MA 724
Homework 1, Comments and some solutions.

Problem 0

It is enough to prove the following:

Claim 1. *If $\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5 \in \mathbb{R}^3$ are not coplanar, then there exist $i \neq j \in [5]$ such that the line segment $[\vec{x}_i, \vec{x}_j]$ is an edge of $\text{conv}(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5)$.*

One proof. If the claim fails, then we have a 3-polytope with 5 vertices, all pairs of which form an edge. The boundary of the polytope is homeomorphic to a sphere, and the vertices and edges constitute an embedding of the complete graph with five vertices into the sphere. The edges in this embedding don't cross, because the intersection of two edges would be a vertex. However, the complete graph on five vertices has no embedding, without crossings, into the sphere. □

If you aren't familiar with the result that the complete graph is not planar, then you can take a different route.

Sketch of another proof. If $\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5 \in \mathbb{R}^3$ are not coplanar, then without loss of generality $\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4$ are also not coplanar, so their convex hull is a 3-simplex (tetrahedron). Each of the four facets of the tetrahedron is contained in some hyperplane. For each of these hyperplane, we will say a point is "inside" the hyperplane if it is on the same side of the hyperplane as the tetrahedron is.

We want to show that, no matter where $\vec{x}_5$ is, the claim holds. Up to symmetry, we need only consider how many of the four planes x_5 is inside. If it's inside all four, the claim is easy. If it's inside three of them, $\text{conv}(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5)$ is a bipyramid over a triangle, so the claim holds. If it's inside two of them, name the hyperplanes it is **outside** as H_1 and H_2. Then the intersection $H_1 \cap H_2 \cap \text{conv}(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5)$ is an edge of the tetrahedron that is in the interior of $\text{conv}(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5)$. If it's inside one of them, the the three hyperplanes it is **outside** intersect $\text{conv}(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5)$ in a vertex, without loss of generality $\vec{x}_1$, and the line segment $[\vec{x}_1, \vec{x}_5]$ is in the interior of $\text{conv}(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5)$. It's impossible for a point to be outside of all four hyperplanes. □

Problem 5

I will do this in a way that looks forward to the more explicit definition of faces that we will soon have. For each nonempty face F of a polytope P, there exists one or more linear functionals whose maximum on P is attained on precisely the set F. Conversely, given a linear functional f, the subset of P where f attains its maximum on P is a face of P. (Why is this definition of faces as sets maximizing linear functionals the same as the "supporting halfspaces" definition from class?)

A linear functional is a row vector $\vec{c} = [c_1, c_2, \dots, c_d]$, interpreted as a function from $\mathbb{R}^d$ to $\mathbb{R}$ given by matrix multiplication $\vec{x} \mapsto \vec{c}\vec{x}$. The set of points in the cube that maximize $\vec{c}$ is the set of points $\left\{\vec{x} : x_i = \frac{c_i}{|c_i|} \text{ when } c_i \neq 0 \text{ and } -1 \leq x_i \leq 1 \text{ when } c_i = 0\right\}$. To find faces, we may as well replace each nonzero c_i by $c_i/|c_i|$, so that $\vec{c} \in \left(\{-1,0,1\}^d\right)^*$. (Ziegler didn't bother with the * here, but I've got it since c is a row vector.) Each choice of $\vec{c}$ gives a different set of maximum points. This puts the nonempty faces of C_d in bijection with the vectors $\vec{c} \in \left(\{-1,0,1\}^d\right)^*$. Specifically, the bijection sends $\vec{c}$ to $\{\vec{x} \in C_d : x_i = -1 \text{ if } c_i = -1 \text{ and } x_i = 1 \text{ if } c_i = 1\}$.

We also answered the optimization problem. For the separation problem, a point $\vec{y}$ is not in C_d if and only if it violates some inequality $y_i \leq 1$ or $y_i \geq -1$. We can see which by just examining the coordinates of $\vec{y}$. (**Note:** The separation asks for a *linear* inequality, so $|y_i| \leq 1$ is not the answer. The book should have asked you to give a *linear* inequality that is violated by $\vec{y}$.)

You did not have to write anything about other polytopes, but I hope you thought about it. The main general ideas I want you to take away from this problem are:

(1) The separation problem is easy when you know a small set of inequalities defining the polytope (We haven't discussed this yet, but a smallest possible set of inequalities has exactly one for each facet); and

(2) For any linear functional, there is always a vertex (or more than one) that obtains the maximum. (We'll see when we consider faces more carefully.) Thus, the optimization problem is easy when you have a short list of all vertices, and their coordinates.

By the way, optimization is easy on a cyclic polytope: To maximize $[c_1\, c_2\, \cdots\, c_d]$, you just have to evaluate $c_1 t + c_2 t^2 + \cdots + c_d t^d$ for each of the values $t_1, \ldots, t_n$ defining $C_d(t_1, \ldots, t_n)$. Separation may also be not too bad: You can use Gale's Evenness Criterion to find formulas for all of the facet-defining inequalities. (It's that polynomial that came from the determinant, the one we used to prove the Criterion.)

Additional problem 1

We need to show that the set

$$\operatorname{conv}(K) = \left\{\lambda_1 \vec{x}_1 + \cdots + \lambda_k \vec{x}_k :\ \{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq K,\ \lambda_i \geq 0,\ \sum_{i=1}^{k} \lambda_i = 1\right\}$$

is convex. Take $\vec{x} = \lambda_1 \vec{x}_1 + \cdots + \lambda_n \vec{x}_n$ and $\vec{y} = \mu_1 \vec{y}_1 + \cdots + \mu_m \vec{y}_m$ in $\operatorname{conv}(K)$ (and in particular, take each $\vec{x}_i$ and $\vec{y}_i$ in K) and let $0 \leq \nu \leq 1$. We need to show that $\nu\vec{x} + (1-\nu)\vec{y} \in \operatorname{conv}(K)$. But this is just $(\nu\lambda_1)\vec{x}_1 + \cdots + (\nu\lambda_n)\vec{x}_n + (1-\nu)\mu_1\vec{y}_1 + \cdots + (1-\nu)\mu_n\vec{x}_n$. Each of these coefficients is between 0 and 1, and it's easy to see that they sum to 1, so $\nu\vec{x} + (1-\nu)\vec{y} \in \operatorname{conv}(K)$.

Comment 1: You have to start with two completely general points in the set. Specifically, you can't assume immediately that $\vec{x} = \lambda_1 \vec{x}_1 + \cdots + \lambda_n \vec{x}_n$ and $\vec{y} = \mu_1 \vec{x}_1 + \cdots + \mu_n \vec{x}_n$ for the same set $\{x_1, \ldots, x_n\}$. (You can, however, write $\vec{x}$ as a convex combination of vectors in $\{x_1, \ldots, x_n\} \cup \{y_1, \ldots, y_m\}$, but then you would probably finish the argument differently, in a way that relates to Comment 2.)

Comment 2: If $\{x_1, \ldots, x_n\} \cap \{y_1, \ldots, y_m\}$ is nonempty, then the expression $(\nu\lambda_1)\vec{x}_1 + \cdots + (\nu\lambda_n)\vec{x}_n + (1-\nu)\mu_1\vec{y}_1 + \cdots + (1-\nu)\mu_n\vec{x}_n$ is not really a convex combination of vectors, because some vectors appear twice. But this is easily fixed by "combining like terms", and that doesn't hurt the fact that the coefficients sum to 1.

Additional problem 2

a. If $\vec{z} \in S$, then $\vec{z} = \vec{x} + \vec{w}$ for some $\vec{w} \in L$. Since $\vec{y} \in S$, we also have $\vec{y} = \vec{x} + \vec{u}$. Thus $\vec{z} = (\vec{y} - \vec{u}) + \vec{w} = \vec{y} + (-\vec{u} + \vec{w}) \in \vec{y} + L$. Thus $\vec{x} + L \subseteq \vec{y} + L$. In particular, since $\vec{x} = \vec{x} + \vec{0} \in \vec{x} + L$, we have $\vec{x} \in \vec{y} + L$, so the symmetric argument shows that $\vec{y} + L \subseteq \vec{x} + L$.

b. Given $\{\vec{x}_1, \ldots, \vec{x}_n\}$, we can write $\operatorname{aff}(\{\vec{x}_1, \ldots, \vec{x}_n\}) = \vec{v} + L$ for a vector $\vec{v}$ and linear subspace L. Since $\vec{x}_n \in \operatorname{aff}(\{\vec{x}_1, \ldots, \vec{x}_n\})$, part a says that $\operatorname{aff}(\{\vec{x}_1, \ldots, \vec{x}_n\}) = \vec{x}_n + L$.

The property of being an affine subspace is preserved by translation (i.e. adding a fixed vector), so if we add $-\vec{x}_n$ thoughout, we see that $\operatorname{aff}(\left\{\vec{x}_1 - \vec{x}_n, \ldots, \vec{x}_{n-1} - \vec{x}_n, \vec{0}\right\}) = L$.

Now, recall that $\operatorname{aff}(K)$ is the smallest affine subspace containing a set K, and recall also that every linear subspace is in particular an affine subspace. Since L is a linear subspace and also the smallest affine subspace containing $\left\{\vec{x}_1 - \vec{x}_n, \ldots, \vec{x}_{n-1} - \vec{x}_n, \vec{0}\right\}$, we conclude that L is the smallest *linear* subspace containing $\left\{\vec{x}_1 - \vec{x}_n, \ldots, \vec{x}_{n-1} - \vec{x}_n, \vec{0}\right\}$. Thus L has dimension $\leq n-1$, so $\vec{v} + L$ has dimension $\leq n-1$.

c. Recall that a set $\{\vec{x}_1, \ldots, \vec{x}_n\}$ is affinely independent if and only if its affine hull is $(n-1)$-dimensional. Thus, continuing from the argument in part b, (and leaving out the unnecessary $\vec{0}$ in the set), we see that $\{\vec{x}_1, \ldots, \vec{x}_n\}$ is affinely independent if and only if $\{\vec{x}_1 - \vec{x}_n, \ldots, \vec{x}_{n-1} - \vec{x}_n\}$ is linearly independent.

The set $\{\vec{x}_1 - \vec{x}_n, \ldots, \vec{x}_{n-1} - \vec{x}_n\}$ is linearly ***de***pendent if and only if there exist constants $c_1, \ldots, c_{n-1}$, *not all zero*, such that $\sum_{i=1}^{n-1} c_i(\vec{x}_i - \vec{x}_n) = \vec{0}$. Equivalently, if and only if there exist constants $c_1, \ldots, c_{n-1}$, *not all zero*, such that $\sum_{i=1}^{n-1} c_i \vec{x}_i + (-\sum_{i=1}^{n-1} c_i)\vec{x}_n = \vec{0}$.

We see that this is equivalent to the existence of constants $c_1, \ldots, c_n$, *not all zero*, such that $\sum_{i=1}^{n} c_i \vec{x}_i = \vec{0}$ and $\sum_{i=1}^{n} c_i = \vec{0}$. (Given constants $c_1, \ldots, c_{n-1}$ as in the previous paragraph, set $c_n = -\sum_{i=1}^{n-1} c_i$. Given constants $d_1, \ldots, d_n$, *not all zero*, such that $\sum_{i=1}^{n} d_i \vec{x}_i = \vec{0}$ and $\sum_{i=1}^{n} d_i = \vec{0}$, note that $d_n = -\sum_{i=1}^{n-1} d_i$ and thus that $d_1, \ldots, d_{n-1}$ are also not all zero.)