Solving linear recurrence relations

I will first describe a method of solving a homogeneous linear recurrence relation with constant coefficients, by giving a closed form for the sequence in terms of what I call **exponomial functions**. I will then describe a method of solving an inhomogeneous linear recurrence relation with constant coefficients in which the right hand side of the relation is an exponomial function.

An exponomial function of n is a sum of finitely many terms, each of which is equal to an **expo**nential function of n times a poly**nomial** function of n. For example, $f(n) = 2^n + n^2 + n2^n + 3$ is an exponomial function, since it can be written as $(2)^n \cdot (1) + (1)^n \cdot (n^2) + (2)^n \cdot (n) + (1)^n \cdot 3$. We express it in **standard form** by gathering together those terms $(r)^n \cdot p(n)$ with the same value of r; for example, the standard form of $2^n + n^2 + n2^n + 3$ is $(2)^n \cdot (n+1) + (1)^n \cdot (n^2+3)$. The values of r that occur in an exponomial function $f(n)$ are called **bases**, and each base r has **multiplicity** equal to 1 more than the degree of the polynomial $p(\cdot)$, where $r^n \cdot p(n)$ is the term in the standard form of $f(n)$ involving r. Thus, the bases in $(2)^n \cdot (n+1) + (1)^n \cdot (n^2+3)$ are 2 (with multiplicity $\deg(n+1)+1 = 1+1 = 2$) and 1 (with multiplicity $\deg(n^2+3)+1 = 2+1 = 3$). We don't allow expressions of the form $(0)^n \cdot p(n)$, so 0 can never be a characteristic value of an exponomial function. Note also that terms involving the base 1 don't use contain the expression $(1)^n$ explicitly, since 1^n equals 1 for all n.

Theorem 1: If $f(n)$ satisfies a homogeneous linear recurrence relation with constant coefficients, whose characteristic polynomial has roots $r_1, \dots, r_k$ with respective multiplicities $m_1, \dots, m_k$, then $f(n)$ can be expressed as an exponomial function of n with characteristic values $r_1, \dots, r_k$ with respective multiplicities $m_1, \dots, m_k$ (or smaller).

Example: Suppose $f(n)$ satisfies the recurrence relation

$$f(n) = 3f(n-1) - 4f(n-3),$$

whose characteristic polynomial $x^3 - 3x^2 + 4$ factors as $(x-2)(x-2)(x+1)$; then $f(n)$ can be written as $(2)^n \cdot (An+B) + (-1)^n \cdot (C)$ for suitable constants A, B, and C. (Note that if, say, A turns out to be 0, then the characteristic value 2 has multiplicity 1 rather than 2; this is why I wrote "(or smaller)" in the statement of Theorem 1. For that matter, if A and B both turn out to be zero, then the characteristic value 2 has multiplicity 0, which is smaller still.)

For purposes of the next result, we have to change perspective a bit. If some real number r is NOT a root of some polynomial, call it a "root of multiplicity 0". Likewise, if r is NOT a base of some exponomial function, call it a "base of multiplicity 0". Thus, 3 is a characteristic value of multiplicity 0 in the polynomial $(x-1)(x-2)$, while 2 is a base of multiplicity 0 in the exponomial $3^n + n$.

Theorem 2: If $f(n)$ satisfies an inhomogeneous linear recurrence relation with constant coefficients, whose left hand side has characteristic polynomial $p(x)$ and whose right hand side is the exponomial function $g(n)$, then $f(n)$ can be expressed as an exponomial function of n in which, for every non-zero real number r, the multiplicity of r as a base for $f(n)$ is equal to the multiplicity of r as a root of $p(x)$ PLUS the multiplicity of r as a base for $g(n)$.

(Note that if r is neither a root of the polynomial $p(x)$ nor a base for the inhomogeneous term $g(n)$, then the sum of the two multiplicities is zero, so there is no r^n term in the formula for $f(n)$.)

Example: Suppose $f(n)$ satisfies the recurrence relation

$$f(n) - 3f(n-1) + 4f(n-3) = 2^n + 1$$

(note that the LHS is the same as in the previous example). The roots of the characteristic polynomial of the LHS are -1 and 2, with respective multiplicities 1 and 2, and the bases for the RHS are 2 and 1, each with multiplicity 1. So $f(n)$ can be expressed as an exponomial function with bases -1, 2, and 1. The multiplicity of -1 (in our exponomial formula for $f(n)$) is $1+0=1$, the multiplicity of 2 is $2+1=3$, and the multiplicity of 1 is $0+1=1$. That is, we must have $f(n) = (-1)^n \cdot (A) + (2)^n \cdot (Bn^2 + Cn + D) + (1)^n \cdot (E)$, for suitable coefficients A, B, C, D, E.

If we want to find a specific solution that satisfies specified initial conditions, we need a way to solve for the undetermined coefficients. One way to do this is to set up a system of simultaneous linear equations. For instance, to solve for A through E in the preceding example, we could plug $n = 1$ through $n = 5$ into the equation for $f(n)$, obtaining five linear relations for the five unknowns (making use of the values of $f(1), \ldots, f(5)$ specified in the initial conditions).

What if we want to follow the standard approach of writing some desired solution to the inhomogeneous recurrence relation as the sum of two functions of n, one of which is a particular solution to the inhomogeneous recurrence relation and the other of which is the general solution to the homogeneous recurrence relation? Here too the "exponomial function" point of view can help us.

Theorem 3: If we are content with finding a closed-form formula for *some* sequence $f(n)$ satisfying an inhomogeneous linear recurrence relation with constant coefficients, whose right hand side is the exponomial function $g(n)$, and we don't have to satisfy any initial conditions, then we can find such a formula in which the only bases r that occur are bases of $g(n)$; however, the multiplicity of such an r in $f(n)$ could still be as great as the sum of the multiplicity of r as a base for $g(n)$ and the multiplicity of r as a root of the characteristic polynomial.

Example: If we want any old solution to the inhomogeneous recurrence relation

$$f(n) - 3f(n-1) + 2f(n-2) = 2^n,$$

we can find one of the form $2^n \cdot (An + B)$. Note that 2 is a root of the characteristic equation with multiplicity 1, and is a base of $g(n) = 2^n$ with multiplicity 1, so the multiplicity of 1 as a characteristic value of $f(n)$ could be as high as $1 + 1 = 2$. Indeed, if we plug $f(n) = 2^n \cdot (An + B)$ into the equation $f(n) - 3f(n-1) + 2f(n-2) = 2^n$ and divide the result by 2^{n-2}, we obtain $4 \cdot (An + B) - 6 \cdot (A(n-1) + B) + 2 \cdot (A(n-2) + B) = 4$, which simplifies to $2A = 4$; so $A = 2$. B drops out entirely, so we may as well set it equal to zero, obtaining $f(n) = 2^n \cdot (2n) = 2^{n+1} \cdot n$.

Final note: Most semesters, there's a student whose approach to solving something like $f(n) - 3f(n-1) + 4f(n-3) = 2$ is to mis-derive the characteristic equation as $x^2 - 3x + 4 = 2$. Don't be that student!