Chapter 6

Sequences and Series of Functions

6.1 Discussion: Branching Processes

The fact that polynomial functions are so ubiquitous in both pure and applied analysis can be attributed to any number of reasons. They are continuous, infinitely differentiable, and defined on all of $\mathbb{R}$. They are easy to evaluate and easy to manipulate, both from the points of view of algebra (adding, multiplying, factoring) and calculus (integrating, differentiating). It should be no surprise, then, that even in the earliest stages of the development of calculus, mathematicians experimented with the idea of extending the notion of polynomials to functions that are essentially polynomials of infinite degree. Such objects are called power series, and are formally denoted by

$$
\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^4 + \cdots.
$$

The basic dilemma from the point of view of analysis is deciphering when the desirable qualities of the limiting functions (the polynomials in this case) are passed on to the limit (the power series). To put the discussion in a more concrete context, let’s look at a particular problem from the theory of probability.

In 1873, Francis Galton asked the London Mathematical Society to consider the problem of the survival of surnames (which at that time were passed to succeeding generations exclusively by adult male children). “Assume,” Galton said, “that the law of population is such that, in each generation, $p_0$ percent of the adult males have no male children who reach adult life; $p_1$ have one such male child; $p_2$ percent have two; and so on... Find [the probability that] the surname will become extinct after $r$ generations.” We should add (or make explicit) the assumption that the lives of each offspring, and the descendants thereof, proceed independently of the fortunes of the rest of the family.
Galton asks for the probability of extinction after \( r \) generations, which we will call \( d_r \). If we begin with one parent, then \( d_1 = p_0 \). If \( p_0 = 0 \), then \( d_r \) will clearly equal 0 for all generations \( r \). To keep the problem interesting, we will insist that from here on \( p_0 > 0 \). Now, \( d_2, \) whatever it equals, will certainly satisfy \( d_1 \leq d_2 \) because if the population is extinct after one generation it will remain so after two. By this reasoning, we have a monotone sequence

\[
d_1 \leq d_2 \leq d_3 \leq d_4 \cdots,
\]

which, because we are dealing with probabilities, is bounded above by 1. By the Monotone Convergence Theorem, the sequence converges, and we can let

\[
d = \lim_{r \to \infty} d_r
\]

be the probability that the surname eventually goes extinct at any time in the future. Knowing it exists, our task is to find \( d \).

The truly clever step in the solution is to define the function

\[
G(x) = p_0 + p_1 x + p_2 x^2 + p_3 x^3 + \cdots.
\]

In the case of producing male offspring, it seems safe to assume that this sum terminates after five or six terms, because nature would have it that \( p_n = 0 \) for all values of \( n \) beyond this point. However, if we were studying neutrons in a nuclear reactor, or heterozygotes carrying a mutant gene (as is often the case with the theory of branching processes), then the notion of an infinite sum becomes a more attractive model. The point is this: We will proceed with reckless abandon and treat the function \( G(x) \) as though it were a familiar polynomial of finite degree. At the end of the computations, however, we will have to again become well-trained analysts and be prepared to justify the manipulations we have made under the hypothesis that \( G(x) \) represents an infinite sum for each value of \( x \).

The critical observation is that

\[
G(d_r) = d_{r+1}.
\]

The way to understand this is to view the expression

\[
G(d_r) = p_0 + p_1 d_r + p_2 d_r^2 + p_3 d_r^3 + \cdots
\]

as a sum of the probabilities for different distinct ways extinction could occur in \( r + 1 \) generations based on what happens after the first generational step. Specifically, \( p_0 \) is the probability that the initial parent has no offspring and so still has none after \( r + 1 \) generations. The term \( p_1 d_r \) is the probability that the initial parent has one male child times the probability that this child’s own lineage dies out after \( r \) generations. Thus, the probability \( p_1 d_r \) is another contribution toward the probability of extinction in \( r + 1 \) steps. The third term represents the probability that the initial parent has two children and that the surnames of each of these two children die out within \( r \) generations. Continuing in this way, we see that every possible scenario for extinction in \( r + 1 \) steps is