RÓBERT FREUD

Linear Algebra: A Problem-Centered Approach
Hints to Selected Exercises

The Reader is advised to consult the hint only if the exercise turns out to be absolutely unmanageable. We suggest rather to return to the same problem later or to solve first some special case of it. Consulting a hint before attempting a problem intensively has a very limited pedagogical effect.

When working on an exercise, it is important to unravel its message and background, to reveal its position and role in the mathematical landscape. Thinking about a generalization or raising new problems is very useful (even if it is not clear how to solve them).

Instead of just thinking in your head about the exercises, it is important to write down their solutions. Working through the details clarifies the argument, reveals its most important features, and helps to identify any errors or gaps.

Note that there can be several other correct approaches that differ significantly from the one suggested by the hint. In this case, however, it is worth checking the validity of the argument very carefully.

CONTENTS

1. Determinants

1.1.

1.1.1 (a) The number of inversions is minimal in the natural order $1, 2, 3, \ldots, n$ and is maximal in the reverse order $n, n-1, \ldots, 2, 1$.
(b) Verify that we can arrive from the natural order $1, 2, 3, \ldots, n$ at the reverse order $n, n-1, \ldots, 1$ by repeated swapping of adjacent elements. The number of inversions changes by one in each step, so it has to assume every integer between 0 and $\binom{n}{2}$.

1.1.8 (a) Group the permutations according to the position of 1.
(b) Combine (a) for k and $k-1$.
(e) Use the pigeonhole principle.

1.2.

1.2.5 In the given k rows, we can choose non-zero elements only from $n-m < k$ columns into the products in the definition the determinant.

1.2.7 The determinant can be written as $D = \alpha_{11}\beta + \gamma$ if we factor out α_{11} from the products containing it in the sum defining D. If $\beta \neq 0$, then replacing α_{11} by $-\gamma/\beta$ works. Otherwise proceed similarly with another element in the first row. If all elements in the first row are 0, then $D = 0$, and no change is needed (or we can change any element in any other row arbitrarily).

1.2.8 The simplest way is to multiply the first equation by α_{21} and the second equation by α_{11}, then subtracting the two equations x_1 gets eliminated, and we can express x_2. We can obtain x_1 similarly. This proves that only the values given in the exercise can provide a solution. We have to check by substitution that these are solutions indeed.
Remark: The generalization of this observation to systems of n linear equations with n variables is Cramer's rule to be discussed in Section 3.2.

1.2.9 The simplest way is to transform the parallelogram into another one of the same area and with a side on a coordinate axis.
Remark: An analogous statement holds for the volume of a parallelopiped in the space and in higher dimensions; see Section 9.8.

1.2.11 There is exactly one odd integer among the products occurring in the definition of the determinant, all others are even, so their (signed) sum is an odd number.

1.3.

1.3.3 Subtracting rows and columns, we get a determinant with smaller and nicer numbers.

1.3.4 We can argue similar to the proof of Theorem 1.3.1/III.

1.3.5 Proceed similar to the proof of Theorem 1.3.3.

1.3.6 Swapping the two rows the determinant remains the same on the one hand, and changes sign on the other hand, so it must be 0. This argument does not work, e.g., for the modulo 2 field where also $1 = -1$.

1.3.10 To handle the divisibility by 3, add every column to the last column. For the divisibility by 23 or any other integer, add 10, 100, etc. times the previous columns (considered right to left) to the last column.

1.3.12 Add every row to the first row, then factor out the common value $\gamma + (n-1)\delta$ form the first row, and subtract δ times the first row from every other row. Another option: Proceeding bottom-up, subtract every row from the row below it, then right-to-left add every column to the previous column.

1.3.13 Reflect through the main diagonal.

1.3.14 If the conjugate of a row occurs more than once, then there are two equal rows, so $D = 0$. If the conjugate of a row is itself, then every element in it is a real number. Applying appropriate subtractions to a conjugate pair of rows we can transform them into a row with real numbers and another row with imaginary numbers. Factoring out i from the s imaginary rows we get $D = i^s D'$ where the elements of D' are real numbers, so $D' \in \mathbf{R}$.

1.3.15 Using that every element is the sum of its upper and left neighbors, we can reduce the problem with subtractions to a similar determinant of smaller size.

1.3.16 (a) Subtract every row from the first row.
(b) Make a last column from the first one by repeated swaps, and subtract the first row from the other rows.

1.3.18 Apply the addition formula for the cosine and write the determinant as a sum of 2^n determinants by the repeated use of Theorem 1.3.2. Each of these determinants is 0 for $n > 2$ by Theorem 1.3.3A.

1.3.19 Add every column to the last column and use Exercise A.2.7.

1.3.20 Argue similarly as in the previous exercise.

1.3.21 If $(i, n) = 1$, then adding the ith and $(n - i)$th rows, we always obtain a row $n, n, \ldots, n, 2n$.

1.4.

1.4.2 The skew expansion along the first two rows is the same as the cofactor expansion along one of these rows.

1.4.3 Expand both the old and new determinants along the first row.

1.4.5 Denote the determinant by D_n and expand it along its last row. After computing the cofactor A_{n1} we get the recursion $D_n = \delta D_{n-1} - \beta\gamma\delta^{n-2}$. Computing D_n for a few small values of n (it is worth using the recursion for these computations, too), we can easily conjecture the formula for D_n and then verify it via induction relying on the recursion. We can solve the exercise without recursion, too. If $\delta = 0$, then $D_n = 0$. Otherwise we can achieve 0 in every position above the main diagonal if we subtract from the first row suitable multiples of the other rows.

1.4.7 Expanding along the first or last row or column, we get the recursion $D_n = (\gamma + \delta)D_{n-1} - \gamma\delta D_{n-2}$.

1.4.8 Write the new determinant as the sum of 2^n determinants. Most terms are 0.

1.4.9 Prove by induction using expansion along the last row.

1.4.11 (a) Verify the equality of the cofactors belonging to the same row or column first. E.g., A_{11} and A_{1j} differ only in one column and possibly in sign. Using the condition, express the elements of the first column with the other columns.
(b) Apply expansion.

1.4.12 Rewrite the last γ in the main diagonal as $\beta + (\gamma - \beta)$ and transform the determinant into a sum of two determinants according to the last row. Now we can express D_n with D_{n-1}. The quickest continuation is to create a second recursion from this based on the symmetry of β and δ. Then it is easy to express D_n from the two equalities.

1.4.15 Modify appropriately the proof of Theorem 1.4.2.

1.5.

1.5.2 If γ_i are all distinct, then all solutions are $x = \gamma_i$, $i = 2, 3, \ldots, n$. For $\delta \neq 0$, $V(x, \gamma_2, \ldots, \gamma_n) - \delta$ is a polynomial (in x) of degree exactly $n-1$, so it cannot have more than $n-1$ roots. There can be fewer than $n-1$ solutions if this polynomial has multiple roots.

1.5.8 Express $\cos(r\phi)$ as a polynomial of degree r of $\cos\phi$. Determine the leading coefficient and apply Exercise 1.5.5(a).

1.5.11 (a) By the pigeonhole principle, to every $k < n$ there exist $i \neq j$ such that a_i and a_j give the same remainder when divided by k.
(b) $V(a_1, \ldots, a_n)$ does not change if we replace a_i^j by $j!\binom{a_i}{j}$.

1.5.12 Use Wilson's Theorem to compute the remainders of the products $k!(p-1-k)!$.

1.5.13 Expand the $(n+1)\times(n+1)$ determinant $f(x) = V(\gamma_1, \gamma_2, \ldots, \gamma_n, x)$ along its last row.

2. Matrices

2.1.

2.1.9 The wrong argument assumes the commutativity of multiplication.

2.1.15 In A^2 also the elements just above the main diagonal are 0, in A^3 also the elements just above these are 0, etc.

2.1.16 Write $A = I + B$, apply the binomial theorem (rightful in this case), and use the previous exercise.

2.1.17 Multiply the equality by $A - I$.

2.2.

2.2.5 Adapt the reasoning in the proof of Theorem 2.2.2 to this case.

2.2.7 (f) If $\det A \neq 0$, then we can express X using A^{-1}. If $\det A = 0$, then $AC = 0$ for some $C \neq 0$, so $AX = A(X+C)$ shows that there cannot be a unique solution.

2.2.8 1.5.5: In general, if $f_i = \beta_{i,0} + \beta_{i,1}x + \ldots + \beta_{i,n-1}x^{n-1}$, and D_1 is the determinant with elements β_{ij}, $(0 \leq i, j \leq n-1)$, then the determinant in question is the product of D_1 and the Vandermonde-determinant $V(a_1, \ldots, a_n)$.

1.5.6: $V(\alpha_1, \ldots \alpha_n)V(\beta_1, \ldots, \beta_n)$.

1.5.7:

$$\begin{vmatrix} 1 & \binom{n-1}{1}\alpha_1 & \binom{n-1}{2}\alpha_1^2 & \ldots & \alpha_1^{n-1} \\ 1 & \binom{n-1}{1}\alpha_2 & \binom{n-1}{2}\alpha_2^2 & \ldots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \binom{n-1}{1}\alpha_n & \binom{n-1}{2}\alpha_n^2 & \ldots & \alpha_n^{n-1} \end{vmatrix} \cdot \begin{vmatrix} \beta_1^{n-1} & \beta_2^{n-1} & \ldots & \beta_n^{n-1} \\ \beta_1^{n-2} & \beta_2^{n-2} & \ldots & \beta_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \ldots & 1 \end{vmatrix}$$

2.2.9 $A^2 = I$ is equivalent to $A = A^{-1}$, and the other condition means $\hat{A} = \pm A$. Use the formula for the inverse obtained in Theorem 2.2.2 and apply Lemma 2.2.3 and the product theorem for determinants.

2.2.10 Apply Lemma 2.2.3 first for A and then for $\hat{A}$. The adjugate of the adjugate is A multiplied by the scalar $(\det A)^{n-2}$.

2.2.11 Apply Lemma 2.2.3 both for A and B.

2.2.14 For $n = 3$, consider a matrix M with elements 0, 1, and 2, such that multiplying any matrix A by M from the left, the rows of A are permuted and one row is multiplied by 2 (cf. with Exercises 2.1.10 and 2.1.11). For a general $3 \mid n$, build an $n \times n$ matrix of such 3×3 blocks centered around the main diagonal whereas all other elements are 0. Finally, use the product theorem for determinants to prove the necessity of $3 \mid n$.

3. Systems of Linear Equations

3.1.

3.1.10 (d) Use the modulo 7 field F_7.

3.1.12 There must be at least $n - k$ free parameters in the REF.

3.1.13 *First proof*: If $|H| > 1$, then there are solutions $\mathbf{x}'$ and $\mathbf{x}''$ with $x_1' \neq x_1''$. If (i) $\lambda_1 + \lambda_2 = 1$, then $\mathbf{x}^* = \lambda_1\mathbf{x}' + \lambda_2\mathbf{x}''$ is a solution and (ii) $x_1^* = \lambda_1 x_1' + \lambda_2 x_1''$. Prove that to any x_1^* there exist λ_1 and λ_2 satisfying (i) and (ii).
Second proof: If $|H| > 1$, then use the parametric representation of x_1 in the REF.
Third proof: Put x_1 as last unknown and then x_1 will be a free parameter if $|H| > 1$.

3.1.15 $A\mathbf{x}'' = A\mathbf{x}' \iff A(\mathbf{x}'' - \mathbf{x}') = \mathbf{0}$ by the properties of matrix operations.

3.1.18 (b) If the coefficients and the constants are rational numbers, then we remain in $\mathbf{Q}$ during the Gaussian elimination.

3.1.19 (b) Multiplying a non-trivial rational solution by the least common multiple of the denominators and then dividing by the greatest common divisor of the resulting integers, we get an integer solution where not all unknowns are divisible by 11. This yields a non-trivial solution over F_{11}.

3.2.

3.2.1 It is worth separating the real and imaginary parts. The simplest application of Cramer's rule is to guess a solution and to show that the determinant of the coefficient matrix is not 0 (see Exercise 1.3.15), so there is a unique solution. It is not too complicated to compute the determinants in the formulas in Cramer's rule either, but in this case elimination offers a quicker algorithm.

3.2.5 The conditions imply that each system has a unique solution.
(a) The if part is obvious. For the converse, the common solution means $A_1\mathbf{x} = A_2\mathbf{x}$ for every $\mathbf{x} \in F^n$. Choose $\mathbf{x}$ as unit vectors, i.e., let one component be 1 and all others 0.
(b) Both conditions mean that $(A_1 - A_2)\mathbf{x} = \mathbf{0}$ has only a trivial solution.

3.2.6 (a) Switching to the modulo 7 field we get a homogeneous system having only a trivial solution.
(b) For a general integer K instead of 7, the correct condition is that the determinant is (not just not a multiple of K, but is) coprime to K. Prove first the case when K is a power of a prime by induction on the exponent. Then show that if the statement holds for two coprime values of K, then it holds also for their product. Another option: Using

the matrix $M = (m_{ij})_{i,j=1}^{n}$ and the vector $\mathbf{v} = (v_j)_{j=1}^{n}$, the condition means $M\mathbf{v} = \mathbf{b}$ where M is invertible and every component of $\mathbf{b}$ is divisible by K, i.e., $\mathbf{b} = K\mathbf{c}$. Then $\mathbf{v} = M^{-1}\mathbf{b} = K\hat{M}\mathbf{c}/\det M$ has integer components and $(K, \det M) = 1$ implies that the cancellation by $\det M$ does not affect K.

3.2.7 We can even guess the polynomials. Besides Theorem 3.2.4 we can apply also the methods described in Exercises 3.2.10 and 3.2.11.

3.2.8 (b) If g is such a polynomial, then every γ_i if a root of $g - f$. Another option: Extend the list of places by a new $\gamma_{n+1} \in F$, prescribe here an arbitrary value β_{n+1}, and consider the resulting interpolation polynomials. This latter method fails if F is a finite field with exactly n elements.

3.2.9 If there were two distinct such polynomials f and g, then $\deg(f - g) \leq n - 1$, but all the n values γ_i are zeros of $f - g$, a contradiction.

3.2.10 Substituting $\gamma_1, \ldots, \gamma_n$ one after the other, we can determine the coefficients $\nu_0, \ldots, \nu_{n-1}$ successively. This proves the *existence* of the interpolation polynomial, and also its uniqueness among the polynomials *of the form given in the exercise*. The *uniqueness* of the interpolation polynomial requires besides the uniqueness of the coefficients ν_i also the demonstration that every polynomial of degree not greater than $n-1$ can be uniquely represented in the form described in the exercise.

3.2.11 (a) The polynomial L_i contains every root factor $x - \gamma_j$, $j \neq i$, so it must be a constant multiple of the product of these root factors due to the bound on the degree. The constant multiplier is obtained from the condition $L_i(\gamma_i) = 1$.

3.2.13 (a) Take, e.g., $x(x + 1)/2$.
(b) Using sufficiently many places and the corresponding values assumed by f, reconstruct f via (any method of) interpolation, and observe that these algorithms operate within the field from where the places and values are taken.

3.2.14 Construct f as an interpolation polynomial using *all* elements of F as places and the values assumed by Φ at these places.

3.2.15 Ali Baba chose a polynomial f of degree 24 where the constant term is the password and whispered $f(i)$ to the ith thief.

3.3.

3.3.11 (a) Use the previous exercise.
(b) Row-equivalent transformations do not influence whether the corresponding homogeneous system of equations has a non-trivial solution or not.

3.3.12 See Theorem 9.3.1.

3.4.

3.4.1 Use Theorem 3.3.5/II for the columns and the cofactor expansion for the minors.

3.4.6 (a) This follows from any of the three rank notions.
(b) Consider, e.g., an identity matrix extended by a column where every element is 1.
(c) The example of (b) can be adapted to this case, too.

3.4.7 Both are exactly the non-zero matrices occurring in Exercise 3.4.4(a).

3.4.10 The rank is equal to both the number of rows and the number of columns.

3.4.12 Elementary row- and column-equivalent transformation do not alter the rank, this justifies the if part. For the converse, we can convert matrices A and B of the same rank r with such transformations into a matrix C where the first r elements of the main diagonal are 1 and all other elements are 0. Applying the steps leading from A to C followed by the inverse steps leading from C to B will convert A into B.

3.4.13 (b) Each further column is a linear combination of any $r-1$ columns from the r independent columns.
(c) We can rely on Exercise 3.4.5.

3.4.14 (a) Construct, e.g., the first three columns so that any three rows in this submatrix are independent and the other five columns are the same as the first column.
(b) Verify the statement first for the further elements in the rows and columns of the non-zero minor.
(c) We can rely on Exercise 1.5.6.

3.4.17 (a) If $\mathrm{rk}(A) \leq n-2$, then $\hat{A} = 0$. If $\mathrm{rk}(A) = n-1$, then the cofactor and skew expansions imply that every row in $\hat{A}$ is a solution of the homogeneous system of equations $A\mathbf{x} = \mathbf{0}$, and the solutions can be characterized by $n - \mathrm{rk}(A) = 1$ free parameters.
(b) It follows from (a) for $n > 2$ and can be verified directly for $n = 2$.

3.5.

3.5.4 To prove $\det A = 0$, add all columns to the last column of A. Equation $AB = 0$ is satisfied, e.g., if B is a *repunit* matrix, i.e., every element in B is 1.

3.5.6 (a) Adapt the new proof of Theorem 2.2.2 given in this section.
(b) Use the previous two exercises to construct A.

3.5.7 We can construct a counterexample from Exercise 3.5.4 or 3.5.5.

3.5.8 (a) See, e.g., Exercise 3.5.3(b).
(b) Let $\mathbf{x} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$ be non-trivial solutions of the systems $A\mathbf{x} = \mathbf{0}$ and $A^T\mathbf{y} = \mathbf{0}$, respectively. If the elements of B are $\beta_{ij} = \lambda_i\mu_j$, then $AB = BA = 0$.

4. Vector Spaces

4.1.

4.1.5 One way is to check each axiom. Another option is to show that we "labeled" the elements of a well-known vector space differently and defined the operations accordingly.

4.1.6 See the hint to the previous exercise.

4.1.8 (a) E.g., $\lambda = 1/2, v = 3$ causes a problem.
(b) The argument in (a) works except for fields where $1+1 = 0$, so 2 has no multiplicative inverse. In this case Axiom (MS1) is false.
(c) Use a set of the same cardinality as $\mathbf{Z}$ that is a well-known vector space.
(d) Verify that the cardinality of V cannot be smaller than the cardinality of F.
(e) Interpret the real sequences as complex sequences.
(f) By more advanced arguments we can prove the surprising fact that $\mathbf{R}$ and $\mathbf{C}$ have essentially the same structure for addition: their additive groups are *isomorphic*. Indeed, thinking of $\mathbf{R}$ and $\mathbf{C}$ as vector spaces over $\mathbf{Q}$, cardinality considerations imply that their bases have the same cardinality, so their dimensions are equal, and thus the two vector spaces are isomorphic. See also the remarks at the end of Sections 4.5 and 5.2.

4.1.10 (ii) First apply (MS1) on the left-hand side of $(0+\lambda)\mathbf{v} = \lambda\mathbf{v}$.
(iii) Start with $\big(1+(-1)\big)\mathbf{v} = 0\mathbf{v}$.
(iv) If $\lambda \neq 0$, then multiply both sides of $\lambda\mathbf{v} = \mathbf{0}$ by the scalar λ^{-1}.

4.1.12 (a) (MS4) implies (a) trivially. For the converse, apply (MS3) on the right-hand side of $1\mathbf{v} = 1(\lambda\mathbf{v})$.
(b) The validity of (b) in a vector space is statement (iv) in Theorem 4.1.2. For the converse, evaluate $1\big(1\mathbf{v}+(-\mathbf{v})\big)$.

4.1.13 Write $(1+1)(\mathbf{u}+\mathbf{v})$ as the sum of four terms in two different ways.

4.1.14 Several examples are contained in Exercise 4.1.9.
Concerning the independence of (A1)–(A4), the following observation can be useful: If $\mathbf{v}+\mathbf{v} = \mathbf{v}$ and $\lambda\mathbf{v} = \mathbf{v}$ for every $\mathbf{v}$ and λ, then (MS1)–(MS4) are all true.
Independence of (MS2): Let $V = \mathbf{C}^2$ with the usual addition, $F = \mathbf{C}$, and $\lambda\mathbf{v}$ means the usual multiplication of $\mathbf{v}$ by λ or $\overline{\lambda}$ depending on certain properties of $\mathbf{v}$.

Remark: The independence of (A3) raises *conceptual* problems. If (A3) is false, then (A4) makes no sense, since we cannot speak about a negative if there is no zero element. However, *formally* we can still ask the independence of (A3): There is no zero element but there exists an element *denoted* by $\mathbf{0}$ appearing only in (A4) so that (A4) and all other axioms hold except (A3).

It is even more artificial to interpret the other axioms if (A) or (MS) is false since if there is a problem with the operation itself, then we generally do not investigate its properties. Still, in some cases, there is a unique assignment to most pairs of elements, e.g., at the division of real numbers. For such partial operations, it is not totally absurd to keep the axioms alive by saying that the equalities hold if both sides make sense.

4.2.

4.2.10 To start from a concrete example, consider the lines through the origin in the usual vector space of plane vectors. We can generalize this to an arbitrary vector space: If $\mathbf{a} \neq \mathbf{0}$ and $\mathbf{b}$ is not a scalar multiple of $\mathbf{a}$, then all scalar multiples of $\mathbf{c}_\mu = \mathbf{a} + \mu\mathbf{b}$ provide distinct subspaces for every $\mu \in F$.

4.2.12 (g) This follows for F_2^2 from Exercise 4.2.11. We can argue similarly for F^2 over any finite F. The general case can be reduced to this by taking a big subspace W in V so that V is the extension of W by a "two-dimensional" subspace.

4.2.15 The arguments (a)–(c) assume $\mathbf{0}_V \in W$.

Remark: A (correct) proof of $\mathbf{0}_W = \mathbf{0}_V$ — as the one in (d) — *must use* that some element $\mathbf{v}$ has a *negative in* V. Note that the pure existence of an identity element of an operation does not guarantee that a "nice" subset has the same identity element. E.g., consider the positive integers for the operation of taking the maximum of two numbers. Here the identity element is 1 as $\max(1, n) = n$ for every $n > 0$. However, in the subset of integers greater than 5, the identity element will be 6! See also Exercise A.6.11.

4.2.16 (b) Show that if two translates share a vector, then they mutually contain each other.

(c) Let $\mathbf{v}$ be an element of both translates. Verify that $\mathbf{w}$ has the same property if and only if $\mathbf{w}-\mathbf{v}$ is in the intersection of the two translates.

(d) If $\mathbf{a} = \mathbf{v} + \mathbf{w}_1$, $\mathbf{b} = \mathbf{v} + \mathbf{w}_2$, $\mathbf{c} = \mathbf{v} + \mathbf{w}_3$, $\mathbf{w}_i \in W$, then using that W is a subspace, we obtain that $\mathbf{a} + \lambda(\mathbf{b} - \mathbf{c}) = \mathbf{v} + \mathbf{w}_4$. For the converse, try to generate the suitable subspace W from the elements of T and verify that it is a subspace using the condition.

4.2.17 The main difficulty is that the translate $\mathbf{v} + W$ does not determine the vector $\mathbf{v}$ uniquely, i.e., a translate can be represented by several vectors $\mathbf{v}$. Therefore, first we have to show that the operations depend only on the translates themselves and are independent of their representations. In other words, the result of an operation has to be the same translate whichever vectors were taken as representatives in the operands.

4.3.

4.3.5 It is enough to show that the generators on either side are elements of the span on the other side.

4.3.8 If two subspaces satisfy these conditions, then they mutually contain each other by (iii), so they are equal.

4.3.9 Check that this intersection satisfies requirements (i)–(iii) in Theorem 4.3.4.

4.3.11 Argue similar to the proof of Theorem 4.3.4.

4.3.16 (c) Rephrase the problem as whether a system of infinitely many linear equations with rational coefficients and finitely many unknowns is solvable. Show using Gaussian elimination that this does not depend on whether the unknowns are rational or real numbers.

4.3.17 (b) We cannot generate, e.g., a sequence of distinct powers of a *transcendental* number (for the definition see Section A.10). To prove this, write the problem as a system of infinitely many equations with finitely many unknowns, and apply Gaussian elimination or elementary properties of algebraic extensions (see Section A.10).

4.4.

4.4.3 If **0** is a non-trivial combination, then multiply it by any scalar.

4.4.4 Argue similarly to the proof of statement III. in Theorem 4.4.3.

4.4.12 Use the unique prime factorization theorem in (a) and the definition of a transcendental number in (b).

4.5.

4.5.3 Both are equivalent to the notion of a basis. (Rely on Theorem 4.5.3.)

4.5.4 Apply Theorem 4.5.4.

4.5.9 It is sufficient to check independence.

4.5.10 Investigate independence and use Theorem 3.2.3.

4.5.11 Express the vectors $\mathbf{v}_i$ as in the previous exercise. Prove that to every i there exists a j such that $\beta_{ij} \neq 0$ and also the cofactor $A_{ij} \neq 0$ in the determinant formed from the elements β_{rs}. Then $\mathbf{v}_j$ satisfies the requirements.

4.5.12 (b) Express the elements of the vector space as linear combinations of a basis.
(c) Prove that every finite field contains a field F_p and is a vector space over this field (cf. Section A11).
(d) Use part (c).

4.5.13 Apply, e.g., Theorem 4.5.7.

4.5.14 (a) The first element of a basis in F_p^n can be any non-zero vector, the second element can be any vector except the scalar multiples of the first one, etc.
(b) This is equivalent to the general part of (a).
(c) Use the result of part (a) and observe that permuting the elements in a basis yields a new basis.

4.6.

4.6.1 (e) Use Exercise 3.2.14.
(f) Try to find a general combinatorial argument.
(g) Try to find a general combinatorial argument.

4.6.2 Substitute suitable numbers into the polynomials to handle the linear combinations in (a), (b), and (c). In (d), already the first 8 polynomials are linearly dependent; which powers of x occur in them after expanding the expressions?

4.6.3 Consider the span of k independent vectors.

4.6.4 Use that the same condition holds for each column of A.

4.6.5 To start, observe that the union of the bases in W_1 and W_2 is a dependent system.

4.6.6 (a) The union of the bases in W_1 and W_2 is a spanning set in $\langle W_1, W_2\rangle$.
(b) The union of the bases in W_1 and W_2 is a basis in $\langle W_1, W_2\rangle$.
(c) Extend the basis in $W_1 \cap W_2$ to bases in W_1 and W_2.

4.6.8 Answers to the hints in the exercise: (i) 2; (ii) Search for geometric sequences. See the details in the first proof of Theorem 9.2.1 and the remark after it.

4.6.9 The relations prescribed for the nine unknowns form a system of linear equations with 3 free parameters, e.g., α_{11}, α_{12}, and α_{22} can be chosen arbitrarily. We get a basis if one of these elements is 1, the other two are 0, and the remaining six elements are uniquely determined from the conditions.

4.6.11 The pairwise sums of the vectors are dependent for $k \geq 4$.

4.6.13 Investigate the spans of the column vectors.

4.6.14 If $r > 0$, then r independent vectors span an r-dimensional subspace. But we get the same subspace several times, it is spanned by any of its bases.

4.6.15 Clearly, it is sufficient to deal with the case $0 < r \leq \min(k, n)$. We can minimize the computation by choosing an r-dimensional subspace in F_p^k and then specify a spanning set of n elements in this subspace.

4.7.

4.7.3 We need essentially the inverse of the matrix formed of the three basis vectors.

5. Linear maps

5.1.

5.1.6 (b) Use $F = \mathbf{C}$.

5.1.10 Apply Theorem 4.6.6.

5.1.13 $\mathbf{u}_i - \mathbf{u}_1$ are independent vectors in $\operatorname{Ker} \mathcal{A}$, $i = 2, 3, \ldots, k$.

5.1.14 If for some basis $\mathbf{c}_1, \ldots, \mathbf{c}_n$ (say) $\mathcal{A}\mathbf{c}_1 \neq \mathbf{0}$ and $\mathcal{A}\mathbf{c}_i = \mathbf{0}$, then replace $\mathbf{c}_i$ by $\mathbf{c}_i + \mathbf{c}_1$.

5.2.

5.2.2 4.1.5 is isomorphic to the usual vector space of real numbers over itself, a suitable isomorphism is $v \mapsto \ln v$. 4.1.6 is isomorphic to the vector space of complex numbers over $\mathbf{Q}$ with the usual operations, based on the isomorphism $v \mapsto v + 1$.

5.2.4 Apply Theorem 5.2.5.

5.3.

5.3.1 Extend a basis of W to a basis of V and define suitable linear transformations on this basis.

5.3.6 Using the bases, we can define an $\mathcal{A}$ with $\operatorname{Ker} \mathcal{A} = \mathbf{0}$ if $\dim V_1 \leq \dim V_2$ and with $\operatorname{Im} \mathcal{A} = V_2$ if $\dim V_1 \geq \dim V_2$.

5.4.

5.4.1 We get a necessary condition from the Dimension Theorem. To prove its sufficiency, define a suitable linear map on a basis or work in F^n by Theorem 5.2.4.

5.4.5 (i) implies $\dim \operatorname{Im} \mathcal{A} \leq 3$ and (ii) implies $\dim \operatorname{Ker} \mathcal{A} \leq 5$.

5.4.6 Write the Dimension Theorem both for $\mathcal{A}$ and $\mathcal{B}$, and apply Theorem 4.6.4 on the dimension of a subspace in a finite-dimensional vector space.

5.4.7 Apply the Dimension Theorem and Exercise 4.6.6(b) about the dimension of a direct sum.

5.5.

5.5.6 Verify that every $\mathcal{A} \in \mathrm{Hom}\,(V_1\,,\,V_2)$ is a unique linear combination of the maps $\mathcal{C}_{ij}$. Use that $\mathcal{A}$ can be characterized by the images of the basis vectors $\mathbf{a}_j$ and the images are unique linear combinations of the basis vectors $\mathbf{b}_i$. (Cf. Section 5.7.)

5.5.7 This follows from the previous exercise. (Cf. Theorem 5.7.5.)

5.6.

5.6.6 We verify the first distributive law. Either side of $\mathcal{A}(\mathcal{B}+\mathcal{C}) = \mathcal{A}\mathcal{B}+\mathcal{A}\mathcal{C}$ makes sense exactly if $\mathcal{A} \in \mathrm{Hom}\,(V_2\,,\,V_3)$, $\mathcal{B}$, $\mathcal{C} \in \mathrm{Hom}\,(V_1\,,\,V_2)$. Applying the left-hand side to $\mathbf{x} \in V_1$, we get

$$[\mathcal{A}(\mathcal{B}+\mathcal{C})]\mathbf{x} = \mathcal{A}[(\mathcal{B}+\mathcal{C})\mathbf{x}] = \mathcal{A}(\mathcal{B}\mathbf{x}+\mathcal{C}\mathbf{x}) = \mathcal{A}(\mathcal{B}\mathbf{x}) + \mathcal{A}(\mathcal{C}\mathbf{x})\,,$$

we used the linearity of $\mathcal{A}$ in the last step. Applying the right-hand side to $\mathbf{x} \in V_1$, we obtain the same result using the definitions of addition and multiplication of linear maps.

5.6.7 Apply the Dimension Theorem to the restriction of $\mathcal{B}$ to the subspace $\mathrm{Im}\,\mathcal{A}$.

5.6.8 Verify first $\mathrm{Ker}\,\mathcal{A}^2 \supseteq \mathrm{Ker}\,\mathcal{A}$ and $\mathrm{Im}\;\mathcal{A}^2 \subseteq \mathrm{Im}\,\mathcal{A}$ (cf. Exercise 5.7.5). Then apply the Dimension Theorem both to $\mathcal{A}$ and $\mathcal{A}^2$. As the dimension is finite, these imply the equivalence of the first two conditions. The equivalence of the first and third conditions follows from the definitions directly (or the equivalence of the second and third conditions can be deduced from the previous exercise).

5.6.9 Let $V = \mathbf{R}^5$ and $\mathcal{A}\mathbf{x} = A\mathbf{x}$ for every $\mathbf{x} \in V$. Clearly $\mathrm{Im}\,\mathcal{A}^{k+1} \subseteq \mathrm{Im}\,\mathcal{A}^k$, and if this holds with equality for some k, then we have equality for all larger powers, too:

$$U = \mathrm{Im}\,\mathcal{A}^k = \mathrm{Im}\,\mathcal{A}^{k+1} = \mathrm{Im}\,\mathcal{A}^{k+2} = \ldots$$

By the condition of the exercise this equality will get stabilized with $U = \mathbf{0}$. Since the dimension in the chain

$$V \supseteq \mathrm{Im}\,\mathcal{A} \supseteq \mathrm{Im}\,\mathcal{A}^2 \supseteq \mathrm{Im}\,\mathcal{A}^3 \supseteq \ldots$$

decreases by at least one in each step until we reach U, it must get stabilized latest in the 5th step. Therefore $\mathcal{A}^5 = \mathcal{O}$, so $A^5 = 0$. (We can get another solution using the *minimal polynomial*; see Exercise 6.3.8.)

5.6.15 Extend a basis of $\operatorname{Im}\mathcal{A}$ to a basis of V and define $\mathcal{B}$ on this basis.

5.6.18 (e) Verify $\operatorname{Ker}(\mathcal{P}+\lambda\mathcal{E}) = \mathbf{0}$ or $\operatorname{Im}(\mathcal{P}+\lambda\mathcal{E}) = V$, or find the inverse in the form $\alpha\mathcal{P}+\beta\mathcal{I}$.

(f) The corresponding subspaces are $U_1 = \operatorname{Im}\mathcal{P}$ and $U_2 = \operatorname{Ker}\mathcal{P}$. The "if" part can be verified directly. For the converse, use the decomposition $\mathbf{v} = \mathcal{P}\mathbf{v} + (\mathbf{v} - \mathcal{P}\mathbf{v})$.

5.6.19 Extend a basis of $\operatorname{Im}\mathcal{A}$ to a basis of V and define $\mathcal{B}$ on this basis. We have to assign to each basis vector of $\operatorname{Im}\mathcal{A}$ one of its inverse images.
Remark: It can even be achieved that both $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$ and $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$ hold simultaneously. Such a $\mathcal{B}$ is called a *generalized inverse* of $\mathcal{A}$. (Why?)

5.6.24 If the quaternion v is not a real number, then the quaternions $\alpha+\beta v$ with $\alpha,\ \beta \in \mathbf{R}$ form a field isomorphic to $\mathbf{C}$.

5.6.25 Let $c \neq 0$ be an arbitrary but fixed element in the algebra A and consider the multiplication by c as a linear transformation of the vector space A, i.e., $\mathcal{C} : x \mapsto cx$ for every $x \in A$; $\mathcal{C} \in \operatorname{Hom} A$. There are no zero divisors, so we have $\operatorname{Ker}\mathcal{C} = 0$. Since $\dim A < \infty$, this implies $\operatorname{Im}\mathcal{C} = A$, i.e., the equation $cx = d$ is solvable for every $d \in A$. We get the solvability of $yc = d$ similarly.

5.7.

5.7.6 Extend a basis of $\operatorname{Ker}\mathcal{A}$ to a basis of V_1 and extend the non-zero images of this basis to a basis of V_2.

5.7.8 If $\mathbf{c} \notin \operatorname{Ker}\mathcal{A}$, then the basis $\mathcal{A}\mathbf{c}$, $\mathbf{c}$ works.

5.7.10 By Theorems 2.2.5 and 3.5.2, a non-zero square matrix is a one- or two-sided zero divisor if and only if its determinant is 0. This is equivalent to the condition $\operatorname{Ker}\mathcal{A} \neq \mathbf{0}$ by the isomorphism between linear transformations and matrices and Theorem 3.2.3 about non-trivial solutions of a homogeneous system of linear equations.

5.7.11 The subspace spanned by the column vectors is $\operatorname{Im}\mathcal{A}$.

5.7.12 Switch to the corresponding linear maps, apply the Dimension Theorem to $\mathcal{A}$ in part (b) and to the restriction of $\mathcal{A}$ to $\operatorname{Im}\mathcal{B}$ in part (a), and use the previous exercise.

5.7.13 There exist non-identity linear transformations of the plane with this property, e.g., among rotations.

5.7.14 Over all other fields, we get the same matrix if we multiply every basis vector by a scalar $\lambda \neq 0$ or 1. Over the modulo 2 field, let, e.g., $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ be a matrix of a linear transformation; this matrix determines the basis uniquely.

5.8.

5.8.3 Apply Theorem 5.8.1A for $\mathcal{A} = \mathcal{S}$.

5.8.4 Apply Theorems 5.8.1A and 2.2.4 about the multiplication of determinants.

5.8.5 A counterexample for both (a) and (b) is differentiation in the vector space of polynomials of degree at most one.

5.8.8 Any basis works for $\mathcal{A} = \lambda\mathcal{I}$. Otherwise, try $\mathcal{S} = \mathcal{A} - \lambda\mathcal{I}$ as an accompanying transformation.

6. Eigenvalue, Minimal Polynomial

6.1.

6.1.4 (c) If $\mathcal{A}^2\mathbf{v} = \mu^2\mathbf{v}$, then $(\mathcal{A} + \mu\mathcal{I})\big((\mathcal{A} - \mu\mathcal{I})\mathbf{v}\big) = \mathbf{0}$ implies that either $\mathbf{v}$ is an eigenvector with eigenvalue μ or $(\mathcal{A} - \mu\mathcal{I})\mathbf{v}$ is an eigenvector with eigenvalue $-\mu$.

6.1.9 Prove by induction on k.

6.1.10 It follows from the previous exercise.

6.1.11 The dimension of the kernel is 2 and one can easily find an eigenvector with eigenvalue 3.

6.1.12 (a) 0 is an eigenvalue if and only if the corresponding transformation has no inverse and this holds for $\mathcal{AB}$ and $\mathcal{BA}$ at the same time. Otherwise, apply $\mathcal{B}$ to the equality $(\mathcal{AB})(\mathbf{x}) = \lambda\mathbf{x}$.
(b) It follows from (a) by $\mathcal{A} = \mathcal{AB}\mathcal{B}^{-1}$.

6.2.

6.2.4 6.1.10: A polynomial cannot have more roots than its degree.
6.1.9: If they were dependent than they would span a subspace of dimension less than k. Restricting the transformation onto this subspace, it would have at least k eigenvalues contradicting to Exercise 6.1.10.

6.2.5 The characteristic polynomial has a root by the fundamental theorem of algebra.

6.2.9 Use the Vieta formulas for the characteristic polynomial.

6.2.10 Compare the traces and determinants of the matrices, the characteristic polynomials (hence the eigenvalues), and check whether the matrices are similar to diagonal matrices.

6.2.11 Use Theorem 1.3.6.

6.3.

6.3.4 Multiply $m_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$ by $\mathcal{A}^{-1}$.

6.3.7 The fundamental theorem of algebra is true for the minimal polynomial.

6.3.8 The minimal polynomial divides x^{1000} and has degree 5 or less, so it is x^k for some $k \leq 5$.

6.3.9 (a) (a) Consider, e.g., the matrix of a suitable rotation around the origin.
(b) The minimal polynomial would have degree 2 and would divide $x^5 - 1$ over $\mathbf{Q}$.

6.3.10 If $\alpha_0\mathcal{I} + \alpha_1\mathcal{AB} + \alpha_2(\mathcal{AB})^2 + \ldots + \alpha_k(\mathcal{AB})^k = \mathcal{O}$, then multiplying this equality by $\mathcal{B}$ from the left and by $\mathcal{A}$ from the right, we obtain $\alpha_0\mathcal{BA} + \alpha_1(\mathcal{BA})^2 + \alpha_2(\mathcal{BA})^3 + \ldots + \alpha_k(\mathcal{BA})^{k+1} = \mathcal{O}$.

6.3.11 If $m_{\mathcal{A}} = f$ and $m_{\mathcal{B}^{-1}\mathcal{AB}} = g$, then $f(\mathcal{B}^{-1}\mathcal{AB}) = \mathcal{B}^{-1}f(\mathcal{A})\mathcal{B} = \mathcal{O}$ implies $g \mid f$. We get the reverse divisibility similarly.

6.3.12 If the minimal polynomial has degree $j(\leq n)$, then every power of $\mathcal{A}$, thus also $\mathcal{A}^k$ is a linear combination of $\mathcal{I}, \mathcal{A}, \ldots, \mathcal{A}^{j-1}$.

6.3.13 If j is the degree of the minimal polynomial, then $\mathcal{I}, \mathcal{A}, \ldots, \mathcal{A}^{j-1}$ is a basis in this subspace. (Cf. Theorem 6.5.4.)

6.3.14 Let $\beta_0 + \beta_1 x + \ldots + \beta_s x^s$ be the minimal polynomial of $\mathcal{A}^2$. Substituting $\mathcal{A}^2$, we see that $\mathcal{A}$ is a root of $\beta_0 + \beta_1 x^2 + \ldots + \beta_s x^{2s}$, so $k \leq 2s$. Turning to the other direction, observe that $\mathcal{A}^j \in \langle \mathcal{I}, \mathcal{A}, \ldots, \mathcal{A}^{k-1} \rangle$ for every j, so

any $k+1$ powers of $\mathcal{A}$, consequently also of $\mathcal{A}^2$ are linearly dependent. This proves $s \leq k$. To verify that all these values of s do appear, consider transformations with k distinct eigenvalues such that also the negatives of some of them are eigenvalues.

6.3.15 Let $r^*(x) = r(x^2)$ for any polynomial r. Then $r(\mathcal{A}^2) = \mathcal{O} \iff m_{\mathcal{A}} \mid r^*$, and if $\lambda \neq 0$, then the multiplicities of the roots λ^2 in r and λ in r^* are the same.

6.3.16 If $(h, m_{\mathcal{A}}) = d \neq 1$, then $h(\mathcal{A})(m_{\mathcal{A}}/d)(\mathcal{A}) = \mathcal{O}$ implies that $h(\mathcal{A})$ is (zero or) a zero divisor, so it cannot have an inverse. If $(h, m_{\mathcal{A}}) = 1$, then $1 = hr + sm_{\mathcal{A}}$ with suitable polynomials r and s, and $h(\mathcal{A})r(\mathcal{A}) = \mathcal{I}$.

6.3.18 If $f = \alpha_0 + \alpha_1 x + \ldots + \alpha_k x^k$, then let $\mathbf{b}_1, \ldots, \mathbf{b}_k$ be a basis in V and

$$\mathcal{A}\mathbf{b}_1 = \mathbf{b}_2, \ldots, \mathcal{A}\mathbf{b}_{k-1} = \mathbf{b}_k, \mathcal{A}\mathbf{b}_k = \sum_{i=0}^{k-1} (-\alpha_i/\alpha_k)\mathbf{b}_{i+1}\,.$$

6.3.19 T6.3.5 does not follow from the T6.2.1 and T6.3.4 as one can set the multiplicities of the roots in two polynomials so that none of them divides the other though they have exactly the same roots.
None of T6.3.4 and T6.2.1 follows from the other two since if $f \mid g$, then not necessarily all roots of g are roots of f or, in a reverse perspective, g can also have further roots besides the roots of f.

6.3.20 Verify $S^{-1}f(A)S = f(S^{-1}AS)$ for any polynomial $f \in F[x]$.

6.3.21 To decide whether or not A is a root of a non-zero polynomial $f \in F[x]$ of degree not greater than k we have to check whether the matrices $I, A, A^2, \ldots, A^k$ are linearly dependent or independent in $F^{n\times n}$. Equivalently, we have to determine whether or not the corresponding homogeneous system of linear equations has a non-trivial solution. As we remain in the field of the coefficients during the Gaussian elimination, we get the same reduced echelon form if the real coefficients are considered as complex numbers.
Remark: Also in general, if K is a subfield of F, i.e., $K \subseteq F$ and F is a field under the operations of F, then the minimal polynomial of a matrix $A \in K^{n\times n}$ over K is its minimal polynomial over F, too. Also the analogous statement for the characteristic polynomial is true, this follows directly from the definition.

6.3.22 (a) $m_A \mid k_A$, so every irreducible factor of m_A must occur in the decomposition of k_A. For the converse, let f be an irreducible factor of k_A.

Consider A as a matrix A' with complex entries, then $k_{A'} = k_A$ and $m_{A'} = m_A$ (see the remark after the hint to the previous exercise). By the fundamental theorem of algebra, f is the product of root factors (apart from a constant multiplier) and these roots are distinct due to irreducibility (see Exercise A.7.10). All these are roots of $k_{A'} = k_A$, so they are eigenvalues of A', hence they are roots of $m_{A'} = m_A$, too. This means that the same root factors occur in m_A, i.e., $f \mid m_A$.

(b) Necessity: $A \in \mathbf{Q}^{n\times n}$ is a root of the irreducible polynomial $x^3 - 2$, so this is the minimal polynomial. By part (a), the characteristic polynomial must be a power of it implying $3 \mid n$. Sufficiency: For $n = 3$, we can construct a matrix from the hint of Exercise 6.3.18. For an arbitrary $n = 3k$ we can assemble the matrix from such 3×3 blocks.

6.4.

6.4.2 (d) The kernel of the restriction of $\mathcal{A}$ to U is $\mathbf{0}$, so its image is the entire U.

6.4.4 (a) We have to check that a scalar multiple of such a transformation and the sum and product of two such transformations have this property.

(b) Instead of the transformations consider their matrices in a basis where the first k elements are in U.

6.4.5 (b) Use Exercise 6.4.1 to prove that every subspace is invariant.

6.4.6 For any U eigenspace of $\mathcal{A}$: If $\mathbf{u} \in U$, i.e., $\mathcal{A}\mathbf{u} = \lambda\mathbf{u}$ for some λ, then $\mathcal{A}(\mathcal{B}\mathbf{u}) = \mathcal{B}(\mathcal{A}\mathbf{u}) = \mathcal{B}(\lambda\mathbf{u}) = \lambda(\mathcal{B}\mathbf{u})$, so $\mathcal{B}\mathbf{u} \in U$.

6.4.7 (b) Consider, e.g., $[\mathcal{A}] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $[\mathcal{B}] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

6.4.8 Verify that if a polynomial is in an invariant subspace, then also every polynomial of the same or smaller degree is an element of this subspace.

6.4.9 (b) Let $(f, m_{\mathcal{A}}) = d$. Verify first $\operatorname{Ker} f(\mathcal{A}) = \operatorname{Ker} d(\mathcal{A})$; here one of the containments is obvious and the other follows from the representation $d = sf + tm_{\mathcal{A}}$ of the gcd. Continue by proving that if $\operatorname{Ker} d_1(\mathcal{A}) = \operatorname{Ker} d_2(\mathcal{A})$ for two divisors d_1 and d_2 of the minimal polynomial, then d_1 is a constant multiple of d_2. Using $(d_1\,,\, d_2)$, we can reduce it to the case $d_1 \,|\, d_2$. If here $m_{\mathcal{A}} = d_2 h = r d_1 h$, then the condition implies

$d_1(\mathcal{A})h(\mathcal{A}) = \mathcal{O}$, so r must be a constant by the definition of the minimal polynomial.

(c) By part (b), this is the number of distinct subspaces of the form $\operatorname{Ker} f(\mathcal{A})$.

6.4.10 Apply similar arguments as in the previous exercise.

6.4.11 (b) Consider, e.g., $[\mathcal{A}] = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $f = g = x$.

6.5.

6.5.2 Adapt the proofs of Theorems 6.3.2 and 6.3.3.

6.5.4 Use the result of the previous exercise and the connection between the minimal polynomial and the eigenvalues.

6.5.7 If $i > 0$, then $\mathcal{A}^i\mathbf{u} \in \operatorname{Im}\mathcal{A}$.

6.5.8 Let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be eigenvectors belonging to the eigenvalues $\lambda_1, \ldots, \lambda_n$, they form a basis. Every subspace spanned by one of the 2^n subsets of this basis is invariant (the zero subspace is spanned by the empty subset). We have to prove that there are no more invariant subspaces. Verify that if a vector $\mathbf{v} = \beta_1\mathbf{u}_1 + \ldots + \beta_n\mathbf{u}_n$ is an element of an invariant subspace U and $\beta_i \neq 0$, then $\mathbf{u}_i \in U$. This follows if we apply the transformation $f_i(\mathcal{A})$ to $\mathbf{v}$ where $f_i = m_{\mathcal{A}}/(x - \lambda_i)$.

6.5.9 Every divisor of the minimal polynomial is the order of some vector $\mathbf{u}$ and the subspaces $\langle \mathbf{u}, \mathcal{A}\rangle$ belonging to pairwise non-associate divisors are distinct.

6.5.10 Adapt the ideas used in verifying (i) in the proof of Lemma 6.5.7.

6.5.11 Use the pattern seen at Exercise 6.3.14. Or we can simply refer to the results for minimal polynomials, by Theorem 6.5.6. A third option is to apply Theorem 6.5.4.

6.6.

6.6.1 To prove sufficiency, use Theorem 6.6.2.

6.6.2 (a) Eigenvectors belonging to the distinct eigenvalues of $\mathcal{A}$ form a basis and these are eigenvectors of $\mathcal{B}$, too.

(b) The transformations commuting with $\mathcal{A}$ form an n-dimensional subspace in $\operatorname{Hom} V$. This contains the subspace of the polynomials of $\mathcal{A}$. The latter is n-dimensional, as well, so the two subspaces are equal. The easiest way to verify all this is to work with matrices in the basis of the eigenvectors of $\mathcal{A}$. Another option is to use the interpolation polynomial (see Theorem 3.2.4).

6.6.4 (a) Use Theorem 6.5.3.

(b) If, e.g., $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ are linearly independent eigenvectors with eigenvalues 1, 1, and 2 and $U_1 = \langle \mathbf{b}_1, \mathbf{b}_3 \rangle$, $U_2 = \langle \mathbf{b}_2, \mathbf{b}_3 \rangle$, then $m_\cap = (x-2) \neq (m_1, m_2) = (x-1)(x-2)$.

6.6.5 (a) and (b): There is no equality even for $\mathcal{A} = \mathcal{I}$ in general.

(c) Extend a basis of $U_1 \cap U_2$ to bases of U_1 and U_2, their union is a basis of $\langle U_1, U_2 \rangle$. Write the matrix of $\mathcal{A}$ in this basis to compute the characteristic polynomial.

6.6.6 We show first that if $\deg m_\mathcal{A} < \dim V$, then $\mathcal{A}$ has infinitely many invariant subspaces. Since $\dim\langle \mathbf{u}, \mathcal{A} \rangle = \deg o(\mathbf{u}) \leq \deg m_\mathcal{A} < \dim V$, so $\langle \mathbf{u}, \mathcal{A} \rangle$ is a non-trivial subspace for every $\mathbf{u} \neq \mathcal{A}$. The union of these subspaces is V. Over an infinite field, the union of finitely many non-trivial subspaces cannot be equal to V (Exercise 4.2.12(e)), so there must be infinitely many invariant subspaces already of type $\langle \mathbf{u}, \mathcal{A} \rangle$. (For a finite field we can just say that V is not a subspace of the form $\langle \mathbf{u}, \mathcal{A} \rangle$, so, by Exercise 6.5.9, the number of invariant subspaces is bigger than the number of pairwise non-associate divisors of the minimal polynomial.)

Next we show that if $\deg m_\mathcal{A} = \dim V$, then (both over finite and infinite fields) every invariant subspace is of the form $\operatorname{Ker} f(\mathcal{A})$. Then we are done since, by Exercise 6.4.9, the number of such subspaces is equal to the number of pairwise non-associate divisors of the minimal polynomial.

Let $\mathbf{v}$ be a vector satisfying $o(\mathbf{v}) = m_\mathcal{A}$. Then $\dim\langle \mathbf{v}, \mathcal{A} \rangle = \deg o(\mathbf{v}) = \deg m_\mathcal{A} = \dim V$, so $\langle \mathbf{v}, \mathcal{A} \rangle = V$.

Let $f \mid m_\mathcal{A}$, i.e., $m_\mathcal{A} = fg$. Then $\operatorname{Im} f(\mathcal{A}) = \langle f(\mathcal{A})\mathbf{v}, \mathcal{A} \rangle$, so $\dim \operatorname{Im} f(\mathcal{A}) = \deg o[f(\mathcal{A})\mathbf{v}] = \deg g$. By the Dimension Theorem, $\dim \operatorname{Ker} f(\mathcal{A}) = \deg f$.

Consider now an arbitrary invariant subspace and let f denote the minimal polynomial of the restriction of $\mathcal{A}$ to U. We show that $U = \operatorname{Ker} f(\mathcal{A})$.

Let $\mathbf{u} \in U$ be a vector satisfying $o(\mathbf{u}) = f$. Then $\langle \mathbf{u}, \mathcal{A} \rangle \subseteq U \subseteq \operatorname{Ker} f(\mathcal{A})$. Further, $\dim \langle \mathbf{u}, \mathcal{A} \rangle = \dim \operatorname{Ker} f(\mathcal{A}) = \deg f$, so $U = \operatorname{Ker} f(\mathcal{A})$.

6.6.7 Combine the previous exercise with Exercises 6.4.9, 6.4.10, and 6.5.9.

6.6.8 (d) Consider, e.g., $\mathcal{D} = -\mathcal{A}$ and $\mathcal{D} = \mathcal{A}^{-1}$.

(f) Concerning the converse, e.g., the minimal and characteristic polynomials of $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ are equal, but the matrices are not similar since they differ in the dimensions of the eigenspaces belonging to the eigenvalue 1.

6.6.9 Combine the results of parts (a) and (f) in the previous exercise.

6.6.14 Use the Jordan form.

6.6.16 Use that the Jordan form is similar to its transpose. This latter follows from a simple permutation of the basis elements.

6.6.17 (a) Consider A as a matrix of a linear transformation $\mathcal{A}$. Computing the characteristic polynomial, we see that there exists an eigenbasis. The matrix of $\mathcal{A}$ in this basis is a diagonal matrix C. By Theorem 5.8.1A, $C = S^{-1}AS$ where S is the accompanying matrix in switching the bases. Hence $A = SCS^{-1}$, so $A^n = SC^nS^{-1}$. The exponentiation of a diagonal matrix happens per elements, further S and S^{-1} are easy to determine.

(b) In this case, we have no eigenbasis. So we use the Jordan form as C and compute its powers according to Exercise 6.6.12.

7. Bilinear Functions

7.1.

7.1.5 Verify that the vector space is isomorphic to $F^{n\times n}$.

7.1.8 (c) Apply part (b).

7.2.

7.2.1 If $\mathbf{A}$ is antisymmetric, then applying $\mathbf{A}(\mathbf{u},\mathbf{v}) = -\mathbf{A}(\mathbf{v},\mathbf{u})$ with $\mathbf{u} = \mathbf{v} = \mathbf{x}$ we get $\mathbf{A}(\mathbf{x},\mathbf{x}) = -\mathbf{A}(\mathbf{x},\mathbf{x})$, so $\mathbf{A}(\mathbf{x},\mathbf{x}) = 0$. To prove the converse, expand $\mathbf{A}(\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v})$ (see the formula at the beginning of the second proof of Theorem 7.2.3).

7.2.2 Let $\mathbf{B}$ be an arbitrary bilinear function and assume that $\mathbf{B}(\mathbf{u},\mathbf{v}) = \mathbf{S}(\mathbf{u},\mathbf{v}) + \mathbf{A}(\mathbf{u},\mathbf{v})$ where $\mathbf{S}$ is symmetric and $\mathbf{A}$ is antisymmetric. Swapping $\mathbf{u}$ and $\mathbf{v}$ and using the symmetric and antisymmetric properties, we

get another equation. Solving the system of these two equation, we can express $\mathbf{S}$ and $\mathbf{A}$ with $\mathbf{B}$ uniquely: $\mathbf{S}(\mathbf{u},\mathbf{v}) = (1/2)\big(\mathbf{B}(\mathbf{u},\mathbf{v}) + \mathbf{B}(\mathbf{v},\mathbf{u})\big)$ and $\mathbf{A}(\mathbf{u},\mathbf{v}) = (1/2)\big(\mathbf{B}(\mathbf{u},\mathbf{v}) - \mathbf{B}(\mathbf{v},\mathbf{u})\big)$. This means that only this $\mathbf{S}$ and $\mathbf{A}$ can satisfy the conditions. Finally we have to verify that these functions are symmetric and antisymmetric, indeed.

7.2.3 (c) This follows from the previous exercise.

7.2.8 $\mathbf{w} = \mathbf{u} + \lambda\mathbf{v}$ works for a suitable scalar λ. Another option: By the conditions, the main diagonal of a suitable diagonal matrix contains both 1 and -1. The sum of the corresponding basis vectors satisfies the requirement on $\mathbf{w}$.

7.3.

7.3.1 (a) It follows from Exercise 7.2.1.

(b) Expand $\mathbf{A}(\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v})$. If $\mathbf{A}$ is symmetric, then we can express $\mathbf{A}(\mathbf{u},\mathbf{v})$ from the quadratic form as $\mathbf{A}(\mathbf{u},\mathbf{v}) = (1/2)\big(\tilde{\mathbf{A}}(\mathbf{u}+\mathbf{v}) - \tilde{\mathbf{A}}(\mathbf{u}) - \tilde{\mathbf{A}}(\mathbf{v})\big)$. This shows that every quadratic form is derived from at most one symmetric bilinear function. To complete the proof, we have to verify that the bilinear function obtained above is symmetric.

7.3.9 Multiplying a basis element by λ, the determinant is multiplied by λ^2. To show the invariance of the sign, observe that the elementary equivalent transformations in the third proof of Theorem 7.2.3 do not alter the sign of the determinant, and these transformations lead to a diagonal matrix where the sign of the determinant is uniquely determined by the law of inertia.

7.4.

7.4.1 If the coordinates of $\mathbf{u}$ and $\mathbf{v}$ are $u_1,\dots,u_n$ and $v_1,\dots,v_n$, then $\mathbf{u}\cdot\mathbf{v} = \sum_{j=1}^{n} \overline{u_j} v_j$. This is a positive definite Hermitian bilinear function.

7.4.4 The only changes compared to the real case are that we have to use the conjugate of the scalar for the rows in (b) and (c):

(b) The third row is multiplied by $\overline{\lambda}$ and the third column is multiplied by λ (so α_{33} gets multiplied by $|\lambda|^2$).

(c) We add $\overline{\mu}$ times the second row to the third row and add μ times the second column to the third column (so the third element in the new third row is $\alpha'_{33} = \alpha_{33} + \overline{\mu}\alpha_{23} + \mu\alpha_{32} + |\mu|^2\alpha_{22}$).

7.4.5 Replace symmetry by the Hermitian property and use the conjugate when needed.

7.4.7 (b) Use diagonal matrices for the examples.

7.4.8 (b) Check that, on the one hand, i times an Hermitian function is skew-Hermitian, and on the other hand, $1/i$ times a skew-Hermitian function is Hermitian.

(d) Argue similarly as in Exercise 7.2.2.

7.4.10 Check directly that the scalar multiples of Hermitian functions satisfy this condition. To prove the converse, select a δ satisfying $\mathbf{A}(\mathbf{u}, \mathbf{v} - \delta\mathbf{u}) = 0$, and apply the property. After some transformations we get $\mathbf{A}(\mathbf{u}, \mathbf{v}) = \theta\overline{\mathbf{A}(\mathbf{v}, \mathbf{u})}$ where θ does not depend on $\mathbf{v}$. Further maneuvers yield that it does not depend on $\mathbf{u}$ either. Finally, $|\theta| = 1$ implies that $(1/\sqrt{\theta})\mathbf{A}$ is Hermitian. A more systematic argument for the converse is based on the trick $\mathbf{A}\big(\mathbf{u}, \mathbf{A}(\mathbf{u}, \mathbf{v})\mathbf{u} - \mathbf{A}(\mathbf{u}, \mathbf{u})\mathbf{v}\big) = 0$. A third way to prove the converse is to perform a (modified) Gaussian elimination on a matrix of $\mathbf{A}$. The condition implies that obtaining zeros in a column will automatically yield zeros in the corresponding row. So, we can get a diagonal matrix. Factoring out a non-zero element of the main diagonal, we have to show that the remaining entries will be real numbers.

8. Euclidean Spaces

8.1.

8.1.1 The simplest way is to show that the second four-tuple of vectors is an orthonormal basis with respect to the inner product defined by the first four-tuple.

8.1.2 The inner product of $\sum_{j=1}^{k} \lambda_j \mathbf{a}_j = \mathbf{0}$ and $\mathbf{a}_j$ yields $\lambda_j = 0$.

8.1.3 It follows, e.g., from the Gram–Schmidt orthogonalization.

8.1.6 It follows from the first proof of Theorem 8.1.7.

8.1.8 The second condition implies $W \subseteq U^\perp$ and the first condition implies $\dim W \geq \dim U^\perp$.

8.1.9 (c) Apply (b) with $U_i^\perp$ instead of U_i and take the orthogonal complement of both sides.

8.1.10 (b) Let the coordinates of $\mathbf{v}_j$ be, e.g., $j, j^2, \ldots, j^n$ in an orthonormal basis.

8.1.11 The condition $\mathbf{b}_i \cdot \mathbf{c}_j = 0$ $(i \neq j)$ implies that $\mathbf{c}_j$ is in the one-dimensional subspace $\langle \mathbf{b}_1, \ldots, \mathbf{b}_{j-1}, \mathbf{b}_{j+1}, \ldots, \mathbf{b}_n \rangle^\perp$. Then $\mathbf{b}_j \cdot \mathbf{c}_j = 1$ determines $\mathbf{c}_j$ uniquely. Finally, we can show similar to Exercise 8.1.2 that the vectors $\mathbf{c}_j$ are linearly independent.

8.1.12 Define the inner product by a basis of V that is the union of bases of U and W.

8.1.13 (b) Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis. Then $\mathbf{c}$ satisfies the requirement if and only if its coordinates are $\Psi(\mathbf{e}_i)$.

8.1.14 We know that the equality of dimensions is equivalent to the isomorphism of the vector spaces. So we only have to show that in the case of equal dimensions there exists a vector space isomorphism that preserves the inner product, too. A linear map satisfies this property if and only if it carries an orthonormal basis into an orthonormal basis.

8.1.15 (a) The only difficulty lies in showing that the sum of two such sequences is in V: check $\sum_{j=1}^{\infty}(\alpha_j + \beta_j)^2 \leq 2\sum_{j=1}^{\infty} \alpha_j^2 + 2\sum_{j=1}^{\infty} \beta_j^2$.

(b) Do not forget to show that $\sum_{j=1}^{\infty} \alpha_j \beta_j$ is convergent.

8.1.16 (b) Direction $\Rightarrow$ is false; see, e.g., Exercise 8.1.15(c).

(c) Equality fails, e.g., for U in Exercise 8.1.15(c).

(d) One containment follows by taking the orthogonal complements in (c). To get the converse containment, apply (c) with $U^\perp$ instead of U.

8.1.17 Theorem 8.1.7: false. Exercise 8.1.9: one of the directions in (a) is false; (b) is true; (c) is false. Counterexamples can be obtained from Exercise 8.1.15(c).

8.2.

.

8.2.4 (e) To prove the triangle inequality, verify first

$$(\dagger) \qquad a_1 b_1 + \ldots + a_n b_n \leq (a_1^3 + \ldots + a_n^3)^{\frac{1}{3}} \cdot (b_1^{3/2} + \ldots + b_n^{3/2})^{\frac{2}{3}}$$

for non-negative values of a_j and b_j. Multiplying every a_j or b_j with a positive number, we get an equivalent inequality, so we can

assume that both factors on the right-hand side are 1. Applying the inequality of arithmetic and geometric means to $a^3, b^{3/2}$, and $b^{3/2}$, we get $ab \le a^3/3 + 2b^{3/2}/3$. Adding it up for all pairs a_j, b_j, we arrive at the desired inequality $a_1b_1 + \ldots + a_nb_n \le 1$. Turning to the triangle inequality, it suffices to prove

$$S = \left(\sum_{j=1}^{n}(c_j + d_j)^3\right)^{\frac{1}{3}} \le \left(\sum_{j=1}^{n} c_j^3\right)^{\frac{1}{3}} + \left(\sum_{j=1}^{n} d_j^3\right)^{\frac{1}{3}} = T + U$$

for non-negative values of c_j and d_j. We write the cube of the left-hand side as $S^3 = \sum_{j=1}^{n} c_j(c_j + d_j)^2 + \sum_{j=1}^{n} d_j(c_j + d_j)^2$ and apply (†) to both sums with $b_j = (c_j + d_j)^2$ and $a_j = c_j$ or d_j. This yields $S^3 \le TS^2 + US^2$, i.e., $S \le T + U$ as stated.

Remark. We can define a norm similarly for any exponent $p > 1$ instead of cubes. The corresponding triangle inequality is *Minkowski's inequality*, and the general form of (†) is *Hölder's inequality*

$$a_1b_1 + \ldots + a_nb_n \le (a_1^p + \ldots + a_n^p)^{\frac{1}{p}} \cdot (b_1^q + \ldots + b_n^q)^{\frac{1}{q}}$$

with $1/p + 1/q = 1$. In the special case $p = 2$, Hölder's inequality turns into CBS. The proof for $p = 3$ sketched above can be adapted to any rational value of p.

8.2.5 (a) See the examples of the previous exercise.

(b) One of the directions follows from Exercise 8.2.3(c). To prove the converse, express the inner product of two vectors with the norm: $\mathbf{x}\cdot\mathbf{z} = (1/4)(\|\mathbf{x} + \mathbf{z}\|^2 - \|\mathbf{x} - \mathbf{z}\|^2)$, and verify that this defines an inner product if the condition holds. The difficult part is to prove (bi)linearity. Applying the condition to the pairs of vectors $\mathbf{x} = \mathbf{u}+\mathbf{v}$, $\mathbf{z} = \mathbf{w}$, and $\mathbf{x} = \mathbf{u} - \mathbf{v}$, $\mathbf{z} = \mathbf{w}$, and subtracting the two equalities we obtain that addition is preserved in the first variable. This implies preserving the multiplication by an integer scalar and, more generally, by a rational scalar. To extend it to any real λ, we have to show that $f(\lambda) = (\lambda\mathbf{x})\cdot\mathbf{z} - \lambda(\mathbf{x}\cdot\mathbf{z})$ is continuous everywhere. This is obvious for the second term. The first term can be written with the norm as $(\lambda\mathbf{x})\cdot\mathbf{z} = (1/4)(\|\lambda\mathbf{x} + \mathbf{z}\|^2 - \|\lambda\mathbf{x} - \mathbf{z}\|^2)$. We show that $\|\lambda\mathbf{x} + \mathbf{z}\|$ is continuous (the argument is the same for the other term). Using the triangle inequality and factoring out the scalar, we get

$$\|\lambda\mathbf{x} + \mathbf{z}\| - \|\mu\mathbf{x} + \mathbf{z}\| \le \|(\lambda - \mu)\mathbf{x}\| = |\lambda - \mu| \cdot \|\mathbf{x}\|\,.$$

8.2.8 Put $z_j = 1$ in the second proof of CBS.

8.2.12 Show that, in accordance with our experience from geometry, the closest vector to $\mathbf{v}$ in a subspace U is the orthogonal projection of $\mathbf{v}$.

8.2.13 We are looking for the solutions of $A\mathbf{z} = \mathbf{b}'$ where $\mathbf{b}'$ is the closest vector to $\mathbf{b}$ in $\operatorname{Im} A$. So, $\mathbf{b}'$ is the orthogonal projection of $\mathbf{b}$ into $\operatorname{Im} A$. Substituting the solutions of $A\mathbf{z} = \mathbf{b}'$ into the original system of equations will give the minimal sum of squares of differences from the values on the right-hand side, i.e., from the corresponding components of $\mathbf{b}$. In the given system of equations, $\mathbf{b}' = \begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix}$, $\mathbf{z} = \begin{pmatrix} -3+\mu+2\nu \\ 3-2\mu-3\nu \\ \mu \\ \nu \end{pmatrix}$ where μ are ν arbitrary real numbers, and substituting these values, we get 6 as the minimal sum of squares of differences.

8.2.14 (a) The inner product of $\mathbf{x} = \sum_{j=1}^{n} \lambda_j \mathbf{e}_j$ and $\mathbf{e}_i$ yields $\lambda_i = \mathbf{x}\cdot\mathbf{e}_i$.
(b) Form the inner product $\mathbf{x}\cdot\mathbf{z}$ using the representation in (a).
(c) Apply (b) with $\mathbf{z} = \mathbf{x}$.
Each of the three formulas can also be directly obtained using coordinates.

8.2.15 Inequality follows from Exercise 8.2.14(c).

8.2.16 The second and third proofs of CBS can easily be adapted to the semidefinite case. Note that equality can also hold for linearly independent vectors.

8.2.18 The first and third proofs work for the infinite-dimensional case, too.

8.3.

8.3.1 The proofs given in the real case are generally valid without any modification. Note, that in Exercises 8.2.14–8.2.15 the order of the factors in the (generally complex-valued) inner products is important and the squares of the *absolute values* of the inner product occur.

8.3.2 (a) Use $i\mathbf{x} + \mathbf{z} = i(\mathbf{x} - i\mathbf{z})$.
(b) Expanding the inner product $(\mathbf{x} + i\mathbf{z})\cdot(i\mathbf{x} + \mathbf{z})$, separate the real and imaginary parts.

8.3.4 Adaptation of the second proof of Theorem 8.2.8: The square of the inequality expressed with coordinates in an orthonormal basis is

$$|\overline{x_1}z_1 + \ldots + \overline{x_n}z_n|^2 \le (|x_1|^2 + \ldots + |x_n|^2)(|z_1|^2 + \ldots + |z_n|^2).$$

Applying $|w|^2 = \overline{w}w$, this can be transformed into

$$0 \le \sum_{1 \le i < j \le n} |x_i z_j - x_j z_i|^2.$$

Adaptation of the third proof: We start from the inequality

$$0 \le \|\lambda\mathbf{x} + \mathbf{z}\|^2 = (\lambda\mathbf{x} + \mathbf{z})\cdot(\lambda\mathbf{x} + \mathbf{z}) = |\lambda|^2(\mathbf{x}\cdot\mathbf{x}) + 2\mathrm{Re}[\lambda(\mathbf{z}\cdot\mathbf{x})] + \mathbf{z}\cdot\mathbf{z}$$

valid for every (complex) λ. Choose $\lambda = \mu\rho$ where μ is an arbitrary real number and ρ is a complex number on the unit circle so that $\rho(\mathbf{z}\cdot\mathbf{x})$ is a positive real number. This can always be achieved except if $\mathbf{z}\cdot\mathbf{x} = 0$ but then CBS is true trivially. Using our inequality for this λ, we get $0 \le \mu^2(\mathbf{x}\cdot\mathbf{x}) + 2\mu|\mathbf{z}\cdot\mathbf{x}| + \mathbf{z}\cdot\mathbf{z}$ for every real μ. From here, we can finish the proof as in the real case.

8.3.5 (a) If $\mathbf{z} \neq \mathbf{0}$, then factoring out a non-zero component of $\mathbf{z}$, we get $\mathbf{z} = \alpha\mathbf{w}$ where $\mathbf{w}$ has a component 1. If $\mathbf{w}$ and $\overline{\mathbf{w}}$ are scalar multiples of each other, then this scalar has to be 1, so $\mathbf{w} = \overline{\mathbf{w}}$, i.e., $\mathbf{w}$ is a real vector.

(b) In checking the scalar multiple, be aware of $\overline{\mu\mathbf{x}} = \overline{\mu} \cdot \overline{\mathbf{x}} (\neq \mu\overline{\mathbf{x}})$.

(d) Sufficiency: represent the elements of U in this basis of real vectors. Necessity: For any $\mathbf{0} \neq \mathbf{z} \in U$, we have $\overline{\mathbf{z}} \in U$, so the real vector $\mathbf{v} = \overline{\mathbf{z}} + \mathbf{z} \in U$. If $\mathbf{v} = \mathbf{0}$, then $i\mathbf{z}$ is a real vector. So, we found anyway a non-zero real vector $\mathbf{b}$ in U, this will be the first element of the basis. Let W be the orthogonal complement of $\langle\mathbf{b}\rangle$ in U. We show $\overline{W} = W$, so repeating the previous procedure on W, finally we arrive at a real (orthonormal) basis of U. Conjugating the representation $U = \langle\mathbf{b}\rangle \oplus W$, we get $\overline{U} = \overline{\langle\mathbf{b}\rangle} \oplus \overline{W}$. Here $\overline{U} = U$ and $\mathbf{b}$ is real, so we can omit the conjugate sign on the left-hand side and in the first term of the right-hand side. The orthogonal complement is unique, so $\overline{W} = W$.

(e) Using the relation between the dimensions, we get that n cannot be odd. For n even, we choose the jth basis element of U with $2j-1$st component 1, $2j$th component i, and all other components 0, $j = 1, 2, \ldots, n/2$.

8.4.

8.4.1 Proof of the law for the product: $((\mathcal{A}\mathcal{B})\mathbf{x})\cdot\mathbf{z} = (\mathcal{A}(\mathcal{B}\mathbf{x}))\cdot\mathbf{z} = (\mathcal{B}\mathbf{x})\cdot(\mathcal{A}^*\mathbf{z}) = \mathbf{x}\cdot((\mathcal{B}^*(\mathcal{A}^*\mathbf{z})) = \mathbf{x}\cdot((\mathcal{B}^*\mathcal{A}^*)\mathbf{z})$. Comparing the first and last terms, we

obtain $(\mathcal{AB})^* = \mathcal{B}^*\mathcal{A}^*$ from the definition and uniqueness of $(\mathcal{AB})^*$. The other parts of the exercise can be verified similarly. Another option is to switch to matrices in an orthonormal basis and apply the similar laws for matrices (Exercise 2.1.20).

8.4.2 The simplest approach is to work with matrices in an orthonormal basis.

8.4.4 (a) Let $h = \beta_0 + \beta_1 x + \beta_2 x^2$. Due to linearity, $\mathcal{A}^*h = \beta_0\mathcal{A}^*1 + \beta_1\mathcal{A}^*x + \beta_2\mathcal{A}^*x^2$, so it is sufficient to determine the images of the basis vectors. To compute $\mathcal{A}^*1 = r$, we apply

$$\int_{-1}^{1} f''(t)dt = (\mathcal{A}1)\cdot f = f\cdot(\mathcal{A}^*1) = \int_{-1}^{+1} f(t)r(t)dt$$

for $f = 1, x$, and x^2. This means

$$0 = \int_{-1}^{1} r(t)dt, 0 = \int_{-1}^{1} tr(t)dt, \text{ and } 4 = \int_{-1}^{1} t^2 r(t)dt.$$

The first integral yields $r = \alpha(-3x^2+1)$ and this r satisfies the second integral, too. Plugging it into the third integral, we get $\alpha = -15/2$. So, $\mathcal{A}^*1 = 15(3x^2-1)/2$. A similar argument yields $\mathcal{A}^*x = 0$ and $\mathcal{A}^*x^2 = 5(3x^2-1)/2$. Combining these, we get $\mathcal{A}^*h$ for the general h above.

8.4.5 $(\mathcal{A}^*\mathbf{x})\cdot(\mathcal{A}\mathbf{x}) = \mathbf{x}\cdot(\mathcal{A}^2\mathbf{x}) = 0$. Concerning the converse, if the quadratic form $\mathbf{x}\cdot(\mathcal{A}^2\mathbf{x})$ is everywhere zero, then over $\mathbf{C}$, by Theorem 7.4.3, also the bilinear function $\mathbf{u}\cdot(\mathcal{A}^2\mathbf{v})$ is everywhere zero, so $\mathcal{A}^2 = \mathcal{O}$. Over $\mathbf{R}$, however, only $\mathbf{u}\cdot(\mathcal{A}^2\mathbf{v}) = -\mathbf{v}\cdot(\mathcal{A}^2\mathbf{u})$ follows by Exercise 7.3.1(a), and this holds, e.g., in the plane if $\mathcal{A}$ is the rotation by 45 degrees around the origin. In a real Euclidean space, $\mathcal{A}^2 + (\mathcal{A}^*)^2 = \mathcal{O}$ is a necessary and sufficient condition to have $\mathcal{A}\mathbf{x} \perp \mathcal{A}^*\mathbf{x}$ for every $\mathbf{x}$.

8.4.6 Let $\mathcal{A}\mathbf{x} = \mu\mathbf{x}$ and $\mathcal{A}^*\mathbf{z} = \nu\mathbf{z}$. Then $\overline{\mu}(\mathbf{x}\cdot\mathbf{z}) = (\mu\mathbf{x})\cdot\mathbf{z} = (\mathcal{A}\mathbf{x})\cdot\mathbf{z} = \mathbf{x}\cdot(\mathcal{A}^*\mathbf{z}) = \mathbf{x}\cdot(\nu\mathbf{z}) = \nu(\mathbf{x}\cdot\mathbf{z})$. So, $\mathbf{x}\cdot\mathbf{z} = 0$ or $\overline{\mu} = \nu$.

8.4.7 Write the characteristic polynomial in an orthonormal basis. For the minimal polynomial, use the definition and Exercise 8.4.1. The assertion for eigenvalues follows from either result on the characteristic and minimal polynomials. Note that there is no close relation between the eigenvectors of $\mathcal{A}$ and $\mathcal{A}^*$ in general.

8.4.8 The proof of $\operatorname{Ker}\mathcal{A}^* = (\operatorname{Im}\mathcal{A})^\perp$:
$\mathcal{A}^*\mathbf{z} = \mathbf{0} \iff \forall \mathbf{x}\ \ \mathbf{x}\cdot(\mathcal{A}^*\mathbf{z}) = 0 \iff \forall \mathbf{x}\ \ (\mathcal{A}\mathbf{x})\cdot\mathbf{z} = 0 \iff \mathbf{z} \perp \operatorname{Im}\mathcal{A}$.

8.4.9 Use the previous exercise or write the matrices in an orthonormal basis and compare the ranks.

8.4.10 Use the first equality in Exercise 8.4.8.

8.4.11 (b) If $\mathcal{A}\mathbf{x} = \mathbf{0}$, then $(\mathcal{A}^*\mathcal{A})\mathbf{x} = \mathcal{A}^*\mathbf{0} = \mathbf{0}$. Conversely, $(\mathcal{A}^*\mathcal{A})\mathbf{x} = \mathbf{0}$ implies $0 = \mathbf{x}\cdot((\mathcal{A}^*\mathcal{A})\mathbf{x}) = (\mathcal{A}\mathbf{x})\cdot(\mathcal{A}\mathbf{x})$, so $\mathcal{A}\mathbf{x} = \mathbf{0}$. The assertion for the images follows from the containment in one direction and the equality of dimensions deduced from the assertion for the kernels.

8.4.12 (a) $(\mathcal{A}\mathbf{x})\cdot(\mathcal{B}\mathbf{z}) = \mathbf{x}\cdot((\mathcal{A}^*\mathcal{B})\mathbf{z})$ proves the statement and its converse.

(b) If $\mathcal{A}\mathbf{x} = \mathcal{B}\mathbf{x} = \mathbf{0}$, then clearly $(\mathcal{A}+\mathcal{B})\mathbf{x} = \mathcal{A}\mathbf{x} + \mathcal{B}\mathbf{x} = \mathbf{0}$. To verify the other containment, $(\mathcal{A}+\mathcal{B})\mathbf{x} = \mathbf{0}$ implies $\mathcal{A}\mathbf{x} = -\mathcal{B}\mathbf{x}$. The left-hand side is in $\operatorname{Im}\mathcal{A}$ and the right-hand side is $\operatorname{Im}\mathcal{B}$, so part (a) implies that both vectors are zero. The converse is false; the condition $\operatorname{Ker}(\mathcal{A}+\mathcal{B}) = \operatorname{Ker}\mathcal{A} \cap \operatorname{Ker}\mathcal{B}$ is independent of the inner product.

8.4.13 The condition $\mathcal{A}^*\mathcal{B} = \mathcal{O}$ implies $\operatorname{Im}\mathcal{A} \cap \operatorname{Im}\mathcal{B} = \mathbf{0}$ by the previous exercise. We have to show $\operatorname{Im}(\mathcal{A}+\mathcal{B}) = \langle \operatorname{Im}\mathcal{A}, \operatorname{Im}\mathcal{B}\rangle$. One containment is obvious. To prove $\operatorname{Im}\mathcal{A}, \operatorname{Im}\mathcal{B} \subseteq \operatorname{Im}(\mathcal{A}+\mathcal{B})$, we observe first that the adjoint of $\mathcal{B}\mathcal{A}^* = \mathcal{O}$ gives $\mathcal{A}\mathcal{B}^* = \mathcal{O}$, so the conditions are symmetric in $\mathcal{A}$ and $\mathcal{B}$. Thus it is sufficient to deal with $\operatorname{Im}\mathcal{A}$. If $\mathbf{x} \in \operatorname{Im}\mathcal{A}^*$, then $\mathcal{B}\mathbf{x} = \mathbf{0}$, so $(\mathcal{A}+\mathcal{B})\mathbf{x} = \mathcal{A}\mathbf{x}$. This implies $\operatorname{Im}(\mathcal{A}+\mathcal{B}) \supseteq \operatorname{Im}(\mathcal{A}\mathcal{A}^*) = \operatorname{Im}\mathcal{A}$ (the last equality follows from Exercise 8.4.11b).

8.5.

8.5.1 (a) The eigenvalues are the elements of the main diagonal of a matrix in an orthonormal eigenbasis.

(b) It is possible that the transformation has no diagonal matrix, or even if it has one, it is not in an orthonormal basis.

8.5.2 If $\mathcal{A}^s = \mathcal{A}^t$ for some $0 < s < t$, then $\mathcal{A}$ is a root of the polynomial $x^t - x^s$, so its roots (that are real numbers according to the previous exercise) can only be 0 and ± 1. Using a matrix in an orthonormal eigenbasis, we obtain $\mathcal{A}^3 = \mathcal{A}$. (We could have switched immediately from the self-adjoint transformation to a diagonal matrix with real entries.)

8.5.3 See the hints to Exercise 8.5.1.

8.5.4 A transformation becomes normal for some inner product if and only if it has a diagonal matrix.

8.5.5 The simplest way is to check the conditions for the matrices in the usual orthonormal basis.

8.5.6 $\lambda\mathcal{A}$ is self-adjoint if and only if $\lambda \in \mathbf{R}$ or $\mathcal{A} = \mathcal{O}$, and $\mathcal{AB}$ is self-adjoint if and only if $\mathcal{AB} = \mathcal{BA}$ (see Exercise 8.5.7).

8.5.7 (a) The argument works if the two transformations share a *common* orthonormal eigenbasis, but this is not true in general. It is easy to construct counterexamples to the claim using (b).

(b) If $\mathcal{A}$ and $\mathcal{B}$ are self-adjoint, then $(\mathcal{AB})^* = \mathcal{B}^*\mathcal{A}^* = \mathcal{BA}$.

8.5.8. Use the existence of an orthonormal eigenbasis. Concerning the converse, it can happen that there is no eigenbasis.

8.5.10 (a) $\|\mathcal{A}\mathbf{x}\|^2 = (\mathcal{A}\mathbf{x})\cdot(\mathcal{A}\mathbf{x}) = \mathbf{x}\cdot(\mathcal{A}^*\mathcal{A}\mathbf{x})$. Similarly, $\|\mathcal{A}^*\mathbf{x}\|^2 = \mathbf{x}\cdot(\mathcal{A}\mathcal{A}^*\mathbf{x})$. Both functions are quadratic forms in $\mathbf{x}$, so they are equal if and only if the corresponding bilinear functions are equal, as we are over the complex field. The argument at the end of the proof of Theorem 8.4.1 shows that equality of the bilinear functions is equivalent to the condition $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$ of normality.

(b) If $\mathcal{A}$ and $\mathcal{A}^*$ have the same eigenvectors, then we can construct an orthonormal eigenbasis similar to the proof of Theorem 8.5.2. For the converse, we can see the eigenvalues from the matrices of $\mathcal{A}$ in an orthonormal eigenbasis.

(c) This follows from (b) and the relation between the corresponding eigenvalues.

(d) and (e) follow from (c) by Exercise 8.4.8.

8.5.11 $\mathcal{AB} = \mathcal{O} \Rightarrow \operatorname{Im}\mathcal{B} \subseteq \operatorname{Ker}\mathcal{A} \Rightarrow \operatorname{Im}\mathcal{B}^* \subseteq \operatorname{Ker}\mathcal{A}^* \Rightarrow (\operatorname{Ker}\mathcal{B})^\perp \subseteq (\operatorname{Im}\mathcal{A})^\perp \Rightarrow \operatorname{Ker}\mathcal{B} \supseteq \operatorname{Im}\mathcal{A} \Rightarrow \mathcal{BA} = \mathcal{O}$.
We used parts (c) and (d) of the previous exercise for $\lambda = 0$ in the second step and applied Exercise 8.4.8 in the third step.

8.5.12 Modify the argument of the proof of Theorem 8.5.2 and also rely on Exercise 8.5.10(b).

8.5.13 Apply the previous exercise. The converse is false, it is easy to find counterexamples among unitary transformations.

8.5.14 Use the previous exercise to prove sufficiency. To verify necessity, consider a diagonal matrix of a normal transformation in an orthonormal

eigenbasis. Write each element in the main diagonal as a product of a real number and a complex number of absolute value 1, and factor the matrix into the product of two diagonal matrices accordingly.

8.5.15 We need an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ where $\langle \mathbf{e}_1, \ldots, \mathbf{e}_k \rangle$ is an invariant subspace of $\mathcal{A}$ for every k. Let $\mathbf{e}_n$ be an eigenvector of $\mathcal{A}^*$ with norm 1. Then $U_n = \langle \mathbf{e}_n \rangle^\perp$ is an invariant subspace of $\mathcal{A}$. To obtain $\mathbf{e}_{n-1}$, repeat the procedure for U_n, etc. (Note that we need $\mathcal{A}^*$ according to the restriction of $\mathcal{A}$ onto U_n and this is not the restriction of the original $\mathcal{A}^*$ in general as U_n is not necessarily an invariant subspace of the original $\mathcal{A}^*$.)

8.5.16 (a) $\mu_1, \ldots, \mu_n$ are the eigenvalues of $\mathcal{A}^*\mathcal{A}$. If $\mathbf{x} \neq \mathbf{0}$ and $\mathcal{A}^*\mathcal{A}\mathbf{x} = \mu\mathbf{x}$, then $\overline{\mu}\|\mathbf{x}\|^2 = (\mu\mathbf{x})\cdot\mathbf{x} = (\mathcal{A}^*\mathcal{A}\mathbf{x})\cdot\mathbf{x} = \|\mathcal{A}\mathbf{x}\|^2$. So, $\mu \geq 0$.

(b) Consider an orthonormal basis providing an upper triangle matrix of $\mathcal{A}$. The elements in the main diagonal are the eigenvalues λ_j. The matrix of $\mathcal{A}^*$ in this basis is a lower triangle matrix with the values $\overline{\lambda_j}$ in the main diagonal. The matrix of $\mathcal{A}^*\mathcal{A}$ is the product of the two matrices. Its trace, i.e., the sum of elements in the main diagonal is the sum of squares of absolute values of the elements in $[\mathcal{A}]$, so it is not less than $\sum_{j=1}^n |\lambda_j|^2$. On the other hand, this trace is $\sum_{j=1}^n \mu_j$.

(c) The previous argument shows that equality holds if and only if every element is 0 outside the main diagonal of the upper triangle matrix.

8.5.17 Sufficiency: If $\mathcal{A}$ is unitary and $\mathbf{u} \perp \mathbf{v}$, then

$$((\lambda\mathcal{A})\mathbf{u})\cdot((\lambda\mathcal{A})\mathbf{v}) = (\lambda(\mathcal{A}\mathbf{u}))\cdot(\lambda(\mathcal{A}\mathbf{v})) = |\lambda|^2(\mathcal{A}\mathbf{u}\cdot\mathcal{A}\mathbf{v}) = |\lambda|^2(\mathbf{u}\cdot\mathbf{v}) = 0.$$

Necessity: Let $\mathcal{A}$ preserve orthogonality and $\mathbf{u}$ and $\mathbf{v}$ be orthogonal unit vectors. We prove $\|\mathcal{A}\mathbf{u}\| = \|\mathcal{A}\mathbf{v}\|$. Indeed, $\mathbf{u} \perp \mathbf{v}$ and $\|\mathbf{u}\| = \|\mathbf{v}\|$ imply $\mathbf{u}+\mathbf{v} \perp \mathbf{u}-\mathbf{v}$, so $0 = (\mathcal{A}(\mathbf{u}+\mathbf{v}))\cdot(\mathcal{A}(\mathbf{u}-\mathbf{v})) = \|\mathcal{A}\mathbf{u}\|^2 - \|\mathcal{A}\mathbf{v}\|^2$ (we used $(\mathcal{A}\mathbf{u}) \perp (\mathcal{A}\mathbf{v})$ in the last equality). So, if $\mathbf{e}_j$ is an orthonormal basis, then the norms $\|\mathcal{A}\mathbf{e}_j\|$ are all equal, let their common values be λ. If $\lambda \neq 0$, then $(1/\lambda)\mathcal{A}$ is unitary. If $\lambda = 0$, then $\mathcal{A} = \mathcal{O}$ is 0 times any unitary transformation.

8.5.18 (a) and (b) These are the matrix versions of $\mathcal{A}^*\mathcal{A} = \mathcal{I}$ and $\mathcal{A}\mathcal{A}^* = \mathcal{I}$.

(c) $1 = \det I = (\det[\mathcal{A}]) \cdot (\det[\mathcal{A}]^*) = (\det[\mathcal{A}]) \cdot \overline{(\det[\mathcal{A}])} = |\det[\mathcal{A}]|^2$.

(d) Apply (c) and the formula with cofactors for the inverse matrix.

8.6.

8.6.1 $\mathcal{A}^2 = \mathcal{I}$ follows from $\mathcal{A}^* = \mathcal{A} = \mathcal{A}^{-1}$. To disprove the converse, consider two non-parallel vectors of different lengths in the plane and let $\mathcal{A}$ be the linear transformation mapping them into each other. A correct statement is that any two conditions below imply the third one: (i) $\mathcal{A}$ is symmetric; (ii) $\mathcal{A}$ is orthogonal; (iii) $\mathcal{A}^2 = \mathcal{I}$.

8.6.2 (a) It follows from Exercise 8.4.6.
(b) Apply a similar argument and use that only ± 1 can be eigenvalues.

8.6.4 The condition leads to a contradiction:
$0 \leq (\mathcal{A}\mathbf{x})\cdot(\mathcal{A}\mathbf{x}) = \mathbf{x}\cdot(\mathcal{A}^*\mathcal{A}\mathbf{x}) = \mathbf{x}\cdot(-\mathbf{x}) < 0$ (for $(\mathbf{x} \neq \mathbf{0})$.

8.6.6 The conditions imply $(\mathcal{A}^*\mathcal{A})^k = \mathcal{I}$. So, the eigenvalues of $\mathcal{A}^*\mathcal{A}$ have absolute value 1. Also, $\mathcal{A}^*\mathcal{A}$ is symmetric with non-negative eigenvalues. Hence each eigenvalue is 1, thus $\mathcal{A}^*\mathcal{A} = \mathcal{I}$. A counterexample for symmetric transformations: $\mathcal{A}$ is a rotation by 90° in the plane and $k = 4$.

8.6.7 We have $\mathcal{A} = (\mathcal{A}^m)^* = \mathcal{A}^{m^2}$, so $\operatorname{Ker}\mathcal{A} = \mathbf{0}$ implies $\mathcal{A}^{m^2-1} = \mathcal{I}$. Thus $(\mathcal{A}^{m-1})^* = \mathcal{A}^{m^2-m} = (\mathcal{A}^{m-1})^{-1}$. Apply now the previous exercise.

8.6.8 For the only if part, represent the gcd as $1 = (k,t) = kq - tr$ with integers $q, r > 0$.

8.6.9 Argue similarly to the complex case.

8.6.11 Adapt the proof of Theorem 8.6.4. A counterexample for the converse: the plane itself is a subspace with the prescribed property, but for the transformation $\mathcal{A}$ with matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ in the usual orthonormal basis, the adjoint is not a polynomial of $\mathcal{A}$.

8.6.12 It is enough to prove the equivalence for the matrices in an orthonormal basis. Considering these real matrices as complex ones, the matrix version of Theorem 8.5.3 verifies the equivalence. We only have to show that if A is a real matrix and A^T is a polynomial of A with complex coefficients, then this also holds with a polynomial with real coefficients. The required representation means a system of linear equations where the unknowns are the coefficients of the polynomial, the coefficients of the system are the entries in the powers of A, and the constants on the right-hand side of the system are the entries in A^T. The coefficients and the constants on the right-hand side of the equations are real numbers. Thus Gaussian elimination remains within the real numbers. So, if there exists a complex solution, there must be a real solution, too.

9. Combinatorial Applications

9.1.

9.1.2 Verify the statement first for positive integer weights by induction, then for arbitrary integers and next for rational numbers. The main difficulty lies in switching from rationals to real numbers. Here are some hints to four distinct proofs; see the details at the solutions.

First proof: Consider the subspace spanned by the 13 real numbers in the usual vector space of real numbers over the rationals, and switch to coordinates (in an arbitrary basis).

Second proof: The condition can be interpreted as a homogeneous system of linear equations (with coefficients 0 and ± 1). It is worth to fix the value of (say) the first variable as 0. Then the statement is equivalent to the system having only a trivial solution. Verify that if this is true among the rational numbers, then the same holds for the reals, as well.

Third proof: Consider the previous system of equations over the modulo 2 field, then turn from here to the real numbers.

Fourth proof: Verify and use that real numbers can be approximated well by rational numbers in the following sense: To any finite set of real numbers and to any $\varepsilon > 0$, there exist fractions with a common denominator such that the difference between each real number and the corresponding fraction is at most ε times the reciprocal of the denominator.

9.1.3 There are many counterexamples, e.g., 12 weights of 1 and 1 weight of 11.

9.1.6 (a) Linear algebra is useful in constructing continuum many sets.

9.1.7 (a) E.g., the vertices of a regular tetrahedron with edges of unit length meet the requirements.

(b) For a proof by contradiction, let the origin be one of the points and let $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2)$, and $\mathbf{c} = (c_1, c_2)$ be the vectors from the origin to the other three points. Expressing the distances with the inner product, show that each of the inner products $\mathbf{a}\cdot\mathbf{a}, \mathbf{b}\cdot\mathbf{b}, \mathbf{c}\cdot\mathbf{c}, 2(\mathbf{a}\cdot\mathbf{b}), 2(\mathbf{a}\cdot\mathbf{c})$, and $2(\mathbf{b}\cdot\mathbf{c})$ is an integer of the form $8k+1$. Verify

$$M = \begin{pmatrix} \mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} & \mathbf{a}\cdot\mathbf{c} \\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b} & \mathbf{b}\cdot\mathbf{c} \\ \mathbf{c}\cdot\mathbf{a} & \mathbf{c}\cdot\mathbf{b} & \mathbf{c}\cdot\mathbf{c} \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}.$$

By Exercise 5.7.12(a), the rank of M is not greater than 2 implying $\det M = 0$. Thus $\det(2M) = 8 \det M = 0$. But

$$\det(2M) \equiv \begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 4 \pmod 8,$$

a contradiction.

9.1.8 (a) Clearly, $c = f(1)$. Verify the statement first for positive integers x, then for positive rationals, for 0, and finally for negative rationals.

(b) Prove that f is continuous everywhere and use the result of (a).

(c) Show that there exists an interval around 0 where f is bounded. This, combined with $f(x/n) = f(x)/n$, implies that f is continuous in 0.

(d) Considering $\mathbf{R}$ as a vector space V over $\mathbf{Q}$, the condition implies that f is a linear transformation on V. The characterization of a linear map by the image of a basis also holds for infinite dimensional vector spaces. Thus we can prescribe the values of f arbitrarily on a (Hamel) basis, this uniquely determines a suitable function defined on all real numbers. (Of course, we cannot visualize these weird functions f.)

9.1.9 (a) If $r_1 = a_1/b_1$ and $r_2 = a_2/b_2$ are any elements in the two subsets, then the condition implies that $a_1a_2 = a_2b_1r_1 = a_1b_2r_2$ is a common element in them.

(b) Considering $\mathbf{R}$ as a vector space over $\mathbf{Q}$, we can express the real numbers as linear combinations of the elements of a(n infinite Hamel) basis. Any such representation involves only finitely many basis vectors, but we can agree that the missing basis vectors appear with 0 coefficient in it. We fix a basis vector $\mathbf{b}$ and partition the positive numbers into two subsets depending on whether or not the coefficient of $\mathbf{b}$ is non-negative in the above representation of the number. Or, another option is to use the sign of the sum of the coefficients (instead of the sign of a single coefficient).

9.1.10 (a1) Two periodic functions defined on $\mathbf{Q}$ share a common period, so their sum is periodic.

(a2) In the vector space V of the real numbers over $\mathbf{Q}$, 1 and $\sqrt{2}$ are linearly independent, hence they can be extended to a basis of V. So, every real number s has a unique representation $s = r_1 + r_2\sqrt{2} + t$, where $r_1, r_2 \in \mathbf{Q}$ and t is an element of the subspace spanned by

the other basis vectors. Then $h_1(s) = r_1$ and $h_2(s) = r_2\sqrt{2} + t$ are periodic functions with periods $\sqrt{2}$ and 1 and their sum is s.

(b1) Prove the impossibility by induction on the degree of the polynomial. Observe that if a polynomial $f(x)$ is the sum of k periodic functions, then $g(x) = f(x) - f(x-1)$ having a degree by one less is the sum of $k-1$ periodic functions.

(b2) Generalizing the construction in (a2), let $b_1, \dots, b_{n+1}$ be linearly independent elements in the (infinite dimensional) vector space of $\mathbf{R}$ over $\mathbf{Q}$, and extend them into a basis. Then every $s \in \mathbf{R}$ has a unique representation $s = \sum_{i=1}^{n+1} r_i b_i + t$, where $r_i \in \mathbf{Q}$ and t is an element of the subspace spanned by the other basis vectors. Raising this equality to the kth power, s^k is the sum of products of k factors where each factor is some $r_i b_i$ or t. If $k \le n$, then at least one $r_i b_i$ is missing from each factor. Expressing the given polynomial $f(s)$ of degree n as the sum of such products, let $h_1(s)$ be the sum of terms containing no $r_1 b_1$, let $h_2(s)$ be the sum of the remaining terms without $r_2 b_2$, etc. Finally, let $h_{n+1}(s)$ be the sum of terms without $r_{n+1} b_{n+1}$. Then $h_i(s)$ is periodic with a period b_i and $f(s) = \sum_{i=1}^{n+1} h_i(s)$.

9.2.

9.2.2 (c) The (complex) roots of the polynomial

$$f = x^4 - x^3 + x^2 - x + 1 = \big(x^{10} - 1\big)\big/\big((x^5 - 1)(x + 1)\big)$$

are the primitive 10th roots of unity (so f is the 10th cyclotomic polynomial), therefore the sequence is periodic with a period 10 for any initial values.

9.2.3 We can adapt any of the three proofs seen at the Fibonacci numbers to verify the statement. The hints below help in handling the extra difficulties caused by multiple roots.

First proof: Note that λ_i is also a root of the first $s_i - 1$ derivatives of f.

Second proof: Use the Jordan-form instead of the diagonal matrix.

Third proof: If the denominator of a partial fraction is a higher power of a root factor, we can expand it into power series using the first or a higher derivative of the geometric series.

9.2.4 Use the previous exercise.

9.2.5 We can eliminate the problematic term 2 by translating the sequence β_n with a suitable constant and thus get a Fibonacci type sequence.

9.2.6 The other term in the formula of φ_n tends to 0.

9.2.7 (b) Let R_{2k} denote a $3 \times (2k)$ rectangle and M_{2k-1} be a $3 \times (2k-1)$ rectangle with one of the corners missing. Let ϱ_{2k} and μ_{2k-1} be the number of tilings of R_{2k} and M_{2k-1} with 2×1 dominoes. Then $\varrho_{2k} = 2\mu_{2k-1} + \varrho_{2k-2}$ and $\mu_{2k-1} = \varrho_{2k-2} + \mu_{2k-3}$. A repeated application of the second equality yields $\mu_{2k-1} = \varrho_{2k-2} + \varrho_{2k-4} + \ldots + \varrho_2 + 1$. Substituting back into the first equation, we get a recursion formula for ϱ_{2k}. Subtracting these expressions for ϱ_{2k} and ϱ_{2k-2}, we arrive at a simple recursion $\varrho_{2k} = 4\varrho_{2k-2} - \varrho_{2k-4}$.

(c) We get the recursion $\psi_n = \psi_{n-1} + \psi_{n-3}$, so the corresponding polynomial is $f = x^3 - x^2 - 1$. Standard calculus methods reveal that f has a single real root $\rho_1 > 1$, thus the other two roots ρ_2 and ρ_3 are complex conjugates. The product of the three roots is 1, so $|\rho_2| = |\rho_3| < 1$. Therefore the last two terms in the formula $\psi_n = \gamma_1\rho_1^n + \gamma_2\rho_2^n + \gamma_3\rho_3^n$ tend to 0 if n tends to infinity. Thus the statement follows from $1.46 < \rho_1 < 1.47$.

9.2.8 Establish a recursion on whether or not the largest element is missing from the subset (cf. Exercise 9.2.10(h)).

9.2.9 Use the greedy algorithm: as a next term, always pick the largest Fibonacci number available.

Remark: More generally, the statement holds for any infinite sequence of positive integers containing 1 where no element is greater than the double of the previous one.

9.2.10 (a) Use induction on n.

(b) It follows from (a) by a suitable substitution.

(c) Use induction or add the equalities $\varphi_{k+1} - \varphi_k = \varphi_{k-1}$ for $k = 1, 2, \ldots, n+1$.

(d) Use induction or add the equalities $(k-2)\varphi_{k+1} - (k-2)\varphi_k = (k-2)\varphi_{k-1}$ for $k = 3, 4, \ldots, n+1$ and reorder the terms.

(e) Use induction or add the equalities $\varphi_k^2 = \varphi_k\varphi_{k+1} - \varphi_k\varphi_{k-1}$ for $k = 1, 2, \ldots, n$.

(f) Derive $\varphi_{2n-1} = \varphi_{n-2}\varphi_n + \varphi_{n-1}\varphi_{n+1}$ from (a) and combine it with (b) to obtain $\varphi_n^2 - \varphi_{n-1}\varphi_{n+1} = -(\varphi_{n-1}^2 - \varphi_{n-2}\varphi_n)$. Then apply this for $n-1, n-2, \ldots$ instead of n.

(g) Apply (a) to decompose $\varphi_{n+2n}, \varphi_{n+n}$, and $\varphi_{(n+1)+n}$ to express φ_{3n} with φ_n and φ_{n-1} and use (f) to obtain the final form of the equality.

(h) Use induction or rely on Exercise 9.2.8.

(i) For any $k \le n$, $\varphi_{2n} = \sum_{j=0}^{k} \binom{k}{j} \varphi_{2n-2k+j}$ follows from a repeated application of $\varphi_t = \varphi_{t-1} + \varphi_{t-2}$. The statement is the special case $k = n$.

9.2.12 Let r_k denote the remainder of φ_k under division by m. The pairs (r_k, r_{k+1}) can assume m^2 distinct values, so there exist $t > s$ satisfying $(r_t, r_{t+1}) = (r_s, r_{s+1})$. Verify $(r_k, r_{k+1}) = (r_{k+t-s}, r_{k+t-s+1})$ for every k, i.e., the remainders r_n are periodic with a period $t - s$. Since $r_0 = 0$, we get $r_{j(t-s)} = 0$ for every j, so $m \,|\, \varphi_{j(t-s)}$.

9.2.13 Apply Exercise 9.2.10(a). Verify $k \,|\, n \Rightarrow \varphi_k \,|\, \varphi_n$ by induction on n/k. To prove the converse and the statement on the gcd, show that if $a = bq + r$, then $(\varphi_a, \varphi_b) = (\varphi_b, \varphi_r)$. Another option is to verify that the indices of the Fibonacci numbers divisible by m are exactly the multiples of the index of the smallest Fibonacci number with this property.

9.2.15 Let $\rho = (1 + \sqrt{5})/2$, then $\varphi_n = \big(\rho^n - (-1/\rho)^n\big)/\sqrt{5}$. The sum of the first m terms of the series can be written as a telescoping sum:

$$\sqrt{5} \sum_{k=1}^{m} \frac{\rho^{2^k}}{\rho^{2^{k+1}} - 1} = \sqrt{5} \sum_{k=1}^{m} \Big(\frac{1}{\rho^{2^k} - 1} - \frac{1}{\rho^{2^{k+1}} - 1} \Big) = \sqrt{5} \Big(\frac{1}{\rho^2 - 1} - \frac{1}{\rho^{2^{m+1}} - 1} \Big) .$$

As the second term tends to 0, the sum of the series is $\sqrt{5}/(\rho^2 - 1)$.

9.2.16 Here are some hints to the three distinct proofs described in the Solutions.

First proof: We have to perform $n - 1$ multiplications of two factors. For each multiplication, write $+1$ for every opening parenthesis and -1 for every closing parenthesis. This transforms the problem to sequences having $n - 1$ elements $+1$ and $n - 1$ elements -1 so that the sum of the first k elements is non-negative for every $1 \le k \le 2n - 2$. Inserting an extra term $+1$ before each sequence, we get sequences having n elements $+1$ and $n-1$ elements -1 so that the sum of the first k elements is strictly positive for every $1 \le k \le 2n-1$. Verify that the number of bad sequences with n elements $+1$ and $n - 1$ elements -1, i.e., the ones not meeting the last condition, is the double of the sequences with first element -1.

Second proof: In the last step, we multiply the products formed of the first k factors and the last $n - k$ factors for $k = 1, 2, \ldots, n - 1$, so we get the recursion $\alpha_n = \sum_{k=1}^{n-1} \alpha_k \alpha_{n-k}$, $\alpha_1 = 1$. This means the equation $A^2(z) = A(z) - z$ for the power series $A(z) = \sum_{n=1}^{\infty} \alpha_n z^n$.

Third proof: We get a simpler recursion if we also take the permutations of the factors into consideration. If β_n is the solution of this recursion, then clearly $\alpha_n = \beta_n/n!$.

9.2.17 Fix one edge of the n-gon and group the partitions according to the position of the third vertex of the triangle containing this edge. This yields essentially the same recursion as used in the second proof of the previous exercise.

9.3.

9.3.1 (a) Take, e.g., the first and fourth powers of the primes p_i.

(b) By the pigeonhole principle, there are three distinct integers c_j where the mod 3 remainders of the exponents are the same for every p_i. So, the product of these three integers c_j is a cube.

(c) Any of the two proofs of Theorem 9.3.1 can be adapted to verify the statement.

9.3.2 In a proof by contradiction, Fermat's little theorem implies

$$\prod_{i=1}^{k}\bigl(1 - f_i^{p-1}(x_1, x_2, \ldots, x_t)\bigr) \equiv \prod_{j=1}^{t}(1 - x_j^{p-1}) \pmod{p}$$

for every $x_1, \ldots, x_t$. Using the condition on the sum of degrees, show that this is impossible.

If every polynomial is linear, then we get a homogeneous system of linear equations over the field F_p where the number of variables is greater than the number of equations. Thus there exists a non-trivial solution. Chevalley's theorem can be considered as a generalization of this well-known result.

9.3.3 Let γ_{ij} be the exponent of the prime p_i in c_j $(1 \le i \le k, 1 \le j \le t)$. Apply Chevalley's theorem for the polynomials $f_i(x_1, \ldots, x_t) = \sum_{j=1}^{t} \gamma_{ij} x_j^2$ and $p = 3$.

9.3.4 The result follows from Chevalley's theorem similar to the previous exercise.

9.3.5 (a) Consider the modulo n remainders of $c_1, c_1 + c_2, \ldots, c_1 + c_2 + \ldots + c_n$.

(b) Prove first that if the statement is true for $n = r$ and $n = s$, then it also holds for $n = rs$. Pick any $2r - 1$ elements from the $2rs - 1$

numbers. By the statement with $n = r$, we can select r of them whose sum is divisible by r. Repeat the procedure for the remaining $2rs - 1 - r$ numbers, etc., and verify that we get $2s - 1$ groups of r numbers where the sum of elements in each group is divisible by r. Let these sums be $ru_1, \dots, ru_{2s-1}$. Finally apply the statement with $n = s$ for the integers $u_1, \dots, u_{2s-1}$.

So, it is sufficient to deal with the case when n is a prime p. Let $f_1 = \sum_{j=1}^{2p-1} c_j x_j^{p-1}$, $f_2 = \sum_{j=1}^{2p-1} x_j^{p-1}$, and apply Chevalley's theorem.

9.3.6 (a) By the Chinese remainder theorem, it is sufficient to solve the problem for prime power moduli p^m. If $m > 1$, then $x_1 = x_2 = x_3 = p^{\lceil m/2 \rceil}$ works. If $m = 1$, then, e.g., by Chevalley's theorem, the congruence $x_1^2 + x_2^2 + x_3^2 \equiv 0 \pmod{p}$ has a non-trivial solution. We may assume here $|x_i| \le p/2$, so $0 < x_1^2 + x_2^2 + x_3^2 < p^2$. Therefore the sum $x_1^2 + x_2^2 + x_3^2$ is divisible by p but not by p^2.

(b) We only have to refine the procedure in (a) if $m > 1$ is odd: let $x_i = p^{(m-1)/2} y_i$ and apply the argument in (a) with $m = 1$ for the numbers y_i.

9.3.7 Compute the approximative (real) values of the kth roots of the integer N in question for every k. Let n_k be the closest integer to this kth root and check whether $n_k^k = N$ holds. Since the smallest base of a power is 2, therefore $k \le \log_2 N$. So this is a fast algorithm.

9.3.8 (a) Each prime sieves independently roughly the half of the numbers. Therefore the ratio of definitely wrong numbers is about $(2^s - 1)/2^s$ in the case of s primes. Thus the ratio of the potential integers x to be tested is $1/2^s$.

(b) The favorable cases are when d and e are close to each other.

(c) In the search for $N = 86519 = x^2 - y^2$, we have $x \ge \sqrt{86519}$, so the minimal potential value of x is 295. Considering $y^2 = x^2 - N$ modulo 8, the possible values of the left-hand side are $0, 1$, and 4, those of the right-hand side are $0 - 7, 1 - 7$, and $4 - 7$. The only common value is $1 = 0 - 7$, so $8 \mid x^2$. Therefore x is divisible by 4. Modulo 3 we get similarly $3 \mid x$. Thus the smallest integer to be tested is $x = 300$. This works as $y = \sqrt{300^2 - 86519} = 59$ is an integer. For $N = 584189$, we get that x is odd and a multiple of 3, so the potential values are $x = 765, 771, \dots$. Testing these, $x = 783$ yields a solution.

Remark: Before trying to decompose a big number, one should verify by a quick primality test that this number is composite. Note that no

really good factorization algorithm is known, and also the method of this exercise is desperately slow for general (say) 300-digit numbers.

9.4.

9.4.1 (a) For each of the first $k-1$ elements, we can decide freely whether to insert it into the subset and, for the last element, it is uniquely determined whether or not it has to be included. Another option is to count the subsets of $0, 2, 4, \ldots$ elements and use the binomial theorem: $\sum_{i=0}^{\lfloor k/2 \rfloor} \binom{k}{2i} = \big((1+1)^k + (1-1)^k\big)/2$.

(b) In a k-element set, let α_k, β_k, and γ_k be the number of subsets of size congruent to 0, 1, and 2 mod 3. We want to compute α_k. Clearly, $\alpha_i + \beta_i + \gamma_i = 2^i$ for every i. Also, considering whether or not the first element is in the subset, we get $\alpha_k = \gamma_{k-1} + \alpha_{k-1}$. Therefore $\beta_{k-1} = 2^{k-1} - (\alpha_{k-1} + \gamma_{k-1}) = 2^{k-1} - \alpha_k$. Similarly, $\beta_{k-1} = 2^{k-2} - \gamma_{k-2}$ and $\gamma_{k-2} = 2^{k-3} - \alpha_{k-3}$. Combining the last three equalities, we get $\alpha_k + \alpha_{k-3} = 3 \cdot 2^{k-3}$ leading to the recursion $\alpha_k = \alpha_{k-6} + 21 \cdot 2^{k-6}$. We can expand it to the formula $\alpha_k = 21 \cdot 2^{k-6} + 21 \cdot 2^{k-12} + \ldots$ which, apart from its last term, is the sum of a geometric sequence with a quotient 2^6. Another option is to use

$$\alpha_k = \sum_{i=0}^{\lfloor k/3 \rfloor} \binom{k}{3i} = \big((1+1)^k + (1+\omega)^k + (1+w^2)^k\big)/3,$$

where ω is a primitive third complex root of unity.

9.4.3 For a proof by contradiction, assume a non-trivial linear combination $\delta_1 \mathbf{h}_1 + \ldots + \delta_n \mathbf{h}_n = \mathbf{0}$ with rational coefficients. Multiplying by the least common multiple of the denominators and dividing by the gcd of the (new) integer coefficients, we can achieve that the (new) coefficients δ_j are coprime integers. Forming the inner product with $\mathbf{h}_j$, we obtain $\delta_1(\mathbf{h}_1 \cdot \mathbf{h}_j) + \ldots + \delta_j(\mathbf{h}_j \cdot \mathbf{h}_j) + \ldots + \delta_n(\mathbf{h}_n \cdot \mathbf{h}_j) = 0$. Since $|H_j|$ is odd, but $|H_t \cap H_j|$ is even for $t \neq j$, every inner product $\mathbf{h}_t \cdot \mathbf{h}_j$ is even except for $\mathbf{h}_j \cdot \mathbf{h}_j$ which is odd. Therefore δ_j must be even. This holds for every j but then the integers δ_j are not coprime, a contradiction.

9.4.4 (b) Oddtown: The product $B = A^T A$ is the $n \times n$ identity matrix, so $n = \mathrm{rk}(B) \leq \mathrm{rk}(A) \leq k$, by Exercise 5.7.12(a).

Eventown: The product $B = A^T A$ is the $n \times n$ zero matrix. By Exercise 5.7.12(b), $k \geq \mathrm{rk}(A) + \mathrm{rk}(A^T) = 2\mathrm{rk}(A)$, thus $\mathrm{rk}(A) \leq \lfloor k/2 \rfloor$. The number of columns of A is n, so $n \leq 2^{rk(A)} \leq 2^{\lfloor k/2 \rfloor}$.

9.4.5 (a) The one-element subsets always work. For any even $k \geq 4$, also the complements of the one-element subsets satisfy the requirement. For any odd $k \geq 5$, just combine the two types.

(b) The combinations indicated in (a) guarantee this, too.

(c) The upper bound is roughly the number of selecting any k subsets. The best way to obtain the lower bound is to use the method in Exercise 9.4.4. For convenience, assume that k is even: $k = 2t$. Let C be any $t \times t$ symmetric 0–1 matrix and I_t the $t \times t$ identity matrix. Then for the $k \times k$ matrix $A = \begin{pmatrix} C + I_t & C \\ C & C + I_t \end{pmatrix}$, $B = A^T A$ is the $k \times k$ identity matrix, so A is the incidence matrix of a suitable set system H_j. There are $2^{k(k+2)/8}$ such matrices A (or C). We have to divide by $k!$ due to the set systems differing only in the order of columns. If we also want to disregard from isomorphic set systems that only differ in the order of inhabitants, this corresponds to the permutations of the rows, and then we have to divide by $k!$ once again. This is still a very big number, it is bigger than $2^{k^2/9}$ for k large enough. This shows that there are very many ways to form k clubs in Oddtown.

9.4.6 The subsets of $k-1$ elements work for k odd and the two-element subsets containing x_1 are suitable for k even. We can verify that this is the maximum by adapting the original proof of the Oddtown theorem or the one indicated in Exercise 9.4.4. For k even, use that the rank (over the field F_2) of the incidence matrix is at most $k-1$ since the sum of the rows is 0.

9.4.7 (a) Any proof of the Oddtown theorem can be adapted by using the field F_3 instead of F_2.

(b) Only the proof suggested in Exercise 9.4.3 works.

(c) $|H_t \cap H_j|$ is divisible by 6 for $t \neq j$, so it is a multiple of both 2 and 3. On the other hand, $|H_j|$ is not divisible by 6, so $|H_j|$ cannot be divisible by both 2 and 3. If there are $2k+1$ clubs in Sixvillage, then there are either at least $k+1$ Oddtown clubs, or at least $k+1$ Triple City clubs, but both are impossible.

Remark: In general, let $s > 1$ be a fixed integer. What is the maximal number of subsets H_j in a set X of k elements X satisfying $s \nmid |H_j|$ and $s \,\big|\, |H_t \cap H_j|$ for $t \neq j$?

The k one-element subsets always satisfy the condition. We saw in the Oddtown theorem that this is the maximum for $s = 2$. The same applies

if s is a prime. The statement also remains valid if s is a power of a prime, the proof can be performed along the hint to Exercise 9.4.3. The problem is *unsolved* for every other s. We get similar to (c) that the maximum is not greater than $\omega(s) \cdot k$, where $\omega(s)$ is the number of distinct prime factors of s. This upper bound was slightly improved, e.g., for $s = 6$, the best result is $n \leq 2k - 2\log_2 k$ instead of $n \leq 2k$. On the other hand, no more than k subsets were found for any s.

9.4.9 Follow the original proof of the Oddtown theorem. To verify the independence of the vectors $\mathbf{b}_j$ corresponding to the blue subsets B_j, form the inner products of the linear combination $\delta_1\mathbf{b}_1 + \ldots + \delta_n\mathbf{b}_n = \mathbf{0}$ with the vectors $\mathbf{r}_1, \ldots, \mathbf{r}_n$ corresponding to the red subsets R_t.

9.4.10 To prove an upper bound, verify that the corresponding vectors are linearly independent over $\mathbf{R}$.

9.4.11 Use similar linear algebra argument as in the previous exercise.

9.4.13 (b) We first give an example of such a two-dimensional subspace. Let $0 < r < p$ be a quadratic non-residue modulo p, i.e., $n^2 - r$ is not a multiple of p for any integer n. Then let $V = F^{p-r+1}$ and $U = \left\langle \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\rangle$. (We note that even F^3 contains such a subspace for every p, moreover, also F^2 works if $p \equiv 3 \pmod 4$; see Exercise 9.4.14(a).)

Now we show that in any subspace of dimension at least 3, there exists a non-zero vector orthogonal to itself. We construct three pairwise orthogonal vectors with the orthogonalization algorithm, let these be $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ and $\mathbf{b}_i \cdot \mathbf{b}_i = \beta_i, \quad i = 1, 2, 3$. A vector $\mathbf{v} = \sum_{i=1}^{3} \gamma_i \mathbf{b}_i$ is orthogonal to itself if and only if

$$0 = \mathbf{v} \cdot \mathbf{v} = (\sum_{i=1}^{3} \gamma_i \mathbf{b}_i) \cdot (\sum_{i=1}^{3} \gamma_i \mathbf{b}_i) = \sum_{i=1}^{3} \gamma_i^2 \beta_i.$$

Considering the values γ_i as unknowns, this equation has a non-trivial solution by Chevalley's theorem (see Exercise 9.3.2).

9.4.14 (a) Show that the congruence $z_1^2 + z_2^2 \equiv 0 \pmod p$ has a non-trivial solution if and only if $p \not\equiv 3 \pmod 4$ and the congruence $z_1^2 + z_2^2 + z_3^2 \equiv 0 \pmod p$ has a non-trivial solution for every p.

9.4.15 (a) E.g., let $p = k = 2$ and $U = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$.

(b) A vector $\mathbf{x} \in V$ is orthogonal to a subspace U if and only if $\mathbf{x}$ is orthogonal to every element in a basis of U. This condition is a homogeneous system of $\dim U$ linear equations with $\dim V$ unknowns. Its rank is $\dim U$ since the rows are the coordinates of the basis vectors of U, so they are linearly independent. Therefore the number of free parameters in the solutions is $\dim U^\perp = \dim V - \dim U$.

(c) It follows from (b) and Exercise 4.6.6.

9.4.16 $\dim U + \dim U^\perp = \dim V$ implies that k must be even. Rely also on Exercise 9.4.14(a).

9.4.17 (a) Take, e.g., a large subset with an odd number of elements and the one-element subsets disjoint to it.

(b) In the proof of the Eventown theorem, consider the subspace U spanned by the vectors $\mathbf{h}_j$ corresponding to the subsets H_j. If not every element of U occurs among the vectors $\mathbf{h}_j$, then we can extend the system with such a missing element. If every element of U occurs as $\mathbf{h}_j$ but the number of elements is not maximal, then $\dim U < \lfloor \dim V/2 \rfloor$ and so $\dim U^\perp \geq \dim U + 2$. Extend the basis of U to a basis of $U^\perp$ with the vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots$. Then each of $\mathbf{w}_1, \mathbf{w}_2$, and $\mathbf{w}_1 + \mathbf{w}_2$ is perpendicular to U and at least one of them is perpendicular to itself. Thus, we can extend the system with this vector.

9.4.18 Separate the clubs of even and odd membership and show that the two subspaces spanned by the corresponding vectors are disjoint. If their dimensions are e and o, then $2e + o \leq k$, so $n \leq o + 2^{\lfloor (k-o)/2 \rfloor}$.

9.4.19 Verify first that the maximum is 4 for a 6-element set. Turning to a 9-element set, show that if the subsets H_i satisfy the conditions, then so do their complements, too. Therefore we can assume $|H_1| = 3$. Then $H_i \cap H_1 = \emptyset$ or $H_i \supseteq H_1$ for every H_i. Thus, there are two options for $H_i \cap H_1$ and four options for $H_i \setminus H_1$ by the result on a 6-element set.

9.5.

9.5.1 (b) The simplest verification is to consider the special case $d = 3$ in the proof of Theorem 9.5.1.

9.5.2 It follows from the definitions of adjacency matrix and eigenvector.

9.5.3 The eigenvalues are the zeroes of the characteristic polynomial of the adjacency matrix. We can reduce the amount of calculations in certain cases by using the previous exercise.

9.5.4 Let A and A' be the adjacency matrices of the graphs G and its complement $\overline{G}$. Regularity implies that the vector $\mathbf{j}$ with all components 1 is an eigenvector of both G and $\overline{G}$ with eigenvalues d and $n-1-d$. Let $\mathbf{j}, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be an orthonormal eigenbasis of A in the Euclidean space $\mathbf{R}^n$ with eigenvalues $d, \lambda_2, \ldots, \lambda_n$. Since the eigenvectors of J independent from $\mathbf{j}$ are the non-zero elements in $\langle \mathbf{j} \rangle^\perp$ with eigenvalue 0, thus also every $\mathbf{v}_i$ is an eigenvector of J with eigenvalue 0. By $A' = J - I - A$, every $\mathbf{v}_i$ is an eigenvector of A', too, with eigenvalue $-1-\lambda_i$. So the eigenvalues of A' are $n-1-d, -1-\lambda_2, \ldots, -1-\lambda_n$.

9.5.5 Show that an equivalent property to both conditions is that every eigenvector of A is an eigenvector of J.

9.5.9 If $\mathbf{j}$ is the vector with all coordinates 1, then the ith coordinate of $A\mathbf{j}$ is the degree d_i of the ith vertex. Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be an orthonormal eigenbasis with eigenvalues $\lambda_1, \ldots, \lambda_n$ and $\mathbf{j} = \sum_{i=1}^n \rho_i \mathbf{b}_i$. Then

$$\delta n \le d_1 + \ldots + d_n = \mathbf{j} \cdot (A\mathbf{j}) = \sum_{i=1}^n \lambda_i \rho_i^2 \le \Lambda \sum_{i=1}^n \rho_i^2 = \Lambda(\mathbf{j} \cdot \mathbf{j}) = \Lambda n,$$

so $\delta \le \Lambda$. To verify $\Lambda \le \Delta$, let $\mathbf{v}$ be an eigenvector belonging to Λ such that its maximal coordinate is 1, let this be (say) the first coordinate. The relation $\mathbf{x} \le \mathbf{z}$ should mean that $x_i \le z_i$ for every coordinate. Then $\Lambda\mathbf{v} = A\mathbf{v} \le A\mathbf{j} \le \Delta\mathbf{j}$ and comparing the first coordinates, we obtain $\Lambda \le \Delta$.

9.5.10 Prove two lemmas first. Let A and A' be symmetric non-negative matrices with maximal eigenvalues Λ and Λ'.

(i) If the coordinates of a non-zero vector $\mathbf{x}$ are non-negative (i.e., $\mathbf{x} \ge \mathbf{0}$) and $A\mathbf{x} \ge \tau\mathbf{x}$, then $\Lambda \ge \tau$.

(ii) If $\alpha_{ij} \ge \alpha'_{ij}$ for every pair of entries in A and A', then $\Lambda \ge \Lambda'$.

We can prove the statement by induction on the number of vertices; delete a vertex of minimal degree with the edges on it and apply (ii).

Remark: It can be proved that there always exists a non-negative eigenvector to the maximal eigenvalue, moreover, this holds for any (not necessarily symmetric) non-negative matrices (Frobenius–Perron theorem).

9.5.11 (a) Adding the sums for all lines, we obtain that the sum of the numbers on all points is 0. On the other hand, adding the sums for all lines containing a given point, the number on the given point appears $n+1$ times in this sum, whereas the numbers on the other points occur only once.

Another option is to use the incidence matrix defined in Exercise 9.4.4 where the jth element in the ith row is 1 or 0 according to whether or not the ith point is on the jth line. Verify for this $(n^2+n+1)\times(n^2+n+1)$ matrix A that every element in the main diagonal of A^TA is $n+1$ and all other elements are 1. By Exercise 1.3.12, $0 \neq \det(A^TA) = (\det A)^2$, so $A^T\mathbf{v} = \mathbf{0} \Rightarrow \mathbf{v} = \mathbf{0}$. Let the number on the ith point be the ith coordinate of $\mathbf{v}$. Then the jth coordinate of $A^T\mathbf{v}$ is the sum of the numbers on the jth line which is 0 by the condition. Therefore $\mathbf{v} = \mathbf{0}$, i.e., all numbers on the points are 0.

(b) Let R and B be the number of red and blue points, so $R+B = n^2+n+1$. Similarly, let R_i and B_i be the number of red and blue points on the ith line, thus $R_i + B_i = n+1$. It is sufficient to verify that the average of the differences $|R_i - B_i|$ on the n^2+n+1 lines is at least $\sqrt{n}$:

$$\sum_{i=1}^{n^2+n+1} (R_i - B_i)^2 \geq n(n^2+n+1). \tag{H.9.1}$$

The square of the left-hand side is

$$\sum_{i=1}^{n^2+n+1} R_i^2 + \sum_{i=1}^{n^2+n+1} B_i^2 - 2\sum_{i=1}^{n^2+n+1} R_iB_i. \tag{H.9.2}$$

The first sum in (H.9.2) is the number of ordered pairs of collinear red points. This is 1 for distinct points and $n+1$ if the two points are the same. Therefore, the first sum equals $R(R-1)+(n+1)R = R^2+nR$. Similarly, the second sum is B^2+nB. The third sum is the number of ordered pairs of collinear bicolor points, this is just RB. Thus (H.9.2) can be rewritten as $R^2+B^2+n(R+B)-2RB = (R-B)^2+n(n^2+n+1)$ proving (H.9.1).

Another option is to use a modified $(n^2+n+1)\times(n^2+n+1)$ incidence matrix B where $b_{ij} = 1$ or -1 if the ith red or blue point

is on the jth line, and is 0 otherwise. Show that every entry in the main diagonal of B^TB is $n+1$ and every other entry is 1. Let $\mathbf{w}$ be the vector with every coordinate 1. The statement is that $\mathbf{z} = B^T\mathbf{w}$ has a coordinate of absolute value at least $\sqrt{n}$. This follows if the average of the squares of coordinates is at least n, i.e., the inner product $\mathbf{z} \cdot \mathbf{z} \geq n(n^2+n+1)$. The inner product is

$$\mathbf{z} \cdot \mathbf{z} = (B^T\mathbf{w})^T(B^T\mathbf{w}) = \mathbf{w}^TBB^T\mathbf{w}. \tag{H.9.3}$$

Show that every entry in the main diagonal of BB^T is $n+1$ and the other entries are 1 or -1 depending on whether the points corresponding to the row and column indices have the same or different colors. This means that $BB^T = nI + H$, where the elements in the main diagonal of H are 1 and all other elements are ± 1. So (H.9.3) can be rewritten as $\mathbf{z} \cdot \mathbf{z} = n\mathbf{w}^T\mathbf{w} + \mathbf{w}^TH\mathbf{w}$. Here the first term is $n(n^2+n+1)$. Verify that the second term is $(R-B)^2$. This latter is non-negative, so $\mathbf{z}^T\mathbf{z} \geq n(n^2+n+1)$ indeed.

9.6.

9.6.1 We apply the greedy algorithm, and pick always the first element that does not violate the Sidon property. Assume that we have already chosen $a_1 < a_2 < \ldots < a_s < n$. We cannot choose d as a_{s+1} if $d + a_i = a_j + a_k$, i.e., $d = a_j + a_k - a_i$ for some $i, j, k \leq s$. (The case $d + d = a_j + a_k$ cannot occur, since then $d < a_k$ and so we would have chosen d into the sequence earlier instead of a_k.) This excludes at most s^3 (in fact, less than $s^3/2$) elements. This means that if $s < n^{1/3}$, then we can still find a new element $a_{s+1} \leq n$.

9.6.2 To verify the Sidon property, assume $a_i + a_j = a_k + a_l$, i.e.,

$$2p(i+j-k-l) + (\langle i^2 \bmod p\rangle + \langle j^2 \bmod p\rangle - \langle k^2 \bmod p\rangle - \langle l^2 \bmod p\rangle) = 0.$$

The second term is divisible by $2p$ and has absolute value less than $2p$, so it must be 0. Then also the first term is 0. This means $i - k = l - j$ and $i^2 - k^2 \equiv l^2 - j^2 \pmod p$. A simple calculation shows that either $i = k$ and $j = l$, or $i = l$ and $j = k$.

9.6.3 Apply (a simplified version of) the proof of Theorem 9.6.2 for the field of p^2 elements and its subfield of p elements.

9.6.4 Let g be a primitive root modulo p and let a_i be the modulo $p(p-1)$ solution of the system of congruences $x \equiv i \pmod{p-1}$, $x \equiv g^i \pmod p$, $i = 1, 2, \ldots, p-1$.

9.6.5 We can take a Sidon set S_1 between 1 and n_1 having about $\sqrt{n_1}$ elements by Theorem 9.6.1. Let n_2 be much larger than n_1. We leave the interval $(n_1, n_1+n_2]$ empty, choose a Sidon set in the interval $(n_1+n_2, n_1+2n_2]$ of about $\sqrt{n_2}$ elements, delete (at least one member of) those pairs whose difference is less than $< n_1$, and denote the remaining set by S_2. By the Sidon property, we deleted less than $2n_1$ elements. Therefore we selected altogether about $\sqrt{n_2}+\sqrt{n_1}-2n_1 \approx \sqrt{n_2}$ elements up to n_1+2n_2. Verify that $S_1 \cup S_2$ is a Sidon set. Choose now an n_3 much bigger than n_1+2n_2, place a Sidon set of size about $\sqrt{n_3}$ between $n_1+2n_2+n_3$ and $n_1+2n_2+2n_3$, delete the elements with differences less than n_1+2n_2, etc. Continuing the procedure, we obtain an infinite Sidon set meeting the requirements.

9.6.6 (a) Generalize the method of Exercise 9.6.3 to the field of p^h elements.

(b) The h-fold sums are all distinct and fall between 1 and nh.

9.6.7 (b) How many sums do arise and what are natural lower and upper bounds for them?

(c) Consider the sums as the values of a classical random variable and apply Chebyshev's inequality.

9.6.8 Let C be the set of integers between 1 and $n^{2/3}$, and let D be the union of C and the primes not exceeding n. First verify that every number up to n can be represented as $n = cd$ where $c \in C$, $d \in D$ (the representation is generally not unique). Then fix such a representation $a_i = c_i d_i$ for every a_i, and construct a bipartite graph with $|C| + |D| \le \pi(n) + 2n^{2/3}$ vertices where the two groups of vertices are C and D, and the number a_i is represented by the edge between vertices c_i and d_i. If the number of edges is not less than the number of vertices, then the graph contains a cycle. Since the graph is bipartite, this cycle has an even number of edges, and by the construction, the product of numbers a_i corresponding to every second edge is equal to the product of numbers a_i corresponding to the other edges in the cycle (as both products are equal to the product of all numbers appearing in the vertices of the cycle).

9.6.9 We use the number system with base d where we shall specify d later. Consider those positive integers up to n where every digit is less than $d/2$ and the sum of the squares of digits is a given q. Show that such a set contains no three-term arithmetic progression, further, we can choose q and d so that the set should be as large as required in the exercise.

9.6.10 *First proof*: Establish an upper bound for the number of wrong colorings and show that this is less than the number of all colorings.

Second proof: Take the field F of 2^p elements, where p is a prime, and let Δ be a generator of its multiplicative group. Considering F as a vector space over $\mathbf{Z}_2$, let W be a $p-1$-dimensional subspace in it. We color an integer red if $\Delta^k \in W$. Coloring the numbers $1, 2, \ldots, p(2^p-1)$, there do not arise $p+1$-term monochromatic arithmetic progressions.

Third proof: We color red the integers divisible by exactly one of 7 and 17.

Fourth proof: We color red the integers divisible by an odd number of the primes 2, 3, 5, and 7.

Fifth proof: We denote by A any block of 17 consecutive integers where the first 16 numbers are red and the last one is blue. Similarly, B is a dual of A, i.e., the first 16 numbers are blue and the last one is red. A good coloring is to take alternately 15 blocks A and 15 blocks B altogether 16 times.

9.7.

9.7.1 *Step 1.* Any polygon can be cut into triangles.

Step 2. We can equidissect any triangle into a parallelogram. Cut a triangle ABC at the segment DE connecting the midpoints D and E of edges AB and AC into a small triangle ADE and a trapezoid $BCED$. Rotating the small triangle ADE by 180 degrees around D, we get a triangle BDF. The union of the trapezoid $BCDE$ and the triangle BDF is the parallelogram $BCEF$.

Step 3. We can equidissect any parallelogram into a rectangle keeping one edge unaltered. In a parallelogram $ABCD$, let DE be the altitude perpendicular to the edge AB. If E is an inner point of the segment AB, then translating the triangle ADE into the triangle BCF, we get a rectangle $EFCD$ where $EF = AB$. If E lies outside the segment AB, then cut the parallelogram with lines parallel to AB into several flatter parallelograms for which we can apply the previous procedure and then heap up the obtained flat rectangles into a big rectangle.

Step 4. We can equidissect any rectangle into a rectangle with an edge of given length x. If necessary, we cut the rectangle into thinner stripes and glue them together at their shorter sides to obtain a rectangle $ABCD$ where $BC < x < AB$. Then the circle with center B and radius x intersects CD in an inner point E. We get a parallelogram $ABEF$ with a suitable point F on the line CD satisfying $EF = CD$. Applying Step 3,

we can equidissect the rectangle $ABCD$ into this parallelogram $ABEF$. Finally, applying Step 3 again, we can equdissect this parallelogram into a rectangle with an edge $BE = x$.

Step 5. We can equidissect any polygon into a rectangle with an edge of given length x. Applying Steps 1–4, we cut the polygon into triangles, equidissect each triangle into a parallelogram, equidissect each parallelogram into a rectangle, then equidissect each rectangle into a rectangle with an edge of length x, and finally heap up these rectangles into a big rectangle with an edge x. If two polygons have the same area, then the above procedure yields the same rectangle with an edge x. So equidissecting the first polygon into this rectangle and then equidissecting this rectangle into the second polygon, we get an equidissecting of the first polygon into the second one.

9.7.2 (a) Observe that $\delta/\pi \in \mathbf{Q}$ if and only if $n\delta/(2\pi) \in \mathbf{Z}$, i.e., $\cos(n\delta) = 1$ for some integer n. Using

$$\cos(n\alpha) = 2\cos\big((n-1)\alpha\big)\cos\alpha - \cos\big((n-2)\alpha\big),$$

prove by induction that $\cos(n\alpha)$ is a fraction with a denominator 3^n in its reduced form, so it cannot be equal to 1.

(b) Similar to (a), prove by induction that $2\cos(n\gamma)$ is a monic polynomial of $2\cos\gamma$ with integer coefficients. Thus, if $\cos(n\gamma) = 1$, then $2\cos\gamma$ is a root of a monic polynomial with integer coefficients. The rational roots of such a polynomial can only be integers, so $2\cos\gamma \in \mathbf{Z}$. Since $|\cos\gamma| \le 1$, we get the statement.

9.7.3 Assume that the ratio of the sides a and b in a rectangle R is a rational number: $a/b = r/s$ with integers $r, s > 0$. Then we can cut R into rs congruent squares of side $a/r = b/s$. To prove the converse, we define a function F on the set of all rectangles similar to the area: Let f be a real function satisfying Cauchy's equation $f(x+y) = f(x)+f(y)$ (see Exercise 9.1.8), and if a rectangle U has sides c and d, then $F(U) = f(c)f(d)$. Specifically, if U is a square, then $F(U) \ge 0$. Show that F is additive: cutting a rectangle U into rectangles U_i, we have $F(U) = \sum_i F(U_i)$. To obtain a contradiction, assume that the sides a and b of a rectangle R are not rational multiples of each other but still we could cut R into squares: $R = \cup S_i$ (where the squares S_i share no common inner points). Then the right-hand side of $F(R) = \sum F(S_i)$ is non-negative. On the other hand, we can achieve $F(R) < 0$ by defining $f(a) = 1$ and $f(b) = -1$; such an f exists as a and b are linearly independent over $\mathbf{Q}$.

9.7.4 (a) This can be verified along the proof of the Bolyai–Gerwien theorem (see the hint to Exercise 9.7.1).

(b) The tetrahedron $ABCC'$ is equidissectible into a prism, thus also into a cube. But this is false for the tetrahedron $ABCB'$. The angles on its edges are $\pi/2$ and ϑ satisfying $\cos\vartheta = 1/\sqrt{3}$. Show that ϑ/π is irrational and follow the line of the proof of Theorem 9.7.1.

9.7.5 Going around through all edges of a polygon, use the signed sum of the lengths of edges parallel to a given direction as an invariant for equidissectibility with only translations.

9.7.6 (a) If $n = 2k > 2$, then dissect a square of side k into a square of side $k-1$ and $2k-1$ unit squares in the remaining stripes. If $n = 2k+3 > 5$, then use the previous construction and cut one of the resulting squares into four congruent parts. To prove the converse, observe that there must be a small square in every corner of the original square, so 2 and 3 are impossible. For $n = 5$, there would be 3 small squares on one side of the big square and 2 on all other sides which leads to a contradiction after considering all possible cases.

(b) We can easily cut a cube into 8 or 27 small (congruent) cubes, so with a repeated application of these steps, we can cut a cube into $1+7x+26y$ cubes, where x and y are arbitrary non-negative integers. Using that 7 and 26 are coprime, show that every sufficiently large n can be represented in this form.

(c) Cutting a cube into 8 cubes, we can always increase the number of small cubes by 7. Hence it suffices to verify the statement for $48 \le n \le 54$.

48: $48 = 27 + 3 \cdot 7$, so we cut the cube into 27 cubes, and then we cut each of 3 small cubes into 8 parts.

49: For brevity, let us write Cuk for a cube if the length of its edge is k. We cut the lower half of Cu6 into four Cu3, the top row into thirty-six Cu1, and the remaining two rows into nine Cu2.

50: $50 = 7 \cdot 7 + 1$.

51: In Cu6, we form five Cu3 from the lower half plus one quarter in the upper half, select five Cu2 from the remaining part, and there are forty-one Cu1 left.

52: We cut a Cu3 from Cu4, and partition two of the remaining thirty-seven Cu1 into eight parts.

53: Using $53 = 1 + 2 \cdot 19 + 2 \cdot 7$, it is enough to show a procedure which increases the number of cubes by 19; we cut Cu3 into a Cu2 and nineteen Cu1.

54: We cut Cu8 into six Cu4, two Cu3, four Cu2, and forty-two Cu1.

9.7.7 (a) If $n \neq 2$, 3, and 5, then we can apply a similar construction as in Exercise 9.7.6(a). The converse is simple for $n = 2$, but for $n = 3$ and 5 it is more complicated than at the squares since the small triangles may also cut the angles of the original triangle. An observation of independent interest can help: If the angles of a triangle T are linearly independent over $\mathbf{Q}$ and T is dissected into similar triangles, then these are similar to T and do not cut the angles of T.

(b) If $n = k^2$, then dividing every side of a triangle into k equal parts, we dissect the triangle into n congruent small triangles. To prove the converse, consider a triangle where both the angles and the sides are linearly independent. Verify the existence of such a triangle and show that it cannot be dissected into n congruent triangles if $\sqrt{n}$ is irrational.

(c) If $n = k^2 + m^2$, then consider a right triangle with legs k and m. If $n = 3k^2$, then the half of a regular triangle works.

9.8.

9.8.1 If, e.g., $\mathbf{a}_n = \sum_{j=1}^{n-1} \lambda_j \mathbf{a}_j$, then (i) implies

$$D(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}, \mathbf{a}_n) = \sum_{j=1}^{n-1} \lambda_j D(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}, \mathbf{a}_j),$$

and every term in the sum is 0 by (iv).

9.8.2 Similar to the proof of Theorem 9.8.1, we infer that $F_i(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ determines F_i uniquely.

9.8.3 Use the previous exercise and the equality $F(\mathbf{e}_1, \ldots, \mathbf{e}_n) = G(\mathbf{e}_1, \ldots, \mathbf{e}_n)$.

9.8.4 Rely on the previous exercise.

9.8.5 (a) The determinant remains the same if we subtract the third column from the first two columns and expand by the third row. The resulting determinant $\begin{vmatrix} \gamma_{11} - \gamma_{13} & \gamma_{12} - \gamma_{13} \\ \gamma_{21} - \gamma_{23} & \gamma_{22} - \gamma_{23} \end{vmatrix}$ is the signed area of the parallelogram spanned by the vectors from P_3 to P_1 and P_2. This parallelogram is degenerate if and only if the three points are collinear.

(c) We can choose an arbitrary basis in any coordinate system by Theorem 9.8.1.

9.8.6 Orthogonality: For any $\mathbf{z} \in \mathbf{R}^n$, we obtain the inner product $\mathbf{z}\cdot\mathbf{d}$ by replacing the unit vectors with the coordinates of $\mathbf{z}$ in the first row of the determinant defining $\mathbf{d}$. If $\mathbf{z} = \mathbf{v}_i$, then this new determinant has two equal rows, so it is 0.

Length: Use that the length of $\mathbf{d}$ is the absolute value of a function from systems of $n-1$ vectors into $\mathbf{R}$ satisfying the conditions of Theorem 9.8.1 (with $n-1$ instead of n). Another option is to apply Gaussian elimination to the 2nd to nth rows of the matrix.

10. Codes

10.1.

10.1.5 Let A and B be the sets of places where the digits of $\mathbf{u}$ and $\mathbf{v}$ are equal or unequal. By the condition, $|A| = k - d, |B| = d$. Let j be the number of places in A where the digit of $\mathbf{z}$ differs from the other two vectors. Then the number of places in B is $q - j$ where $\mathbf{u}$ and $\mathbf{z}$ differ and is $r - j$ where $\mathbf{v}$ and $\mathbf{z}$ differ. Clearly, $(q - j) + (r - j) = d$, so $j = (r + q - d)/2$. Hence, $r + q - d$ must be even. In this case, we can choose j places from A and $q - j$ places from B arbitrarily.

10.1.6 Apply similar arguments as in the previous exercise.

10.1.8 We adapt the argument applied in the special case $t = 1$. Let $N(\mathbf{c})$ denote the neighborhood of radius t of a codeword $\mathbf{c}$; this means the set of vectors in F^k of distance at most t from $\mathbf{c}$ including $\mathbf{c}$ itself. The distance of exactly $\binom{k}{i}$ vectors is i from a given vector, so $|N(\mathbf{c})| = \sum_{i=0}^{t} \binom{k}{i}$. By the condition the 2^n sets $N(\mathbf{c})$ are pairwise disjoint in F^k, so $2^n|N(\mathbf{c})| \le 2^k$.

10.1.9 We sketch the solution of the 2-error correction for $n = 2$.

Six check digits are sufficient, e.g., consider $\varphi : \alpha_1\alpha_2 \mapsto \alpha_1\alpha_2\alpha_1\alpha_2\alpha_1\alpha_2\beta\beta$ where $\beta = \alpha_1 + \alpha_2$. In this combination of a 3-repetition and two parity checks, the distance is at least 5 between any two of the four codewords.

Five check digits are not sufficient since in F^7 not even three vectors can have pairwise distances 5 or more. If both $\mathbf{c}_2 - \mathbf{c}_1$ and $\mathbf{c}_3 - \mathbf{c}_2$ contain at least five digits 1, then at least three of them must be at the same place. But then there are 0s at these places in $\mathbf{c}_3 - \mathbf{c}_1 = (\mathbf{c}_3 - \mathbf{c}_2) + (\mathbf{c}_2 - \mathbf{c}_1)$. So $\mathbf{c}_3 - \mathbf{c}_1$ contains at most four digits 1.

10.2.

10.2.4. A linear map $\mathcal{A} : F^n \to F^k$ is injective if and only $\dim \operatorname{Im} \mathcal{A} = n$ and $\dim \operatorname{Im} \mathcal{A} = \operatorname{rk}[\mathcal{A}]$.

10.2.8 (b) We have to check injectivity. Let $d = (g, h)$, then

$$h \,|\, gf_1 - gf_2 = g(f_1 - f_2) \iff \frac{h}{d} \,\Big|\, \frac{g}{d}(f_1 - f_2) \iff \frac{h}{d} \,\Big|\, f_1 - f_2 \,.$$

Since $\deg f_i \leq n-1$, this implies $f_1 = f_2$ if and only if $n \leq \deg(h/d) = k - \deg d$, i.e., $\deg d \leq s$.

(c) Let D be the set of polynomials of degree at most $k-1$ divisible by $d = (g, h)$. Then $|D| = 2^{k - \deg d} = 2^n$, so D is a polynomial code generated by d and clearly $d \,|\, h$. The remainder of gf in the long division by h is divisible by d, so $C \subseteq D$. Since $|C| = |D|$, this implies $C = D$.

10.2.9 If G is a generator matrix of C, then M has to satisfy $MG = I_{n \times n}$. This means n systems of linear equations in k variables where each system consists of n equations and its coefficient matrix is G^T. The rank of G is n, so these systems have solutions that can be quickly determined, e.g., by Gaussian elimination. As $k > n$, the solutions are not unique, but any solution provides a suitable matrix M. We applied the same procedure in the proof of Theorem 3.5.3 to compute the inverse of a square matrix. We can interpret the present situation, too, as computing a left inverse of the matrix G.

10.3.

10.3.1 Assuming the independence of the rows of a matrix, the columns are dependent if and only if there are more columns than rows. The rank of a parity-check matrix is the number of its rows, so the rows are independent. As there are more columns than rows, the columns are dependent. To prove the converse, consider an $s \times k$ matrix P where $k - s = n > 0$ and $\operatorname{rk}(P) = s$. By the Dimension Theorem, the kernel $\operatorname{Ker} P$ is a $k - s = n$-dimensional subspace C in F^k, so P is a parity-check matrix of the linear (k, n) code C.

10.3.3 (a) The second property means $\operatorname{Ker} \mathcal{P} \supseteq C$. Equality holds if and only if $n = \dim C = \dim \operatorname{Ker} \mathcal{P} = \dim F^k - \operatorname{rk}(P) = k - \operatorname{rk}(P)$, i.e., if $r(P) = s$.

(b) The second property in (a) is equivalent to $PG = 0$.

(c) The two properties in (b) mean that the rows of P are independent and orthogonal to C. Equivalently, they form a basis in $C^\perp$ as $\dim C^\perp = \dim F^k - \dim C = k - n = s$.

10.3.4 (a) Use Exercise 10.3.3(c). How many non-zero elements are in $C^\perp$?

(b) Combine Exercises 10.3.3(c) and 4.5.14(a).

(c) Show: (i) $MP\mathbf{z} = \mathbf{0} \iff P\mathbf{z} = \mathbf{0}$; (ii) $M_1P = M_2P \iff M_1 = M_2$; and (iii) the number of such matrices M is equal to the number of all parity-check matrices.

10.3.5 (b) We use the criterion in Exercise 10.3.3(b). Concerning the rows of P, the condition $PG = 0$ means s homogeneous systems of n linear equations with k unknowns and with the same coefficient matrix G^T. There are $k - \mathrm{rk}(G^T) = k - n = s$ free parameters, so the solutions form an s-dimensional subspace in F^k. After solving the system with Gaussian elimination, we get s independent solutions, e.g., by choosing the ith parameter as 1 and the others as 0 for every $i = 1, 2, \ldots, s$. These s independent solutions provide suitable rows for a parity-check matrix P as $PG = 0$ and $\mathrm{rk}(P) = s$.

10.3.6 If a row is the linear combination of other rows, then deleting this row does not alter the kernel of the matrix.

10.3.8 Argue similar to the proof of Theorem 10.3.2.

10.3.9 Use the previous exercise. Another option is to use the generator matrix instead of the parity-check matrix. If the code is systematic, i.e., every codeword starts with the corresponding message word, then the weight of the codeword belonging to a unit vector is not greater than $1 + s$, so we are done. For a general generator matrix, choose any n independent rows and verify that there exist codewords the first n digits of which in these rows are unit vectors.

10.3.11 Use Exercises 10.3.10 and 10.3.4(c).

10.3.12 It cannot be less than 3 due to 1-error correcting. And it cannot be greater than 3 either since the disjoint neighborhoods with radius 1 of the codewords fill the entire F^k as was shown at the end of Section 10.1.

10.3.13 First prove that if, in a linear code, there exist complement codewords, then the complement of every codeword is a codeword. Therefore it is sufficient to verify that the complement of $\mathbf{0}$ is a codeword. This complement is the vector $\mathbf{j}$ with all coordinates 1. It is a codeword as every row of the parity-check matrix contains an even number of 1s.

10.3.14 As the encoding function is a linear map, this is a linear code. Its generator matrix is of the form $G = \begin{pmatrix} I_{n\times n} \\ B_{s\times n} \end{pmatrix}$ where the columns of B are binary representations of the positive integers less than 2^s that are not powers of two. Thus the columns of B are all distinct vectors with at least two coordinates 1. By Exercise 10.3.5(a), $P = (\, B_{s\times n} I_{s\times s}\,)$ is a parity-check matrix. This means that the last s columns in P are unit vectors, so the columns of P altogether are all distinct non-zero vectors in F^s. This was the definition of a Hamming code.

A direct proof without a parity-check matrix is to verify that the code is 1-error-correcting. We have to show that every non-zero codeword contains at least three coordinates 1. If this holds for the message word, we are done. If the message word contains a single coordinate $1 = \alpha_m$, then the values of γ corresponding to the at least two binary digits of m are 1, too. Finally, if a message word has exactly two coordinates $1 = \alpha_m = \alpha_q$, then m and q differ it at least one binary digit, so the corresponding $\gamma = 1$, too.

10.4.

10.4.1 The multiplicative group of a $2^5 = 32$-element field has 31 elements. As 31 is a prime, every non-identity element is a generator of the group. So, a generator can be a root of any irreducible polynomial of degree 5. Choosing $x^5 + x^2 + 1$, we have to apply $\Delta^5 = 1 + \Delta^2$ repeatedly. As an illustration, we compute the 3rd column of the parity-check matrix. The upper and lower parts contain the components of Δ^2 and $\Delta^6 = \Delta + \Delta^3$. So the upper and lower five entries are 0,0,1,0,0 and 0,1,0,1,0.

10.4.2 We know that $s = \deg g_t$ where $g_t = [m_1, m_3, \ldots, m_{2t-1}]$. Every m_i is irreducible, so g_t is the product of the distinct ones among the polynomials m_i. Further, $\deg m_i \le q$. Thus $s = \deg g_t = tq$ holds exactly if every m_i is distinct and has degree q.

10.4.3 (a) Use that the field F^q has exactly one 2^v-element subfield for every $v \,|\, q$ and the non-zero elements in this subfield are the powers of Δ with exponents $j(2^q - 1)/(2^v - 1)$.

(b) Verify that these are distinct roots of m_i. Use that two powers of Δ are equal if and only if the difference of the exponents is a multiple of $2^q - 1$.

10.4.4 (a) We can work in the 16-element field similar to the Example after Theorem 10.4.1. Observe $(\Delta^3)^5 = (\Delta^5)^3 = 1$.

(b) Show that also Δ^3 and Δ^5 generate the multiplicative group of the field F^q, so their minimal polynomials must have degree q. To show that the minimal polynomials are distinct, use Exercise 10.4.3(b).

10.4.5 Similar to the previous exercise, we obtain $m_3 \neq m_1$, so $s = \deg g = \deg m_1 + \deg m_3 = q + \deg(\Delta^3)$. As $\deg(\Delta^3) \mid q$, it is either equal to q or it is not greater than $q/2$. The latter would imply $s \leq 3q/2$ but this contradicts $s \geq 2q - 1$.

10.4.6 Show that the conditions of Exercise 10.4.2 are satisfied. Rely on Exercise 10.4.3, too.

10.4.7 (a) A vector $\mathbf{r} = \varrho_0 \dots \varrho_{k-1} \in F^k$ has an even weight if and only if $\varrho_0 + \varrho_1 + \ldots + \varrho_{k-1} \equiv 0 \pmod 2$. This means that 1 is a root of the polynomial $R = \varrho_0 + \varrho_1 x + \ldots + \varrho_{k-1} x^{k-1}$ which is equivalent to $x - 1 = x + 1 \,|\, R$. As the code C consists of the multiples of g, the condition on the weights is equivalent to $1 + x \,|\, g$.

(b) There are $|F^k|/2$ vectors of even weight, so the number of check digits is $s = 1$. Combining this with (a), we obtain the statement.

10.4.8 Raising any element in the multiplicative group of the field F^q to the $|F^q| - 1 = 2^q - 1 = k$th power, we obtain the identity. So every element of the multiplicative group is a root of $x^k - 1$. This implies that every minimal polynomial m_i divides $x^k - 1$. So this also holds for their least common multiple which is the generator polynomial.

10.4.9 (a) Verify that a cyclic code is a principal ideal in the factor ring $R_k = F[x]/(x^k - 1)$ generated by a polynomial that divides $x^k - 1$.

(b) It follows from (a) and the previous exercise.

10.4.10 We have to construct a quasi-parity-check matrix where any $d-1$ columns are independent. After preparing the first j columns, no linear combination of $d-2$ or fewer of them can be the $j+1$st column. By the condition, there are enough free vectors to select a $j+1$st column even for $j = k-1$.

10.4.11 Prove by induction on q in (a) and on $q + m$ in (b).

A. Basic Algebra

A.1.

A.1.1 (a) Induction step: The sum on the left-hand side for $n+1$ differs only in the last term from the sum for n. So, assuming the formula for n, we have to verify $\dfrac{n}{n+1} + \dfrac{1}{(n+1)(n+2)} = \dfrac{n+1}{n+2}$. For a direct proof, use $\dfrac{1}{k(k+1)} = \dfrac{1}{k} - \dfrac{1}{k+1}$.

(b) Similar to (a), we have to verify $\dfrac{n(n+1)(2n+1)}{6} + (n+1)^2 = \dfrac{(n+1)(n+2)(2n+1)}{6}$ in the induction step. For a direct proof, add the equalities $(k+1)^3 - k^3 = 3k^2 + 3k + 1$ for $1 \le k \le n$.

A.1.2 To establish the upper bound, in the induction step it suffices to verify $2\sqrt{n} - 1 + (1/\sqrt{n+1}) \le 2\sqrt{n+1} - 1$. Apply the identity $\sqrt{n+1} - \sqrt{n} = 1/(\sqrt{n+1} + \sqrt{n})$. The lower bound follows similarly. For a direct proof, compare the sum with $\int_1^n dx/\sqrt{x}$.

A.1.3 Proceed similar to Example E1. For a direct proof, show that the remainder of 5^n upon division by 8 is 1 if n is even, and 5 if n is odd.

A.1.4 Computing the first few values, we conjecture $a_n = 2^n - 1$. In the induction step, we have to assume its truth for $n-1$ and n, and deduce it for $n+1$: $a_{n+1} = 3a_n - 2a_{n-1} = 3(2^n - 1) - 2(2^{n-1} - 1) = 2^{n+1} - 1$. For direct methods to handle such recursions; see Section 9.2.

A.1.5 The induction step from $n = 1$ to $n + 1 = 2$ is false.

A.1.6 Measure the terms one after the other onto a segment of length 2000.

A.1.7 The elegant proofs use the combinatorial interpretation.

(a) A subset of k elements defines the complementary subset of $n - k$ elements uniquely.

(b) We classify the k-element subsets of $\{1, 2, \ldots, n+1\}$ whether or not they contain the element $n + 1$.

(c) We classify the $k + 1$ element subsets of $\{1, 2, \ldots, n+1\}$ according to their largest element.

(d) We have n red and n blue balls, and select n balls from them.

A.1.8 (a) Apply the binomial theorem for $(1-1)^n$ and combine it with Example E2.

(b) Apply the binomial theorem for $(1-2)^n$.

A.1.9 (a) For (a3), apply the inclusion-exclusion formula with the properties that certain digits are missing from the number.

(b) By (a), both sides represent the number of k-digit integers containing all digits $1, 2, \ldots, n$. The lesson is that a complicated count may sometimes be useful because it leads to a nice identity if combined with the simple direct count.

A.1.11 The argument fails if a is in relation with no element.

A.2.

A.2.1 $(2s+1)^2 - (2t+1)^2 = 4(s(s+1) - t(t+1))$, and the product of two consecutive integers is always even. Another option: The remainder of a^2 depends only on the remainder of a, as $(8k+r)^2 = 8(8k^2+2kr)+r^2$. Since $(8k\pm1)^2 = 8L+1$ and $(8k\pm3)^2 = 8M+9 = 8N+1$, the remainder of an odd square is always 1 upon division by 8. This method can be generalized from 8 to any integer and from squares to arbitrary powers. Congruences help to formalize this procedure even better.

A.2.3 To obtain (i), verify and use the equality $a^n - b^n = (a-b)(a^{n-1}+a^{n-2}b+a^{n-3}b^2+\ldots+b^{n-1})$. Substituting here $-b$ instead of b, we get (ii) and (iii) for n odd and even, resp. Another option: Raise the true congruences $a \equiv b \pmod{a-b}$ and $a \equiv -b \pmod{a+b}$ to the nth power.

A.2.4 (a) $10 \equiv 1 \pmod 9$, so $10^k \equiv 1 \pmod 9$, hence

$$\begin{aligned}\overline{a_s a_{s-1}\ldots a_1 a_0} =& a_s 10^s + a_{s-1}10^{s-1} + \ldots + a_1 10 + a_0 \\ \equiv& a_0 + a_1 + a_2 + \ldots + a_s \pmod 9.\end{aligned}$$

Note that we proved the more general statement that an integer and the sum of its digits give the same remainder upon division by 9.

(b) Use $10 \equiv -1 \pmod{11}$, and proceed as in (a).

(c) By (a), we have to determine the remainder of 2^{2019} modulo 9: raising $2^3 \equiv -1 \pmod 9$ to the 673rd power, we get $2^{2019} \equiv -1 \pmod 9$.

(d) We can prove by induction, but it is simpler to use congruences: $36 \equiv -7$ and $49 \equiv 6 \pmod{43}$, so

$$6^{n+2} + 7^{2n+1} = 36 \cdot 6^n + 7 \cdot 49^n \equiv -7 \cdot 6^n + 7 \cdot 6^n = 0 \pmod{43}.$$

A.2.6 The statement holds for any integer $n > 0$ instead of 11111. There are n possible remainders upon division by n, so there are infinitely many a_i in the sequence giving the same remainder, by the pigeon-hole principle. The difference of any two of them is divisible by n.

A.2.7 There are n possible remainders, so the sum of a maximal set of numbers with distinct remainders is

$$\begin{aligned} S =& k_0 n + (k_1 n + 1) + (k_2 n + 2) + \ldots + (k_{n-1} n + n - 1) \\ =& kn + (1 + 2 + \ldots + (n-1)) = kn + n(n-1)/2. \end{aligned}$$

If n is odd, then $(n-1)/2$ is an integer, so S is divisible by n. If n is even, then

$$n(n-1)/2 = n^2/2 - n/2 = n((n/2) - 1) + (n/2),$$

so the remainder is $n/2$.

A.2.8 (a) This happens for $n = 99$, e.g., when the serial and ranking numbers coincide for everyone.

(b) Let r_i and s_i be the ranking and serial numbers of contestant i, and $t_i = r_i + s_i$. By Exercise A.2.7, $\sum_{i=1}^{100} r_i \equiv \sum_{i=1}^{100} s_i \equiv 50 \pmod{100}$, so $\sum_{i=1}^{100} t_i \equiv 0 \pmod{100}$. If the remainders of t_i are all distinct, then $\sum_{i=1}^{100} t_i \equiv 50 \pmod{100}$, a contradiction.

A.2.9 (a) Devise natural algorithms if m is odd or is a multiple of 4. If $m = 4k + 2$, then the total number of jumps needed to meet on a tree would be odd, so this is impossible.

(b) For m odd, there is a natural algorithm. To show the impossibility for m even, label the adjacent trees by $1, 2, \ldots, m$, and consider the modulo m remainder S_k of the sum $j_1 + \ldots + j_m$ where j_i is the number of the tree where squirrel i is after step k. Show that S_k is constant during the process, but its initial value is different from the value when all squirrels are on one tree.

A.2.10 (a) If $n = p^k$ where $p > 0$ is a prime, then $(c, n) = 1$ if and only if c is not a multiple of p. The multiples of p among the integers $1, 2, \ldots, p^k$ are the p^{k-1} numbers $p, 2p, 3p, \ldots, p^k = p^{k-1}p$.

A.2.11 The last two digits mean a congruence modulo 100. As $(1357, 100) = 1$, $1357^{k\varphi(100)} \equiv 1 \pmod{100}$, by the Euler–Fermat Theorem.

A.2.12 If $n^2 \equiv -1 \pmod{p}$ for a prime $p = 4k - 1$, then raising the congruence to the power $(p-1)/2$, the result contradicts the Euler–Fermat Theorem.

A.2.14 Instead of a congruence modulo 1000, investigate the corresponding simultaneous system modulo 125 and modulo 8. To reduce the exponent of c^k modulo a prime power p^s, use the Euler–Fermat theorem if $(c, p^s) = 1$, and observe that $c^k \equiv 0 \pmod{p^s}$ if $p \mid c$ and $k \geq s$.

A.2.15 Instead of $x^2 \equiv x \pmod{10^{20}}$, consider the system of congruences with the corresponding prime power moduli. Show that the congruence $x(x-1) \equiv 0$ has 2 solutions modulo a prime power.

A.2.16 Theorem A.2.8: If (x, y) is a solution, then (A, B) divides both A and B, so $(A, B) \mid Ax + By = C$. For the converse, use that the Euclidean algorithm implies $(A, B) = Au + Bv$ with some integers u and v. This gives also an efficient algorithm to find a solution (x_0, y_0). Then check that the formula in the theorem yields further solutions. To prove that these are all solutions, pick a solution (x, y), so $Ax + By = Ax_0 + By_0$, i.e., $A(x - x_0) = B(y_0 - y)$, or $\frac{A}{(A,B)}(x - x_0) = \frac{B}{(A,B)}(y_0 - y)$, and use that $A/(A, B)$ and $B/(A, B)$ are coprime.

Theorem A.2.9: Transform the linear congruence into a linear Diophantine equation as shown after the statement of the theorem.

Theorem A.2.10: The conditions are equivalent to $x = m_1y_1 + c_1 = m_2y_2 + c_2$. Apply Theorem A.2.8 to the linear Diophantine equation induced by the second equality.

A.3.

A.3.1 (c) $5 - 12i = (x + yi)^2 \iff x^2 - y^2 = 5, 2xy = -12$ where $x, y \in \mathbf{R}$. (This method does not work for nth roots in general, since there exist no formulas to express the zeroes of an equation of high degree.)

A.3.2 Writing $z = a + bi$, we transfer the question to a problem in analytic geometry. One has to be careful to perform equivalent steps in solving the resulting equations or inequalities. But it is often simpler and more elegant to use the geometric interpretation of the condition directly.

(a) The condition is $3 - 5b = 7$, or $b = -4/5$ yielding a line.

(b) The condition is $-2b + 5 > 9$ yielding a half-plane.

(c) The distance of z from the point $(3, -8)$ is less than or equal to 1 yielding a disc.

(d) If $z \neq 0$, then the condition is $\sin(\arg z) = 1/2$, so $\arg z = \pi/6$ or $5\pi/6$ yielding two half-lines.

(e) The distance of z from $5i$ is greater than or equal to its distance from $-i$ yielding a closed half-plane.

(f) If $z \neq 0$, then using the trigonometric form, we obtain $\pi/4 - \arg z = \arg z + 2k\pi$, $k \in \mathbf{Z}$, or $\arg z = \pi/8 + k\pi$, which is a line through the origin.

(g) The condition is that the real part of the fraction is 0. Writing $z = a + bi$ and performing the division, we get $\dfrac{(a+1)(a+3)+b^2}{(a+3)^2+b^2} = 0$. The numerator is 0 exactly if $(a+2)^2 + b^2 = 1$ yielding a circle with center $(-2, 0)$ and radius 1. The point $z = -3$, or $(-3, 0)$, has to be excluded as the denominator cannot be 0. Another option: The argument of the fraction, i.e., the difference of the arguments of the numerator and the denominator is $\pm\pi/2$, so the vector from -1 to z is perpendicular to the vector from -3 to z. This means that the triangle with vertices $A = (-1, 0), B = (-3, 0)$, and $C = z$ has a right angle at C. By Thales' theorem, the points C form a circle. We have to include A, but exclude B.

A.3.3 (b) and (c): use the quadratic formula.

A.3.4 Use $z_j = a_j + b_j i$ or apply $|w|^2 = w \cdot \overline{w}$. Geometric meaning: In a parallelogram, the sum of squares of the four sides equals the sum of squares of the diagonals.

A.3.5 $a^2 + b^2 = |a + bi|^2$ and $|z| \cdot |w| = |zw|$. Another option: Perform the multiplication $(a^2 + b^2)(c^2 + d^2) = (ac)^2 + (bd)^2 + (ad)^2 + (bc)^2$, add a new term to the first two terms and subtract it from the last two terms so that the resulting three-three terms should be complete squares.

Counterexample for the converse: $3 \neq a^2 + b^2$, but $3 \cdot 3 = 3^2 + 0^2$ or $3 \cdot 75 = 9^2 + 12^2$.

Counterexample for three squares: $3 = 1^2 + 1^2 + 1^2$ and $5 = 2^2 + 1^2 + 0^2$, but $15 \neq a^2 + b^2 + c^2$.

A.3.6 Multiply $3 + 4i$ by i and $e^{i\pi/3}$. In general, the geometric transformation corresponding to the multiplication by w is a rotation by $\arg w$ plus an enlargement by $|w|$ from the origin.

A.3.7 Working with $z_j = a_j + b_j i$, the real part of the fraction is 0 exactly if $|z_1| = |z_2|$. A simple geometric proof: The diagonals of a parallelogram are perpendicular if and only if its sides have equal lengths (i.e., it is a rhombus).

A.3.8 Let $z = \cos x + i \sin x$. Compute z^5 in two ways: using the trigonometric form and by the binomial theorem. Compare the imaginary parts and replace $(\cos x)^2$ by $1 - (\sin x)^2$.

A.3.9 (b) Compare the real parts in the two forms obtained in (a).

A.3.11 (a) If $z^n = w^k = 1$, then $(zw)^{[n,k]} = (z/w)^{[n,k]} = 1$. Another option: Use the characterization of roots of unity by their modulus and argument.

(b) Observe that the shorter diagonal and the two sides of the rhombus spanned by the two roots of unity form an equilateral triangle.

A.3.12 If $w = e^{2\pi i/n}$, then $1, w, w^2, \ldots, w^{n-1}$ are the nth roots of unity. Sum: $S_n = \sum_{k=0}^{n-1} w^k = (w^n - 1)/(w-1)$. Product: $P_n = \prod_{k=0}^{n-1} w^k = w^{n(n-1)/2}$. To simplify the results, apply $w^n = 1$. The answer for the product depends on the parity of n.

Another option: If z is an nth root of unity, then so is $1/z$. Thus, we can group the factors of the product into pairs $z \cdot (1/z) = 1$. We have to consider separately if $z = 1/z$, or $z^2 = 1$, i.e., $z = \pm 1$. Observe that -1 is an nth root of unity if and only if n is even.

A.3.13 (a) $(z^t)^s = 1 \iff z^{ts} = 1 \iff k = o(z) \mid ts \iff k/(k,t) \mid s$, so the smallest s is $k/(k,t)$.

(b) See the hint to Exercise A.3.11(a).

A.3.14 (a) Form pairs $z, 1/z$.

A.3.15 Let A, B, C, and D be the complex numbers corresponding to the vertices of the quadrangle labeled counter-clockwise. Let O_1 be the center of the square leaning on side AB. We get the vector $\overrightarrow{AO_1}$ by rotating $\overrightarrow{AB}$ with angle $-\pi/4$ and dividing the length by $\sqrt{2}$. This corresponds to the multiplication by $e^{-i\pi/4}/\sqrt{2} = (1-i)/2$. Using $\overrightarrow{AB} = B - A$ and $\overrightarrow{AO_1} = O_1 - A$, we obtain

$$O_1 = (O_1 - A) + A = \frac{(B-A)(1-i)}{2} + A = \frac{B(1-i)}{2} + \frac{A(1+i)}{2}.$$

Expressing the midpoints of the other three squares similarly, we have to verify $O_4 - O_2 = i(O_3 - O_1)$.

A.3.16 Putting $z = \cos x + i \sin x$, the sum is the real part of the geometric series $\sum_{k=1}^{n} z^k$. To obtain a nicer form, rewrite the result using $w = \cos(x/2) + i \sin(x/2)$. Another option: Multiply the sum by $\sin(x/2)$ and apply $\sin\beta \cos\alpha = \dfrac{\sin(\alpha+\beta) - \sin(\alpha-\beta)}{2}$.

A.4.

A.4.1 In (d1) and (d2) the assignment depends on the representative chosen from the residue class, so it is not unique: e.g., for $m = 10$, 3 and 13 are in the same class but $\max(8, 13) = 13$ and $\max(8, 3) = 8$ are in different

classes, similarly $8 \equiv 8^{13} \not\equiv 8^3 \equiv 2 \pmod{10}$. In (e2), the composition of two rotations around two distinct points with angles of, e.g., 90 and -90 degrees is a translation.

A.4.3 We can visualize an operation and its properties with its *Cayley table*. Let $a_1, \ldots, a_n$ be the elements of X. We draw an $n \times n$ table, list the elements a_i in the top and left margins, and let the jth element in the ith row be $a_i a_j$ (the product of the elements heading the corresponding row and column).

To define an operation we can write any of the n elements of X into any of the n^2 little squares of the table.

Commutativity means that the table is symmetric to its main diagonal. Thus, we can write any of the n elements into the squares below and on the main diagonal and these determine the other entries.

The identity of an operation can be any of the n elements and it determines all squares in its row and column. We can write any element into the remaining $(n-1)^2$ squares.

Note that we counted two operations as distinct if a permutation of the elements in X transforms one operation into the other; these are essentially the same operations, just we ordered the elements of X differently. In this case, the algebraic structures created by the two operations are *isomorphic*(=have the same form).

A.4.4 $(ab)^{-1} = b^{-1}a^{-1}$. Use, e.g., Exercise A.4.2 to construct a counterexample to the converse.

A.4.6 (a) See, e.g., Exercise A.4.2.

(b) Let a be arbitrary and b a left inverse of a, so $ba = e$. It is enough to show that b is a right inverse of a, as this implies $b = a^{-1}$, by the associative law, and a cannot have another left inverse. To prove $ab = e$, let c be a left inverse of b, and compute $cbab$ in two different ways.

(c) The only left or right inverse of the identity is itself.

(d) See, e.g., Exercise A.4.1(h).

A.4.8 Apply first (i) and then (ii) in Theorem A.4.7.

A.4.9 Parts (h) and (g) in Exercise A.4.1 yield counterexamples to (i) and (ii), resp.

A.5.

A.5.4 (a) No two of the finite fields F_p are isomorphic as they differ in the number of elements. Clearly, they cannot be isomorphic to an infinite field either. The cardinality of the rational numbers is countable, whereas the real and the complex numbers are of the power of continuum, so there is no bijection between $\mathbf{Q}$ and $\mathbf{R}$, or $\mathbf{Q}$ and $\mathbf{C}$. The fields $\mathbf{R}$ and $\mathbf{C}$ are not isomorphic (though there exists a bijection between them) since they differ in many properties of multiplication, e.g., one can take a square root from the negative of the identity in $\mathbf{C}$, but this is not the case in $\mathbf{R}$.

A.5.5 (a) Take, e.g., the subfields $U_p = \{a + b\sqrt{p} \,|\, a, b \in \mathbf{Q}\}$ where p is a prime.

(b) Dividing a non-zero element in the subfield with itself we obtain the identity, and applying the four basic operations to the identity we get every element in $\mathbf{Q}$ or F_p.

(c) Applying the four basic operations to the identity we always get a subfield that is isomorphic to $\mathbf{Q}$ or an F_p.

(d) Adding any element in F_p with itself p times we get the zero, but 0 is the only element with this property in $\mathbf{R}$.

Remark: This is a good illustration that a field can have a subset that is a field but not a subfield. The set $H = \{0, 1, 2, \ldots, p-1\}$ is a subset of $\mathbf{R}$, and if we define the addition and multiplication on H as the modulo p remainder of the sum and product of the numbers, then we get a field isomorphic to F_p. So, H is a subset of $\mathbf{R}$ that forms a field with this addition and multiplication. But it is not a subfield in $\mathbf{R}$ since the operations are different than those in $\mathbf{R}$: e.g., for $p = 7$ we have $3 + 5 = 1$ in H, whereas $3 + 5 = 8$ in $\mathbf{R}$.

We could not apply the argument directly to F_p and needed the mediation of the set H as F_p cannot be considered as a subset $\mathbf{R}$. The elements of F_p are not numbers but residue classes, so each element of F_p is an infinite set of integers.

A.5.6 (a) Since multiplication is the usual, 2 will not have a multiplicative inverse.

(b) If $1 \odot 1 = c$, then verify $a \odot b = abc$. This implies that there is no identity if $c \neq \pm 1$, and most elements have no inverse if $c = \pm 1$.

(c) Use that there exists a bijection between the integers and the rational numbers.

A.6.

A.6.2 (c) We have to solve the linear congruence $37x \equiv 1 \pmod{100}$, or equivalently, the linear Diophantine equation $37x - 100y = 1$.

A.6.4 Add the negative of ab to the equality $ab = a(b + 0) = ab + a0$ and use the associative property of addition.

A.6.8 Multiplying a fixed non-zero element by all elements from either side, we get distinct elements by Exercise A.6.6. Hence, all elements of the ring will occur among them, as the ring is finite. So, the equations $xb = a$ and $by = a$ have a (unique) solution x and y for every $b \neq 0$ and a.

A.6.9 Let c be the only left identity. We have to prove $bc = b$, i.e., $bc - b = 0$ for every b. Verify that $c + bc - b$ is a left identity and use the uniqueness of the left identity.

A.6.10 Let 0_R and 0_S be the zeroes of the ring R and of the subring S. Picking any $s \in S$, we have $s = 0_R + s = 0_S + s$. Adding the negative of s in R(!) to both sides of the last equality (from the right) and applying the associative law, the right-hand side becomes

$$(0_S + s) + (-s) = 0_S + \big(s + (-s)\big) = 0_S + 0_R = 0_S,$$

and similarly, the left-hand side becomes 0_R.

Remark: Note that the proof makes a strong use of the negative.

A.6.11 Each (sub)ring in Example E3 in Section A.6 is a counterexample for (a), and each of parts (b1), (d2), and (d3) in Exercise A.5.1 provides a counterexample for (b) and (d). To prove (c), let e_R and e_S be the identities of R and S, and pick any $s \neq 0$ from S. Then $s = e_R s = e_S s$ and as there are no zero divisors in R(!), we can cancel s.

A.6.12 Show first that there are no zero divisors. Then pick some $b \neq 0$ and verify that a solution $x = e$ of the equation $xb = b$ is a right(!) identity. Now starting from $be = b$, we get similarly that e is a left identity, too. Next, a solution $x = c$ of the equation $xb = e$ is a left inverse of b, by definition. To prove $bc = e$, demonstrate $(e - bc)b = 0$ and use that there are no zero divisors.

A.7.

A.7.1 (a) There are infinitely many polynomials but only finitely many (polynomial) functions.

(b) If $f \neq g$ induce the same polynomial function, then h and $h + t(f - g)$ induce the same polynomial function for any polynomials h and t.

Another option: h and $h+t\prod_{\gamma\in F}(x-\gamma)$ induce the same polynomial function.

A.7.2 If zero divisors are allowed, we can easily find counterexamples for III. among polynomials over residue classes modulo a composite integer, i.e., among congruences with composite moduli.

A.7.3 (b) E.g., the product of $f = x$ and $g = \prod_{0\neq\gamma\in F}(x-\gamma)$ is the zero polynomial *function*.

(c) Use also Exercise 3.2.14.

A.7.4 E.g., any function that has infinitely many zeroes but is not identically zero.

A.7.5 Consider G as a polynomial over the field F_{11}.

A.7.6 (b) Apply the fundamental theorem of algebra.

A.7.7 (a) Substitute any integer j into the polynomial function and modify a_0 to satisfy $f(j) = 0$.

(b) Substitute the reciprocal of an integer.

(c) On the one hand, we can achieve 1 to be a root, and on the other hand, the rational root test allows only finitely many rational roots.

(d) Substitute now an arbitrary rational number.

(e) We can achieve 1 to be a root in infinitely many ways.

(f) By part (c) it is enough to deal with $i = 0$ and $i = n$. The substitution $x \mapsto 1/x$ reduces the latter case to the first one. We prove $i = 0$ by contradiction. By the rational root test, the potential rational roots are fractions with denominator a_n. Substituting these into $g = a_1x + a_2x^2 + \ldots + a_nx^n$ we have to obtain all integers with finitely many exceptions. If x is large enough, then the order of magnitude of $g(x)$ is a_nx^n. So, to get all values $|g(x)| < M$ we can use only the numbers $|x| < cM^{1/n}$ (where c is a constant independent of M). The number of suitable fractions $x = k/a_n$ is thus less than $2ca_nM^{1/n}$. This is not enough to make $g(x)$ represent nearly all integers of absolute value less than M if M is large enough.

A.7.8 Show that f and f' have no common roots.

A.7.9 The simplest algorithm is to determine the roots of f' (having degree 4), one of these is a root of f, too. Dividing f by this root factor (or even by its square), the quotient has degree less than 5, so we can compute its roots. It is, however, quicker and more elegant if we determine the

greatest common divisor of f and f' by the Euclidean algorithm, since the roots of $d = (f, f')$ are just the multiple roots of f with a multiplicity by one less. Therefore, the roots of f/d are the same as the roots of f but with multiplicity one. So the roots of $g = (d, f/d)$ are the multiple roots of f with multiplicity one. Computing the roots successively of g, d/g, and, e.g., $f/(dg)$, we obtain the roots of f and their multiplicities.

A.7.10 Since $d = (f, f') \mid f$, $\deg d < \deg f$, so d can only be a constant due to irreducibility.

A.7.11 Apply similar arguments as in the previous two exercises.

A.7.14. (a) A "universal" counterexample for certain relations is $f = 3x$ and $g = 5x$.

A.7.15. Show that the (complex) roots of the first polynomial are roots of the second polynomial with the same or greater multiplicity. Another option: rewrite the second polynomial as $(x^{3m} - 1) + x(x^{3n} - 1) + x^2(x^{3k} - 1) + (x^2 + x + 1)$. A third method is to apply induction on (say) $m + n + k$.

Remark: By the previous exercise, this divisibility is equally valid in $\mathbf{Q}[x]$, $\mathbf{C}[x]$, $\mathbf{Z}[x]$, or $F_2[x]$.

A.7.16. *First proof*: Working over the complex field we have to find the common (complex) roots.

Second proof: The steps of the Euclidean algorithm with the two polynomials correspond to the steps of the Euclidean algorithm with the exponents (in the integers).

Third proof: We can verify directly that $x^{(n,k)} - 1$ is a common divisor. On the other hand, if h is a common divisor, then $h \mid (x^{un}-1)-(x^{vk}-1) = x^{vk}(x^{un-vk} - 1)$, and choosing u and v suitably, we obtain $h \mid x^{(n,k)} - 1$.

Fourth proof: Use cyclotomic polynomials.

Remark: Similar to Exercise A.7.14, we have the same gcd over the rational, real, or complex field, moreover over the integers and F_2 (though the last two statements are not obvious).

A.7.17 Subtracting the hypothetic common remainder from the polynomial of degree 10 we get another polynomial of degree 10 that is a common multiple of the two given polynomials. But the two polynomials are coprime, so their least common multiple has degree 11.

A.7.18. Due to long division, there is a perfect analogy between the number theory of integers and of polynomials over a field F. So the answer and its

proof are the same as for the "genuine" Diophantine equations for integers (Theorem A.2.8 and Exercise A.2.16).

A.7.21 Find integers c such that the polynomial $x^4 + c$ is reducible over $\mathbf{Z}$. We have to exlude those values of c when one of the factors assumes ± 1, and the other assumes a (positive or negative) prime for some integer $n > 0$.

A.7.23 Let $f = (x - a_1) \cdot \ldots \cdot (x - a_k) - 1 = gh$. We can assume that g and h have integer coefficients by Gauss's lemma II. So $g(a_i)h(a_i) = -1$ implies $g(a_i) + h(a_i) = 0$, $i = 1, 2, \ldots, k$. If the factorization is non-trivial, then $\deg(g + h) < k$, therefore only $g + h = 0$, i.e., $g = -h$ is possible. But then the leading coefficient of $f = gh = -g^2$ cannot be 1 yielding the contradiction.

A.7.25 Let w be an mth root of unity. Show that if $o(w) < m$, then the root factor $(x - w)$ gets canceled on the right-hand side.

A.7.27 Use complex numbers.

A.7.28 The given negative number is the sum of squares of the roots.

A.7.29 Consider the sum of the roots. Do not forget to prove *both* directions.

A.7.30 Studying the product of the roots yields a root. After dividing the polynomial by this root factor, repeat the procedure.

A.8.

A.8.1 Examine the possible images of two adjacent vertices and how far these determine the complete position of the solid. We have to show that all these possibilities can be achieved by suitable congruences.

Remark: The symmetry groups of the cube and the octahedron have not just the same number of elements, but the groups themselves are isomorphic. This can be verified by simple geometric considerations as the face centers of a cube are vertices of a regular octahedron. Also, the symmetry groups of the dodecahedron and the icosadedron are isomorphic for similar reasons.

A.8.2 Squares: Multiplying the equality $abab = aabb$ by a^{-1} from the left and by b^{-1} from the right, we obtain an equivalent equality $ba = ab$.

Fourth powers: In a commutative group we can put the factors of a product in any order, so $(ab)^4 = a^4b^4$, but, e.g., D_4 shows that the converse is false.

A.8.3 Verify $(ab)^{[o(a),o(b)]} = e$.

A.8.4 E.g., the product of two reflections in the plane is a rotation that has infinite order if the angle of the two axes is not a rational multiple of π. (By the previous exercise, this phenomenon can occur only in infinite non-Abelian groups.)

A.8.6 Apply the divisibility $o(g) \,|\, |G|$ for an appropriate element g in an appropriate group G.

A.8.7 (a) Pair every element with its inverse. If more than one element equals its inverse, construct another pairing for these elements.

(b) Apply part (a) to the multiplicative group of the non-zero residue classes modulo p.

A.8.8 We can distinguish the non-isomorphic groups by some difference in the properties of the operation (commutative or non-commutative, the number of elements of a given order). To "identify" the isomorphic ones we have to show that the two groups obey exactly the same computational rules (e.g., by comparing the 8×8 operational tables).

A.8.9 (a) Verify first that such a group must be commutative. Then show that its elements can be listed as $e, a, b, ab, c, ac, bc, abc, \ldots$. Finally, check that such a group has exactly the same structure as the additive group of a finite dimensional vector space over F_2.

(b) Counterexample: G_1 is the additive group of the three-dimensional vector space over F_3 and G_2 is the multiplicative group of the 3×3 upper triangle matrices over F_3 where each element of the main diagonal is 1.

A.8.10 (a) Consider cyclic subgroups to exclude most groups and apply Lagrange's theorem to show that the remaining ones have only trivial subgroups.

(b) Show that any infinite group has infinitely many subgroups. Distinguish two cases whether or not the group contains an element of infinite order, and argue with cyclic subgroups.

A.8.11 To exclude groups of odd size use Lagrange's theorem.

A.8.12 (a) Prove that every subgroup is cyclic with a generator g^d where $d \,|\, n$ (here g denotes a fixed generator of the original group).

A.8.13 If $C = gH$ for some subgroup H and $a, b, c \in C$, then

$$ab^{-1}c = (gh_1)(gh_2)^{-1}(gh_3) = gh_1h_2^{-1}g^{-1}gh_3 = gh_1h_2^{-1}h_3 = gh_4 \in H.$$

To prove the converse, verify that if the condition holds, then the set $\{b^{-1}c \,|\, b, c \in C\}$ is a subgroup.

A.9.

A.9.2 If m has two distinct prime divisors, then 1 can be written as their combination.

A.9.3 (a) $i_j \in I_j \Rightarrow \prod_j i_j \in \bigcap_j I_j$.

(b),(c) Use Exercises A.9.1–A.9.2 to construct examples.

A.9.4 (a) If $a \neq 0$, then every element in the field can be represented as ra.

(b) $I_b = \{rb \,|\, r \in R\}$ is an ideal for every $b \in R$, so $I_b = 0$ or $I_b = R$ by the condition. Verify that the elements $b \in R$ satisfying $I_b = 0$ form an ideal I, thus $I = 0$ or $I = R$. The latter implies that the product of any two elements of R is 0, a contradiction, so $I = 0$. This means that $I_b = R$ for every $b \neq 0$, i.e., division is possible, so R is a field.

Remark: The essential use of commutative law was that I_b is an ideal. We could not write (b) immediately instead of I_b since we did not assume the existence of an identity, so it could have occurred $b \notin I_b$ (as this is the case, e.g., if every product is 0).

(c) Let I be a non-zero ideal. We have to show that I contains every matrix. Take an arbitrary matrix $A \neq 0$, $A \in I$ and compute all possible products BAC where B and C are matrices with only one non-zero entry. Prove that every matrix is the sum of such products. As an ideal is closed for these steps, I must be the complete ring $F^{n\times n}$ of matrices.

A.9.5 (a) Identifying $\mathbf{Z}_m$ with the remainders $0, 1, \ldots, m-1$, every non-zero ideal is generated by its smallest positive element.

(c) Verify that the cosets mod the principal ideal (k) can be uniquely characterized by the remainders $0, 1, \ldots, k-1$, and we perform the operations with them exactly as computing modulo k.

A.9.6 Theorem A.9.3: We have to use the commutative law and the identity to prove (i) and (ii).

Theorem A.9.4: A generator of a non-zero ideal will be a non-zero element of minimal absolute value at the integers, and of minimal degree at the polynomials. Use the division algorithm in the proof.

Theorem A.9.5: The main difficulty is to show that though the operations for the cosets are defined with the help of representatives, the resulting coset is independent of the choice of the representatives. The operational laws and the existence of special elements follow from the corresponding properties of the original ring.

Theorem A.9.6: If $g = 0$, then $F[x]/(g)$ is isomorphic to $F[x]$, and if g is a unit, then $F[x]/(g)$ contains only the zero coset, so the factor ring is not a field in these cases. If g is reducible, $g = rs$ where $\deg r < \deg g$, $\deg s < \deg g$, then none of the cosets $r+(g)$ and $s+(g)$ is the zero coset, but their product is $\big(r+(g)\big)\big(s+(g)\big) = g+(g) = 0+(g)$. Thus the factor ring contains zero divisors, so it cannot be a field. Finally we verify that if g is irreducible, then the factor ring is a field. The multiplication is commutative, $1+(g)$ is an identity, so we have to show that every non-zero coset $h+(g)$ has an inverse. The condition means $h \notin (g)$, i.e., $g \not| h$. The inverse of $h+(g)$ is a coset $u+(g)$ satisfying $\big(h+(g)\big)\big(u+(g)\big) = 1+(g)$. This is equivalent to $hu + vg = 1$ with some polynomial v. Since g is irreducible and does not divide h, the polynomials g and h are coprime. Therefore this Diophantine equation for polynomials has a solution u, v (see Exercise A.7.18).

A.9.8 (b) If I is an ideal, then $I = (A)$ where $A = \bigcup_{B\in I} B$.

(c) Part (a) implies that this is not a principal ideal. It cannot be finitely generated either as we can show similar to part (b) that every finitely generated ideal of R_S is a principal ideal.

(d) Verify that two elements of the ring R_S, i.e., two subsets B and C of S are in the same coset mod (A) if and only if $B \setminus A = C \setminus A$. Accordingly, every coset can be uniquely represented by a subset of $S \setminus A$. We have to check that this map preserves the operations.

A.9.9 (c) We show that I is not a principal ideal. To prove by contradiction, assume $I = (g)$. Then $g \,|\, 2$ and $g \,|\, x$, so g is a common divisor (in $\mathbf{Z}[x]$) of 2 and x, i.e., $g = \pm 1$. However, $\pm 1 \notin I$, as the constant term of ± 1 is odd. The contradiction guarantees that I is not a principal ideal. The factor ring $\mathbf{Z}[x]/I$ is obtained essentially by taking the remainders of polynomials with integer coefficients both mod 2 and mod x. So only the cosets represented by 0 and 1 are distinct.

A.9.10 (c) By part (a), $(a) \subseteq (a,b) = (d)$ implies $d \,|\, a$, and similarly $d \,|\, b$, so d is a common divisor of a and b. Let now c be an arbitrary common divisor, i.e., $c \,|\, a$ and $c \,|\, b$. Since $d \in (d) = (a,b)$, we have $d = au + bv$ for some u, v, so $c \,|\, d$.

(d) Use that the gcd of a and b can be written as $au + bv$.

(e) We get a counterexample, e.g., from Exercise A.9.9(c).

A.9.12 (b) Two elements of the ring R, i.e., two real functions g and h are in the same coset mod (f) if and only if $g(c) = h(c)$ for every $c < 5$ (the values at other places do not matter). Accordingly, every coset can be uniquely characterized by the values assumed at places less than 5. Also the operations are preserved, so $R/(f)$ is isomorphic to the usual ring of the real functions defined on real numbers less than 5. Finally, the latter is isomorphic to the ring R of all real functions, because there is a bijection between the two domains (i.e., between the set of real numbers less than 5 and the set of all real numbers).

A.9.14 Use Theorem A.9.6.

A.10.

A.10.1 To show $\deg\Theta \le n$, use that the elements $1, \Theta, \Theta^2, \dots, \Theta^n$ must be linearly dependent. To prove $\deg\Theta \mid n$, apply the Tower Theorem A.10.3 and Theorem A.10.11.

A.10.2 The field M is zero divisor free, whereas $\operatorname{Hom} V$ contains zero divisors.

A.10.3 Theorem A.10.3: Let $\Theta_1, \dots, \Theta_m$ be a basis of M over L, and $B_1, \dots, B_n$ a basis of N over M. Verify that the elements $\Theta_i B_j$ form a basis of N over L. Elaborate the proof also for the case when at least one of the degrees is infinite.

Theorem A.10.5: (i) We have to check that $L(\Theta)$ is closed under addition, multiplication, and forming negatives and reciprocals. Let us see, e.g., addition:

$$g_1(\Theta)/h_1(\Theta) + g_2(\Theta)/h_2(\Theta) = [(g_1h_2 + g_2h_1)(\Theta)]/[(h_1h_2)(\Theta)].$$

(ii) If $g = x$, $h = 1$, then $g(\Theta)/h(\Theta) = \Theta$. We can prove $L \subseteq L(\Theta)$ similarly.
(iii) Since F is a field and contains Θ and the elements L, it must also contain the elements obtained from these via the four basic operations.

Theorem A.10.8: (ii) If $f = m_\Theta g$, then $f(\Theta) = m_\Theta(\Theta)g(\Theta) = 0 \cdot g(\Theta) = 0$. To prove the converse, assume $f(\Theta) = 0$, and perform the long division of f by m_Θ: $f = m_\Theta h + r$ where $\deg r < \deg m_\Theta$ or $r = 0$. Then $r(\Theta) = f(\Theta) - m_\Theta(\Theta)h(\Theta) = 0 - 0 = 0$, so $r = 0$ by the definition of the minimal polynomial.

(iii) We prove by contradiction. If $m_\Theta = gh$ where $\deg g < \deg m_\Theta$ and $\deg h < \deg m_\Theta$, then $0 = m_\Theta(\Theta) = g(\Theta)h(\Theta)$. M has no zero divisors, so $g(\Theta) = 0$ or $h(\Theta) = 0$, but both contradict the definition of the minimal polynomial.
(iv) By (ii), $m_\Theta \mid f$, and m_Θ is not a constant, so f must be a constant multiple of m_Θ as f is irreducible.

Theorem A.10.10: If an element has two such representations, then subtracting them we obtain that Θ is a root of a polynomial of degree at most $n-1$, a contradiction. We prove the existence of such a representation in two steps: (i) $g(\Theta)/h(\Theta) = f(\Theta)$ for some $f \in L[x]$; (ii) one can achieve $\deg f < n$.
To verify (i), check that $g(\Theta)/h(\Theta) = f(\Theta)$ is equivalent to the "Diophantine" equation $g = hf + m_\Theta u$. This is solvable as h and m_Θ are coprime.
(ii) Let r be the remainder in the long division of f by m_Θ, then $f(\Theta) = r(\Theta)$.

Theorem A.10.12: The elements $\alpha_0 + \alpha_1\Theta + \ldots + \alpha_{n-1}\Theta^{n-1}$ obey exactly the same computational rules as the remainders modulo the polynomial m_Θ.

A.10.4 The roots of a polynomial $g \neq 0$ with rational coefficients are algebraic numbers by definition, so this also holds for any divisor of g. Conversely, if every root of f is an algebraic number, then the product of the minimal polynomials of the roots of f provides an appropriate g.

A.10.5 $\sqrt[k]{\Theta}$ is a root of $f(x) = m_\Theta(x^k)$.

A.10.6 We have to verify that the sum and product of two algebraic numbers are algebraic and so are also the negative and the reciprocal of an algebraic number. The latter two can be shown using just the definition. For the sum and product, let Θ and Ψ be algebraic, and consider the extensions $M = \mathbf{Q}(\Theta)$ and $N = M(\Psi)$. Then both $\Theta + \Psi$ and $\Theta\Psi$ are in N. Apply now Theorems A.10.3 and A.10.11.

A.10.8 (a) (iv) Express the original numbers from S and P.

A.10.9 (a) If a and b are algebraic, then using that also i is algebraic and the sum and product of algebraic numbers are algebraic, we get that $a + bi$ is algebraic. To prove the converse, show that if $z = a + bi$ is algebraic, then so is also $\overline{z} = a - bi$ (it has the same minimal polynomial as z), and express a and b from z and $\overline{z}$.

(b) Apply part (a) and the fact that if one of $\cos\varphi$ and $\sin\varphi$ is algebraic, then so is the other.

A.10.10 We need only a minimal computation in (b) and (e) if we apply the Tower Theorem. For (j), argue similarly as in the hint to Exercise A.10.12 below.

A.10.11 The minimal polynomials of the roots of unity are the cyclotomic polynomials. These are monic and have integer coefficients. Find a complex number of absolute value 1 that has no such minimal polynomial.

A.10.12 If $z = \cos\varphi + i\sin\varphi \notin \mathbf{R}$, then $\overline{z} = 1/z, \cos\varphi, i\sin\varphi \in \mathbf{Q}(z)$. Let $M = \mathbf{Q}(\cos\varphi)$ and $N = M(i\sin\varphi)$, then $N = \mathbf{Q}(z)$ and $\deg(N:M) = 2$.

A.10.13 $\mathbf{Q} \subseteq \mathbf{Q}(\Theta^2) \subseteq \mathbf{Q}(\Theta)$ and the degree of the second extension is not greater than two.

A.10.14 (a) $\sqrt{18}/\sqrt{8}$ is a rational number.

(b) Show that the right-hand side is a subset of the intersection and apply the Tower Theorem.

A.10.15 Let $\Psi = 1 + 3\sqrt[7]{25} + 11\sqrt[7]{125} + 1000\sqrt[7]{625}$. Then $\Psi \in \mathbf{Q}(\sqrt[7]{5})$ but $\Psi \notin \mathbf{Q}$. Since 7 is a prime, $\mathbf{Q}(\Psi) = \mathbf{Q}(\sqrt[7]{5})$, so $\sqrt[7]{5} \in \mathbf{Q}(\Psi)$.

A.10.16 Use $\operatorname{Re} z = \big(z + (1/z)\big)/2$ for $|z| = 1$. Prove the statement for transcendental z, too.

A.10.17 Extend $\mathbf{Q}$ with the coefficients and a root of the polynomial one after the other.

A.10.18 Theorem A.10.12 suggests to take $M = L[x]/(f)$. Then M is a field by Theorem A.9.6, the cosets constant$+(f)$ form L^*, and the coset $x + (f)$ yields Θ.

A.11.

A.11.1 Use the binomial theorem. Note that the binomial coefficients mean addition of the same term that many times and every binomial coefficient greater than 1 is divisible by p.

A.11.2 **Sum**:

First proof: Write the elements as powers of a generator and use the summation formula for this geometric series.

Second proof: Pair every element with its negative if $p \neq 2$. For $p = 2$, form pairs whose sum is a fixed $c \neq 0$.

Third proof: Multiplying every element by some $c \neq 0$ we get again all non-zero elements, so $S = cS$.

Fourth proof: Use Exercise A.11.7 and the appropriate Vieta formula (see part 11 in Section A.7).

Fifth proof: Represent the field as a vector space or as a set of polynomials; this requires slightly more computation.

Product: Apply the suitable modifications of the first, second, and fourth proofs.

A.11.3 $o(\Theta) \mid p^k - 1$ for every Θ, and Θ is a root of $x^m - 1$ if and only if $o(\Theta) \mid m$. Another option: Representing the elements as powers of a generator, the exponents belonging to the roots satisfy a certain divisibility condition.

A.11.4 Use the previous exercise.

A.11.5 (a) Multiplying the k-fold sum $0 = a + a + \ldots + a$ by any b we obtain $0 = (a + a + \ldots + a)b = ab + ab + \ldots + ab = a(b + b + \ldots + b)$. Since $a \neq 0$, the second factor is 0. Therefore adding any non-zero element with itself k times yields 0. Consider the smallest positive k with this property. If it were composite, $k = rs$ where $r < k, s < k$, then breaking the sum of k terms of $a \neq 0$ into groups of length r, either the sum of a group is 0, or the sum of the s non-zero groups is 0, and both contradict the minimality of k. This implies that the smallest k is a prime p. This prime is unique as every other k must be a multiple of it.

(b) Consider the quotients of polynomials (algebraic fractions) over the field F_p.

A.11.6 Find an irreducible polynomial of degree 2 and 4 over F_{13} and F_3, resp.

A.11.7 Apply Lagrange's theorem for the multiplicative group and deduce that the non-zero elements are roots of this polynomial (0 is a root obviously). This means as many roots as the degree of the polynomial, so there cannot be more roots.

A.11.8 To prove the only if part, use the previous exercise and that the degree of any element divides k in a field of size p^k. For the converse, consider the field $F_p[x]/(f)$.

A.11.9 The statement is essentially equivalent to the previous exercise.

A.11.10 Consider the multiplicative group of the field and use that distinct subgroups in a finite cyclic group cannot have the same number of elements.

A.11.11 Consider A as a matrix of a linear transformation $\mathcal{A}$ and show that the minimal polynomial of $\mathcal{A}$ is f (cf. with Exercise 6.3.18). Therefore the powers of the matrix A behave exactly as the powers of x in the factor ring $F_p[x]/(f)$.

A.11.12 (b) Denote this number by I_k. By Exercise A.11.8, $x^{p^k} - x$ is the product of all irreducible polynomials over F_p of degree $d \mid k$. Comparing the degrees of the two sides in this polynomial equality, we get a recursion for I_k. We can express I_k via the Möbius inversion formula.

Remark: More generally, the same formula applies for the number of irreducible polynomials of degree k over a finite field of q elements if we replace p by q. Also the proof is analogous to the case $q = p$.

A.11.13 (a) A common point of two given lines is a one-dimensional subspace in the intersection of the two two-dimensional subspaces. Since the vector space is three-dimensional, the intersection cannot be 0. So the intersection itself is a one-dimensional subspace. Similarly, the line containing two given points will be the unique two-dimensional subspace generated by the two one-dimensional subspaces.

(b) Points: The intersection of any two one-dimensional subspaces is just the zero vector, so the other $p - 1$ elements in the two subspaces are disjoint. Hence the number of one-dimensional subspaces is $(p^3 - 1)/(p - 1) = p^2 + p + 1$. Lines: there is a bijection between the two-dimensional subspaces U and the one-dimensional subspaces $U^\perp$ (cf. with Exercise A.11.14). Another option is a direct count using the bases. (Both assertions are special cases of Exercise 4.6.14.)

(c) Apply similar argument as in part (b).

A.11.14 This can be deduced from the previous exercise by representing every line with its normal vectors, so instead of a two-dimensional subspace U we take the one-dimensional subspace $U^\perp$.