ERRATA FOR KUNG, ROTA, AND YAN,
Combinatorics: The Rota Way

* denotes mathematical errors, including serious typographical errors.

** denotes very substantial errors.

1. *Maximal vs. maximum chains, et al. The difference between these concepts is important in combinatorics and not discussed in our book. A *maximum chain* is a chain of maximum size.

   Proposition. A maximum chain is maximal, but a maximal chain need not be maximum.
2. Stirling numbers (note by Tom Zaslavsky): They are in the index under "numbers" (not "Stirling numbers").

   An error: the book uses the notation $s(n, k)$ for the *unsigned* Stirling number of the first kind, but this is not standard. The usual and traditional notation is that $s(n, k)$ is the *Stirling number of the first kind*. If we write $c(n, k)$ for the unsigned Stirling number, then $s(n, k) = (-1)^{n-k}c(n, k)$. It was a mistake to use the Stirling number notation for a different number.

   The book seems to omit the inverse relationship of the Stirling numbers (which is where they began, historically). This relationship is (using correct notation)

   $$x^n = \sum_{k\geq 0} S(n, k)(x)_k, \qquad (x)_n = \sum_{k\geq 0} s(n, k)x^k.$$

   Here $(x)_n$ denotes the falling factorial.
3. **P. 8, problem 1.1.9(b) (note by Laura Anderson): I believe that this is incorrect. For instance, let $\mathcal{G}$ be the collection of all finite unions and intersections of sets of the following forms:
   (a) $(-\infty, b)$,
   (b) $[b, \infty)$,
   (c) the single set $S = \bigcup_{i=1}^{\infty}[\frac{1}{2^i+1}, \frac{1}{2^i})$,
   (d) S^c.

   $\mathcal{G}$ satisfies the hypotheses of the problem, but it's not isomorphic to $\mathcal{H}$. $\mathcal{H}$ has the property that there is a chain C such that every element of $\mathcal{H}$ is a finite union/intersection of elements of C and $\{X^c : X \in C\}$. $\mathcal{G}$ does not have this property.
4. *P. 9, "maximal chain": The text says a chain $x_0 < x_1 < \cdots < x_n$ is "maximal or saturated" if all the $<$ are covering relations. This is a correct definition of *saturated* but it is incorrect for "maximal". A chain is *maximal* if it is not a proper subset of any other chain.

Proposition. A maximal chain is saturated, but a saturated chain need not be maximal.

5. *P. 12, problem 1.2.2b. If "suborder" means that $Q \subseteq P$ and $\leq_Q$ is the restriction of $\leq_P$ to Q, then this is correct. If "suborder" means the sets P and Q are the same but the relation $\leq_Q \subseteq \leq_P$, then this is incorrect; for instance, take Q to be an antichain and P to be a chain.
6. *Exercise 1.2.4(c): $(-1)^n$ should be $(-1)^{|P|}$.
7. **P. 21 (note by Tom Zaslavsky): Exercise 1.3.4(b) is misstated. A counterexample is a chain of length > 2. I'm not sure what they mean.
8. P. 44, first line of §1.5: "from S *to* X".
9. *P. 45, displayed equation about h: $\cup$ should be $\vee$.
10. *P. 46, displayed equation: $c \geq a$ should be $c \geq x$.
11. *P. 47, Proposition 1.5.3b: "the order dual $Q^{\downarrow}$" should be "Q".
12. **Problem 1.5.4 (note by Tom Zaslavsky): The "improvement" is mistaken. I looked in the cited papers and did not find a reasonable formula to serve as a correction. For Part (b), prove only the lower bound.
13. **P. 53: The Putnam problem is not stated correctly. The last sentence should read:

    > Show that there exists a permutation π of $\{1, 2, \ldots\}$, matching the red points with the blue points, such that no pair of finite line segments $\overline{r_i b_{\pi(i)}}$ and $\overline{r_j b_{\pi(j)}}$ crosses—that is, no such pair intersects at exactly one point in their interior.

    (Note by Kung–Yan)
14. **P. 56: The statement of Marshall Hall's Theorem is incorrect. Notice that if $|S| = 1$ then R has only k matchings. The correct statement is that R has at least $(k)_{|S|}$ matchings if $|S| \leq k$ and at least $k!$ matchings if $|S| \geq k$. (Exercise: Prove it.) (Thanks to Laura Anderson, Thomas Galvin, and Alireza Salahshoori.)
15. *Exercise 2.2.3 (note by Tom Zaslavsky): There is a missing assumption. Without assuming $|X| \geq |S|$, the conclusion is not provable. (Exercise!)

    If you solve the problem for the case $|X| = |S|$, that is good work.

    If you prove it is not true for the case $|X| > |S|$ (I don't know whether that is so, but maybe it is), that would be good work.
16. P. 61, Exercise 2.2.2. The problem may intend to assume G is finite. If it is infinite, the index is $|G{:}H|$, the number of cosets.

    In part (a): Don't use König's theorem; use a direct method so you can answer part (b).

17. *P. 63, determinant formula: There should be a sign factor, $\operatorname{sgn}(\sigma)$, in the sum.
18. *Corollary 2.3.2. In part (a), "matching" should be "partial matching". Part (b) should read "rank $\geq |S| - d$." There is no upper bound in this corollary.

    *In part (b), "rank $\geq |S| - d$."
19. *Exercise 2.3.2: In the top line of page 66, "and columns in A" should be "and columns in B".
20. **P. 71, Theorem 2.4.4: In Equation (FF), τ should be $\max \tau$. The formula gives the maximum possible size of a common transversal.

    Note that a transversal is not the same as a matching; it is only the S-ends of a matching.
21. P. 82, line 2: "row vector" should be "vector". This is standard notation for a vector and is equivalent to a column matrix.
22. *P. 82, Theorem 2.6.4: Here is a proof that is more complete and easier to follow. (Contributed by Jake Zukaitis.)

    Proof. We may assume $\mathbf{r}$ and $\mathbf{s}$ are in weakly decreasing order. This is because any reordering of $\mathbf{s}$ or $\mathbf{r}$ corresponds, respectively, to rearranging the rows or columns of D, which keeps the matrix doubly stochastic and preserves the relation $\mathbf{r} = \mathbf{s}D$. So, we let $\mathbf{r} = (r_1, \ldots, r_n)$ and $\mathbf{s} = (s_1, \ldots, s_n)$ with both in weakly decreasing order.

    Proof of Necessity. The proof is by induction on the number of corresponding entries that differ in $\mathbf{r}$ and $\mathbf{s}$. Suppose that $\mathbf{r} \preceq \mathbf{s}$. Then either $\mathbf{r} = \mathbf{s}$ and D can be taken to be the identity matrix, or there exist indices j and k such that $r_j < s_j$ and $r_k > s_k$. Choose k to be the smallest value with $r_k > s_k$ and choose j to be the largest value smaller than k with $r_j < s_j$. (j exists because $r_1 + \cdots + r_k \leq s_1 + \cdots + s_k$.) This means that $j < k$ and for all i in this interval, if there are any, $r_i = s_i$. Let $\delta = \min\{s_j - r_j, r_k - s_k\}$ and let $\lambda = \frac{\delta}{s_j - s_k}$; thus, $0 < \lambda < 1$. Let Q be the permutation matrix for the transposition (jk). Then $T_1 = (1-\lambda)I + \lambda Q$ is a transfer matrix and $\mathbf{s}' = \mathbf{s}T_1$ has entries $(s_1, \ldots, s_j - \delta, \ldots, s_k + \delta, \ldots, s_n)$. Either $\delta = s_j - r_j$, so that $s_j - \delta = r_j$, or $\delta = r_k - s_k$, so that $s_k + \delta = r_k$. Either way, $\mathbf{s}'$ will have more entries in common with $\mathbf{r}$ than $\mathbf{s}$ does. We still must show that $\mathbf{r} \preceq \mathbf{s}'$.

    First we show that $r_j \leq s'_j$. If $\delta \leq s_j - r_j$, equality holds; but if $\delta = r_k - s_k \leq s_j - r_j$, rearranging the terms gives $r_j \leq s_j + s_k - r_k = s_j - \delta = s'_j$. Similarly, $r_k \geq s'_k$.

    We next show that $\mathbf{s}'$ is in weakly decreasing order; that is, $s'_j \geq s'_{j+1}$ and $s'_k \leq s'_{k-1}$. Suppose $s'_{j+1} > s'_j$. Since $s'_j \geq r_j$, this implies

$s'_{j+1} > r_j$. If $j+1 < k$, then $s'_{j+1} = s_{j+1} = r_{j+1} \le r_j$, which is impossible. If $j+1 = k$, then $r_j \le s'_j < s'_k \le r_k$, contradicting weak decrease of $\mathbf{r}$. Similarly, $s'_{k-1} \ge s'_k$. Thus $\mathbf{s}'$ is in weakly decreasing order.

Finally we show that $\mathbf{r} \preceq \mathbf{s}'$. Clearly, $\sum_{i=1}^n r_i = \sum_{i=1}^n s'_i$. To show that $\sum_{i=1}^l r_i \le \sum_{i=1}^l s'_i$ for $l < n$, we consider cases. First, if $l < j$ or $l \ge k$, the sum for $\mathbf{s}'$ equals that for $\mathbf{s}$. Now consider the case $l = j$. Since $r_j \le s'_j$,

$$\sum_{i=1}^{j} r_i = \sum_{i=1}^{j-1} r_i + r_j \le \sum_{i=1}^{j-1} s_i + r_j = \sum_{i=1}^{j-1} s'_i + r_j \le \sum_{i=1}^{j-1} s'_i + s'_j = \sum_{i=1}^{j} s'_i.$$

The inequality then holds for $j < l < k$ by the fact that $r_i = s_i$ on this interval. Thus $\mathbf{r} \preceq \mathbf{s}' = \mathbf{s}T_1$.

This process can be continued until some $\mathbf{r} = \mathbf{s}^{(m)} = \mathbf{s}T_1 \cdots T_m$. Letting $D = T_1 \cdots T_m$ finishes the proof of necessity.

Proof of Sufficiency. Suppose that there is a doubly stochastic matrix $D = (d_{ij})$ such that $\mathbf{r} = \mathbf{s}D$. Then $r_1 + \cdots + r_n = \mathbf{r}\mathbf{e}^T = \mathbf{s}D\mathbf{e}^T = \mathbf{s}\mathbf{e}^T = s_1 + \cdots + s_n$, where $\mathbf{e} = (1, 1, \ldots, 1)$ (see page 78). Also, $r_j = \sum_{i=1}^n s_i d_{ij}$.

To show that $\sum_{i=1}^k r_i \le \sum_{i=1}^k s_i$ for $1 \le k < n$, we calculate:

$$\begin{aligned}\sum_{j=1}^{k} r_j &= \sum_{j=1}^{k}\sum_{i=1}^{n} s_i d_{ij} = \sum_{i=1}^{k-1}\sum_{j=1}^{k} s_i d_{ij} + \sum_{i=k}^{n}\sum_{j=1}^{k} s_i d_{ij} \\ &\le \sum_{i=1}^{k-1}\sum_{j=1}^{k} s_i d_{ij} + s_k \sum_{j=1}^{k}\sum_{i=k}^{n} d_{ij}\end{aligned}$$

because $s_k \ge s_i$ for $i > k$,

$$\begin{aligned}&= \sum_{i=1}^{k-1}\sum_{j=1}^{k} s_i d_{ij} + s_k \sum_{j=1}^{k}(1 - \sum_{i=1}^{k-1} d_{ij}) \\ &= \sum_{i=1}^{k-1}\sum_{j=1}^{k} s_i d_{ij} + k s_k - s_k \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} d_{ij} \\ &= \sum_{i=1}^{k-1}(s_i - s_k)\sum_{j=1}^{k} d_{ij} + k s_k \le \sum_{i=1}^{k-1}(s_i - s_k) + k s_k = \sum_{i=1}^{k} s_i.\end{aligned}$$

With this, $\mathbf{r} \preceq \mathbf{s}$, completing the proof. □

23. P. 83, line 2: “matrices” should be “entries”.

24. **Pp. 82–83, Theorem 2.6.4: The proof sketch is confusing. It needs to be substantially rewritten.
25. *P. 93, Exercise 2.6.17: The inequality should be $\geq$.
26. P. 94, conjugate: It is important that $l \geq \max c_i$ $(= c_1)$. Also, if $l > \max c_i$, the effect is to add a string of 0's to the end of the conjugate sequence. You are always allowed to add 0's at the end if you need a larger l. (Thanks to Thomas Galvin and Jake Zukaitis.)

    *Also, in the definition of c_k^*, the range of k is $1 \leq k \leq l$. (Thanks to Alireza Salahshoori.)
27. P. 99, Exercise 2.7.1: Recall that the conjugate must be long enough, as in the note about p. 94.
28. P. 100, Exercise 2.7.3: There seem to be four errors. (Thanks to Jake Zukaitis.)

    The length of r^* must satisfy $n \geq \max r_i$ (as was stated in the definition of a conjugate on p. 94), or else the row sums cannot be attained.

    *The vector s ends with s_n, not s_m.

    At the end of the same line, "a" should be "any".

    **Most seriously, the stated problem is false. Ryser's paper proves only the "generalization".
29. *P. 110: In the displayed formula at the bottom, 0 should be $\delta(x, y)$.

    Also, since y is closed, $\bar{y} = y$ and the bar should be ignored.
30. *P. 112, last line of Thm 3.1.7: "closure operator" should be "coclosure operator".
31. P. 113, line 7: We are now assuming A is a lower crosscut because it is the A of the definition on page 112. (Thanks to Thomas Galvin.)
32. *P. 114, line 5 and the summation in Theorem 3.1.10: A should be C. (Thanks to Thomas Galvin.)
33. P. 115, Philip Hall's theorem: If we write $\mu(x, y) = c_0 - c_1 + c_2 - \cdots$ we cover $x \leq y$, not only $x = y$.
34. *P. 116, Eq. (Mf2): z should be b and (a, b) should be (a, y).

    **(Note by Tom Zaslavsky) The book's suggestion for a proof makes no sense to me. There is a simple direct proof. Use the definition of $\mu(a, b)$ twice: in $[a, b]_P$ it gives

$$0 = \sum_{z \in [a,b]_P} \mu_P(a, z) = \sum_{\substack{z \in [a,b]_P \\ f(z) < f(b)}} \mu_P(a, z) + \sum_{\substack{z \in [a,b]_P \\ f(z) = f(b)}} \mu_P(a, z)$$

    and in $[a, b]_f$ it gives

$$0 = \sum_{z \in [a,b]_f} \mu_f(a, z) = \sum_{\substack{z \in [a,b]_f \\ z < b}} \mu_f(a, z) + \mu_f(a, b).$$

Now observe that any $x \in [a,b]_P$ such that $f(x) < f(b)$ is in the interval $[a,b)_f$ and vice versa, so that if $z \in [a,b]_P$ satisfies $f(z) < f(b)$, then the intervals $[a,z]_P$ and $[a,z]_f$ are identical. Thus, $\mu_P(a,z) = \mu_f(a,z)$. It follows that

$$\mu_f(a,b) = -\sum_{\substack{z\in[a,b]_f \\ z<b}} \mu_f(a,z) = -\sum_{\substack{z\in[a,b]_P \\ f(z)<f(b)}} \mu_P(a,z) = \sum_{\substack{z\in[a,b]_P \\ f(z)=f(b)}} \mu_P(a,z).$$

35. *P. 118, Ex. 3.1.1: In PM3, "image of f" should be "the image of σ". Also, $\mathbb{A}$ must have characteristic 0 for equivalence to hold.
36. P. 121, Ex. 3.1.8: There may be some confusion about the definition. The zeta function $Z(P;n)$ = the number of multichains of *size* n in P.

    **In (a): The correct formula is $Z(P;n) = \sum_{i=2}^{d+2} b_i \binom{n-2}{i-2}$, where b_i is the number of *chains* of *length* $i-2$ in P. (Note that $d+2$ can be replaced by ∞.)

    In (b): Prove that $\zeta^n(\hat{0},\hat{1})$ = the number of multichains of *length* n from $\hat{0}$ to $\hat{1}$ in P (hint). Then prove that $Z(P;n) = \zeta^n(\hat{0},\hat{1})$ for $n \in \mathbb{Z}$, starting with $n > 0$ (hint).

    *In (c): $2n-1$ should be $2n$. (Thanks to Jake Zukaitis.)

    Also in (c), the zeta and order polynomials should have semicolons, not commas.

    Also in (c), the zeta polynomial should be $Z(\text{Interval}(P);n)$ with Z, not ζ. (Thanks to Roberto Gonzalez.)
37. *P. 123, Exercise 3.1.10b: This is not true for general closure operators. One needs to assume the property mentioned in the hint, which is not true in general. (Steven Collazos pointed this out.)
38. *P. 123, Exercise 3.1.10c: The Frattini subgroup is the intersection of all the maximal *proper* subgroups of G.
39. *P. 126, Exercise 3.1.15: The displayed union has a mistake. The union should be over $C \in \mathcal{C} : C \subseteq A$ and $C \subseteq B$.

    The word "interval" has two different meanings here. First meaning: an integer interval ($[a, a+1, \ldots, b]$ in $\mathbb{Z}$). Second meaning: a lattice interval $[A,B]$ in $L(\mathcal{C})$. Be careful to distinguish.

    In (b), by "is an interval lattice" they meant "is isomorphic to an interval lattice".
40. *P. 127, Cor. 3.2.2: "a chain of *size* $m+1$".
41. *P. 131, Pretzel's proof: The book appears to be defining a "saturated" chain but this definition is wrong. If we change the name to "orange" (a randomly chosen word), the proof should be valid.
42. P. 132, Exercise 3.2.1: Remember that the length of a sequence is the number of elements.

43. *P. 132, Exercise 3.2.2: "maximum length" should be "maximum size". (Thanks to Jake Zukaitis.)
44. P. 136: The other "Sperner's theorem" is generally known as Sperner's lemma.
45. P. 137: "Mešalkin" is Russian and is normally spelled Meshalkin in English.
46. P. 137: The quotient $\mathcal{C}/\mathcal{A}_m$ should be $|\mathcal{C}|/|\mathcal{A}_m|$.
47. **P. 137, end of Sperner's proof: "the possibility of constructing an *equally large* antichain by taking a mixture ...". (We already ruled out a larger antichain.) It is not clear how (NM) gives an easy proof of this.
48. P. 139: A symmetric chain decomposition of a (ranked) poset implies a Sperner-type theorem. In particular, Lemma 3.3.3 gives a fifth proof of Sperner's theorem.
49. **P. 145, Ex. 3.3.2: The setup for Meshalkin's theorem is wrongly stated.

    An *ordered weak partition* of $[n]$ is an s-tuple $A = (A_1, A_2, \dots, A_s)$ of pairwise disjoint subsets of $[n]$ whose union is $[n]$, not excluding the empty set (hence called "weak"). (Any number of the subsets may be empty.)

    We have a set $\mathcal{O}$ of ordered weak partitions of $[n]$. Let $\mathcal{O} = \{A^1, A^2, \dots, A^p\}$ be a family of ordered weak partitions of $[n]$; Meshalkin's condition is that, for each $j < p$, the sets A^i_j for $1 \le i \le p$ form an antichain. (We don't care about the p-th sets in each ordered weak partition.) Meshalkin's formula for the maximum size of $\mathcal{O}$ is stated correctly in the book.
50. P. 145, Ex. 3.3.6: "matchin." should be "matchin'".
51. *P. 146, Ex. 3.3.7, HR2: "max" should be "min". (Thanks to Thomas Galvin.)
52. *P. 155, after proof of Dubreil–Jacotin Thm.: "A lattice L is linear if there exists an <u>in</u>jective lattice homomorphism..."
53. P. 157, line −15: "Whether this equality holds for $n \ge 4$ is unknown." In fact, this is known and it is false. (Note by Kung–Yan)
54. P. 157, line −4: We should have noted that Haiman's proof theory does not imply that the universal Horn theory of linear lattices is decidable. It is not. See, for example, G. Hutchinson, Recursively unsolvable word problems of modular lattices and diagram-chasing, *J. Algebra* 26 (1973) 385–399; L. Lipshitz, The undecidability of the word problems for projective geometries and modular lattices, *Trans. Amer. Math. Soc.* 193 (1974) 171–180. (Note by Kung–Yan)

55. P. 162, second paragraph and Lemma 3.5.2: The Jordan–Dedekind chain condition is unnecessary. Since we're restricting to finite lattices, Jordan–Dedekind is equivalent to ranked.
56. *P. 162, proof of 3.5.2: The statement that $x_1 \vee y_1$ covers x is true if $x_1 \neq y_1$. If $x_1 = y_1$, the conclusion follows from the shorter interval $[x_1, y]$.

    In line -2, $z_{k-1} < z_k$.
57. *P. 167, Theorem 3.5.9: J_k and M_k are reversed.
58. **P. 180: The proof of Theorem 4.1.1 has a major but fixable flaw—a "gaffe" as Rota would say. To fix this, replace the second sentence in the proof by

    > The number of ways to choose a partition $B_1, B_2, \ldots, B_c$ of $\{1, 2, \ldots, n\}$ such that $|B_i| = n_i$ is
    >
    > $$\frac{n!}{a_1!a_2!\cdots a_n!n_1!n_2!\cdots n_c!}$$
    >
    > where a_i is the number of parts n_j such that $n_j = i$. On the block B_i, there are $|C_{n_i}|$ ways of putting an atom. Hence, the total number of molecules on $\{1, 2, \ldots, n\}$ with c components is
    >
    > $$\sum n_1, n_2, \ldots, n_c \frac{n!}{a_1!a_2!\cdots a_n!n_1!n_2!\cdots n_c!}|C_{n_1}||C_{n_2}|\cdots|C_{n_c}|.$$
    >
    > The product formula now follows from the multinomial theorem.

    (Note by Kung–Yan)
59. *P. 181: In the formula for $f(\mathcal{C}; t)$, the summation should begin at $n = 1$. (Thanks to Jake Zukaitis.)
60. *P. 181 (bottom) and p.182 (top): The $x_{(n)}$ at the bottom of p. 181 should be $x^{(n)}$ (rising factorial), and rising factorial should replace falling factorial at the top of p. 182 in the generating function for the unsigned Stirling numbers of the first kind. The formula given is the correct formula for the (*signed*) Stirling numbers of the first kind.
61. P. 182, Ex. 4.1.1: In line 2, A should be S.
62. P. 186: "rigorous proof of Proposition 4.1.1" means Proposition 4.2.1.
63. **P. 187, proof of Theorem 4.2.2: (Note by Laura Anderson.) There's something here that is copied from Rota's original paper, but I think it's wrong there too. (Actually, Tom Zaslavsky pointed it out.) The function M is defined on $\mathbb{Q}[\beta]$, but then the proof applies M to things that aren't in $\mathbb{Q}[\beta]$ – specifically, $M(e^{\beta t})$. The fix I came up

with was to define

$$\begin{aligned} \tilde{M} : \mathbb{Q}[\beta][[t]] &\rightarrow \mathbb{Q}[[t]] \\ \tilde{M}(\sum_{i=0}^{\infty} p_i(\beta)t^i) &= \sum_{i=0}^{\infty} M(p_i(\beta))t^i. \end{aligned}$$

It's then easy to check that for any$P(\beta, t)$,

$$\frac{d}{dt}\tilde{M}(P(\beta,t)) = \tilde{M}(\frac{d}{dt}P(\beta,t))$$

$$p(t)\tilde{M}(q(\beta,t)) = \tilde{M}(p(t)q(\beta,t)).$$

This makes the rest of the argument work.

64. *P. 188: The statement of Dobinski's formula is incorrect. The numerator $(j+1)^n$ should be j^n. The expansion given immediately afterwards has the same problem.
65. P. 187: The D in Boole's formula is the differentiation operator.
66. *P. 190, before Theorem 4.3.3: In the definition of a basic sequence for the delta operator Q, the correct relation is that $Qp_n(x) = np_{n-1}(x)$, not $p_{n-1}(x)$.
67. P. 352: The Ex. 3.5.2(a) hint is for Ex. 3.5.3(a).