we first present a simplified version of the $QR$ algorithm which we call basic $QR$ algorithm. We prove a convergence theorem for the basic $QR$ algorithm, under the rather restrictive hypothesis that the input matrix $A$ is diagonalizable and that its eigenvalues are nonzero and have distinct moduli. The proof shows that the part of $A_k$ strictly below the diagonal converges to zero and that the diagonal entries of $A_k$ converge to the eigenvalues of $A$.

Since the convergence of the $QR$ method depends crucially only on the fact that the part of $A_k$ below the diagonal goes to zero, it would be highly desirable if we could replace $A$ by a similar matrix $U^*AU$ easily computable from $A$ and having lots of zero strictly below the diagonal. It turns out that there is a way to construct a matrix $H = U^*AU$ which is almost triangular, except that it may have an extra nonzero diagonal below the main diagonal. Such matrices called, *Hessenberg matrices*, are discussed in Section 18.2. An $n \times n$ diagonalizable Hessenberg matrix $H$ having the property that $h_{i+1i} \neq 0$ for $i = 1, \ldots, n-1$ (such a matrix is called *unreduced*) has the nice property that its eigenvalues are all distinct. Since every Hessenberg matrix is a block diagonal matrix of unreduced Hessenberg blocks, *it suffices to compute the eigenvalues of unreduced Hessenberg matrices.* There is a special case of particular interest: symmetric (or Hermitian) positive definite tridiagonal matrices. Such matrices must have real positive distinct eigenvalues, so the $QR$ algorithm converges to a diagonal matrix.

In Section 18.3, we consider techniques for making the basic $QR$ method practical and more efficient. The first step is to convert the original input matrix $A$ to a similar matrix $H$ in Hessenberg form, and to apply the $QR$ algorithm to $H$ (actually, to the unreduced blocks of $H$). The second and crucial ingredient to speed up convergence is to add shifts.

A shift is the following step: pick some $\sigma_k$, hopefully close to some eigenvalue of $A$ (in general, $\lambda_n$), $QR$-factor $A_k - \sigma_k I$ as

$$A_k - \sigma_k I = Q_k R_k,$$

and then form

$$A_{k+1} = R_k Q_k + \sigma_k I.$$

It is easy to see that we still have $A_{k+1} = Q_k^* A_k Q_k$. A judicious choice of $\sigma_k$ can speed up convergence considerably. If $H$ is real and has pairs of complex conjugate eigenvalues, we can perform a double shift, and it can be arranged that we work in real arithmetic.

The last step for improving efficiency is to compute $A_{k+1} = Q_k^* A_k Q_k$ without even performing a $QR$-factorization of $A_k - \sigma_k I$. This can be done when $A_k$ is unreduced Hessenberg. Such a method is called $QR$ iteration with implicit shifts. There is also a version of $QR$ iteration with implicit double shifts.

If the dimension of the matrix $A$ is very large, we can find approximations of some of the eigenvalues of $A$ by using a truncated version of the reduction to Hessenberg form due to Arnoldi in general and to Lanczos in the symmetric (or Hermitian) tridiagonal case. *Arnoldi iteration* is discussed in Section 18.4. If $A$ is an $m \times m$ matrix, for $n \ll m$ ($n$ much smaller