Oscillation theorems for $\psi(x)$ and $\theta(x)$ and $\pi(x)$. Let's start by recalling Landau's theorem (Theorem 1.7): if $\alpha(s) = \sum_{n=1}^\infty a_n n^{-s}$ is a Dirichlet series with $a_n \ge 0$ for all n, then the abscissa of convergence σ_c is a singularity of $\alpha(s)$ (that is, $\alpha(s)$ cannot be analytically continued to $s = \sigma_c$). On the other hand, equation (1.10) tells us that for $\sigma > \sigma_c$,

$$\alpha(s) = s \int_1^\infty A(x) x^{-s-1}\, dx; \tag{$*$}$$

thus Landau's theorem can be interpreted as a statement about functions defined by "Dirichlet integrals" of the form $(*)$ where $A(x)$ is an increasing step function. It turns out that Landau's theorem can be generalized to a broader class of Dirichlet integrals

$$F(s) = \int_1^\infty A(x) x^{-s}\, dx,$$

where the function $A(x)$ doesn't have to be a summatory function of a sequence a_n but can simply be any nonnegative function (satisfying some technical conditions of local boundedness and Riemann integrability).

Read Theorem 15.1. Notice that the book merely says that it can be proved in the same way as Theorem 1.7; working though that adapted proof is a good exercise for you.

Landau's theorem and this generalization seem like ways to start with information about a real-valued function $A(x)$ and deduce information about an analytic function $F(s)$. However, there is surprising power in taking the contrapositive of Landau's theorem: if the function $F(s)$ does not have a singularity at a particular point, then the corresponding real-valued function $A(x)$ cannot be eventually positive. For example, we can take

$$A(x) = x^{\Theta-\varepsilon} - \big(\psi(x) - x\big)$$

(where Θ as usual is the supremum of the real parts of zeros of $\zeta(s)$), for which it's simple to find

$$F(s) = \int_1^\infty A(x) x^{-s}\, dx = \frac{1}{s-(\Theta-\varepsilon)} - \frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1}$$

in some half-plane; and Theorem 15.1 implies that if $A(x)$ were eventually nonnegative, then this identity of analytic functions would in fact hold for $\sigma > \Theta - \varepsilon$, which contradicts the definition of Θ as there is a $\zeta(s)$ in the denominator of the right-hand side. Therefore there must be arbitrarily large values of x such that $A(x) < 0$; and this is equivalent to the statement $\psi(x) - x = \Omega_+(x^{\Theta-\varepsilon})$. A similar method (changing $\psi(x) - x$ to $x - \psi(x)$) gives the corresponding Ω_- result, and further modifications address our other prime-counting functions $\theta(x)$ and $\pi(x)$.

Read Theorem 15.2 and Corollary 15.3 and their proofs. The proof of Theorem 15.2 is a proof by contradiction inside a contrapositive, so make sure you appreciate the nuances of the logical structure. Notice that Theorem 15.2 is a slightly weaker statement than problem 4(a) on Homework #4; on the other hand, the proof of Theorem 15.2 doesn't require a zero-free region or explicit formula, and this Landau's-theorem method has quite a broad range of applications, some of which appear in the remainder of Section 15.1.

Oscillation theorems of Littlewood and Skewes. You might have noticed that Corollary 15.3 gives that $\psi(x) - x = \Omega_\pm(x^{1/2})$ (in other words, there are reasonably large oscillations, both positive and negative, for the error term for $\psi(x)$) but only that $\theta(x) - x = \Omega_-(x^{1/2})$ and $\pi(x) - \mathrm{li}(x) = \Omega_-(x^{1/2}/\log x)$ (that is, there are large negatives values of these error terms, but possibly not large positive values). And indeed, emperically we see that $\pi(x) - \mathrm{li}(x) < 0$ as far up as we care to calculate primes. Why the asymmetry?

The explicit formula (see for example the Fourier-like expansion from problem 6 of Homework #4) shows that $\psi(x)-x$ is, assuming the Riemann hypothesis, equal to $x^{1/2}$ times an infinite sum where each summand is just as likely to be positive as negative. Roughly speaking, the explicit formula tells us that $(\psi(x) - x)/\sqrt{x}$ is "centred at 0". However, we know that

$$\theta(x) - x = \big(\psi(x) - \theta(x^{1/2}) + O(x^{1/3})\big) - x = \big(\psi(x) - x\big) - x^{1/2} + o(x^{1/2}) :$$

the first equality can be proved in the same way as problem 2(a) of Group Work #2 except singling out the contribution of the squares of primes, while the second equality is the prime number theorem applied to $\theta(x^{1/2})$. In other words, $(\theta(x) - x)/\sqrt{x}$ is "centred at -1", and so it is harder to prove that $\theta(x) - x$ is often positive than it is to prove that it is often negative. Similar comments apply to $\pi(x) - \mathrm{li}(x)$, which is most directly accessed by partial summation from $\theta(x) - x$. (Note, however, that while partial summation preserves O-results, it does not necessarily preserve Ω-results, basically because the quantifier in the definitions of O and Ω are different in a way that interacts with the integrals in such proofs. Indeed, this is the reason why the proof of Theorem 15.2 essentially treated the $\pi(x)$ case from scratch.) Indeed, for a while at the beginning of the 20th century, it was a widely believed conjecture that $\pi(x) < \mathrm{li}(x)$ for all x.

Enter Littlewood, who in 1914 announced a *disproof* of this conjecture by showing that

$$\psi(x) - x = \Omega_\pm(x^{1/2} \log\log\log x) \quad \text{and} \quad \theta(x) - x = \Omega_\pm(x^{1/2} \log\log\log x)$$

and

$$\pi(x) - \mathrm{li}(x) = \Omega_\pm\left(\frac{x^{1/2} \log\log\log x}{\log x}\right).$$

(Note that these statements are in fact weaker than Theorem 15.2 if the Riemann hypothesis is false, so one might as well assume that the Riemann hypothesis is true if trying to prove these statements.) Littlewood started with a Fourier-like expansion similar to the one from problem 6 of Homework #4, except not for $\psi(x) - x$ itself but an average of $\psi(x) - x$ over a relatively small interval containing x. It's reasonable to try to show that an infinite series is quite positive by showing that lots of the terms of the series are simultaneously positive; Littleword used (homogeneous) Diophantine approximation to do just that.

Start from the middle of page 476 (after the proof of Theorem 15.8) and read through the proof of Theorem 15.11.

An interesting aspect of these theorems of Littlewood is that they are "ineffective", in that it is not possible even in theory to extract a bound for the smallest x such that $\pi(x) > \mathrm{li}(x)$. (It would be annoying, but possible in theory, to extract such a bound if the Riemann hypothesis is true; but in the case where the Riemann hypothesis is false, the bound depends upon a specific hypothetical nontrivial zero of $\zeta(s)$ off the critical line.) Decades later, Skewes overcame this obstacle to prove

there exists a real number $x < 10^{10^{10^{10^3}}}$ for which $\pi(x) > \mathrm{li}(x)$.

Indeed, this bound ("Skewes's number") was infamous at the time for being the largest number ever to appear non-gratuitously in a serious mathematical result.

Random model for these error terms. In problem 6 of Homework #4 you showed that

$$\psi(x) = x - \sqrt{x} \sum_{\substack{\rho\in\mathbb{C} \\ 0<\gamma<x^{1/2+\varepsilon}}} 2\Re\biggl(\frac{x^{\rho-1/2}}{\rho}\biggr) + O_\varepsilon(x^{1/2-\varepsilon/2});$$

assuming the Riemann hypothesis, this statement can be written as

$$\frac{\psi(x)-x}{\sqrt{x}} = \sum_{\substack{0<\gamma<x^{1/2+\varepsilon} \\ \zeta(1/2+i\gamma)=0}} 2\Re\biggl(\frac{x^{i\gamma}}{1/2+i\gamma}\biggr) + o(1).$$

Each quantity $x^{i\gamma} = e^{i\gamma\log x}$ runs around the unit circle, at a constant rate if we consider $y = \log x$ to be the fundamental variable; indeed, in the limit, for any given $\gamma > 0$, the distribution of $e^{i\gamma y}$ is uniform on the unit circle in the complex plane. This suggests replacing $x^{i\gamma} = e^{i\gamma\log x}$ with a random variable X_γ that is exactly uniform on this unit circle; in other words, we might expect that the complicated random variable

$$\sum_{\substack{0<\gamma<x^{1/2+\varepsilon} \\ \zeta(1/2+i\gamma)=0}} 2\Re\biggl(\frac{X_\gamma}{1/2+i\gamma}\biggr) \quad \text{is a good model for} \quad \frac{\psi(x)-x}{\sqrt{x}}.$$

We haven't said anything about the relationships between the (identically distributed) random variables X_γ as γ ranges over positive ordinates of zeros of $\zeta(s)$. Of course we would like this collection of random variables to be independent; a little thought reveals that the condition on the set of ordinates γ that would lead to such independence is that the γ should be linearly independent over the rational numbers. This linear independence of ordinates is conjectured to be true on general principles, although there is a scarcity of solid evidence supporting it.

Optional reading: you have all the background you need to read and appreciate the article "Prime number races" that I wrote with Andrew Granville; it can be found on my web site under Research.

Using this sort of approach, together with data on a large number of actual zeros of $\zeta(s)$, we can see that while there are arbitrarily large values of x for which $\pi(x) > \mathrm{li}(x)$, such values are quite rare: Rubinstein and Sarnak conditionally proved that

the relative density of those x for which $\mathrm{li}(x) > \pi(x)$ is approximately 99.999973%.

All the material in today's lecture is part of an area of research called "comparative prime number theory", which focuses on detailed examination of error terms for prime-counting functions, particularly counting primes in arithmetic progressions. Anyone who is interested in learning more about comparative prime number theory should keep their eyes open for the topics course I plan to teach in Fall 2020. I also have a large collaborative research project—annotating a bibliography of papers in comparative prime number theory—that students can join for the coming summer if they are interested.