Appendices

Exercises: A1.1, A1.2, A1.3, A1.4, A1.5, A1.6, A2.1, A2.2, A2.3, A2.4, A2.5, A2.6, A2.7, A2.8, A2.9, A2.10, A2.11, A2.12, A2.13, A2.14, A2.15, A2.16, A2.17, A2.18, A2.19, A2.20, A2.21, A2.22, A2.23, A2.24, A3.1, A3.2, A3.3, A3.4, A3.5, A3.6, A4.1, A4.2, A4.3, A4.4, A4.5, A4.6, A4.7, A4.8, A4.9, A4.10, A4.11, A4.12, A4.13, A4.14, A4.15, A4.16, A4.17, A4.18, A4.19, A5.1, A5.2, A6.1, A6.2, A6.3, A6.4, A6.5, A6.7, A6.8.

Notes on basic probability theory

A1.1

From the definition of conditional probability we have

$$p(y|x) = \frac{p(x,y)}{p(x)} \quad \text{and} \quad p(x|y) = \frac{p(x,y)}{p(y)}.$$

From the first equality the joint probability for $X = x$ and $Y = y$ is $p(x, y) = p(y|x)p(x)$. Substituting in the second equality we obtain Bayes' rule

$$p(x|y) = p(y|x)\frac{p(x)}{p(y)}.$$

A1.2

From the definition of conditional probability we have that $p(x, y) = p(y|x)p(x)$, summing both sides over all possible values for the random variable X we get

$$\sum_x p(x,y) = \sum_x p(y|x)p(x) \quad \Longrightarrow \quad p(y) = \sum_x p(y|x)p(x)$$

A1.3

Let us suppose that the statement is false, that is, for all x such that $p(x) > 0$ we have $x < \mathbf{E}(X)$. Then we conclude that

$$\sum_{x<\mathbf{E}(X)} p(x)x = \sum_x p(x)x < \sum_x p(x)\mathbf{E}(X) = \mathbf{E}(X),$$

which is a contradiction to the definition of $\mathbf{E}(X)$. Therefore the premise must be wrong and we conclude that there exists a value $x \geq \mathbf{E}(X)$ such that $p(x) > 0$.

A1.4

$$\mathbf{E}(X+Y) = \sum_{x,y} p(x,y)(x+y)$$

$$
\begin{aligned}
&= \sum_{x,y} p(x,y)x + \sum_{x,y} p(x,y)y \\
&= \sum_{x} x \sum_{y} p(x,y) + \sum_{y} y \sum_{x} p(x,y) \\
&= \sum_{x} xp(x) + \sum_{y} yp(y) \\
&= \mathbf{E}(X) + \mathbf{E}(Y).
\end{aligned}
$$

A1.5

$$
\begin{aligned}
\mathbf{E}(XY) &= \sum_{x,y} p(x,y)xy \\
&= \sum_{x,y} p(x)p(y)xy \\
&= \sum_{x} p(x)x \sum_{y} p(y)y \\
&= \mathbf{E}(X)\mathbf{E}(Y).
\end{aligned}
$$

A1.6

Let us define an indicator function $\mathbf{1}(X)$ for the random variable X, defined as

$$
\mathbf{1}(X) = \begin{cases} 1 & \text{if } |X - \mathbf{E}(X)| \geq \lambda\Delta(X); \\ 0 & \text{if } |X - \mathbf{E}(X)| < \lambda\Delta(X), \end{cases}
$$

where $\lambda > 0$ is a real number. From the definition of expectation it follows immediately that

$$
\mathbf{E}(\mathbf{1}(X)) = p(|X - \mathbf{E}(X)| \geq \lambda\Delta(X)).
$$

Now let us analyze the two possible values the indicator function can assume. When $|X - \mathbf{E}(X)| \geq \lambda\Delta(X)$ we may write

$$
\frac{|X - \mathbf{E}(X)|}{\lambda\Delta(X)} \geq 1 = \mathbf{1}(X) \quad \Longrightarrow \quad \mathbf{1}(X) \leq \frac{|X - \mathbf{E}(X)|^2}{\lambda^2\Delta(X)^2}
$$

And when $|X - \mathbf{E}(X)| < \lambda\Delta(X)$ we have $\mathbf{1}(X) = 0$, meaning it is also true for this case that

$$
\mathbf{1}(X) \leq \frac{|X - \mathbf{E}(X)|^2}{\lambda^2\Delta(X)^2}.
$$

So since this relation is true for all possible cases, we may take the expectation on both sides, yielding

$$
p(|X - \mathbf{E}(X)| \geq \lambda\Delta(X)) \leq \frac{1}{\lambda^2\Delta(X)^2}\mathbf{E}\big(|X - \mathbf{E}(X)|^2\big).
$$

But the standard deviation is defined as $\Delta(X) \equiv \left[\mathbf{E}\left(|X - \mathbf{E}(X)|^2\right)\right]^{1/2}$. Using this fact we obtain Chebyshev's inequality

$$p(|X - \mathbf{E}(X)| \geq \lambda\Delta(X)) \leq \frac{1}{\lambda^2}.$$

Group theory

A2.1

Because of the *closure* property, if $g \in G$ then $g^n \in G$ for all integers n. If G is a finite group then it has a finite number of elements $|G|$. So if we take the set $\left\{g, g^2, \cdots, g^{|G|}, g^{|G|+1}\right\}$ at least two of the elements must be the same, that is, there exist a and b such that $g^a = g^b$. Without loss of generality we may consider that $b > a$ and thus $g^{b-a} = e$. So there always exists a positive integer $r \coloneqq b - a$ such that $g^r = e$.

A2.2

Let $g_1, g_2 \in G$. We may define two distinct sets given by $g_1H \equiv \{g_1h | h \in H\}$ and $g_2H \equiv \{g_2h | h \in H\}$. It is straightforward that $|g_1H| = |g_2H| = |H|$. Now consider $g \in G$ such that $g \in g_1H$ and $g \in g_2H$ simultaneously. This means that $g = g_1h_1$ for some $h_1 \in H$ and $g = g_2h_2$ for some $h_2 \in H$, thus

$$g_1h_1 = g_2h_2 \quad \Longrightarrow \quad g_1 = g_2h_2h_1^{-1}.$$

Since H is also a group, we have $h_2h_1^{-1} \in H$, meaning $g_1 \in g_2H$ and therefore $g_1H \subseteq g_2H$. Analogously, we could write

$$g_1h_1 = g_2h_2 \quad \Longrightarrow \quad g_2 = g_1h_1h_2^{-1},$$

and conclude $g_2H \subseteq g_1H$, leading to $g_1H = g_2H$, which contradicts the fact that these two sets are distinct. Therefore, all $g \in G$ must belong exclusively in one set of the form g_iH for $g_i \in G$, and since G is a finite group, there is a finite number n of distinct sets g_iH, which form a partition of G. The union of these n sets of size $|H|$ must result in G, so we have $|G| = n|H|$, meaning $|H|$ divides $|G|$, proving Lagrange's theorem.

A2.3

If the order of $g \in G$ is r, then the set $H \equiv \{e, g, \cdots, g^{r-1}\}$ is a subgroup of G because it satisfies all properties of a group, and all its elements are in G. Using Lagrange's theorem and the fact that $|H| = r$, we find that r divides $|G|$.

A2.4

If $y \in G_x$, it can be written as $y \equiv h^{-1}xh$ for some $h \in G$. The conjugacy class of y is

$$G_y \equiv \left\{k^{-1}yk | k \in G\right\} = \left\{k^{-1}h^{-1}xhk | k, h \in G\right\}.$$

By the *closure* property we have that $g \equiv hk \in G$ and $g^{-1} \equiv k^{-1}h^{-1} \in G$, therefore

$$G_y = \left\{g^{-1}xg | g \in G\right\} = G_x.$$

A2.5

If G is Abelian then for all $g \in G$ we have $xg = gx$, meaning $g^{-1}xg = g^{-1}gx = x$. Thus $G_x = \{x\}$.

A2.6

Let G be a group of order $p \geq 2$, where p is a prime number. Consider an element $a \neq e \in G$, then we must have $\langle a \rangle \leq G$, that is, $\langle a \rangle$ is a subgroup of G. By Lagrange's theorem $|\langle a \rangle|$ divides $|G|$, but since $|G| = p$ it follows that either $|\langle a \rangle| = 1$ or $|\langle a \rangle| = p$. The first would only be true if $a = e$, which is not the case, thus $|\langle a \rangle| = p$, which implies $\langle a \rangle = G$, meaning G is cyclic.

A2.7

Let G be a cyclic group. Its elements can be written as $\left\{e, a, \cdots, a^{|G|-1}\right\}$, where $e = a^{|G|}$. Any subgroup must have the identity element e and may have elements of the form a^k, k integer. From Lagrange's theorem, any subgroup will have order $|G|/n$ for some integer n that divides $|G|$, meaning its elements will be given by $\left\{e, a^{|G|/n}, \cdots, a^{(n-1)|G|/n}\right\}$. From this, we conclude that any subgroup will always be given by $\left\langle a^{|G|/n} \right\rangle$, and therefore be cyclic.

A2.8

If $g \in G$ has order r, then $e = g^{kr}$ for any integer k, and by extension $g^m = g^m g^{kr} = g^{m+kr}$. So if $g^m = g^n$, we must have $n = m + kr$, which means $m = n \pmod{r}$. The converse is immediate.

A2.9

Consider the coset gH for some $g \in G$. If both g_1 and g_2 are in this same coset then we have $g_1 = gh_1$ and $g_2 = gh_2$ for some $h_1, h_2 \in H$. But then $g = g_1 h_1^{-1} = g_2 h_2^{-1}$, meaning $g_2 = g_1 h_1^{-1} h_2$. since H is a group $h \equiv h_1^{-1} h_2 \in H$, thus we can always write $g_2 = g_1 h$. Conversely, if $g_2 = g_1 h$ then evidently $g_2 \in g_1 H$. But since $e \in H$, we have $g_1 \equiv g_1 e \in g_1 H$, thus g_1 and g_2 are in the same coset.

A2.10

Any coset of H has $|H|$ elements. Since an element $g \in G$ belongs exclusively in a single coset, the union of all cosets must result in G, that is, $|G| = n|H|$, where n is the number of cosets. Therefore, the number of cosets of H in G is $|G|/|H|$.

A2.11

Property (1):

Since I is the $n \times n$ identity matrix, we have $\chi(e) = \mathrm{tr}(I) = n$.

Property (2):

Since G is finite, all elements have a finite order, meaning all elements must satisfy $g^r = e$ for some integer r. Thus, all eigenvalues of the representation must correspond to possible r-th roots of the number 1, that is, all eigenvalues of the elements $\rho(g)$ have the form $\lambda = \exp(i2\pi k/r)$ for integers k ranging from 0 to $r-1$. The norm of the character is then calculated to be

$$|\chi(g)| = |\mathrm{tr}(\rho(g))| = \left|\sum_{j=1}^{n} \lambda_j\right| \leq \sum_{j=1}^{n} |\lambda_j| = n.$$

Property (3):

If $|\chi(g)| = n$, then $|\mathrm{tr}(\rho(g))| = \left|\sum_{j=1}^{n} \lambda_j\right| = \sum_{j=1}^{n} |\lambda_j|$, meaning all its eigenvalues must correspond to the same value $e^{i\theta}$ and thus $\rho(g)$ must be proportional to the identity. Therefore, $\rho(g) = e^{i\theta} I$.

Property (4):

Consider $x \in G$. For any element in the conjugacy class G_x of this element we have $\chi(g^{-1}xg) = \mathrm{tr}(\rho(g^{-1}xg)) = \mathrm{tr}(\rho(g^{-1})\rho(x)\rho(g)) = \mathrm{tr}(\rho(g)\rho(g^{-1})\rho(x)) = \mathrm{tr}(\rho(x)) = \chi(x)$.

Property (5):

We may write any group element as $g = h^{-1}dh$, where $h, d \in G$ and d is such that $\rho(d)$ is a diagonal matrix. Analogously we have $g^{-1} = h^{-1}d^{-1}h$. From property (4) we have $\chi(g) = \chi(d)$ and $\chi(g^{-1}) = \chi(d^{-1})$. Also, note that since all eigenvalues lie on the complex unit circle, we have $\rho(d^{-1}) = \rho(d^*)$. Therefore, $\chi(g^{-1}) = \chi(d^*) = \chi^*(d) = \chi^*(g)$.

Property (6):

Since for $g \in G$ all eigenvalues of $\rho(g)$ are r-th roots of the number 1 for some r, they all satisfy the equation $\lambda^r - 1 = 0$, which is a polynomial equation with rational coefficients. Thus all eigenvalues are algebraic numbers, and since the character is calculated as a sum of such eigenvalues, $\chi(g)$ is also an algebraic number.

A2.12

Let us define the matrix

$$P \equiv \frac{1}{|G|} \sum_{g \in G} \rho^\dagger(g)\rho(g).$$

For all $h \in G$ we have

$$\rho^\dagger(h) P \rho(h) = \frac{1}{|G|} \sum_{g \in G} \rho^\dagger(h)\rho^\dagger(g)\rho(g)\rho(h) = \frac{1}{|G|} \sum_{g \in G} \rho^\dagger(gh)\rho(gh),$$

but since $k \equiv gh \in G$, because of the *closure* property, we have that P is left invariant, that is,

$$\rho^\dagger(h) P \rho(h) = \frac{1}{|G|} \sum_{k \in G} \rho^\dagger(k)\rho(k) = P.$$

It is also direct to see that P is Hermitian, thus there exists a unitary matrix U such that $U^\dagger PU = D$, where D is a diagonal matrix. Now, taking any vector $|v\rangle \neq 0$, we get that

$$\langle v|P|v\rangle = \frac{1}{|G|}\sum_{g\in G} |\rho(g)\,|v\rangle|^2$$

is always positive. If we define another vector $|w\rangle \equiv U^\dagger\,|v\rangle$, we also obtain

$$\langle v|P|v\rangle = \langle v|UU^\dagger PUU^\dagger|v\rangle = \langle w|D|w\rangle = \frac{1}{|G|}\sum_i D_{ii}\,\langle w|w\rangle\,.$$

Since $|w\rangle$ is arbitrary and $\langle w|w\rangle > 0$, it must hold that each D_{ii} is real and non-negative, meaning $D^\dagger = D$. We can then define another matrix $d \equiv \sqrt{D}$. Since D is diagonal, it follows that $d_{ii} = \sqrt{D_{ii}}$ and $d_{ij} = 0$ for all $i \neq j$, and also $d = d^\dagger$. We can then write

$$P = U^\dagger DU = U^\dagger ddU = U^\dagger dUU^\dagger dU \equiv QQ,$$

where we set $Q \equiv U^\dagger dU$. Let us then define a new representation given by $\sigma(g) \equiv Q\rho(g)Q^{-1}$, which equivalently means that $\rho(g) = Q^{-1}\sigma(g)Q$. Notice that Q is also Hermitian. Using the fact that $\rho^\dagger(g)P\rho(g) = P$ for all $g \in G$ we obtain

$$\begin{aligned}\rho^\dagger(g)P\rho(g) &= \left(Q^{-1}\sigma(g)Q\right)^\dagger QQ\left(Q^{-1}\sigma(g)Q\right)\\ &= \left(Q\sigma^\dagger(g)Q^{-1}\right)QQ\left(Q^{-1}\sigma(g)Q\right)\\ &= Q\sigma^\dagger(g)\sigma(g)Q.\end{aligned}$$

Since this must result in $P = QQ$, it follows that $\sigma^\dagger(g)\sigma(g) = I$, meaning $\sigma(g)$ is a unitary representation. Therefore, any finite representation $\rho(g)$ is equivalent to a unitary matrix representation.

A2.13

Applying Schur's lemma for $G = H$, we have that for any $g \in G$, a matrix S satisfying $S\rho(g) = \rho(g)S$ is either the zero matrix or a nonsingular square matrix. If G is an Abelian group, S can be any matrix of the representation itself, meaning $S \neq 0$. Considering an eigenvector $|v\rangle$ of S, we have $S\,|v\rangle = \lambda\,|v\rangle$ for some number λ. Notice that $\rho(g)\,|v\rangle$ must also be an eigenvector since

$$S\rho(g)\,|v\rangle = \rho(g)S\,|v\rangle = \lambda\rho(g)\,|v\rangle\,.$$

Since this is true for any $\rho(g)$, it means the entire representation vector space is spanned by a single vector. If this representation is irreducible, then $|v\rangle$ must be one dimensional, otherwise, there would be invariant subspaces with lower dimensions, which would be a contradiction.

A2.14

For irreducible representations, we can use Equation (A2.2) to write

$$\sum_{g\in G}\chi(g)\chi^*(g) = |G|.$$

Instead of denoting the sum over all group elements, we can consider a sum over all conjugacy classes of G because of property (4). Considering the i-th conjugacy class has size r_i, we can write

$$\sum_i r_i\chi_i\chi_i^* = |G|,$$

where the index i runs over all conjugacy classes, and χ_i denotes the value of the character of any element belonging in the i-th conjugacy class. Now, dividing both sides by $\chi(e) = d_\rho$ we get

$$\sum_i \frac{r_i\chi_i}{d_\rho}\chi_i^* = \frac{|G|}{d_\rho}.$$

From the property (6) of characters, we know $\chi(g)$ is always an algebraic number. Furthermore, since all eigenvalues of $\rho(g)$ are i-th roots of unity (they must satisfy $\lambda^r - 1 = 0$), they are algebraic integers, meaning all characters are in fact algebraic integers. so χ_i^* is also an algebraic integer. Now consider that for the i-th conjugacy class $C_i \subseteq G$ we define

$$\sigma \equiv \sum_{g\in C_i}\rho(g).$$

Let us also consider that such conjugacy class has size r_i and the character of any $g \in C_i$ is χ_i. Notice that for any $h \in G$ we have

$$\rho(h^{-1})\sigma\rho(h) = \sum_{g\in C_i}\rho(h^{-1}gh) = \sum_{g\in C_i}\rho(g) = \sigma,$$

where we used the fact that $h^{-1}gh$ is still in the conjugacy class C_i. But then $\rho(h^{-1})\sigma\rho(h) = \sigma$ means that $\sigma\rho(h) = \rho(h)\sigma$ for all $h \in G$. Applying Schur's lemma, σ must be a multiple of the identity, that is $\sigma \equiv \lambda I$ for some number λ. It is possible to determine the value of the eigenvalue λ by taking the trace of σ. First, we have that $\mathrm{tr}(\sigma) = \lambda\,\mathrm{tr}(I) = \lambda d_\rho$, but also

$$\mathrm{tr}(\sigma) = \sum_{g\in C}\mathrm{tr}(\rho(g)) = r_i\chi_i \quad\Longrightarrow\quad \lambda = \frac{r_i\chi_i}{d_\rho}.$$

We have that σ is an integer linear combination of matrices with algebraic integer eigenvalues, thus λ must also be an algebraic integer. This means that the entire sum $\sum_i(r_i\chi_i/d_\rho)\chi_i^*$ is an algebraic integer, and so is $|G|/d_\rho$. But given that both $|G|$ and d_ρ are integers, the only way for this quotient to be an algebraic integer is if $|G|/d_\rho$ is in fact an integer.

A2.15

For the first relation we multiply both sides by $\delta_{ij}\delta_{kl}$, and sum over i, k, j, and l yielding

$$\sum_{g\in G}\sum_{i=1}^{d_\rho}\sum_{k=1}^{d_{\rho^q}}\sum_{j=1}^{d_\rho}\sum_{l=1}^{d_{\rho^q}}\delta_{ij}\delta_{kl}\left[\rho^p(g)\right]^{-1}_{ij}\left[\rho^q(g)\right]_{kl}=\frac{|G|}{d_\rho}\sum_{i=1}^{d_\rho}\sum_{k=1}^{d_{\rho^q}}\sum_{j=1}^{d_\rho}\sum_{l=1}^{d_{\rho^q}}\delta_{ij}\delta_{kl}\delta_{il}\delta_{jk}\delta_{pq}.$$

Since any finite representation is equivalent to a unitary matrix representation, we may consider that $[\rho^p(g)]^{-1}=[\rho^p(g)]^\dagger$, and thus simplifying both sides yields

$$\sum_{g\in G}\sum_{i=1}^{d_\rho}\sum_{k=1}^{d_{\rho^q}}\left[\rho^p(g)\right]^\dagger_{ii}\left[\rho^q(g)\right]_{kk}=\frac{|G|}{d_\rho}\sum_{i=1}^{d_\rho}\sum_{k=1}^{d_{\rho^q}}\delta_{ik}\delta_{pq}.$$

The right-hand side simplifies to $|G|\delta_{pq}$, and using the definition of character, the left-hand side simplifies to

$$\begin{aligned}\sum_{g\in G}\sum_{i=1}^{d_\rho}\sum_{k=1}^{d_{\rho^q}}\left[\rho^p(g)\right]^\dagger_{ii}\left[\rho^q(g)\right]_{kk}&=\sum_{g\in G}\left(\sum_{i=1}^{d_\rho}\left[\rho^p(g)\right]^\dagger_{ii}\right)\left(\sum_{k=1}^{d_{\rho^q}}\left[\rho^q(g)\right]_{kk}\right)\\&=\sum_{g\in G}\left[\chi^p(g)\right]^*\chi^q(g),\end{aligned}$$

thus

$$\sum_{g\in G}\left[\chi^p(g)\right]^*\chi^q(g)=|G|\delta_{pq}.$$

The final step is to notice that, since the character is the same for all elements of a conjugacy class, instead of summing over each element individually, we may sum over the conjugacy classes of g, accounting for their sizes r_i, that is

$$\sum_{g\in G}\left[\chi^p(g)\right]^*\chi^q(g)=\sum_{i=1}^{r}r_i\left(\chi^p_i\right)^*\chi^q_i\quad\Longrightarrow\quad\sum_{i=1}^{r}r_i\left(\chi^p_i\right)^*\chi^q_i=|G|\delta_{pq}.$$

For the second relation, we begin with the first one. Dividing both sides by $|G|$, it is possible to rewrite it as an equality between two $r\times r$ matrices as

$$\frac{1}{|G|}\begin{bmatrix}\sum_{i=1}^r r_i(\chi_i^1)^*\chi_i^1 & \sum_{i=1}^r r_i(\chi_i^1)^*\chi_i^2 & \cdots & \sum_{i=1}^r r_i(\chi_i^1)^*\chi_i^r\\ \sum_{i=1}^r r_i(\chi_i^2)^*\chi_i^1 & \sum_{i=1}^r r_i(\chi_i^2)^*\chi_i^2 & \cdots & \sum_{i=1}^r r_i(\chi_i^2)^*\chi_i^r\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{i=1}^r r_i(\chi_i^r)^*\chi_i^1 & \sum_{i=1}^r r_i(\chi_i^r)^*\chi_i^2 & \cdots & \sum_{i=1}^r r_i(\chi_i^r)^*\chi_i^r\end{bmatrix}=\begin{bmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{bmatrix}.$$

Defining the matrices

$$A \equiv \frac{1}{|G|}\begin{bmatrix} r_1(\chi_1^1)^* & r_2(\chi_2^1)^* & \cdots & r_r(\chi_r^1)^* \\ r_1(\chi_1^2)^* & r_2(\chi_2^2)^* & \cdots & r_r(\chi_r^2)^* \\ \vdots & \vdots & \ddots & \vdots \\ r_1(\chi_1^r)^* & r_2(\chi_2^r)^* & \cdots & r_r(\chi_r^r)^* \end{bmatrix} \quad \text{and} \quad B \equiv \begin{bmatrix} \chi_1^1 & \chi_1^2 & \cdots & \chi_1^r \\ \chi_2^1 & \chi_2^2 & \cdots & \chi_2^r \\ \vdots & \vdots & \ddots & \vdots \\ \chi_r^1 & \chi_r^2 & \cdots & \chi_r^r \end{bmatrix},$$

this equality takes the form $AB = I_{r\times r}$, from which we conclude that $B = A^{-1}$. But then we must also have $BA = I_{r\times r}$, meaning

$$BA = \frac{1}{|G|}\begin{bmatrix} \sum_{p=1}^r r_1(\chi_1^p)^*\chi_1^p & \sum_{p=1}^r r_2(\chi_2^p)^*\chi_1^p & \cdots & \sum_{p=1}^r r_r(\chi_r^p)^*\chi_1^p \\ \sum_{p=1}^r r_1(\chi_1^p)^*\chi_2^p & \sum_{p=1}^r r_2(\chi_2^p)^*\chi_2^p & \cdots & \sum_{p=1}^r r_r(\chi_r^p)^*\chi_2^p \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{p=1}^r r_1(\chi_1^p)^*\chi_r^p & \sum_{p=1}^r r_2(\chi_2^p)^*\chi_r^p & \cdots & \sum_{p=1}^r r_r(\chi_r^p)^*\chi_r^p \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Reverting back to index notation we obtain

$$\frac{1}{|G|}\sum_{p=1}^r r_i(\chi_i^p)^*\chi_j^p = \delta_{ji} \quad \Longrightarrow \quad \sum_{p=1}^r (\chi_i^p)^*\chi_j^p = \frac{|G|}{r_i}\delta_{ij}.$$

A2.16

Let us name the group elements respectively as $\{e, g_-, g_+, g_{12}, g_{23}, g_{13}\}$. The composition table of these elements (row then column) can be written as

	e	g_-	g_+	g_{12}	g_{23}	g_{13}
e	e	g_-	g_+	g_{12}	g_{23}	g_{13}
g_-	g_-	g_+	e	g_{13}	g_{12}	g_{23}
g_+	g_+	e	g_-	g_{23}	g_{13}	g_{12}
g_{12}	g_{12}	g_{23}	g_{13}	e	g_-	g_+
g_{23}	g_{23}	g_{13}	g_{12}	g_+	e	g_-
g_{13}	g_{13}	g_{12}	g_{23}	g_-	g_+	e

The trivial function $\rho^1 : S_3 \to 1$, which maps all elements to 1, is a representation because $\rho^1(g_a)\rho^1(g_b) = 1 = \rho^1(g_a g_b)$ for any $g_a, g_b \in S_3$. And since it is one-dimensional, it is irreducible.

From the table, we can also see that the function $\rho^\pm : S_3 \to \{-1, 1\}$, which maps the subset $S_3^{(1)} \equiv \{e, g_-, g_+\}$ to 1 and the subset $S_3^{(2)} \equiv \{g_{12}, g_{23}, g_{13}\}$ to -1, is also a representation. To see that, notice that we have two possible cases:

- $(g_a, g_b \in S_3^{(1)}) \vee (g_a, g_b \in S_3^{(2)})$:
  from the table $g_a g_b \in S_3^{(1)} \implies \rho^\pm(g_a g_b) = 1$
  we have $\rho^\pm(g_a) = \rho^\pm(g_b) \implies \rho^\pm(g_a)\rho^\pm(g_b) = 1 = \rho^\pm(g_a g_b)$;
- $(g_a \in S_3^{(1)}; g_b \in S_3^{(2)}) \vee (g_a \in S_3^{(2)}; g_b \in S_3^{(1)})$:
  from the table $g_a g_b \in S_3^{(2)} \implies \rho^\pm(g_a g_b) = -1$
  we have $\rho^\pm(g_a) \neq \rho^\pm(g_b) \implies \rho^\pm(g_a)\rho^\pm(g_b) = -1 = \rho^\pm(g_a g_b)$.

Like the trivial representation, since $\rho^{\pm}$ is one-dimensional, it is also irreducible.

One last representation $\rho : S_3 \to M_2$ can be obtained by considering a generating set of S_3, which can be chosen as $\{g_-, g_{12}\}$. The reason is that using the table we can directly verify that

$$\begin{aligned} g_- g_- g_- &= g_{12} g_{12} = e, \\ g_- g_- &= g_+, \\ g_- g_{12} &= g_{13}, \\ g_{12} g_- &= g_{23}, \end{aligned}$$

and since we know this set is closed, we conclude that $S_3 = \langle g_-, g_{12} \rangle$. Thus it suffices to find a function ρ whose compositions satisfy these relations and assign the identity matrix to the identity element. We start by setting

$$\rho(g_-) \equiv \frac{1}{2} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \quad \rho(g_{12}) \equiv \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Direct verification yields

$$\begin{aligned} \rho(g_-)\rho(g_-)\rho(g_-) &= \frac{1}{8} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \equiv \rho(e), \\ \rho(g_{12})\rho(g_{12}) &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \equiv \rho(e), \end{aligned}$$

showing that the identity matrix can be correctly assigned to the identity element, satisfying the first condition. The remaining three give the representation of the other three elements, which are

$$\begin{aligned} \rho(g_-)\rho(g_-) &= \frac{1}{4} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix} \equiv \rho(g_+), \\ \rho(g_-)\rho(g_{12}) &= \frac{1}{2} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix} \equiv \rho(g_{13}), \\ \rho(g_{12})\rho(g_-) &= \frac{1}{2} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} \equiv \rho(g_{23}). \end{aligned}$$

Therefore, the matrices

$$\begin{aligned} \rho(e) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \rho(g_-) = \frac{1}{2} \begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \quad \rho(g_+) = \frac{1}{2} \begin{bmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix}, \\ \rho(g_{12}) &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \rho(g_{23}) = \frac{1}{2} \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \quad \rho(g_{13}) = \frac{1}{2} \begin{bmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix}, \end{aligned}$$

form a representation of the S_3 group. To see that it is irreducible, notice that $|\chi(e)|^2 = 4$, $|\chi(g_-)|^2 = |\chi(g_+)|^2 = 1$, and $|\chi(g_{12})|^2 = |\chi(g_{23})|^2 = |\chi(g_{13})|^2 = 0$, thus $\sum_{g \in S_3} |\chi(g)|^2 = 6 = |S_3|$, so it is an irreducible representation according to Theorem A2.3.

To verify their orthogonality we only need the characters. For the ρ^1 representation we have $\chi^1(g) = 1$ for all $g \in S_3$. For $\rho^\pm$ we have $\chi^\pm(g) = 1$ for $g \in S_3^{(1)}$ and $\chi^\pm(g) = -1$ for $g \in S_3^{(2)}$. And finally, for ρ we have $\chi(e) = 2$, $\chi(g_-) = \chi(g_+) = -1$, and $\chi(g_{12}) = \chi(g_{23}) = \chi(g_{13}) = 0$. Thus

$$\sum_{g\in S_3} [\chi^1(g)]^* \chi^\pm(g) = 1 + 1 + 1 - 1 - 1 - 1 = 0,$$
$$\sum_{g\in S_3} [\chi^1(g)]^* \chi(g) = 2 - 1 - 1 + 0 + 0 + 0 = 0,$$
$$\sum_{g\in S_3} [\chi^\pm(g)]^* \chi(g) = -2 + 1 + 1 + 0 + 0 + 0 = 0,$$

showing that they are orthogonal among them.

A2.17

Since there are $|G|$ distinct matrices and G has $|G|$ elements, it is possible to create a one-to-one map between the elements of G and the permutation matrices, meaning this representation is an isomorphism, thus faithful.

A2.18

A permutation matrix that sends the i-th entry of $\vec{v}$ to the j-th position has a 1 in the i-th row and j-th column, and zeros in all other entries of this row and column. For $g, g_i \in G$, the only way to get $gg_i = g_i$ is if g is the identity element e. This means that, except for the identity, all permutation matrices must rearrange all entries of $\vec{v}$, so the positions containing a 1 are always such that $i \neq j$. Therefore all permutation matrices have null diagonals, meaning $\chi(g) = 0$ for all $g \neq e$. The identity on the other hand has $|G|$ entries with 1 in the diagonal, meaning $\chi(e) = |G|$.

A2.19

From Theorem A2.5 we have that $\rho = \bigoplus_p c_p \rho^p$, meaning any ρ can be put in a block diagonal form containing, for all p, c_p copies of the p-th irreducible representation. By taking the trace in both sides we get

$$\mathrm{tr}(\rho) = \sum_p c_p \,\mathrm{tr}(\rho^p) \quad \Longrightarrow \quad \chi = \sum_p c_p \chi^p.$$

Also according to Theorem A2.5, the multiplicities are given by

$$c_p = \frac{1}{|G|} \sum_{i=1}^{r} r_i (\chi_i^p)^* \chi_i.$$

If the representation is the regular one, R, then we know that χ^R is zero for all group elements except the identity, for which we have $\chi^R(e) = |G|$. The conjugacy class of the identity contains only the

identity itself so its size is 1. Therefore the multiplicity is calculated as

$$c_p = \frac{1}{|G|}\sum_{i=1}^{r} r_i(\chi_i^p)^*\chi_i^R = \frac{1}{|G|}[\chi^p(e)]^*\chi^R(e) = \chi^p(e).$$

But for any representation ρ^p, $\chi^p(e)$ will be the dimension of that representation, so $c_p = d_{\rho^p}$. Thus the regular representation contains d_{ρ^p} instances of each representation ρ^p, and the character relation in this case will be

$$\chi^R = \sum_{\rho\in\hat{G}} d_\rho \chi^\rho.$$

A2.20

Applying the relation $\chi^R = \sum_{\rho\in\hat{G}} d_\rho\chi^\rho$ to some element $g \in G$ we have

$$\chi^R(g) = \sum_{\rho\in\hat{G}} d_\rho \chi^\rho(g).$$

We know that $\chi^R(g) = |G|\delta_{ge}$, so defining $N \equiv |G|$ we get

$$\sum_{\rho\in\hat{G}} d_{\rho^p}\chi^p(g) = N\delta_{ge}.$$

A2.21

Applying the relation $\chi^R = \sum_{\rho\in\hat{G}} d_\rho\chi^\rho$ to the identity element e we have

$$\chi^R(e) = \sum_{\rho\in\hat{G}} d_\rho \chi^\rho(e).$$

We know that $\chi^R(e) = |G|$, and $\chi^\rho(e) = d_\rho$ for all representations, so

$$\sum_{\rho\in\hat{G}} d_\rho^2 = |G|.$$

A2.22

Substituting (A2.10) into (A2.9) yields

$$\begin{aligned}\hat{f}(\rho) &= \sqrt{\frac{d_\rho}{N}}\sum_{g\in G}\left[\frac{1}{\sqrt{N}}\sum_{\sigma\in\hat{G}}\sqrt{d_\sigma}\,\mathrm{tr}\Big(\hat{f}(\sigma)\sigma(g^{-1})\Big)\right]\rho(g)\\ &= \frac{1}{N}\sum_{\sigma\in\hat{G}}\sqrt{d_\rho d_\sigma}\sum_{g\in G}\mathrm{tr}\Big(\hat{f}(\sigma)\sigma(g^{-1})\Big)\rho(g).\end{aligned}$$

It will be convenient to write this relation in terms of the matrix indices. In both sides we set the pair of i-th row and j-th column with $[\hat{f}(\rho)]_{ij}$ and $[\rho(g)]_{ij}$. Now notice that $[\hat{f}(\sigma)\sigma(g^{-1})]_{kl} =$

$\sum_m [\hat{f}(\sigma)]_{km}[\sigma(g)]^{-1}_{ml}$, and taking the trace adds a sum over k with $k = l$. Thus we can write

$$\left[\hat{f}(\rho)\right]_{ij} = \frac{1}{N}\sum_{\sigma\in\hat{G}}\sqrt{d_\rho d_\sigma}\sum_{g\in G}\sum_{k,m}\left[\hat{f}(\sigma)\right]_{km}\left[\sigma(g)\right]^{-1}_{mk}\left[\rho(g)\right]_{ij}.$$

Using the orthogonality relation shown in Equation (A2.3) we get

$$\begin{aligned}\left[\hat{f}(\rho)\right]_{ij} &= \frac{1}{N}\sum_{\sigma\in\hat{G}}\sqrt{d_\rho d_\sigma}\sum_{k,m}\left[\hat{f}(\sigma)\right]_{km}\frac{N}{d_\sigma}\delta_{mj}\delta_{ki}\delta_{\sigma\rho}\\ &= \sum_{\sigma\in\hat{G}}\frac{\sqrt{d_\rho d_\sigma}}{d_\sigma}\delta_{\sigma\rho}\sum_{k,m}\left[\hat{f}(\sigma)\right]_{km}\delta_{ki}\delta_{mj}\\ &= \left[\hat{f}(\rho)\right]_{ij}.\end{aligned}$$

A2.23

From the definition of $\hat{f}$ we have

$$\hat{f}(\rho) = \sqrt{\frac{d_\rho}{N}}\sum_{g\in G} f(g)\rho(g).$$

In this case, we have an additive Abelian group with elements represented by $g \in [0, N-1]$, and the representation given by $\rho_h(g) \equiv \exp[-2\pi i g h/N]$, with dimension $d_\rho = 1$. Since the number h uniquely identifies each representation ρ_h, we can directly use it for identification, instead of using ρ_h. In other words, we can do the relabeling $\rho_h \to h$. Direct substitution yields

$$\hat{f}(h) = \frac{1}{\sqrt{N}}\sum_{g=0}^{N-1} f(g)e^{-2\pi i g h/N}.$$

From the definition of f we have

$$f(g) = \frac{1}{\sqrt{N}}\sum_{\rho\in\hat{G}}\sqrt{d_\rho}\,\mathrm{tr}\Big(\hat{f}(\rho)\rho(g^{-1})\Big).$$

Since the group is Abelian, all its irreducible representations are one-dimensional. So given that $|G| = N$, it must have N representations ρ_h, meaning $h \in [0, N-1]$ exactly like g. With these considerations, we can write

$$f(g) = \frac{1}{\sqrt{N}}\sum_{h=0}^{N-1}\hat{f}(h)\rho_h(g^{-1}).$$

Given that $\rho_h(g^{-1})\rho_h(g) = 1$, we obtain $\rho_h(g^{-1}) = \exp[2\pi i g h/N]$, and since the set g and the indices h are isomorphic, we may relabel $g \leftrightarrow h$ in the sum, yielding

$$f(h) = \frac{1}{\sqrt{N}}\sum_{g=0}^{N-1}\hat{f}(g)e^{2\pi i g h/N}.$$

A2.24

Starting with the trivial representation ρ^1, with $\rho^1(g) = 1$ for all $g \in G$, the Fourier transform is

$$\begin{aligned}\hat{f}(\rho^1) &= \sqrt{\frac{d_{\rho^1}}{|S_3|}} \sum_{g\in S_3} f(g)\rho^1(g) = \frac{1}{\sqrt{6}} \sum_{g\in S_3} f(g) \\ &= \frac{1}{\sqrt{6}} \left[f(e) + f(g_-) + f(g_+) + f(g_{12}) + f(g_{23}) + f(g_{13})\right].\end{aligned}$$

For the representation $\rho^\pm$, with $\rho^\pm(g) = 1$ for $g \in S_3^{(1)} \equiv \{e, g_-, g_+\}$, and $\rho^\pm(g) = -1$ for $g \in S_3^{(2)} \equiv \{g_{12}, g_{23}, g_{13}\}$, we write the Fourier transform as

$$\begin{aligned}\hat{f}(\rho^\pm) &= \sqrt{\frac{d_{\rho^\pm}}{|S_3|}} \sum_{g\in S_3} f(g)\rho^\pm(g) = \frac{1}{\sqrt{6}} \sum_{g\in S_3^{(1)}} f(g) - \frac{1}{\sqrt{6}} \sum_{g\in S_3^{(2)}} f(g) \\ &= \frac{1}{\sqrt{6}} \left[f(e) + f(g_-) + f(g_+) - f(g_{12}) - f(g_{23}) - f(g_{13})\right].\end{aligned}$$

For the two-dimensional representation ρ, for each matrix component we will have

$$\Big[\hat{f}(\rho)\Big]_{ij} = \sqrt{\frac{d_\rho}{|S_3|}} \sum_{g\in S_3} f(g)\Big[\rho(g)\Big]_{ij} = \frac{1}{\sqrt{3}} \sum_{g\in S_3} f(g)\Big[\rho(g)\Big]_{ij},$$

where the $\rho(g)$ are given by

$$\begin{aligned}\rho(e) &= \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \quad \rho(g_-) = \frac{1}{2}\begin{bmatrix}-1 & -\sqrt{3}\\ \sqrt{3} & -1\end{bmatrix}, \quad \rho(g_+) = \frac{1}{2}\begin{bmatrix}-1 & \sqrt{3}\\ -\sqrt{3} & -1\end{bmatrix}, \\ \rho(g_{12}) &= \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}, \quad \rho(g_{23}) = \frac{1}{2}\begin{bmatrix}1 & \sqrt{3}\\ \sqrt{3} & -1\end{bmatrix}, \quad \rho(g_{13}) = \frac{1}{2}\begin{bmatrix}1 & -\sqrt{3}\\ -\sqrt{3} & -1\end{bmatrix}.\end{aligned}$$

So explicitly we have

$$\begin{aligned}\Big[\hat{f}(\rho)\Big]_{11} &= \frac{1}{\sqrt{3}}\left[f(e) - \frac{f(g_-)}{2} - \frac{f(g_+)}{2} - f(g_{12}) + \frac{f(g_{23})}{2} + \frac{f(g_{13})}{2}\right], \\ \Big[\hat{f}(\rho)\Big]_{12} &= \frac{1}{2}\left[-f(g_-) + f(g_+) + f(g_{23}) - f(g_{13})\right], \\ \Big[\hat{f}(\rho)\Big]_{21} &= \frac{1}{2}\left[f(g_-) - f(g_+) + f(g_{23}) - f(g_{13})\right], \\ \Big[\hat{f}(\rho)\Big]_{22} &= \frac{1}{\sqrt{3}}\left[f(e) - \frac{f(g_-)}{2} - \frac{f(g_+)}{2} + f(g_{12}) - \frac{f(g_{23})}{2} - \frac{f(g_{13})}{2}\right].\end{aligned}$$

Ordering the array of values of $f(g)$ as $(f(e), f(g_-), f(g_+), f(g_{12}), f(g_{23}), f(g_{13}))$ and looking at the

obtained coefficients, we can write the Fourier transform as the unitary matrix

$$\hat{f} \equiv \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} \end{bmatrix},$$

which, upon acting on the array of $f(g)$, will yield $(\hat{f}(\rho^1), \hat{f}(\rho^{\pm}), [\hat{f}(\rho)]_{11}, [\hat{f}(\rho)]_{12}, [\hat{f}(\rho)]_{21}, [\hat{f}(\rho)]_{22})$.

The Solovay–Kitaev theorem

A3.1

-

A3.2

-

A3.3

-

A3.4

-

A3.5

-

A3.6

-

Number theory

A4.1

If $a|b$ and $b|c$ there exist integers k_1 and k_2 such that $b = ak_1$ and $c = bk_2$, which implies $c = ak_1k_2$. Since the product k_1k_2 is also an integer we have $a|c$.

A4.2

If $d|a$ and $d|b$ there exist integers k_1 and k_2 such that $a = dk_1$ and $b = dk_2$. So for integers x and y we may write the linear combination

$$\begin{aligned} ax + by &= dk_1x + dk_2y \\ &= d(k_1x + k_2y), \end{aligned}$$

and since $k_1x + k_2y$ is an integer, the linear combnation $ax + by$ is divisible by d .

A4.3

If $a|b$ then there exists a positive integer k such that $b = ak$. Being a positive integer, we must have $k \geq 1$, which implies $a \leq b$. If $b|a$ then $b \leq a$. Combining both conditions yields $a = b$.

A4.4

$$\begin{aligned} 697 &= 17 \times 41, \\ 36300 &= 2^2 \times 3 \times 5^2 \times 11^2. \end{aligned}$$

A4.5

If a number p is prime then it does not share any common factors with any number in the range $[1, p-1]$, that is, for all $x \in [1, p-1]$ we have $\gcd(x, p) = 1$, meaning every x in this range has a multiplicative inverse modulo p.

Now, for numbers in the range $[1, p^2 - 1]$, the ones that do not have a multiplicative inverse modulo p^2 are the ones with p as a prime factor. Those are $\{p, 2p, \cdots, (p-1)p\}$.

A4.6

The inverse a will be the solution to the equation $17a - 24b = 1$ for some integer b such that a is also an integer. That is, a is the smallest integer such that

$$a = \frac{1 + 24b}{17}.$$

The smallest b possible is $b = 12$, which yields $a = 17$, meaning 17 is its own inverse modulo 24.

A4.7

We may write $n^2 = (n+1)(n-1)+1$, which means that $(n+1)(n-1) = -1 \pmod{n^2}$. Multiplying both sides by -1 yields $(n+1)(1-n) = 1 \pmod{n^2}$. So $1 - n$ is the multiplicative inverse of $n + 1$ modulo n^2. If we want only numbers in the range $[1, n^2 - 1]$ we just have to sum n^2, meaning the multiplicative inverse is $n^2 - n + 1$.

A4.8

If both b and b' are multiplicative inverses of a modulo n then $ab \equiv 1 \pmod{n} \equiv ab' \pmod{n}$. Dividing both extremes of the congruence relations by a yields $b \equiv b' \pmod{n}$.

A4.9

If we have $a = p_1^{\alpha_1} \cdots p_n^{\alpha_n}$ and $b = p_1^{\beta_1} \cdots p_n^{\beta_n}$, then the gcd can be calculated as

$$\gcd(a, b) = \prod_{i=1}^{n} p_i^{\min\{\alpha_i, \beta_i\}}.$$

For 6825 and 1430 we have

$$6825 = 3^1 \times 5^2 \times 7^1 \times 13^1,$$
$$1430 = 2^1 \times 5^1 \times 11^1 \times 13^1,$$

thus

$$\gcd(6825, 1430) = 5^1 \times 13^1 = 65.$$

A4.10

The prime factorization yields $187 = 11^1 \times 17^1$, thus

$$\varphi(187) = (11 - 1)(17 - 1) = 160.$$

A4.11

First, consider the case where n is the α-th power of some prime number p, that is $n = p^\alpha$. So all the numbers d that divides n belong in the set $\{1, p, p^2, \cdots, p^\alpha\}$. The Euler totient function for some number p^m $(m > 0)$ in this set is

$$\varphi(p^m) = p^{m-1}(p - 1).$$

If we sum the Euler totient function for all divisors d we get

$$\begin{aligned}\sum_{d|n} \varphi(d) &= \varphi(1) + (p - 1) \sum_{m=1}^{\alpha} p^{m-1} \\ &= 1 + (p - 1)\frac{p^\alpha - 1}{p - 1} \\ &= p^\alpha = n.\end{aligned}$$

The result can be extended immediately for the case of general n since it could be written as $n = \prod_i n_i$, where each n_i is the power of some prime number p_i, that is, $n_i = p_i^{a_i}$. The relation would be valid for each n_i, and because of the property $\varphi(ab) = \varphi(a)\varphi(b)$, it would also be valid for n.

To see this extension explicitly let us consider that $n = p^\alpha q$, where $q \neq p$ is a prime number. Now, besides $\{1, p, p^2, \cdots, p^\alpha\}$, the set $\{q, pq, p^2q, \cdots, p^\alpha q\}$ also contains possible divisors d. The Euler totient function for some number $p^m q$ $(m > 0)$ is

$$\varphi(p^m q) = p^{m-1}(p-1)(q-1),$$

thus summing for all divisors yields

$$\begin{aligned}\sum_{d|n} \varphi(d) &= \varphi(1) + (p-1)\sum_{m=1}^{\alpha} p^{m-1} + \varphi(q) + (p-1)(q-1)\sum_{m=1}^{\alpha} p^{m-1} \\ &= 1 + (p-1)\frac{p^\alpha - 1}{p-1} + (q-1) + (p-1)(q-1)\frac{p^\alpha - 1}{p-1} \\ &= p^\alpha + p^\alpha(q-1) \\ &= p^\alpha q = n.\end{aligned}$$

A4.12

$\mathbf{Z}_n^*$ is the set off all elements in $\mathbf{Z}_n$ that have an inverse modulo n, those are, all numbers $x < n$ such that $\gcd(x, n) = 1$, and there are, by definition, $\varphi(n)$ of these numbers. Now we must only show this set satisfies the group properties:

- *closure*

Let $a, b \in \mathbf{Z}_n^*$. Both, a and b, do not share common factors with n, so $(a \cdot b \mod n) \in \mathbf{Z}_n^*$.

- *associativity*

Naturally satisfied by the properties of modular multiplication.

- *identity*

The number $1 \in \mathbf{Z}_n^*$ is the identity since for any $a \in \mathbf{Z}_n^*$ we have $1 \cdot a = a \cdot 1 = a$.

- *inverses*

An integer a has a multiplicative inverse modulo n if and only if $\gcd(a, n) = 1$, but since by definition all elements of $\mathbf{Z}_n^*$ satisfy this property, all elements have an inverse.

A4.13

If $a \in \mathbf{Z}_n^*$ then a does not share any common factors with n, meaning $a^k \in \mathbf{Z}_n^*$ for any integer k. So if we take the set $S = \{1, a, a^2, \cdots, a^{r-1}\}$, where $a^r \equiv 1 \pmod{n}$, we clearly have a subset of $\mathbf{Z}_n^*$ with size r. We must only show it satisfies the group properties:

- *closure*

Naturally satisfied by the construction of the set.

- *associativity*

Naturally satisfied by the properties of modular multiplication.

- *identity*

The identity $1 \in \mathbf{Z}_n^*$ also belongs in this set.

- *inverses*

Since there exists $r \coloneqq x + y$ such that $a^r \equiv 1 \pmod{n}$ there is always a pair a^x and a^y, with $x < r$ and $y < r$, such that $a^x \cdot a^y = a^{x+y} \equiv 1 \pmod{n}$, that is, all elements have an inverse.

A4.14

If g is a generator of $\mathbf{Z}_n^*$ then all $\varphi(n)$ elements of the group can be written as powers of g, that is, $\mathbf{Z}_n^* = \left\{1, g, g^2, \cdots, g^{\varphi(n)-1}\right\}$, where $g^{\varphi(n)} \equiv 1 \pmod{n}$, meaning that g has order $\varphi(n)$.

A4.15

Consider the element $a \in \mathbf{Z}_n^*$ and the subgroup $\{1, a, a^2, \cdots, a^{r-1}\}$ of the group $\mathbf{Z}_n^*$, where r is some number such that $a^r = 1 \pmod{n}$. Using Lagrange's theorem we have that r must divide the size of $\mathbf{Z}_n^*$, that is, $\varphi(n)/r$ is an integer. Alternatively, we may write $\varphi(n) = \alpha r$, where α is a positive integer. Now, if $a^r = 1 \pmod{n}$ then there exists some integer k such that

$$a^r = kn + 1.$$

If we exponentiate both sides with α we obtain

$$\begin{aligned} a^{\varphi(n)} &= (kn+1)^\alpha \\ &= 1 + \sum_{i=1}^{\alpha} \binom{\alpha}{i} (kn)^i. \end{aligned}$$

The second term is clearly divisible by n, thus we have that $a^{\varphi(n)} = 1 \pmod{n}$, *Q.E.D.*

A4.16

If r is the order of x modulo N then $x^r \equiv 1 \pmod{N}$. Since r is the smallest number that satisfies such relation, any other number that also satisfies it must necessarily be a multiple integer of r. From Theorem A4.9 we have that $x^{\varphi(N)} = 1 \pmod{N}$, so there exists some integer α such that $\varphi(N) = \alpha r$, or equivalently $\varphi(N)/r = \alpha$, meaning $r | \varphi(N)$.

A4.17

If we have an efficient factoring algorithm we can efficiently compute the list $P = \{p_1^{\alpha_1}, \cdots p_m^{\alpha_m}\}$ of prime factors of a number N. We know that the order r of x modulo N is such that $r \leq N$, so we can run Euclid's algorithm to compute $\gcd(x^{s/2} - 1, N)$ and $\gcd(x^{s/2} + 1, N)$ for all even $s \in [1, N]$ until we get a number belonging in the list P. When we do we output the order $r = s$, and in the

worst case scenario we need to run Euclid's algorithm $O(N)$ times. And similarly to the reduction of factoring to order-finding, we would have a probability $p \leq 1/2^m$ of failing, which would occur if r is odd or if $x^{s/2} = -1 \pmod{N}$ for all s.

A4.18

$$\frac{19}{17} = 1 + \frac{1}{\frac{17}{2}} = 1 + \frac{1}{8 + \frac{1}{2}} \quad \Longrightarrow \quad \frac{19}{17} = [1, 8, 2],$$

$$\frac{77}{65} = 1 + \frac{1}{\frac{65}{12}} = 1 + \frac{1}{5 + \frac{1}{\frac{12}{5}}} = 1 + \frac{1}{5 + \frac{1}{2 + \frac{1}{\frac{5}{2}}}} = 1 + \frac{1}{5 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2}}}} \quad \Longrightarrow \quad \frac{77}{65} = [1, 5, 2, 2, 2]$$

A4.19

Let us first check for the case $n = 1$. From the definitions we have the relations: $p_0 = a_0$, $q_0 = 1$, $p_1 = 1 + a_0 a_1$ and $q_1 = a_1$, thus we verify that

$$\begin{aligned} q_1 p_0 - p_1 q_0 &= a_1 a_0 - (1 + a_0 a_1) \\ &= -1 = (-1)^1. \end{aligned}$$

So the relation is clearly true for $n = 1$. Now we must only show that, for $n \geq 2$, we have $q_{n+1} p_n - p_{n+1} q_n = (-1) \times (q_n p_{n-1} + p_n q_{n-1})$. From the definitions we have, for $n \geq 2$, the relations: $p_{n+1} = a_{n+1} p_n + p_{n-1}$ and $q_{n+1} = a_{n+1} q_n + q_{n-1}$, thus

$$\begin{aligned} q_{n+1} p_n - p_{n+1} q_n &= (a_{n+1} q_n + q_{n-1}) p_n - (a_{n+1} p_n + p_{n-1}) q_n \\ &= p_n q_{n-1} - q_n p_{n-1} \\ &= (-1) \times (q_n p_{n-1} - p_n q_{n-1}). \end{aligned}$$

So this relation is true for all $n \geq 1$. Now, this relation can always be written in the form $\alpha p_n + \beta q_n = 1$, and using Theorem A4.2 we conclude that $\gcd(p_n, q_n) = 1$.

Public key cryptography and the RSA cryptosystem

A5.1

-

A5.2

-

Proof of Lieb's theorem

A6.1

-

A6.2

-

A6.3

-

A6.4

-

A6.5

-

A6.6

-

A6.7

-

A6.8

-