Chapter 14

The Boundary Integral Equations for Stokes Flow

14.1 The Setting for Computational Microhydrodynamics

Despite the simplicity of the Stokes equations, the extraction of quantitative information for complex geometries, in which the resolution of the particle and boundary shape is an important part of the mathematical model, requires substantial computational resources. Situations where large-scale computations may be avoided were described in Parts II and III: If the particle geometry is simple, *e.g.*, the boundaries fit a coordinate system in which the Laplacian separates, closed-form analytic solutions can be obtained; on the other hand, even for particles of fairly complex shape, the disturbance far away from the particles can be described by a multipole expansion, although the multipole moments in this expansion must be derived from a numerical solution. With regard to the numerical solution, if each particle fits a convenient coordinate system, the multipole boundary collocation scheme provides an accurate solution, with only moderate demands on the computational facility. The basis functions can be represented as distributions of Stokes singularities (inside the particle), which, in some sense, capture the shape of the particle, so that the number of functions needed in the truncated expansion is greatly reduced. (The ultimate limit of this argument would be the use of the exact solution as a basis element, with an adjustable scalar coefficient determined by setting the boundary condition at one point.)

With more complex geometries, such as those shown in Figure 14.1, the only way to capture the details of the particle shape is to bring the Stokes singularities to the particle surface. The resulting class of numerical techniques are known as the *boundary integral methods* (or some variations of this name with the word *equation* added). Their distinct advantage over spatial methods such as finite elements or finite difference is *reduction of dimensionality.* That is, instead of a three-dimensional PDE, we solve a two-dimensional (boundary) integral equation, in which the unknowns are densities of the Stokes singularities

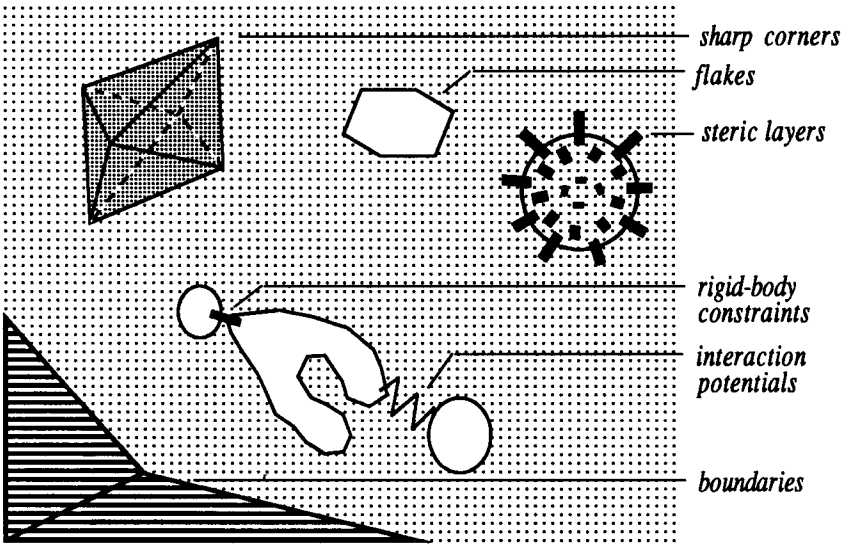

Figure 14.1: Complex microstructures of microhydrodynamics.

distributed over the boundary of the fluid domain, and even infinite fluid domains without a container boundary cause no problems. Furthermore, although the mathematical theory places a smoothness requirement on the surface, this restriction can be loosened in the actual numerical procedure.

By going to two-dimensional formulation, we reduce both the amount of work for discretization and the number of variables needed for getting some required numerical resolution. Especially in time-dependent simulations where the mutual configuration of individual rigid boundaries is changing, no new discretization may be needed after each time step. This is in contrast to spatial methods such as finite elements. Even so, when boundary element methods are applied to problems of interest, we tax even the "state of the art" computers to the fullest. These are, after all, three-dimensional fluid flows with complex geometries and time evolution. Consequently, we would gladly exploit orders of magnitude speedups in computing power.

At first glance, it may appear that our emphasis of boundary methods goes against current trends in computational technology. For example, the emergence of fast, inexpensive processors suggest that the most interesting innovations will occur in parallel computational architectures. Indeed, fundamental physical barriers such as the speed of light and the size of an atom place a limitation on the speed and integration of a single processor. It is unlikely that in the future the speed attainable by using a single processor would be orders of magnitude faster than today. The obvious answer to this dilemma is parallel processing.

It is easy to visualize how some computational problems in areas such as molecular dynamics and operations research can exploit parallel processing. Computational tasks are readily divided into independent pieces. Similarly, in computational fluid dynamics one intuitively senses that the equations from spatial methods, such as the finite element method and the finite difference method, as relations connecting only near neighbors, are also attractive candidates for parallel computers. Boundary methods, on the other hand, lead to dense systems, for the following reason. The (flow) fields emanating from Stokes singularities on one boundary element propagate in all directions with a fairly persistent algebraic decay, so that in some sense, all elements interact with all other elements. The mapping of such dense systems to parallel architectures is a nontrivial problem.

With this environment in mind, the emphasis of Part IV will be the construction of a new boundary integral method, which we call the Completed Double Layer Boundary Integral Equation Method (CDL-BIEM). The name is derived from the fact that Stokes double layer densities are employed, but since the double layers alone do not form a complete basis, a completion procedure is necessary. Using techniques from linear operator theory, we shall prove that the solution of the general mobility problem of N particles in either a bounded or unbounded domain can be cast in terms of a fixed point problem in which the linear operator (or some finite power of it) is a contraction mapping, since the spectral radius is less that one. Furthermore, the three steps in the computational procedure (pre-processing or creation of the system and contraction mapping, iterative solution of the discretized problem, and post-processing to extract the physical variables of interest) can be formulated as parallel algorithms. We will draw upon the developments of Parts II and III to show that for many-body suspension problems, the interactions between double layer elements on different particles are very small, in the sense that the eigenvalue(s) determining the spectral radius of the operator are perturbed only slightly by hydrodynamic interactions. Thus we raise the possibility of attacking large-scale suspension simulations by a network of parallel processor, with each processor powerful enough to handle all boundary elements for a given particle. *In essence we recover the nearest neighbor property of the spatial methods — only now the fundamental entity is not an element or node, but the collection of elements for the particle.* For many-body problems, the resulting iterative strategy will involve asynchronous iterations accelerating the convergence relative to synchronous iterations similarly as the Gauss–Seidel iterations do in comparison with Gauss–Jacobi iterations.

This chapter provides a sufficient background for comprehension of the material covered, to the extent that the reader should be able to apply the theory and algorithms presented for obtaining numerical results. We go over Fredholm integral equations of the first and second kind, and some smoothness requirements for the kernels, which for our application imply smoothness requirements for the surface. The kernels encountered are *weakly singular*, so that an extension of the Fredholm theory, namely, the *Fredholm–Riesz–Schauder* theory

for compact operators, is necessary. The fact that the integral operators are compact will be exploited in two ways. First, it implies that CDL-BIEM is a well-posed problem. Second, the spectrum of a compact operator is discrete, which in our case will be used to show that, even before discretization, the integral equations of CDL-BIEM can be solved by Neumann series expansion, in practice replaced by direct (Picard) iterations (also called successive substitution) corresponding to truncating the Neumann series (see the exercises).

14.2 Integral Operators and Integral Equations

14.2.1 Motivation

Although integral equations seem to be unduly slighted in courses of advanced calculus, they have turned out to be very useful both for theoretical and numerical purposes. In this chapter we shall give a brief introduction (with references for more detailed study) to both the theory and practical numerical solution of integral equations. The theory, roughly speaking, is such that anyone familiar with ordinary matrices should feel comfortable with it.

We summarize here only the main properties of those types of IEs that we shall encounter. These properties may seem abstract and "too mathematical" for engineers, but this should be only a temporary impression. The existence and uniqueness of a solution are significant for practical computation also. It would be a waste of time to look for a solution that does not exist, and multiple solutions can make a linear system hard to treat numerically.

14.2.2 First- and Second-Kind Equations

In an integral equation the unknown function appears within an integral. We shall only be concerned with linear integral equations. The Fredholm equations of the first and second kind are of the form

$$\lambda x(s) - \int_I K(s,t)x(t)\,dt = g(s)\ , \quad \text{for } s \in I, \tag{14.1}$$

where λ is a given constant, g and K are given functions (usually, the *kernel* K is assumed to be continuous), x is to be solved for, and I is a finite interval for the one-dimensional case. If $\lambda = 0$, the equation is of the first kind, otherwise it is of the second kind. In the latter case the equation can be divided by λ, so that we see the identity operator acting on x; this identity operator is the main reason for distinguishing between the first- and second-kind equations. Usually the functions g and x are assumed to be continuous (see Whittaker and Watson [69]). We shall need an extension of this theory to *weakly singular* kernels, and this is provided by the *Fredholm–Riesz–Schauder* theory. It covers the abstract second-kind equation

$$\lambda x - \mathcal{K}x = g\ , \quad (\lambda \neq 0)\ , \tag{14.2}$$

where $\mathcal{K}$ is a *compact* (also called *completely continuous*) linear operator. Henceforth it will always be tacitly assumed that any *operator* under discussion is linear.

We shall limit our attention to the case of operators acting on a Hilbert space, which will be that of square summable functions L_2 in the application of the theory. More general results for Banach spaces, *etc.*, can be found in the references.

14.2.3 Weakly Singular Kernels

Let Ω be a bounded measurable m-dimensional set in Euclidean space, like a region in m-dimensional space or a surface in $m+1$-dimensional space. A *weakly singular kernel* is of the form

$$K(s,t) = \frac{A(s,t)}{r^\alpha}, \text{ for } s,t \in \Omega, \tag{14.3}$$

where r is the Euclidean distance $|s-t|$, A is continuous, and $0 \leq \alpha < m$. Such kernels are also called *polar* or *potential* kernels. Mikhlin [51, 52] shows that the corresponding integral operator, mapping a function x to

$$\int_\Omega K(s,t)x(t)\,dt, \tag{14.4}$$

is compact on $L_2(\Omega)$ and also on $C(\Omega)$ (the space of continuous functions equipped with sup-norm). In the integral above, dt signifies integration with respect to the measure on Ω, and in the case of a surface we have a surface integral. Note that by setting $\alpha = 0$, we recover the usual continuous kernels, and so the theory of weakly singular kernels extends the ordinary theory. Also Petrovskii [55] covers weakly singular kernels in a similar manner as Mikhlin. A more advanced and more demanding presentation is given by Kantorovich and Akilov [36].

The weakly singular operator has the following properties not implied by its compactness. It maps bounded functions to continuous, and for continuous data g, any solution $x \in L_2(\Omega)$ of the second-kind equation is also continuous. The latter also implies that all eigenfunctions in $L_2(\Omega)$ are continuous.

14.2.4 Ill- and Well-Posedness

The definition of this classification for linear problems dates back to Hadamard (1932). He stated that a linear problem $Ax = b$ is well-posed if it is uniquely solvable for any data b, and the solution x depends continuously on the data. More concisely, the operator A must have a well-defined bounded inverse. Linear problems that are not well-posed are called ill-posed.

Originally it was envisioned that any physical problem should have a well-posed formulation, but physically reasonable ill-posed problems have been encountered quite frequently. From a philosophical point of view, well-posed problems are such that the mathematical model has no serious difficulties in dealing

with them. Ill-posed problems, on the other hand, are somewhat "incompatible" with the model used, the solution being excessively sensitive to any variations in the data. As in numerical computations, we may be relying on approximations or measured quantities, and in any case there will be some round-off errors in floating-point numbers; such excessive sensitivity can disturb the calculations badly. Typically the computed result must be considered only intermediate in character — only some projection of it is reliable. Some post-processing that uses just this projection may remove the effects of the "indeterminate part," and give quite accurate results.

There are two standard ways for dealing with ill-posedness. The first is to impose some constraints on the solution x and look for the (best approximate) solution satisfying the constraints. This corresponds to using some extra information not previously utilized to get rid of the indeterminate part of the solution. The second method is to construct a sequence of well-posed problems that approximate, in some sense, better and better the original ill-posed problem. Typically this is done with so-called Phillips–Tikhonov regularization, by adding a "regularizing functional" and thus changing the equations.

In summary, we would like our problems to be well-posed, or at least be aware of the ill-posedness of a formulation, since some difficulties can be expected. An ill-posed problem may suggest that we are computing results that are only intermediate in character, or that the model used is inappropriate.

14.2.5 Compact Operators

In addition to the references in the previous sections, the reader may wish to consult the books by Dieudonné [14], Friedman [21], Ramkrishna and Amundson [60], and Rudin [62]. Both Friedman and Dieudonné give self-contained expositions starting from the fundamentals. In addition, the book by Ramkrishna and Amundson contains applications of linear operator theory to transport phenomena and reactor design problems.

Several properties of compact operators on a Hilbert space H shall be listed here. The set of compact operators is a subspace of the bounded operators, closed with respect to the operator norm. Multiplication with a bounded operator from left or right preserves compactness, as does the involutive operation of taking the adjoint. The adjoint of a bounded operator $\mathcal{K}$ is the unique operator $\mathcal{K}^*$ satisfying the following "flipping rule" within the inner product:

$$\langle \mathcal{K}x, y\rangle = \langle x, \mathcal{K}^* y\rangle, \text{ for all } x, y \in H. \tag{14.5}$$

Inner products, if complex, will be conjugate linear with respect to their *second* argument. In linear algebra the adjoint of a matrix is the complex conjugate of the transpose matrix, and this should not be confused with the *classical adjoint* whose elements are cofactors of the elements in the original matrix.

The identity operator I is compact, iff (if and only if) H is finite-dimensional. From this it follows that a compact $\mathcal{K}$ will not have a bounded inverse if $\dim(H) = \infty$, for assuming the converse, $I = \mathcal{K}\mathcal{K}^{-1}$ would be compact as

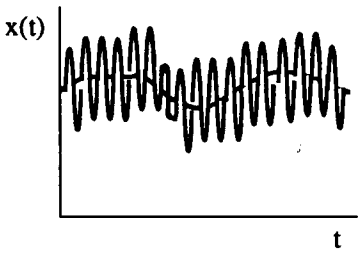

Figure 14.2: Fredholm integral equations of the first kind are ill-posed.

the product of a compact and a bounded operator. This shows why usually little space is devoted to discussing the first-kind equation $\mathcal{K}x = g$. Even if this relation can be inverted to $x = \mathcal{K}^{-1}g$, the unboundedness of $\mathcal{K}^{-1}$ implies that a small change in the data g can result in an arbitrarily large change in the solution x. As this type of excessive sensitivity of the solution to variations in the data is commonly called ill-posedness, *first-kind equations are ill-posed.* The term *compact* describes well the "smallness" of these operators, corresponding to the "largeness" of the inverse operators. In practice the ill-posedness of the first-kind equations manifests itself when a numerical solution is attempted using the discretized linear system. The linear system will be ill-conditioned, increasingly so with finer discretizations.

The situation is illustrated in Figure 14.2. Assume that the smooth curve is the desired solution of the integral equation. Now superimpose on this curve a high-frequency oscillation. The integral operator smooths out such oscillations so both functions get mapped to similar outputs. The inverse problem, of determining the input given the output, must be ill-posed. When very fine discretizations are employed, these high-frequency oscillations can exist in the numerical solution and the corresponding symptom in the discretized problem, *ill-conditioning*, is observed. Later on (Chapter 17) we shall derive explicit examples of this effect as applied to boundary integral equations. On the surface of a sphere, oscillatory inputs of the form $P_n(\cos\theta)$ will map to small outputs that vanish for large n as $O(n^{-1})$.

For second-kind equations the situation is dramatically different. Here the smallness of compact operators is an asset, implying that they only somewhat perturb the identity operator in the equation. As a result, the second-kind equations share several properties with ordinary square matrices, which is the essence of the Fredholm–Riesz–Schauder theory. Often some of the results listed below are termed *Fredholm theorems.* The *Fredholm alternative* states that either the second-kind equation is uniquely solvable for any data g, or else the corresponding homogeneous equation (found by setting $g = 0$) has nontrivial solutions. In the latter case these nontrivial solutions are called *eigenvectors*

or *eigenfunctions* corresponding to the *eigenvalue* λ, and if g happens to allow a solution x, other solutions are found by adding an eigenvector to x; thus solutions are not unique. This alternative is sometimes stated as "1-1 implies onto," meaning that uniqueness of solutions implies solvability for any data. Moreover, if the second-kind operator is invertible, or equivalently λ is not an eigenvalue, the inverse is bounded. Therefore *second-kind equations are well-posed*, although special treatment is needed when an eigenvalue is encountered. Later on we shall introduce a *bordering* scheme to handle these situations.

Suppose now that $\lambda \neq 0$ is an eigenvalue; the statements that follow are not necessarily valid for the case of a first-kind operator with $\lambda = 0$. The corresponding eigenvectors form the *null space* $N(\lambda I - \mathcal{K})$ and are also called *null vectors* (of the second-kind operator). The second-kind equation is solvable iff g can be represented as $\lambda x - \mathcal{K}x$, in other words, iff g is in the *range* $R(\lambda I - \mathcal{K})$. The null space is always finite dimensional, so that there is a finite basis of eigenvectors for each eigenvalue. Just like matrices may be nondiagonalizable, having nontrivial Jordan blocks and *principal vectors* other than eigenvectors (in the terminology used by Householder [31]), a compact operator also can have these; also the term *principal function* may be used. In case a matrix is diagonalizable, its Jordan blocks being trivial, the Jordan canonical form actually gives the diagonal matrix achievable with similarity transformations. Principal vectors are such that repeated application of the second-kind operator on them will at some step give a zero result. All the nonzero vectors that come up during this process form a *Jordan chain*, in which the last vector is an eigenvector. Thus to each eigenvector there corresponds a Jordan chain and Jordan block (in the matrix representation with the principal vectors chosen as the basis), and if the Jordan chain consists of the eigenvector alone, the Jordan block is trivial. The Jordan blocks of a compact operator are all finite, and for a fixed eigenvalue there is only a finite number of these, as the number of linearly independent eigenvectors is finite.

The null spaces of the adjoint operators $\lambda I - \mathcal{K}$ and $\lambda^* I - \mathcal{K}^*$ (λ^* is the complex conjugate of λ) are of equal dimension (the same goes for the Jordan blocks in general), and further they contain all the information about the ranges that are closed subspaces:

$$R(\lambda I - \mathcal{K}) = N(\lambda^* I - \mathcal{K}^*)^{\perp}. \tag{14.6}$$

The second-kind operator and its adjoint can be switched here, since applying the operation of taking the adjoint twice to a bounded operator recovers the original (this, together with conjugate linearity and preserving the norm, makes the adjoint operation involutive), just as with matrices. From the equation above it follows that $g \in R(\lambda I - \mathcal{K})$ and a solution x exists, iff g is orthogonal to the (finite basis of the) null functions of the adjoint $\lambda^* I - \mathcal{K}^*$. This also shows that the *codimension* of the range, meaning the dimension of its orthogonal complement, equals the dimension of the null space, for these second-kind operators. We could also say that the range has a deficiency equal to the dimension of the null space.

The observations above are crucial for the applications to be presented. Firstly we know that even if the second-kind operator is not invertible due to us having encountered an eigenvalue, there will only be a finite number of independent eigenvectors that we need to consider, and the number of these will indicate the deficiency in the range. Secondly, solving the homogeneous adjoint problem will completely describe "the shape of this deficiency," *i.e.*, what is missing from the range.

The term *Jordan block* is used in none of the references listed, but we use it here in order to show the similarity with ordinary matrix algebra and to help interpret the results. The Jordan canonical form of a matrix is explained in Gantmacher [23]. Familiarity with the Jordan form is also useful later on, when it is used to help visualize Wielandt's deflation.

As a compact operator is bounded, so is the set of its eigenvalues. A much stronger statement can be made, namely, there is at most a denumerable set of eigenvalues, and the only possible accumulation point of them is zero. Therefore each nonzero eigenvalue is isolated: It has a neighborhood within which no other eigenvalues are found. Knowledge of this will be useful on considering the spectral radius, which determines whether an iterative solution of a linear system of equations is possible. The set of eigenvalues of a compact operator will also be called the *spectrum* of the operator; for operators in general the spectrum is so defined that it may contain numbers other than eigenvalues. The result above, characterizing the spectrum of a compact operator, can be understood in the following way. Suppose we had a nonzero eigenvalue repeated an infinite number of times. This would mean that in an infinite-dimensional subspace the operator acts as a multiple of the identity, which is not compact. Thus the operator itself would not be compact. The situation remains essentially the same if we have an infinite number of eigenvalues in any neighborhood of one accumulation point; this also is not allowed for a compact operator. Point zero is an exception, as zero multiple of the identity is just zero.

14.2.6 Numerical Solution Methods

In an IE, basically three discretizing approximations are needed to get an algorithmic method ready for computer implementation. Approximate the solution function with another that depends on a finite set of parameters, replace the infinite process of integration with a finite computable approximation, and finally the continuum of equations that resulted has to be dealt with. In practice the first two discretizations here are coupled and, to some extent, performed "in one step."

In general for approximating a function with another of fixed form, say, a polynomial, different criteria may be used to determine the free parameters, such as the coefficients of a polynomial. In *collocation*, a functional equation is enforced exactly at a (finite) set of (collocation) points. In *osculation*, not only are the values of the function enforced, but also some derivatives at some points. The *least squares method* minimizes a number describing the error somehow on

an average, and the *min-max* method strives to minimize the largest error. In accordance with this terminology, the following procedure shall be called *the quadrature collocation method.* Consider a linear IE of the first or second kind, and approximate the integration (infinite process) by a finite quadrature (*e.g.*, Gaussian), as in

$$\int_0^1 u(t)\,dt \approx \sum_{i=1}^{N} \omega_i u(t_i)\ , \tag{14.7}$$

where ω_i and t_i are the quadrature weights and points, respectively. Now the discretization of the first-kind equation is

$$\sum_{i=1}^{N} \omega_i K(s,t_i) X(t_i) = g(s)\ , \tag{14.8}$$

where X is an approximation for x. As only a finite set of values of X is involved, the quadrature has discretized also the unknown function simultaneously with the integration. There are only N values $X(t_i)$ to be determined, so 14.8 cannot be valid for all s (in general). Instead one can enforce the equation exactly at N collocation points:

$$\sum_{i=1}^{N} \omega_i K(s_j,t_i) X(t_i) = g(s_j)\ \ \text{, for } j = 1,\ldots,N\ . \tag{14.9}$$

This is a typical linear system. With similar treatment of the second-kind equation it would be necessary to choose the collocation points to be the quadrature points, $s_j = t_j$, since X comes up both alone with arguments s_j and inside the quadrature with t_j. The choice of collocation points to be the quadrature points is also beneficial if the kernel K is symmetric, $K(t,s) = K(s,t)$, since then the discrete equation system (for the unknowns $\omega_i X(t_i)$) will also be symmetric. Utilization of this symmetry will clearly bring savings in numerical computations, both in storage space and through the use of more efficient routines for this special case.

Consider now an expansion of the solution in terms of some trial (basis) functions, which might be, *e.g.*, the eigensolutions of a related problem, or piecewise polynomials. As numerical integration is fairly expensive computationally, a set of basis functions that allows doing the integrations with the kernel analytically would be preferable. If such basis functions are not known, numerical integrations are necessary, but the way in which these are carried out has no particular effect on the next step here. The discretization of the function leaves us (after the integrations) still with the discretization of the "continuum" of equations, as in going from 14.8 to 14.9 above. This may be done by any of the methods mentioned for approximating a function with another, depending on a finite set of parameters. The "quadrature step" 14.8 above (which discretized both the function and the integral) has been replaced with

$$x(t) \approx \sum_{i=1}^{N} c_i \phi_i(t) = X(t)\ . \tag{14.10}$$

Substitution into the Fredholm equation of the first kind gives

$$\int_0^1 K(s,t) \sum_{i=1}^{N} c_i \phi_i(t)\, dt = g(s)\ . \tag{14.11}$$

Choosing collocation points s_j would not preserve the symmetry of the kernel K in the discretized equations, so the *Galerkin method* shall be introduced here. This means that s is integrated out of Equation 14.11 with weight functions, and, moreover, often the weight functions are chosen to be those in the basis function set, *i.e.*, the ϕ_is. Then

$$\int_0^1 \int_0^1 K(s,t) \sum_{i=1}^{N} c_i \phi_i(t) \phi_j(s)\, dt\, ds = \int_0^1 g(s) \phi_j(s)\, ds\ , \tag{14.12}$$

which, in case K is symmetric, is a symmetric linear system for the unknowns c_i. For unsymmetric kernels the advantage of Galerkin method is lost as far as the symmetry of the resulting system is regarded; it will not create symmetry that was not there to start with. Furthermore, this method requires an extra integration compared with collocation. Thus it may be anticipated that collocation methods will in general be superior to Galerkin methods computationally, unless the choice of basis functions has made the required integrations numerically less difficult.

When one has no prior knowledge of the shape of the solution x, it seems natural to use piecewise polynomial approximation in the basis function approach. Then each basis function will get nonzero values only within some subinterval of the whole range of the integral equation, and this division into subintervals may be compared with the division into elements for a FEM solution; the basis functions can also naturally be called shape functions. The simplest type of approximation would be to use piecewise constant shape functions, and collocation at the centers of the subintervals (or elements). This approach is also the easiest to implement numerically, but by using so crude an approximation to the solution some accuracy will be lost. For a more detailed discussion on this subject, we refer the reader to the papers by Cruse [10, 11].

In fact, a piecewise constant approximation will be used later on with the three-dimensional codes. The reason in that case is that there are three significant advantages in using planar elements for approximating the boundary surfaces. As the approximation of the boundaries is "rough," restricting the accuracy of the solutions, it would be pointless to use very accurate representation of the functions solved for. In addition, matrix generation with such elements is more readily mapped to parallel architectures. Lastly, post-processing for physical quantities such as surface tractions is more straightforward.

14.2.7 Applications to Boundary Integral Equations

In general there will be more than one integration variable in the surface integrals that come up in hydrodynamic problems, although, as will be shown

later, some problems with axisymmetric boundaries can be reduced to systems of integral equations with just one integration variable.

The quadrature collocation scheme introduced for one-variable problems could still be applied in the multi-dimensional case. For surface integrals there is not much choice in choosing the quadrature rule; typically for two-dimensional integrations one uses a product rule found by applying one-dimensional quadratures to the iterated integral. Furthermore, it is not very wise to try using, say, very high order Gaussian quadratures. Instead one should split the interval of integration to subintervals and apply a "decent" order quadrature on each of these separately. However, with the weakly singular kernels here, a simple quadrature collocation scheme would run into difficulties, since at any fixed collocation point there is a singularity that prevents the direct evaluation of the kernel. This situation can be remedied by using the *singularity subtraction* technique.

The basis function approach has so far been overwhelmingly more popular in applications. For surface integral equations, this leads to what is now called the *boundary element method*, a term initially coined by C. Brebbia in applications to solid structures. The procedure is to divide the surface(s) of interest to small elements, over each of which some low-order approximation for the solution can be applied with good conscience; the elements will also give a natural approach to calculating the surface integrals by the use of (perhaps modified) product rule integration on each of these. A modification of the quadrature will be necessary in particular for the element in which the singularity corresponding to the collocation point is located. Naturally it is not necessary to use piecewise polynomials for the basis functions, but use of some "global" approximating functions can only be justified when the shape of the results to be expected is known *a priori*, and still generating a quadrature rule for the surface would have to be accomplished probably by subdivision into elements. Thus the boundary element formulation seems to be the most generally applicable approach in sight.

In one-dimensional cases the subdivision of the integration interval is not a big problem, whereas for surfaces the division into elements would require some algorithmic approach. This is probably the reason why, in spite of its generality, the boundary element method, at least in fluid mechanics, seems to have been applied only to combinations of geometrically simple topologies where some natural coordinate systems allow for easy grid generation. Actually triangulations of surfaces can be refined fairly easily by using a subroutine that generates a surface point approximately "midway between" two given ones; with such a routine available it is straightforward to transform one triangle into four smaller ones fitting the surface more closely by forming new edges that connect the approximate midpoints of the original edges and "twisting" the original edges. Starting from an initial coarse triangulation, it is then possible to generate triangulations with desired resolution of the surface.

14.3 Notation and Definitions

In the following the Stokes equations in fluid domains that are regions in the three-dimensional Euclidean space shall be considered. A *region* is, as usual, an open connected set, meaning that each point of a region has some neighborhood wholly contained within the region, and any two points of the region can be joined by a curve that also lies within the region. The boundary of the fluid domain is denoted by S, and it is further required that each connected component of S be a closed bounded *Lyapunov-smooth surface* (see Günter [27]); this ensures, for example, that the (Gauss) divergence theorem can be applied to bounded fluid domains.

In practice surfaces without edges or corners usually are Lyapunov-smooth. Without going into great detail (see the reference above), it is just required that the surface normal changes its direction in a certain uniformly continuous fashion as we move along the surface, and further that "opposing parts" of the surface are uniformly separated from each other. Sometimes, we will omit the "-smooth" part of the terminology.

The fluid domain will also be called the *interior* and the complement of it and its boundary the *exterior*, although this conflicts with the mathematical term *interior* denoting the largest open set within *any* given set. Each single closed Lyapunov surface divides the space into a bounded and an unbounded region; these are called the *inside* and the *outside*, respectively. A boundary surface restricting the fluid domain to its outside (inside) will sometimes be called a *particle* (*container*). It is clear that for a connected fluid domain there can be at most one container boundary.

The unit normal on S pointing *into the interior fluid domain* will be denoted by $\hat{\boldsymbol{n}}$, and it has the Cartesian components n_i; optionally the surface point determining $\hat{\boldsymbol{n}}$ will be indicated as its argument. To keep the notation explicit, the symbols $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ shall be used for the position vectors of surface points on S, with corresponding components ξ_i and η_i. Vectors corresponding to points not on S will be denoted by $\boldsymbol{x}$, $\boldsymbol{y}$, *etc.* The Euclidean distance of two points, say, $\boldsymbol{x}$ and $\boldsymbol{y}$, is denoted by r_{xy} or $|\boldsymbol{x}-\boldsymbol{y}|$, and the corresponding vector is $\boldsymbol{r}_{xy} = \boldsymbol{y} - \boldsymbol{x}$ (read "vector from x to y"). The Euclidean norm of a single vector $\boldsymbol{x}$ will also be denoted by x.

The *natural inner product* is defined for square integrable vector surface fields on S as the over S integrated pointwise dot product of these fields. With respect to this inner product, the adjoint kernel to any given kernel $\boldsymbol{K}$ is $\boldsymbol{K}^*$ defined by

$$\boldsymbol{K}^*(\boldsymbol{\eta},\boldsymbol{\xi}) = \boldsymbol{K}^t(\boldsymbol{\xi},\boldsymbol{\eta}) \;, \tag{14.13}$$

so that transformation with $\boldsymbol{K}$ inside an inner product may be flipped to the other side of the inner product by changing the transformation to $\boldsymbol{K}^*$. This corresponds to the transpose of a matrix and its relation to the ordinary dot product (a real inner product) of finite vectors. For complex inner products one also needs to complex conjugate the transpose, to create the adjoint. As with matrices, the flipping rule is verified by exchanging the order of finite

summations (Exercise 16.1); with the natural inner product here, one needs to exchange the order of the implied integrations. From now on an inner product will be understood to be the natural inner product; deviations from this practice will be clear from the context.

In Chapter 2 it was shown that the rate of work done by a moving solid particle equals the energy dissipation in the fluid domain. More generally the equations used there show that the dissipation equals the rate of work done on the boundaries of the fluid domain, the latter being defined as the pointwise inner product of the velocity with the surface tractions integrated over the boundary surfaces with sign change:

$$\int_V \boldsymbol{\sigma} : \boldsymbol{e} dV = \oint_S \boldsymbol{v} \cdot (-\boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}}) dS \ . \tag{14.14}$$

Recall that $(\boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}})dS$ is the force by the fluid on the surface element, while here we use the reaction force on the fluid; this causes the sign change. This result, equating the dissipation with a natural inner product of the velocity and traction fields, will be referred to as the *energy relation*, and will be frequently used in the following sections.

14.4 The Boundary Integral Equation in the Primary Variables

Let us now consider the homogeneous Stokes solutions for problems without a body-force field. In Chapter 2 the following integral representations for the Stokes field pair $(\boldsymbol{u}, p)$ in the interior $Q_{(i)}$ were obtained:

$$\begin{aligned} \boldsymbol{u}(\boldsymbol{y}) &= -\frac{1}{8\pi\mu} \oint_S \mathcal{G}(\boldsymbol{y}-\boldsymbol{\xi}) \cdot (\boldsymbol{\sigma}(\boldsymbol{\xi};\boldsymbol{u}) \cdot \hat{\boldsymbol{n}}(\boldsymbol{\xi}))\ dS(\boldsymbol{\xi}) \\ &\quad - \oint_S \{\hat{\boldsymbol{n}}(\boldsymbol{\xi}) \cdot \boldsymbol{\Sigma}(\boldsymbol{y},\boldsymbol{\xi})\} \cdot \boldsymbol{u}(\boldsymbol{\xi})\ dS(\boldsymbol{\xi}) \ , \text{ for } \boldsymbol{y} \in Q_{(i)} \qquad (14.15) \\ p(\boldsymbol{y}) &= -\frac{1}{8\pi\mu} \oint_S \mathcal{P}(\boldsymbol{y}-\boldsymbol{\xi}) \cdot (\boldsymbol{\sigma}(\boldsymbol{\xi};\boldsymbol{u}) \cdot \hat{\boldsymbol{n}}(\boldsymbol{\xi}))\ dS(\boldsymbol{\xi}) \\ &\quad - \frac{1}{4\pi} \oint_S \{\hat{\boldsymbol{n}}(\boldsymbol{\xi}) \cdot \nabla_y \mathcal{P}(\boldsymbol{y},\boldsymbol{\xi})\} \cdot \boldsymbol{u}(\boldsymbol{\xi})\ dS(\boldsymbol{\xi}) \ , \text{ for } \boldsymbol{y} \in Q_{(i)} \ , \qquad (14.16) \end{aligned}$$

and it was noted that these representations give zeroes in the exterior $Q_{(e)}$. We mentioned the jump discontinuity across the boundary, and in the velocity representation 14.15 this must come from the last term. Indeed, $\mathcal{G}(\boldsymbol{y}-\boldsymbol{\xi}) = O(1/r_{y\xi})$ implies that in local polar coordinates about the surface point $\boldsymbol{\eta}$ that $\boldsymbol{y}$ passes through, the first term on the RHS has no essential singularity; so the jump must be from the last term. The analogy with electrostatics has been mentioned before. The electrostatic potential of a surface charge distribution is continuous across the boundary, while that of a surface dipole distribution (the electric double layer) suffers a jump (the idea behind a dry cell battery).

Assume now that the last term exists as an improper integral when we substitute $\boldsymbol{y} \leftarrow \boldsymbol{\eta}$. This is true for sufficiently smooth surfaces such as Lyapunov

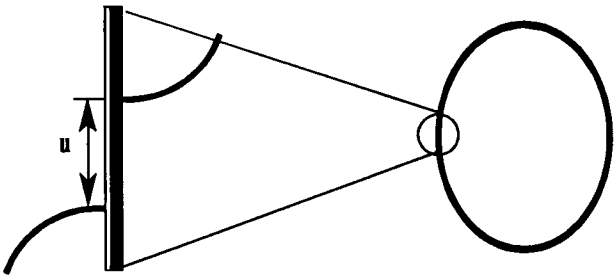

Figure 14.3: Jump in v for a double layer distribution.

surfaces, even though $\boldsymbol{\Sigma}(\boldsymbol{\eta},\boldsymbol{\xi}) = \mathrm{O}(r_{\xi\eta}^{-2})$, because $\hat{\boldsymbol{n}}(\boldsymbol{\xi}) \cdot \boldsymbol{r}_{\xi\eta}^{0} = \mathrm{O}(r_{\xi\eta}^{\alpha})$ as $\boldsymbol{\xi} \to \boldsymbol{\eta}$, with $\alpha \in (0,1]$. Assuming $\boldsymbol{u}$ to be smooth, we see that the integrand is $\mathrm{O}(r_{\xi\eta}^{\alpha-2})$, which is the reason why the kernel in this term is weakly singular and integrable in the local polar coordinates about the singular point.

As $\boldsymbol{y}$ approaches $\boldsymbol{\eta}$ from either side, the surface will locally look planar, and since the jump is due to the immediate neighborhood of $\boldsymbol{\eta}$, it can be claimed by symmetry that the jump is equal in magnitude whether $\boldsymbol{y}$ comes to sit on the surface from $Q_{(i)}$ or $Q_{(e)}$. Therefore, on substituting $\boldsymbol{y} \leftarrow \boldsymbol{\eta}$ in Equation 14.15, the RHS will attain the value symmetrically halfway, as shown in Figure 14.3. This leads to the equation

$$\begin{aligned}\frac{1}{2}\boldsymbol{u}(\boldsymbol{\eta}) \;&=\; -\frac{1}{8\pi\mu}\oint_S \mathcal{G}(\boldsymbol{\eta}-\boldsymbol{\xi}) \cdot (\boldsymbol{\sigma}(\boldsymbol{\xi};\boldsymbol{u}) \cdot \hat{\boldsymbol{n}}(\boldsymbol{\xi}))\, dS(\boldsymbol{\xi}) \\ &\quad - \oint_S \{\hat{\boldsymbol{n}}(\boldsymbol{\xi}) \cdot \boldsymbol{\Sigma}(\boldsymbol{\eta},\boldsymbol{\xi})\} \cdot \boldsymbol{u}(\boldsymbol{\xi}) dS(\boldsymbol{\xi}) \text{ , for } \boldsymbol{\eta} \in S \text{ .} \qquad (14.17)\end{aligned}$$

This is a *boundary integral equation* (BIE) for the surface tractions when the velocity field is known on the surface, or for the velocity field on the surface when the tractions there are known. The kernels, explicitly $\mathcal{G}(\boldsymbol{\eta}-\boldsymbol{\xi})$ and $\hat{\boldsymbol{n}}(\boldsymbol{\xi}) \cdot \boldsymbol{\Sigma}(\boldsymbol{\eta},\boldsymbol{\xi})$, are weakly singular and integrable. For the tractions the integral equation is an *equation of the first kind*, since these are only integrated with the kernel as a weight. For the velocity surface field we have an *equation of the second kind*, since in addition to being integrated with a kernel, $\boldsymbol{u}$ comes up also as such in the equation.

A rigorous derivation of the equation above, for Lyapunov surfaces and $\boldsymbol{u}$ continuous, can be given by considering about $\boldsymbol{\eta}$ a Lyapunov sphere, which in the limit is shrunk to a point. This type of technique is applied to problems of potential theory in Günter [27], and also Ladyzhenskaya [43, Chapter 3] employs it. On a spherical or hemispherical surface S about $\boldsymbol{\eta} = \mathbf{0}$ one can write

$$\begin{aligned}&\oint_S \hat{\boldsymbol{n}}(\boldsymbol{\xi}) \cdot \boldsymbol{\Sigma}(\boldsymbol{\eta},\boldsymbol{\xi})\, dS(\boldsymbol{\xi}) = \oint_S \frac{3}{4\pi}\hat{\boldsymbol{n}}(\boldsymbol{\xi})\boldsymbol{\xi} R^{-3}\, dS(\boldsymbol{\xi}) \\ &= R^{-3}\int_Q \frac{3}{4\pi}\nabla_{\boldsymbol{x}}\, dQ(\boldsymbol{x}) = R^{-3}\int_Q \frac{3}{4\pi}\boldsymbol{\delta}\, dQ(\boldsymbol{x}) = V/\left(\frac{4}{3}\pi R^3\right)\boldsymbol{\delta}\ ,\end{aligned}$$

where R is the radius and V is the volume of the region Q enclosed by this surface. This calculation tacitly utilized the fact that the planar part of the hemispherical surface does not contribute to the surface integrals. So the result is one (half) for a (hemi)sphere. This is the key result required in the limiting procedure explained, and further details shall not be given here.

An extension to infinite fluid domains is possible by assuming that we consider the *disturbance field* of a particle or particles that decays to zero at a sufficient rate as infinity is approached. From Lamb's general solution, valid for a fluid domain outside a sphere (see Happel and Brenner [29]) enclosing the particles, it is seen that a decaying velocity field will have at least $O(|\boldsymbol{x}|^{-1})$ as the decay rate. The corresponding stresses behave as the derivatives of the velocity field and will decay at least as $O(|\boldsymbol{x}|^{-2})$, except for an arbitrary constant base pressure, which is at first assumed so chosen that this decay rate is obeyed. Consider now another larger sphere enclosing the particles, and as the fluid domain use the volume enclosed in between. A simple order of magnitude argument shows that on the particle surfaces the contribution to the integrals from the larger enclosing sphere S_c will become negligible as we let it expand to infinity. Letting R be the radius of this sphere, observe that the surface S_c is $O(R^2)$, and use the estimates above together with similar estimates for $\mathcal{G}$ and $\boldsymbol{\Sigma}$ in 14.17 with $\boldsymbol{\eta}$ on the particles and $\boldsymbol{\xi}$ on S_c. In summary *if a homogeneous Stokes solution is decaying as infinity is approached, then its decay rate is sufficient for neglecting the "boundary at infinity" in the integral equation above.* Since in a bounded domain, like the inside of the larger sphere, we can arbitrarily set the base pressure without affecting the result from the integral involving the tractions in 14.15, we can in fact allow any base pressure with the disturbance fields and still use the above "enclosing sphere expanded to infinity argument" with 14.15 and 14.17.

Later on also, the energy relation 14.14 shall be used in unbounded domains. For this case the base pressure needs more careful consideration. If the pressure is chosen to become zero at infinity, a similar order of magnitude estimate as above shows that the "enclosing sphere expanded to infinity argument" allows us to again neglect the "surface at infinity." Now adding a base pressure to these tractions on the enclosing sphere will result in a term

$$\oint \boldsymbol{u} \cdot \hat{\boldsymbol{n}}\, dS \tag{14.18}$$

evaluated over this container surface, which will vanish only in a particular case. It vanishes iff the same integral evaluated over the totality of particle surfaces vanishes, due to the velocity field being mass-conserving in the fluid domain in between. In summary *the energy equation 14.14 can be applied to decaying homogeneous Stokes solutions in infinite fluid domains, neglecting the "surface at infinity," when either the pressure vanishes at infinity, or the total generation of fluid within the particle surfaces is zero.* The Lorentz reciprocal theorem can be extended to decaying Stokes fields in infinite domains with similar restrictions; now the nonvanishing of the pressure of one flow field at infinity is allowed if *the other* flow field has no net sources or sinks within the

particle surfaces.

We need not consider only disturbance fields in infinite fluid domains, but we can also assume that before any particles were inserted into the space there was some *ambient flow* satisfying the Stokes equations. The change needed to produce the flow field after inserting some particles will then serve as the decaying disturbance field to which the theory above is applicable.

Let $\boldsymbol{u}$ now be the disturbance field cancelling an ambient field $\boldsymbol{v}$ on the particle surfaces. Then $\boldsymbol{u}$ satisfies Equation 14.17 with the normal pointing into the fluid domain as usual, *i.e.*, outwards from the particles. By using the same equation for $\boldsymbol{v}$ *inside* the particles, it is found by observing the sign change for the normal and that $\boldsymbol{v}$ is the negative of $\boldsymbol{u}$ on the surface, that

$$\begin{aligned} -\frac{1}{2}\boldsymbol{u}(\boldsymbol{\eta}) &= \frac{1}{8\pi\mu}\oint_S \mathcal{G}(\boldsymbol{\eta}-\boldsymbol{\xi})\cdot(\boldsymbol{\sigma}(\boldsymbol{\xi};\boldsymbol{v})\cdot\hat{\boldsymbol{n}}(\boldsymbol{\xi}))\,dS(\boldsymbol{\xi}) \\ &\quad -\oint_S \{\hat{\boldsymbol{n}}(\boldsymbol{\xi})\cdot\boldsymbol{\Sigma}(\boldsymbol{\eta},\boldsymbol{\xi})\}\cdot\boldsymbol{u}(\boldsymbol{\xi})dS(\boldsymbol{\xi})\ , \text{ for } \boldsymbol{\eta}\in S\ . \end{aligned} \tag{14.19}$$

By subtracting this from 14.17 one obtains another boundary integral equation where the integration of $\boldsymbol{u}$ over the surface has dropped out.

$$\boldsymbol{u}(\boldsymbol{\eta}) = -\frac{1}{8\pi\mu}\oint_S \mathcal{G}(\boldsymbol{\eta}-\boldsymbol{\xi})\cdot(\{\boldsymbol{\sigma}(\boldsymbol{\xi};\boldsymbol{v})+\boldsymbol{\sigma}(\boldsymbol{\xi};\boldsymbol{u})\}\cdot\hat{\boldsymbol{n}}(\boldsymbol{\xi}))\,dS(\boldsymbol{\xi})\ . \tag{14.20}$$

This equation, which shall be called the *single layer equation* for reasons that become clear in Section 15.2, is useful when the ambient field is given in such a form that its stress field is readily calculable, and we are considering the disturbance created by immersing a finite number of particles into this given flow field. It is obvious that in coding a numerical procedure we benefit by having fewer surface integrals to consider, assuming that we want to use any of the formulations presented above. This equation also allows a simple physical interpretation. The tractions of the final flow field (sum of $\boldsymbol{u}$ and $\boldsymbol{v}$) are such that they prevent any penetration or slip at the particle boundaries, cancelling the ambient velocity there. Here the simple superposition of point forces on the surfaces is visible.

The integral equations shown so far each involve only physical variables of *primary* interest, namely, velocity and surface tractions, which explains the title of this section.

14.4.1 Physically Occurring Boundary Conditions

The two main categories of problems with solid particles in a fluid are mobility and resistance problems. In mobility problems we need to determine the translation and rotation velocities of the particles from the given forces and torques acting on them, while in the resistance problems the roles of the given and solved quantities are inverted. As the Stokes equations are linear, these problems are closely related, and the resistance and mobility matrices are simply each other's inverses. However, construction of such a matrix, say, the resistance matrix, requires solving many resistance problems. Therefore it is desired

to be able to solve mobility problems directly. Moreover, most of the physically occurring problems are of the mobility type, and resistance problems only come up somewhat artificially.

Why is it then that mobility problems seldom are solved directly? The data given in a mobility problem does not really provide us with an ordinary set of boundary conditions for solving partial differential equations. Rather, what we are given are quantities found from the tractions upon integration over the particle surfaces, and that the particles move as rigid bodies. As opposed to this, in resistance problems we know the disturbance velocity field on the particle surfaces, and ordinary solution methods for the Stokes equations can be applied.

Mathematically it is clear that a problem need not fit either of the categories above; we could start from constraints involving both surface velocities and the total forces and torques. Physically such problems are very rare.

Other problems outside this classification include free surface problems and multiphase flows with different fluids. In these cases the boundary shapes, together with the surface tensions, give boundary conditions for the tractions. As an example of such a problem, see the work of Rallison and Acrivos [59] and Pozrikidis [58].

14.4.2 Difficulties with the BIE in the Primary Variables

The boundary integral equation in the primary variables is of the first kind for the tractions, given velocity boundary conditions. This is something of a problem for two reasons.

Firstly there is no general theory for the existence or uniqueness of solutions to integral equations of the first kind. In fact, the BIE is valid only if a Stokes solution with *smooth* tractions exists for the given velocity BCs. Granted the existence of such a solution, multiple solutions, in fact, will occur. This is due to the indeterminacy of the base pressure in a Stokes solution, allowing the addition of a constant multiple of the surface normal to the tractions.

Secondly, first-kind equations are ill-posed. Thus *numerical* solution attempts for the tractions may encounter some problems. Recall that the ill-posedness means excessive sensitivity of the solved tractions for variations in the given velocity data on the boundary. Physically we can understand this, as the tractions should behave as the derivatives of the velocity field, and therefore we are in some sense doing numerical differentiation here. In the next section it is shown that the ill-posedness cannot be removed by restricting the solution space to continuous functions.

14.5 On Solving Problems with Velocity BCs

The boundary integral equation (BIE) 14.17 is of the first kind for the tractions. Assuming that a velocity surface field $\boldsymbol{u}$ has been given, such that allows a Stokes

solution in the whole fluid domain, one can try to solve for the tractions. Since all the integral operators in the BIE are weakly singular and therefore compact, their inverses are unbounded. So $\boldsymbol{u}$ is first affected by a second-kind operator, and the result is mapped to the tractions by an unbounded operator. From this arises the question, whether this unboundedness is an artifact of the formulation used, so that on restricting the surface fields $\boldsymbol{u}$ to "allowed fields" (in the sense of corresponding to a Stokes solution in the interior) one would get a bounded mapping to the tractions. Unfortunately this is not the case.

To give an example of this unboundedness, one can again use Lamb's general solution outside a sphere, which previously was utilized in inspecting the decay rates of disturbance fields. Recall from Chapter 3 (see also Happel and Brenner [29]) that the Stokes velocity field $\nabla\Phi_n$ has the traction field $2(n-1)\mu\nabla\Phi_n/r$ on the surface of a sphere, where r is the radius and Φ_n is a solid spherical harmonic of order n. For a decaying solution any negative integer values are allowed for n, so that these allowed velocity surface fields are mapped to multiples of themselves as tractions, the multiplier attaining arbitrarily large values (see also Exercise 17.3).

Thus, even for a sphere, it is impossible to find a bounded mapping from velocity surface fields to corresponding tractions, provided that the same norm is used both in the domain (the velocities) and the range (the tractions). The reason for this is that tractions are related to the stresses, which essentially are derivatives of the velocity field, and differentiation is an unbounded operation; a small variation in a function can give strong peaks in the derivative. One could use different norms in the domain and range, using also the derivatives of the surface velocity in constructing its norm. This, however, seems to lead to some complications with the integral equation approach adopted here.

Because of the unboundedness observed, it can be stated that *it is impossible to have a second-kind equation for the tractions in which allowed surface velocities would be operated on with a bounded operator.* If, however, we are able to formulate the same problem so that we start from given tractions (different from those of the final solution) instead of velocity BCs, it is possible to have well-posed formulations where the final tractions are obtained from a second-kind equation with no unbounded operators involved.

After these notes we are ready to go on to the secondary variable formulations, in which one of the primary variables of physical interest is replaced by another surface field, namely, an aphysical surface density. This approach will allow a well-posed formulation also for problems with velocity BCs, and it will turn out that a specific set of tractions can be found similarly from a well-posed equation. The reason for this is that ill-posedness is related to the formulation used, and in particular to the information used. For example, although numerical differentiation is not a nice procedure, if we have a recursion formula relating consecutive derivatives, we can differentiate fairly accurately. The basic idea is doing some analytical work before doing numerics. *In the well-posed formulation for tractions given later, the actual velocity field is not used explicitly at all.*

Exercises

Exercise 14.1 Neumann Series and Picard Iterations.
The equation $(1 - \mathcal{K})x = b$ has the solution given by the Neumann series

$$x = b + \mathcal{K}b + \mathcal{K}^2 b + \mathcal{K}^3 b + \dots ,$$

provided that the series on the RHS converges. This is analogous to the geometric series

$$1/(1-k) = 1 + k + k^2 + \dots ,$$

which converges if $|k| < 1$; the Neumann series converges for arbitrary b if the spectral radius of $\mathcal{K}$ satisfies $\rho(\mathcal{K}) < 1$. Show that if S_n is the n-th partial sum of the Neumann series, the recursion relation

$$S_{n+1} = b + \mathcal{K}S_n$$

holds. Here $S_n \to x$ as $n \to \infty$, and if we have a good approximation $x_0 \approx x$ to start with, it makes sense to apply the recursion

$$x_{n+1} = b + \mathcal{K}x_n \ , \ \text{for } n = 0, 1, 2, \dots .$$

Thus we are led to Picard iterations.

Exercise 14.2 Systems of Integral Equations.
Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ be matrices, and $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{a}$, and $\boldsymbol{b}$ vectors, such that the operations in the following equation system are well defined:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} &= \boldsymbol{a} \\ \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{y} &= \boldsymbol{b} . \end{aligned}$$

How would you write this as a single matrix equation? A corresponding system of integral equations, with unknown functions $\boldsymbol{x}$ and $\boldsymbol{y}$, can analogously be transformed to a single integral equation; this just requires accepting a piecewise definition of the kernel function.
Note: For this reason systems of integral equations, with unknown "vector functions," are not treated separately in books on integral equations, even though their occurrence in physical models is quite natural.