in this span. So then $\psi(a)(v_j) \neq w_j$ so we cannot have $\psi(a) = g$ (since $g(v_j) = w_j$). Therefore g is not in the image of ψ; ψ is not surjective; and it is not an isomorphism.