Math 127
Suggested solutions to Homework Set 6

Problem 1. (a) Let a_1, a_2 be elements in A satisfying $f_1(a_1) = f_1(a_2)$; we have to show that $a_1 = a_2$. But, if $f_1(a_1) = f_1(a_2)$, then $g_1\big(f_1(a_1)\big) = g_1\big(f_1(a_2)\big)$, and hence

$$a_1 = \mathrm{id}_A(a_1) = (g_1 \circ f_1)(a_1) = (g_1 \circ f_1)(a_2) = \mathrm{id}_A(a_2) = a_2.$$

Since a_1, a_2 are arbitrary, this shows that f_1 is injective.

(b) Let $b \in B$; we have to show that there is $a \in A$ such that $b = f_2(a)$. We have that

$$b = \mathrm{id}_B(b) = (f_2 \circ h_2)(b) = f_2(h_2(b)),$$

therefore $h_2(b) \in A$ is a preimage of b under f_2.

Since $b \in B$ was arbitrary, this shows that f_2 is surjective.

Problem 2. (i) Let a be an element of $\mathbb{F}_1$. We have that $a = a +_1 0_{\mathbb{F}_1}$, and hence

$$f(a) = f\big(a +_1 0_{\mathbb{F}_1}\big) = f(a) +_2 f\big(0_{\mathbb{F}_1}\big),$$

where the second equality follows by the 1st condition of the definition of a field homomorphism.

In other words we have

$$f(a) +_2 0_{\mathbb{F}_2} = f(a) +_2 f\big(0_{\mathbb{F}_1}\big).$$

By the Cancellation Law for addition in $\mathbb{F}_2$ (recall e.g. HW1, Problem 2), we conclude that

$$0_{\mathbb{F}_2} = f\big(0_{\mathbb{F}_1}\big).$$

(ii) Let a be an arbitrary element of $\mathbb{F}_1$. By part (i), we can write

$$f(a +_1 (-a)) = f\big(0_{\mathbb{F}_1}\big) = 0_{\mathbb{F}_2}.$$

At the same time, $f(a +_1 (-a)) = f(a) +_2 f(-a)$. Thus, we have

$$f(a) +_2 f(-a) = 0_{\mathbb{F}_2} = f(a) +_2 (-f(a)),$$

where the second equality holds because $-f(a)$ is the additive inverse of $f(a)$ in $\mathbb{F}_2$.

Again, by the Cancellation Law for addition in $\mathbb{F}_2$, we can conclude that $f(-a) = -f(a)$.

Since $a \in \mathbb{F}_1$ was arbitrary, this holds true for every $a \in \mathbb{F}_1$.

(iii) Let a be an arbitrary **non-zero** element of $\mathbb{F}_1$. Then we know that a has a multiplicative inverse a^{-1}, and hence we can write

$$f(a) \cdot_2 f(a^{-1}) = f\big(a \cdot_1 a^{-1}\big) = f\big(1_{\mathbb{F}_1}\big) = 1_{\mathbb{F}_2}.$$

We now observe that, since $f(a) \cdot_2 f(a^{-1}) = 1_{\mathbb{F}_2} \neq 0_{\mathbb{F}_2}$, $f(a)$ **cannot** be a zero element of $\mathbb{F}_2$.

From this we see that $f(a)$ has a multiplicative inverse $(f(a))^{-1}$ in $\mathbb{F}_2$, and hence we can write

$$f(a) \cdot_2 (f(a))^{-1} = 1_{\mathbb{F}_2} = f(a) \cdot_2 f(a^{-1}).$$

By the Cancellation Law for multiplication in $\mathbb{F}_2$ (again, recall HW1, Problem 2), we can conclude that $(f(a))^{-1} = f(a^{-1})$.

Since $a \in \mathbb{F}_1 \setminus \{0_{\mathbb{F}_1}\}$ was arbitrary, this holds true for every $a \in \mathbb{F}_1 \setminus \{0_{\mathbb{F}_1}\}$.

Problem 3. Given that μ is an eigenvalue of A, we can find a non-zero vector $\bar{u} \in \mathbb{F}^n$ such that

$$A\bar{u} = \mu \cdot \bar{u}.$$

If we now multiply both sides of this equation by A^{-1}, we obtain

$$\bar{u} = A^{-1}(A\bar{u}) = A^{-1}(\mu \cdot \bar{u}) = \mu \cdot (A^{-1}\bar{u}) \qquad \Rightarrow \qquad A^{-1}\bar{u} = \tfrac{1}{\mu} \cdot \bar{u}.$$

This shows that the vector $\bar{u}$ is an eigenvector of A^{-1} as well, and that it corresponds to eigenvalue μ^{-1}. In other words, μ^{-1} is an eigenvalue of A^{-1}.

Problem 4. (i) By definition we have that

$$\operatorname{span}(u,v,w) = \{\lambda_1 u + \lambda_2 v + \lambda_3 w : \lambda_1, \lambda_2, \lambda_3 \in \mathbb{F}\}$$
$$\text{and } \operatorname{span}(u-v,v,w) = \{\mu_1(u-v) + \mu_2 v + \mu_3 w : \mu_1, \mu_2, \mu_3 \in \mathbb{F}\}.$$

We will show that

$$\operatorname{span}(u,v,w) \subseteq \operatorname{span}(u-v,v,w) \quad \text{and} \quad \operatorname{span}(u-v,v,w) \subseteq \operatorname{span}(u,v,w).$$

To prove the first inclusion, consider a vector $z \in \operatorname{span}(u,v,w)$. Then there exist $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{F}$ so that

$$z = \lambda_1 u + \lambda_2 v + \lambda_3 w.$$

We need to find $\mu_1, \mu_2, \mu_3 \in \mathbb{F}$ so that z can be written as $\mu_1(u-v)+\mu_2 v+\mu_3 w$ too.

If we set $\mu_1 = \lambda_1$, $\mu_2 = \lambda_1 + \lambda_2$ and $\mu_3 = \lambda_3$, then we will have

$$\begin{aligned}
\mu_1(u-v) + \mu_2 v + \mu_3 w &= \lambda_1(u-v) + (\lambda_1 + \lambda_2)v + \lambda_3 w \\
&= \lambda_1 u - \lambda_1 v + \lambda_1 v + \lambda_2 v + \lambda_3 w \\
&= \lambda_1 u + \lambda_2 v + \lambda_3 w = z.
\end{aligned}$$

This shows that $z \in \operatorname{span}(u-v,v,w)$. Since we started with an arbitrary $z \in \operatorname{span}(u,v,w)$, we have proven the first inclusion.

To prove the second inclusion, we similarly start by considering a vector $z' \in \operatorname{span}(u-v,v,w)$. Then there exist $\mu_1', \mu_2', \mu_3' \in \mathbb{F}$ so that

$$z' = \mu_1'(u-v) + \mu_2' v + \mu_3' w.$$

We need to find $\lambda_1', \lambda_2', \lambda_3' \in \mathbb{F}$ so that z' can be written as $\lambda_1' u + \lambda_2' v + \lambda_3' w$ too.

If we set $\lambda_1' = \mu_1'$, $\lambda_1' = \mu_2' - \mu_1'$ and $\lambda_3' = \mu_3'$, then we have

$$\begin{aligned}
\lambda_1' u + \lambda_2' v + \lambda_3' w &= \mu_1' u + (\mu_2' - \mu_1')v + \mu_3' w \\
&= \mu_1' u + \mu_2' v - \mu_1' v + \mu_3' w \\
&= \mu_1' u - \mu_1' v + \mu_2' v + \mu_3' w \\
&= \mu_1'(u-v) + \mu_2' v + \mu_3' w = z'.
\end{aligned}$$

This shows that $z' \in \operatorname{span}(u,v,w)$. Since we started with an arbitrary $z' \in \operatorname{span}(u-v,v,w)$, we have shown the second inclusion too.

Combining the two, we get the equality of the two spans.

(ii) Consider the vector space $\mathbb{R}^3$ over $\mathbb{R}$, and set $u = \bar{e}_1$, $v = \bar{e}_2$ and $w = \bar{e}_3$.

Then $\operatorname{span}(u, v, w) = \mathbb{R}^3$. On the other hand, $\operatorname{span}(u - v, v - u, w) = \operatorname{span}(\bar{e}_1 - \bar{e}_2, \bar{e}_2 - \bar{e}_1, \bar{e}_3)$, so every vector $\bar{x} = \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right)$ in the latter span satisfies $x_1 = -x_2$. This shows that $\bar{e}_1 \notin \operatorname{span}(u - v, v - u, w)$, and thus we obtain that

$$\operatorname{span}(\bar{e}_1, \bar{e}_2, \bar{e}_3) \neq \operatorname{span}(\bar{e}_1 - \bar{e}_2, \bar{e}_2 - \bar{e}_1, \bar{e}_3).$$

(iii) The equality is <u>**not**</u> always true.

To verify this, let us consider again the vector space $\mathbb{R}^3$ over $\mathbb{R}$, and let's set $u = \bar{e}_1$, $v = \bar{e}_1 + \bar{e}_2$ and $w = \bar{e}_1 + \bar{e}_2 + \bar{e}_3$.

Then we can check that all the standard basis vectors of $\mathbb{R}^3$ are in $\operatorname{span}(u, v, w) = \operatorname{span}(\bar{e}_1,\ \bar{e}_1 + \bar{e}_2,\ \bar{e}_1 + \bar{e}_2 + \bar{e}_3)$, and thus $\operatorname{span}(\bar{e}_1,\ \bar{e}_1 + \bar{e}_2,\ \bar{e}_1 + \bar{e}_2 + \bar{e}_3) = \mathbb{R}^3$.

On the other hand, $\operatorname{span}(u - v, v - w, w - u) = \operatorname{span}(-\bar{e}_2,\ -\bar{e}_3,\ \bar{e}_2 + \bar{e}_3)$. Therefore every vector $\bar{y} = \left(\begin{smallmatrix} y_1 \\ y_2 \\ y_3 \end{smallmatrix}\right)$ in this span will have first coordinate $y_1 = 0$. This shows that $\bar{e}_1$ cannot be in this span, and hence

$$\operatorname{span}(\bar{e}_1,\ \bar{e}_1 + \bar{e}_2,\ \bar{e}_1 + \bar{e}_2 + \bar{e}_3) \neq \operatorname{span}(-\bar{e}_2,\ -\bar{e}_3,\ \bar{e}_2 + \bar{e}_3).$$

Problem 5. (a) We show that $\mathcal{P}$ satisfies all the axioms of a vector space over $\mathbb{R}$.

Addition is commutative: Consider two polynomials $p(x)$ and $q(x)$ in $\mathcal{P}$. Then we can find (a large enough) $m \in \mathbb{N}$ such that we can write

$$\begin{aligned} p(x) &= a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \\ \text{and} \quad q(x) &= b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m \end{aligned}$$

for some real coefficients $a_0, a_1, a_2, \ldots, a_m,\ b_0, b_1, b_2, \ldots, b_m$.

But then we have

$$\begin{aligned} p(x) + q(x) &= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_m + b_m)x^m \\ &= (b_0 + a_0) + (b_1 + a_1)x + (b_2 + a_2)x^2 + \cdots + (b_m + a_m)x^m \\ &= q(x) + p(x), \end{aligned}$$

where we used the commutativity of addition in $\mathbb{R}$.

Since $p(x), q(x)$ were arbitrary elements of $\mathcal{P}$, we conclude that addition is commutative.

Addition is associative: Consider three polynomials $p(x), q(x), u(x)$ in $\mathcal{P}$. As before, we can find (a large enough) $m \in \mathbb{N}$ such that we can write

$$\begin{aligned} p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m, \qquad q(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m \\ \text{and} \quad u(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_m x^m \end{aligned}$$

for some real coefficients $a_0, a_1, a_2, \ldots, a_m,\ b_0, b_1, b_2, \ldots, b_m,\ c_0, c_2, c_2, \ldots, c_m$.
But then we have

$$\begin{aligned} &(p(x) + q(x)) + u(x) \\ &= \big((a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_m + b_m)x^m\big) + \big(c_0 + c_1 x + c_2 x^2 + \cdots + c_m x^m\big) \\ &= ((a_0 + b_0) + c_0) + ((a_1 + b_1) + c_1)x + ((a_2 + b_2) + c_2)x^2 + \cdots + ((a_m + b_m) + c_m)x^m \\ &= (a_0 + (b_0 + c_0)) + (a_1 + (b_1 + c_1))x + (a_2 + (b_2 + c_2))x^2 + \cdots + (a_m + (b_m + c_m))x^m \\ &= p(x) + (q(x) + u(x)), \end{aligned}$$

where we used the associativity of addition in $\mathbb{R}$.

Since $p(x), q(x)$ and $u(x)$ were arbitrary elements of $\mathcal{P}$, we conclude that addition is associative.

Neutral element of addition: We check that the zero polynomial $\mathbf{0}$ is the neutral element of addition.

Let $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m \in \mathcal{P}$ be an arbitrary real polynomial. Then

$$\begin{aligned}\mathbf{0} + p(x) &= (0 + 0x + 0x^2 + \cdots + 0x^m) + (a_0 + a_1x + a_2x^2 + \cdots + a_mx^m)\\ &= (0 + a_0) + (0 + a_1)x + (0 + a_2)x^2 + \cdots + (0 + a_m)x^m\\ &= a_0 + a_1x + a_2x^2 + \cdots + a_mx^m = p(x).\end{aligned}$$

Completely analogously we see that $p(x) + \mathbf{0} = p(x)$.

Since $p(x) \in \mathcal{P}$ was arbitrary, the conclusion follows.

Additive inverses: Let $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m \in \mathcal{P}$ be an arbitrary real polynomial. We show that $p(x)$ has an additive inverse. In fact,

$$(-p)(x) = (-a_0) + (-a_1)x + (-a_2)x^2 + \cdots + (-a_m)x^m,$$

where $-a_i$ is the additive inverse in $\mathbb{R}$ of the coefficient a_i. Indeed,

$$\begin{aligned}&p(x) + \big((-a_0) + (-a_1)x + (-a_2)x^2 + \cdots + (-a_m)x^m\big) =\\ &\qquad (a_0 - a_0) + (a_1 - a_1)x + (a_2 - a_2)x^2 + \cdots + (a_m - a_m)x^m = \mathbf{0}.\end{aligned}$$

Multiplicative identity of $\mathbb{R}$ and scalar multiplication: Let $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m \in \mathcal{P}$ be an arbitrary real polynomial. We have that $1_{\mathbb{R}} \cdot p(x) = p(x)$.

Indeed,

$$\begin{aligned}1_{\mathbb{R}} \cdot p(x) &= 1_{\mathbb{R}} \cdot \big(a_0 + a_1x + a_2x^2 + \cdots + a_mx^m\big)\\ &= (1_{\mathbb{R}}a_0) + (1_{\mathbb{R}}a_1)x + (1_{\mathbb{R}}a_2)x^2 + \cdots + (1_{\mathbb{R}}a_m)x^m\\ &= a_0 + a_1x + a_2x^2 + \cdots + a_mx^m\end{aligned}$$

since $1_{\mathbb{R}}$ is the multiplicative identity in $\mathbb{R}$.

Associativity of scalar multiplication: Let $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m \in \mathcal{P}$ be an arbitrary real polynomial, and let $r, s \in \mathbb{R}$ be arbitrary scalars.

We have that $r \cdot (s \cdot p(x)) = (rs) \cdot p(x)$. Indeed,

$$\begin{aligned}r \cdot (s \cdot p(x)) &= r \cdot \big((sa_0) + (sa_1)x + (sa_2)x^2 + \cdots + (sa_m)x^m\big)\\ &= (r(sa_0)) + (r(sa_1))x + (r(sa_2))x^2 + \cdots + (r(sa_m))x^m\\ &= ((rs)a_0) + ((rs)a_1)x + ((rs)a_2)x^2 + \cdots + ((rs)a_m)x^m\\ &= (rs) \cdot p(x),\end{aligned}$$

where we used the associativity of multiplication in $\mathbb{R}$ to go from the second line to the third one.

Since $p(x) \in \mathcal{P}$ and $r, s \in \mathbb{R}$ were arbitrary, the conclusion follows.

Scalar multiplication distributes over vector addition: Let $p(x), q(x) \in \mathcal{P}$ be arbitrary real polynomials, and let $r \in \mathbb{R}$ an arbitrary scalar.

We can find $m \in \mathbb{N}$ such that

$$\begin{aligned} p(x) &= a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \\ \text{and} \quad q(x) &= b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m \end{aligned}$$

for some real coefficients $a_0, a_1, a_2, \ldots, a_m,\ b_0, b_1, b_2, \ldots, b_m$.

Then we have

$$\begin{aligned} & r \cdot (p(x) + q(x)) \\ &= r \cdot \big((a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_m + b_m)x^m\big) \\ &= (r(a_0 + b_0)) + (r(a_1 + b_1))x + (r(a_2 + b_2))x^2 + \cdots + (r(a_m + b_m))x^m \\ &= (ra_0 + rb_0) + (ra_1 + rb_1)x + (ra_2 + rb_2)x^2 + \cdots + (ra_m + rb_m)x^m \\ &= \big((ra_0) + (ra_1)x + (ra_2)x^2 + \cdots + (ra_m)x^m\big) + \big((rb_0) + (rb_1)x + (rb_2)x^2 + \cdots + (rb_m)x^m\big) \\ &= r \cdot p(x) + r \cdot q(x), \end{aligned}$$

where we used the left distributive law in $\mathbb{R}$ to go from the third line to the fourth one.

Since $p(x), q(x) \in \mathcal{P}$ and $r \in \mathbb{R}$ were arbitrary, the conclusion follows.

Scalar multiplication distributes over scalar addition: Let $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \in \mathcal{P}$ be an arbitrary real polynomial, and let $r, s \in \mathbb{R}$ be arbitrary scalars.

We have that $(r + s) \cdot p(x) = r \cdot p(x) + s \cdot p(x)$. Indeed,

$$\begin{aligned} & (r + s) \cdot p(x) \\ &= (r + s) \cdot \big(a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m\big) \\ &= ((r + s)a_0) + ((r + s)a_1)x + ((r + s)a_2)x^2 + \cdots + ((r + s)a_m)x^m \\ &= (ra_0 + sa_0) + (ra_1 + sa_1)x + (ra_2 + sa_2)x^2 + \cdots + (ra_m + sa_m)x^m \\ &= \big((ra_0) + (ra_1)x + (ra_2)x^2 + \cdots + (ra_m)x^m\big) + \big((sa_0) + (sa_1)x + (sa_2)x^2 + \cdots + (sa_m)x^m\big) \\ &= r \cdot p(x) + s \cdot p(x), \end{aligned}$$

where we used the right distributive law in $\mathbb{R}$ to go from the third line to the fourth one.

Since $p(x) \in \mathcal{P}$ and $r, s \in \mathbb{R}$ were arbitrary, the conclusion follows.

Combining all the above, we conclude that $\mathcal{P}$ is a vector space over $\mathbb{R}$.

(b) According to one of the equivalent definitions of the notion of 'dimension' that we gave, the dimension of $\mathcal{P}$ over $\mathbb{R}$ is equal to the largest possible cardinality of a linearly independent subset of $\mathcal{P}$.

Thus, if we show that, for every positive integer n, there exists a linearly independent subset of $\mathcal{P}$ with cardinality **larger than n**, we will be able to conclude that

$$\dim_{\mathbb{R}} \mathcal{P} > n \quad \text{for all } n \in \mathbb{N},$$

and hence the dimension of $\mathcal{P}$ over $\mathbb{R}$ is infinite.

Let us consider an arbitrary positive integer n. Moreover, consider the subset $\{1, x, x^2, \ldots, x^{n-1}, x^n\}$ of $\mathcal{P}$. Clearly this subset has cardinality larger than n, since it contains $n+1$ different elements (in particular, $n+1$ monomials).

Suppose that $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}, \lambda_n$ are real coefficients for which we have

$$\lambda_0 \cdot 1 + \lambda_1 x + \lambda_2 x^2 + \cdots + \lambda_{n-1} x^{n-1} + \lambda_n x^n = \mathbf{0}.$$

If we assumed that not all of these coefficients are zero, then we can find the largest index $k \in \{0, 1, 2, \ldots, n-1, n\}$ such that $\lambda_k \neq 0$. Recall that the polynomial

$$p(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2 + \cdots + \lambda_{n-1} x^{n-1} + \lambda_n x^n$$

will have degree k, and as we have seen it will have at most k real roots (that is, at most k real numbers a will satisfy $p(a) = 0$). But this shows that, in this case, $p(x)$ cannot be equal to the zero polynomial, which has infinitely many roots.

We conclude that, if we have

$$p(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2 + \cdots + \lambda_{n-1} x^{n-1} + \lambda_n x^n = \mathbf{0},$$

then necessarily $\lambda_0 = \lambda_1 = \cdots = \lambda_{n-1} = \lambda_n = 0$. This gives that the set $\{1, x, x^2, \ldots, x^{n-1}, x^n\}$ is a linearly independent subset of $\mathcal{P}$, and hence $\dim_{\mathbb{R}} \mathcal{P} \geqslant n+1$.

Given that we started with an arbitrary $n \in \mathbb{N}$, the desired conclusion follows.